Quarterly Reviews of Biophysics **46**, 2 (2013), pp. 133–180. © Cambridge University Press 2013
doi:10.1017/S0033583513000012 Printed in the United States of America

Structure of viruses: a short history

Michael G. Rossmann*

Department of Biological Sciences, 240 S. Martin Jischke Drive, Purdue University, West Lafayette, IN 47907, USA

Abstract. This review is a partially personal account of the discovery of virus structure and its implication for virus function. Although I have endeavored to cover all aspects of structural virology and to acknowledge relevant individuals, I know that I have favored taking examples from my own experience in telling this story. I am anxious to apologize to all those who I might have unintentionally offended by omitting their work.

The first knowledge of virus structure was a result of Stanley's studies of tobacco mosaic virus (TMV) and the subsequent X-ray fiber diffraction analysis by Bernal and Fankuchen in the 1930s. At about the same time it became apparent that crystals of small RNA plant and animal viruses could diffract X-rays, demonstrating that viruses must have distinct and unique structures. More advances were made in the 1950s with the realization by Watson and Crick that viruses might have icosahedral symmetry. With the improvement of experimental and computational techniques in the 1970s, it became possible to determine the three-dimensional, near-atomic resolution structures of some small icosahedral plant and animal RNA viruses. It was a great surprise that the protecting capsids of the first virus structures to be determined had the same architecture. The capsid proteins of these viruses all had a 'jelly-roll' fold and, furthermore, the organization of the capsid protein in the virus were similar, suggesting a common ancestral virus from which many of today's viruses have evolved. By this time a more detailed structure of TMV had also been established, but both the architecture and capsid protein fold were quite different to that of the icosahedral viruses. The small icosahedral RNA virus structures were also informative of how and where cellular receptors, anti-viral compounds, and neutralizing antibodies bound to these viruses. However, larger lipid membrane enveloped viruses did not form sufficiently ordered crystals to obtain good X-ray diffraction. Starting in the 1990s, these enveloped viruses were studied by combining cryo-electron microscopy of the whole virus with X-ray crystallography of their protein components. These structures gave information on virus assembly, virus neutralization by antibodies, and virus fusion with and entry into the host cell. The same techniques were also employed in the study of complex bacteriophages that were too large to crystallize. Nevertheless, there still remained many pleomorphic, highly pathogenic viruses that lacked the icosahedral symmetry and homogeneity that had made the earlier structural investigations possible. Currently some of these viruses are starting to be studied by combining X-ray crystallography with cryo-electron tomography.

* Author for correspondence: M. G. Rossmann, Department of Biological Sciences, 240 S. Martin Jischke Drive, Purdue University, West Lafayette, IN 47907-2032, USA.
Tel.: (765) 494-4911; Fax: (765) 496-1189; Email: mr@purdue.edu

1. Introduction

Viruses are the facilitators of evolution. Under their influence genes and genomes are shepherded from one organism to another to produce new gene combinations in search of advantageous developments. Viruses are ubiquitous parasites of animals, plants and bacteria that require a host for reproduction. A virus consists of a genome encapsidated into and protected by a protein or proteolipid shell. When a virus recognizes a susceptible host, the virus or the viral genome enters the host and perverts the host cell into synthesizing new viral components for assembly and release of progeny virions (Leiman *et al.* 2003).

Modern biology was founded on the study of bacterial viruses. Since about 1900, information on virus structures has been growing exponentially, hand-in-hand with the development of experimental and computational techniques along with the growth of molecular biology and virology. This review will describe the development of the technology and ensuing insight gained into the 'life' cycle of viruses and their neutralization by vaccines or anti-viral agents. I have unabashedly used extensively examples from my own experience within the framework of the available knowledge. Space and time dictate that I do not describe or give a complete background of every virus that has been studied structurally (Lee & Johnson, 2003). Rather I have picked and chosen topics that have given me especial excitement and encouragement.

1.1 Tobacco mosaic virus

Although viruses had been defined as filterable infectious material in the late 19th century, information on the size, shape, and properties of these agents did not become known until Wendell M. Stanley (B. S. degree in Chemistry from Earlham College, Richmond, Indiana) had isolated tobacco mosaic virus (TMV) from tobacco leaves in 1935 (Kay, 1986) at the Rockefeller Institute for Medical Research. High concentration of the virus gave rise to birefringence in a

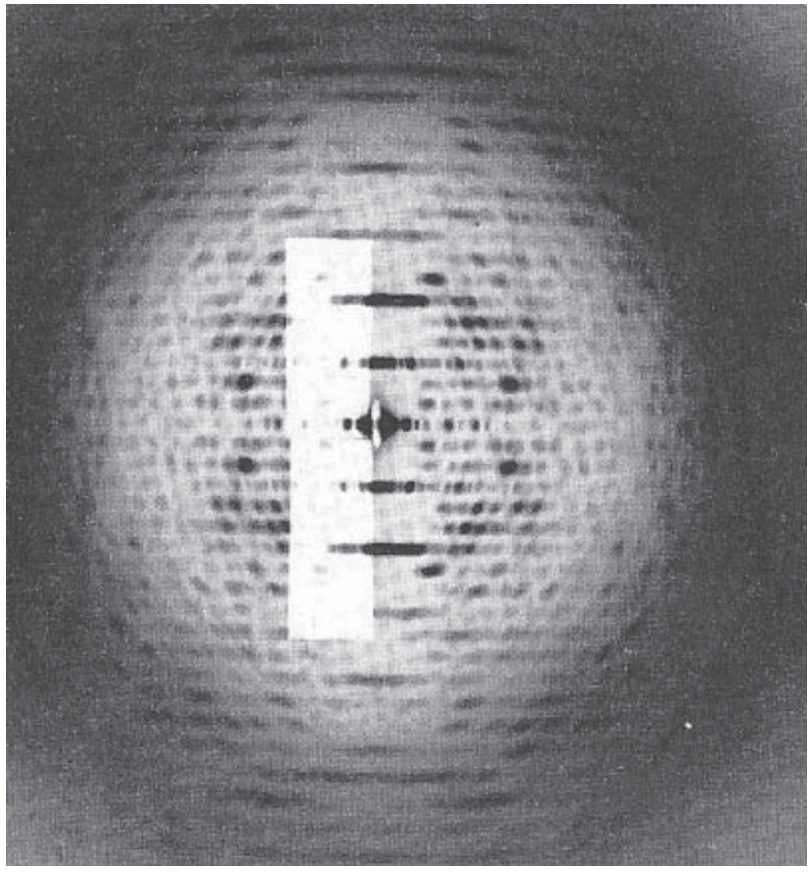

Fig. 1. X-ray diffraction photograph of oriented tobacco mosaic virus gel. Image courtesy of Dr. K.C. Holmes, reproduced from *Diffraction of X-Rays by Proteins, Nucleic Acids and Viruses*, by Herbert R. Wilson (London: Edward Arnold Publishers, Ltd.). Copyright © 1966 Herbert R. Wilson.

polarizing light microscope, indicating the presence of a well-ordered substance. This phenomenon was exploited by Bernal and Fankuchen in Cambridge (Bernal & Fankuchen, 1941) who used X-ray diffraction to examine the TMV suspensions, which enabled them to establish the overall size and shape of the virus. This was the first detailed description of any virus structure.

In 1912 Max von Laue and Peter Ewald, working in München, showed that the X-rays discovered by Röntgen 17 years earlier were waves that could be diffracted by crystals of simple salts. That same year Lawrence Bragg with his father William Bragg, working at Manchester University, formulated a simple expression now known as 'Bragg's Law' to represent the diffraction phenomenon (Bragg & Bragg, 1913) and used this to determine the structure of NaCl and KCl (Bragg, 1913), an achievement that was recognized with a Nobel Prize. This initiated, after the end of World War I, a flurry of structural investigations of inorganic salts and small organic compounds. J. D. Bernal, an Irishman working at Cambridge, was using X-ray diffraction to study crystalline organic compounds of natural products. He and his American student, I. Fankuchen, had heard of Stanley's work and realized that Stanley's TMV suspensions might diffract X-rays. In highly concentrated solutions, the TMV particles were oriented parallel to each other but were randomly rotated about their long axes. Thus, the X-ray diffraction pattern of TMV (Fig. 1) was, in essence, a radially averaged pattern that lacked explicit three-dimensional

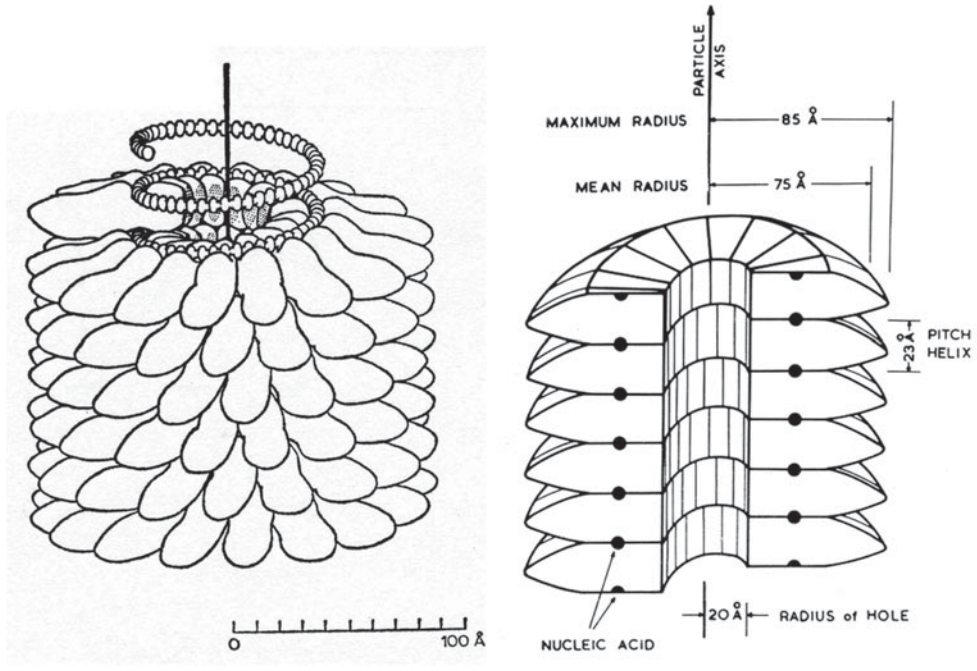

Fig. 2. The structure of tobacco mosaic virus. Reproduced with permission from *X-Ray and crystallographic studies of plant virus preparations*, © 1941 Rockefeller University Press, J. of Gen. Physiol. Vol. 25, pp. 111–165.

information such as is available from X-ray diffraction of fully crystalline material. Nevertheless, the TMV structural investigations established that viruses were homogenous substances with a protein capsid that protected the viral genome, rather than being unseen agents of disease (Bernal & Fankuchen, 1937, 1941).

After the end of World War II, this work was continued by Rosalind Franklin, who showed that TMV had 17–18 protein subunits that covered each turn of the helical RNA genome (Franklin & Holmes, 1958; Franklin *et al.* 1957) (Fig. 2). The TMV analysis was later extended to near-atomic resolution by Holmes and others (Holmes *et al.* 1975; Stubbs *et al.* 1977). Further high-resolution work on crystalline self-assembled discs of TMV protein improved the structural knowledge of TMV (Bloomer *et al.* 1978). The studies of TMV were also among the first to analyze the process of self-assembly of the TMV protein subunits around the RNA genome (Butler & Klug, 1972).

1.2 Crystalline isometric viruses

Early X-ray diffraction studies of crystalline viruses were of RNA plant viruses because it was possible to extract milligrams of virus from a kilogram of leaf material using simple purification procedures based, in part, on Svedberg's centrifugation techniques (Caspar, 1956; Crowfoot & Schmidt, 1945; Finch & Klug, 1959; Huxley & Zubay, 1960; Klug *et al.* 1957; Magdoff, 1960; Nixon & Gibbs, 1960). However, this was mostly in the days before 1957 when even the first protein structures (namely 17 kDa myoglobin (Kendrew *et al.* 1958, 1960) and 64 kDa hemoglobin (Perutz *et al.* 1960) had not yet been determined. The molecular weights of these crystallized RNA viruses were about 5×10^6 Da, or at least 100 times larger than hemoglobin.

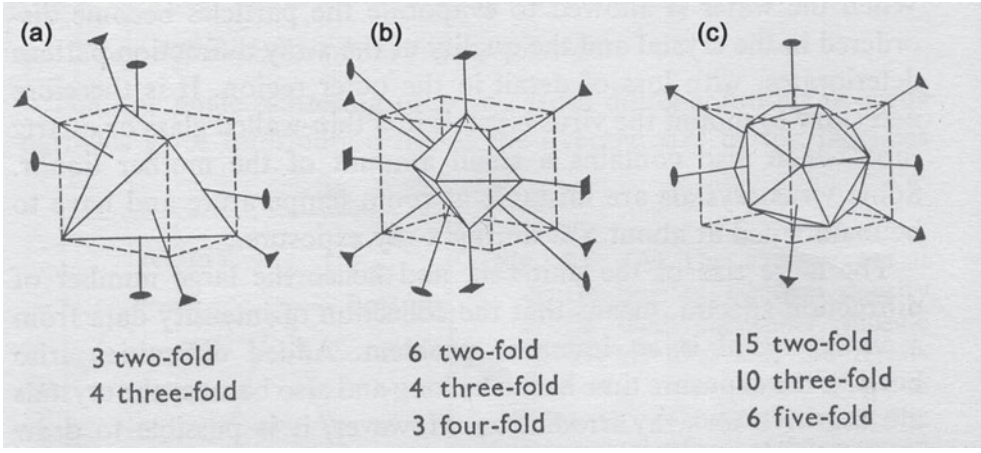

Fig. 3. Diagram showing (*a*) a tetrahedron, (*b*), an octahedron and (*c*) an icosahedron. The number and type of rotation axes of each polyhedron are listed, and some of these are shown in the diagram. Reproduced from *Diffraction of X-Rays by Proteins, Nucleic Acids and Viruses*, by Herbert R. Wilson (London: Edward Arnold Publishers, Ltd.). Copyright © 1966 Herbert R. Wilson.

Some of the major problems in studying crystalline viruses were resolving the individual Bragg reflections because of the large unit cell sizes, measuring the intensities of a million or more reflections usually accomplished by visual inspection of the blackness of spots on a film, having to record all the data by hand in the absence of any electronic computers and no known way to solve the crystallographic phase problem.

These early structural investigations had established that many of these crystalline viruses were roughly spherical with diameters greater than 200 Å and had a protein capsid protecting a cargo of RNA. Then, in 1953, Watson and Crick in Cambridge (UK) discovered the structure of DNA (Watson & Crick, 1953b) and recognized the relationship between the DNA sequence of a gene and the amino acid sequence of a gene product (Watson & Crick, 1953a). This led Crick and Watson (Crick & Watson, 1956, 1957) to suggest that the nucleic acid genome of a virus (such as the RNA of the crystallized yellow mosaic and tomato bushy stunt viruses) would not have the capacity to code for enough protein to cover and protect the genome, but rather that many identical copies of the viral capsid would have to surround the genome. As all the subunits would be identical they would also each have identical environments. Hence, they predicted that the capsid of spherical viruses would have to have the symmetry of regular polyhedra (Fig. 3). This prediction was elegantly verified by Caspar (Caspar, 1956) who showed, in an X-ray study of tomato bushy stunt virus (TBSV) crystals, that the virus had icosahedral symmetry (Fig. 4). These results were further verified by shadow-casting methods using electron microscopy (Horne *et al.* 1959; Williams & Smith, 1958).

Measurements of molecular weights and other data showed that in most spherical viruses the number of protein subunits far exceeded the 60 that would be required for icosahedral symmetry. In 1962, Caspar and Klug (Caspar & Klug, 1962) suggested an extension of the Watson and Crick strict symmetry concept of virus structure by proposing that covalently identical subunits might be able to adapt to slightly different environments, giving rise to a scheme based on the approximate similarity of contacts between six subunits in a hexamer (hexameric 'capsomers' or 'hexones') and five subunits in a pentamer (pentameric 'capsomers' or 'pentons'). The pentones would form the 12 icosahedral vertices, whereas hexons would

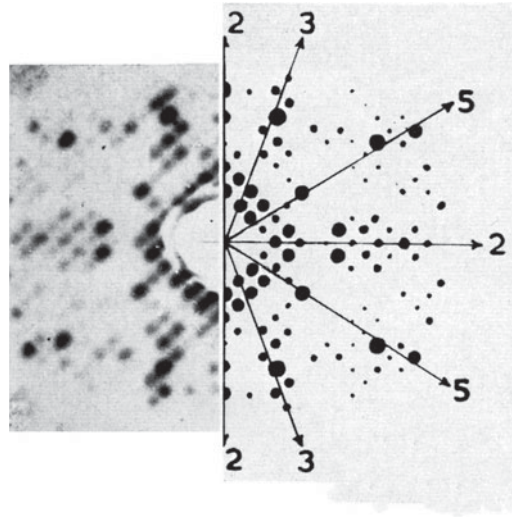

Fig. 4. Tomato bushy stunt X-ray diffraction photograph showing spikes of high intensity in directions that correspond to the angles between symmetry axes of an icosahedron. Reproduced with permission from Caspar (1956).

cover the 20 icosahedral faces. They showed that if T (the 'triangulation number') represents the number of similarly structured subunits per icosahedral asymmetric unit, the possible values of T that would allow each subunit to have a similar environment, would be restricted to $T = h^2 + hk + k^2$. Here h and k count the number of hexons between pentons along the h nd k axes of a hexagonal array. The structure determinations, 16 years later, of TBSV (Harrison *et al.* 1978) and southern bean mosaic virus (SBMV) (Abad-Zapatero *et al.* 1980), the first near-atomic resolution structures of icosahedral viruses, had $T = 3$ quasi-symmetry, consistent with the predictions of Caspar and Klug. Indeed, the predictions of Watson and Crick and then Caspar and Klug were found to be amazingly insightful.

As a result of the early predictive successes, the Caspar and Klug rules became an unbreakable law in the minds of many investigators. Thus, when the structure of polyomavirus (which was thought to have $T = 7$ quasi-symmetry) was found to have pentameric capsomers where there was a hexameric environment (Rayment *et al.* 1982) (Fig. 5), there was consternation and skepticism. The problem was particularly difficult to resolve because the indications of non-Casper and Klug symmetry occurred when the resolution was still poor, making the differentiation between pentameric and hexameric capsomers questionable. However, the preference for pentameric assembly units was supported by the discovery that self-assembled tubes of polyomavirus subunits were formed entirely of pentamers, showing the preference for pentamer formation (Baker *et al.* 1983). Eventually, with the improvement of resolution (Baker *et al.* 1991), it became unquestionably evident that polyomavirus and the structurally similar SV40 (Stehle *et al.* 1996) and papillomaviruses (Modis *et al.* 2002) deviated from the Casper and Klug rules.

Other exceptions to the Casper and Klug rules followed. The internal shell of the dsRNA bluetongue reovirus was found to have two covalently identical subunits in the icosahedral asymmetric unit (Grimes *et al.* 1998). However, $T = 2$ is not allowed (there are no h and k integers that can make $T = 2$ in the above formula for T), implying that the environment of the two subunits cannot be even approximately similar. In fact, the two covalently assembled subunits have some fairly large structural differences after being assembled into the inner shell of the virus.

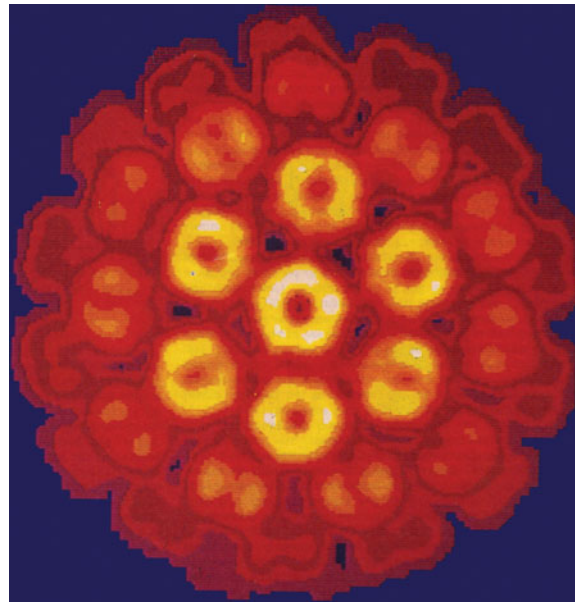

Fig. 5. The structure of polyomavirus at 22 Å resolution. There is one pentameric capsomer at each of 12 vertices and one pentameric capsomer in a general position within the icosahedral environment. However, the latter has a hexagonal environment which produces a different environment for each monomer, in contradiction to the ideas of quasi-equivalence proposed by Caspar & Klug (1962). Reproduced with permission from Rayment *et al.* (1982).

Thus, it is the flexibility of these proteins that circumvents the basic concept of virus assembly, namely that every covalently identical subunit has the same structure. This suggests that the assembly process of viruses might be determined by 'conformational switching' (Kellenberger, 1976). According to this concept, when two proteins bind to each other, they induce a conformational change in each other. Thus, when a third molecule binds to either of the first two molecules, it cannot bind in the same way as do the first two to each other because the original conformations no longer exist. Hence, the path of assembly is controlled by successive events. However, conformational switching does not explain the high symmetry of many viruses.

Notwithstanding the unexpected modification of the Casper and Klug rules, their guidance and insight has been of tremendous benefit for the interpretation of virus structure. Perhaps the strongest confirmation of their proposal to consider the organization of hexameric capsomers as the basis for virus assembly is the structure of the very large dsDNA viruses such as PBCV-1 (Fig. 6). However, even here there is an interesting departure from strict Caspar and Klug logic. The building unit of these viruses is not a hexameric capsomer but only a pseudo-hexameric capsomer, built of three (not six) monomers that each have an approximate repeat structure of a 'jelly-roll' fold along their polypeptide giving six quasi-equivalent jelly-rolls per capsomer (see below). Thus, for these structures, the *T* number represents the number of jelly-rolls per icosahedral asymmetric unit.

Yet another breakdown of the Caspar and Klug rules was found when studying flaviviruses, where there are three heterodimers of the E ('envelope') and M ('membrane') proteins in the icosahedral asymmetric unit. They assemble into dimeric rafts containing six anti-parallel heterodimers producing a herringbone-like pattern. Possibly these dimers are stable intermediates just as are the pentamers in the assembly of polyomavirus discussed above. However, the

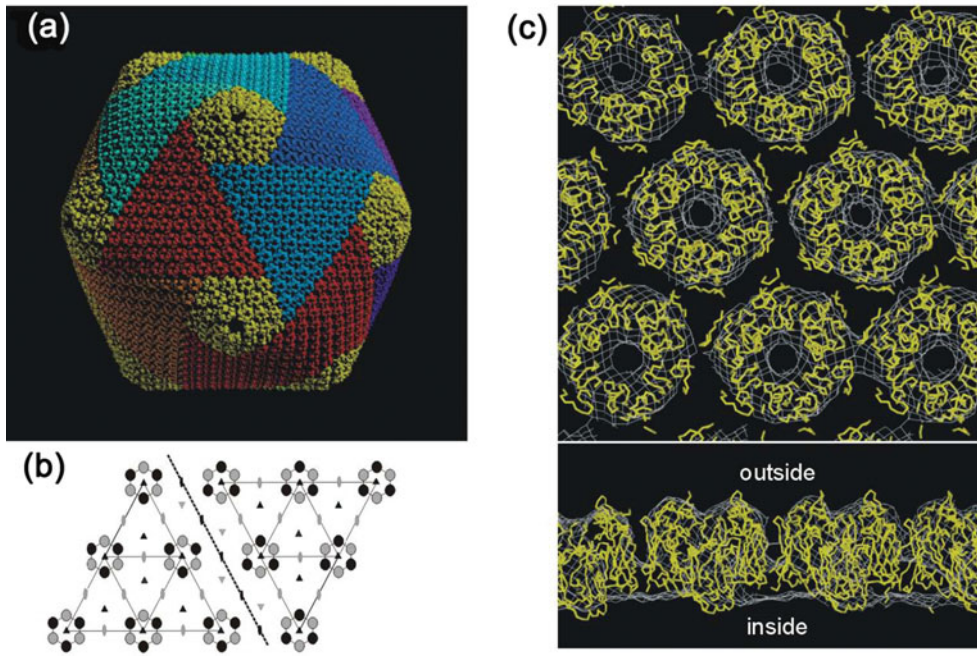

Fig. 6. The structure of PBCV1 (Nandhagopal *et al.* 2002; Simpson *et al.* 2003; Yan *et al.* 2000). (*a*) CryoEM reconstruction with each trisymmetron shown with a different color. (*b*) Diagrammatic representation of the boundary between the red and blue trisymmetrons. Each capsomer is a trimer of a double jelly-roll structure giving a hexameric appearance. (*c*) Fit of the capsomer's crystal structure into the cryoEM density. (top) view from outside the virus. (bottom) side view. Reproduced from Nandhagopal *et al.* (2002).

immature flavivirus assembles into approximate trimers that change reversibly into the herring-bone pattern which makes it difficult to assume that the rafts are especially stable. Publication of the first flavivirus structure (Kuhn *et al.* 2002) was much delayed, in part because the referees were not convinced of the correctness of the technique by which the known structure of the dimeric E protein of the homologous tick-borne encephalitis virus were fitted into the dengue virus cryo-electron microscopy (cryoEM) density map to produce a structure that did not obey Casper and Klug rules.

1.3 X-ray crystallography

The procedures required to determine the structures of icosahedral viruses by X-ray crystallography were hindered by an absence of satisfactory techniques to resolve individual Bragg reflections from each other for unit cells that had dimensions of 500 Å or more, obtaining enough X-ray intensity to observe the diffracted rays when the energy of the diffracted beam was about 100 times weaker than when studying protein crystals (Rossmann, 1984, 1999), automation of collecting intensity data for millions of reflections (Rossmann, 1979; Rossmann *et al.* 1979), determining the orientation of the virions in the crystal unit cell (Rossmann & Blow, 1962; Tong & Rossmann, 1990) (Section 2.1), and solving the phase problem. The last of these was perhaps the most significant barrier to the structure determination of a virus. It was the realization that this problem could be aided by the naturally occurring symmetry of the virus, which made virus structure determinations possible (Rossmann & Blow, 1962).

Whenever an icosahedral virus forms a crystal there has to be at least 5-fold non-crystallographic symmetry (NCS), giving rise to 5-fold 'redundancy', because a 5-fold axis cannot be introduced into an infinite repeating lattice. However, if the whole virus forms the crystallographic asymmetric unit, as is often the case, then there is 60-fold redundancy. In 1962 David Blow and I (Rossmann & Blow, 1962) suggested that the NCS redundancy could be used to solve the crystallographic phase problem, a topic that was central to my interests for many years (Rossmann, 1972). The use of NCS averaging for phase improvement (as opposed to phase extension beyond the resolution of an initial phase resolution) was more quickly recognized as a powerful tool (Buehner *et al.* 1974; Champness *et al.* 1976). Such phase improvement was essential for determining the structure of TBSV (Harrison *et al.* 1978) and SBMV (Abad-Zapatero *et al.* 1980) where initial phases had been determined by isomorphous replacement. In these and other structure determinations that depend on NCS averaging, the phases obtained by Fourier transforming the improved averaged map are applied to the structure amplitudes for the next cycle. In those days the 60-fold averaging of a small virus was still a major computing operation. Thus, these early structure determinations of icosahedral viruses had to be restricted to just one cycle of density averaging. However, as computers improved, it became possible to use density averaging for cycles of iterative phase improvement in which the successive cycles produced smaller and smaller phase changes to eventually converge to a final value.

Gaykema *et al.* (1986) were probably the first to use NCS averaging for phase extension in the structure determination of a crystal that had two hemocyanin molecules in the asymmetric unit although the extension was only from 4·0 to 3·2 Å resolution. These results had been reported at meetings but were not published until 1986. With this encouragement, in 1985 phases we extended phasing knowledge from 6·0 to 3·0 Å resolution in the structure determination of a human common cold virus using the NCS redundancy of 20 (Rossmann *et al.* 1985). The power of NCS averaging combined with solvent flattening in the space between virions had not been fully appreciated until we showed that the quality of the 3 Å resolution rhinovirus map was far better than all other previous structure determinations available to that resolution at that time. Then, In 1991 we showed that phases could be extended from 8·0 to 3·25 Å resolution in the structure determination of canine parvovirus (Tsao *et al.* 1991). Soon other structures were determined using low-resolution cryoEM image reconstructions as a starting model for phase determination (McKenna *et al.* 1992b). However, using a simple spherical shell as an initial model was more difficult as a sphere has a center of symmetry making the initial phase solution centric with the desired structure superimposed on its enantiomorph (Chapman *et al.* 1992; Tsao *et al.* 1992).

The theory for phase extension is best expressed in terms of reciprocal space (Main & Rossmann, 1966; Rossmann, 1989b, 2001; Rossmann *et al.* 1992), which shows that averaging is equivalent to determining the new structure factor, $\mathbf{F_h}$, from

$$\mathbf{F_h} = (1/V)\Sigma_n\Sigma_\mathbf{p}\mathbf{F_p}G_{hp}, \tag{1}$$

where the diffraction function is $G_{hp} = \Sigma_n \int_{Un} [\exp\{2\pi i(\mathbf{h} - [C_n]^T\mathbf{p}) \cdot \mathbf{x}\}]d\mathbf{x}$ integrated over the volume U_n within which the NCS is valid. The points x and x_n are related by the rotation matrix $[C_n]$ in the nth non-crystallographic asymmetric unit such that

$$\mathbf{x}_n = [C_n]\mathbf{x}$$

Equation (1) shows the inter-relationship between the known phases of structure factors, $\mathbf{F_p}$, with the unknown phases immediately outside the previous resolution limit, $\mathbf{F_h}$. For those close

to the previous resolution limit, G_{hp} is only slightly smaller than 1·0, for all other structure factors further away from the current resolution limit, G_{hp} is close to 0·0. Thus it is possible to extend approximate phasing to a resolution that is slightly better than the current limit of resolution. Furthermore, this process can then be repeated after each extension step as we had demonstrated in the structure determination of human rhino virus 14 (Rossmann *et al.* 1985).

1.4 Electron microscopy

Electron microscopes started to be developed in the 1930s and were used for recognizing virus shapes, sizes, and interaction with the host (Brenner & Horne, 1959). Starting in about 1970, computational techniques were developed to use negatively stained images of virus particles to produce three-dimensional images of viruses and, in particular, of RNA plant viruses (Crowther *et al.* 1970). However, these reconstructions were hindered by unequal staining of the specimen as well as distortions produced by the staining. Furthermore, the reconstructions showed only the stained surface features. The situation was radically changed when, in the mid 1980s, Dubochet introduced flash freezing at liquid nitrogen temperatures to embed the sample in vitreous ice (Adrian *et al.* 1984; Dubochet *et al.* 1988). This technique produced essentially undistorted projected images of the sample when viewed in an electron microscope. This cryoEM technique has now come to rival crystallography (Harrison, 2010; Liu *et al.* 2010; Reddy *et al.* 2010; Zhang *et al.* 2012). Today cryoEM, in favorable cases, can produce high-resolution electron density maps of viruses with 3·5 Å resolution. Most importantly, cryoEM can be used where crystals are unlikely to become available, such as viral complexes with antibodies (Hafenstein *et al.* 2009; Llamas-Saiz *et al.* 1997; Rossmann, 1994) or with receptor molecules (Olson *et al.* 1993; Pokidysheva *et al.* 2006), or where the large virus particles cannot pack easily into crystalline arrays (Fokine *et al.* 2004; Tao *et al.* 1998). A significant example of the latter is the structure determination of West Nile virus to a resolution of 9 Å using cryoEM (Zhang *et al.* 2003a), when the best formed and most beautifully shaped crystals of West Nile virus never diffracted X-rays to better than about 16 Å resolution (Kaufmann *et al.* 2010b). Similar attempts to obtain reasonably well-diffracting crystals of Sindbis virus (an alphavirus) were equally unsuccessful (Harrison *et al.* 1992). However, a cryoEM reconstruction of Sindbis virus eventually achieved 7 Å resolution (Tang *et al.* 2011) and an even better resolution was obtained in the study of other alphaviruses (Sun *et al.* 2013; Zhang *et al.* 2011a). The reason for the discrepancy is because the lipid envelope in West Nile virus, as also in many other viruses, is too flexible to produce truly homogenous particles. Thus, crystals will contain particles with a mixture of different conformations, whereas in a cryoEM reconstruction, particles can be individually selected according to their similarity in shape and structure.

The variation in structure in otherwise icosahedral viruses is more extreme in many pleomorphic virus pathogens (e.g. HIV, para- and orthomyxoviruses, bunyaviruses, coronaviruses, influenza viruses, and many more). These do not lend themselves to a cryoEM reconstruction because of their variability. The method of choice to study the three-dimensional structure of these viruses is cryo-electron tomography (cryoET) (Battisti *et al.* 2012; Rossmann *et al.* 2011). In this procedure, the stage is tilted to give a sequence of projections (often at 1° intervals) in different orientations that are then combined to form a three-dimensional image. However, because of radiation damage, the exposure for each projection has to be very limited to minimize the damage before the last exposure. As a result the resolution is

limited and seldom better than maybe 40 Å. Nevertheless, the structure of each particle can be compared locally to improve the resolution and to fit with the crystallographic structure of the appropriate viral component.

The remainder of this review will describe the structures of selected viruses and discuss how the structure determines their various life styles.

2. Small icosahedral non-enveloped viruses: the jelly-roll fold

2.1 RNA plant viruses

By the time enough knowledge had been gained to consider determining the structure of a small spherical virus, Arndt & Wonacott, (1977) had re-introduced the oscillation method of collecting diffraction data on single crystals. Powerful rotating anode generators had to be husbanded and the X-ray beam had to be focused by glancing reflections off silvered mirrors (Harrison, 1968) to resolve individual Bragg reflections and to gain intensity for crystals with unit cells $\sim$50 times larger than anything considered previously. Thus, reflections were about 50 times weaker than for any other previously attempted three-dimensional data collections. Instruments such as the Optronics scanner had become available for digitizing film data. However, computer programs had to be written that had to differentiate between full and partial reflections requiring accurate knowledge of crystal orientations (Rossmann, 1979). Finally, it was necessary to extend and improve on earlier procedures of scaling data where each diffraction pattern had been collected on a different crystal (Rossmann *et al.* 1979). The average type of 0·3° oscillation angle photograph took about 48 h exposure on the available rotating anode X-ray generator. Collecting one, say 2·5 Å resolution, data set for a trial heavy atom compound took somewhere between three to six months. Given that it was thought to be necessary to collect four to seven heavy-atom data sets, the structure determination of the first crystallized icosahedral viruses was expected to take several years with no guarantee of success.

The first step in determining the structure of a crystallized, icosahedral virus is to establish the orientation of the virus or viruses in the crystallographic asymmetric unit relative to the axial directions of the crystal lattice by means of a rotation function (Rossmann & Blow, 1962; Tong & Rossmann, 1990). Although this is now the standard procedure, the first real test of using a rotation function to determine the orientation of a virus ran into a problem (Åkervall *et al.* 1972). Satellite tobacco necrosis virus (STNV) is a small $T=1$ positive-stranded RNA virus (Liljas *et al.* 1982), that crystallized in a monoclinic space group. It turned out that one of the icosahedral 2-fold axes of the virion was oriented at $\sim$45° relative to the crystallographic 2-fold axis. Thus the second virus in the unit cell, generated by the crystallographic 2-fold axis, was related to the first virus by a 90° rotation. As a result there were strong 4-fold peaks arranged as in an octahedron. The initial conclusion (Åkervall *et al.* 1972) was that the virus was octahedral, supported by biochemical evidence regarding the number of labeled cysteine groups and molecular weight determinations. Thanks to aggressive prompting by Klug (1972), the special situation of the rotation function was recognized, drawing attention to the lower, but nevertheless present, icosahedrally arranged rotation function peaks (Åkervall *et al.* 1972). These peaks were dwarfed (and therefore initially overlooked) by the larger octahedral peaks created by the special packing of the virus in the crystal lattice. Although scorn was plentiful in regard to the apparent failure of the rotation function calculations in the analysis of STNV, it is now the basic tool in all analyses of spherical viruses (e.g. Muckelbauer *et al.* 1995; Plevka *et al.* 2012c). One improvement is the

'locked' rotation function (Rossmann *et al.* 1972) in which the presence of all symmetry axes is tested simultaneously for a particular particle orientation, rather looking for each axis independently. This procedure has been elegantly implemented by Liang Tong in his 'General Locked Rotation Function' (GLRF) program (Tong & Rossmann, 1990). However, difficulties do still occur as happened in the determination of the enterovirus 71 (a picornavirus) where there were two independent particles per crystallographic asymmetric unit. Confusion occurred when the trial icosahedral symmetry axes were oriented with one of the icosahedral 2-fold axes parallel to an orthorhombic crystallographic 2-fold axis, resulting in a large but erroneous, locked rotation function peak. This was avoided (Plevka *et al.* 2012a) by testing only the 3- and 5-fold NCS axes when exploring different orientations of the set of icosahedral axes in a locked rotation function search (Rossmann *et al.* 1972; Tong & Rossmann, 1990).

The early icosahedral virus structure determinations depended on heavy atom isomorphous replacement to determine initial low resolution phases. The position of the heavy atoms was determined by using prior determination of the orientation of the viral symmetry axes. A heavy atom test position would generate 60 symmetry related atoms and, in turn, these would generate 60×59 vectors in a $(F_{PH}–F_P)^2$ difference Patterson. Such a cluster of peaks could be checked against the observed difference Patterson for every incremental position of the test heavy-atom site. This would, therefore, establish the position of heavy atoms relative to the icosahedral symmetry sites (Argos & Rossmann, 1974). Given this result, further tests on the cross vectors between the crystallographic symmetrically related heavy atoms were then used to determine the position of the virion in the crystallographic asymmetric unit. The resultant, rather poor, isomorphous replacement phases were significantly improved by icosahedral averaging. In the early icosahedral virus structure determinations, no attempts were made at iterative cycles of phase improvement because of the enormous computing operation, as measured by the standards of the time. Nor was there any attempt at phase extension as such procedures were thought by many to be impossible even if the computing problems could be overcome.

The first icosahedral virus structure to be determined was that of TBSV (Harrison *et al.* 1978). Of major interest was how the three independent subunits (A, B, and C) in the icosahedral asymmetric unit had adapted themselves to their different quasi-equivalent environments. The three subunits had, indeed, organized themselves around a quasi-3-fold axis as predicted by Casper and Klug. The virion assembly probably occurred via the formation of dimers. However, in the assembled particles the AB dimer (Fig. 7) had a disordered amino-terminal β-strand that was ordered in the CC dimer, giving the two types of dimers a different curvature on the viral surface.

The second icosahedral virus structure to be determined was that of SBMV (Abad-Zapatero *et al.* 1980). The great surprise was that it had a structure closely similar to that of TBSV, missing only the dimeric protruding domains associated with each of the AB and CC dimers. Thus, these two virus capsid structures must have diverged from a common ancestral fold, later to be called a 'jelly-roll' fold (Fig. 8). It also became apparent there was no recognizable sequence similarity between the similarly folded TBSV and SBMV capsid subunits. This further reinforced the earlier suggestions derived from the globin structures (Perutz *et al.* 1960) and nucleotide binding structures (Rossmann *et al.* 1974) that three-dimensional tertiary protein structures are more conserved than primary structural amino acid sequences. That is, the three-dimensional structure, but not necessarily the specific amino acids, plays an essential role in determining function.

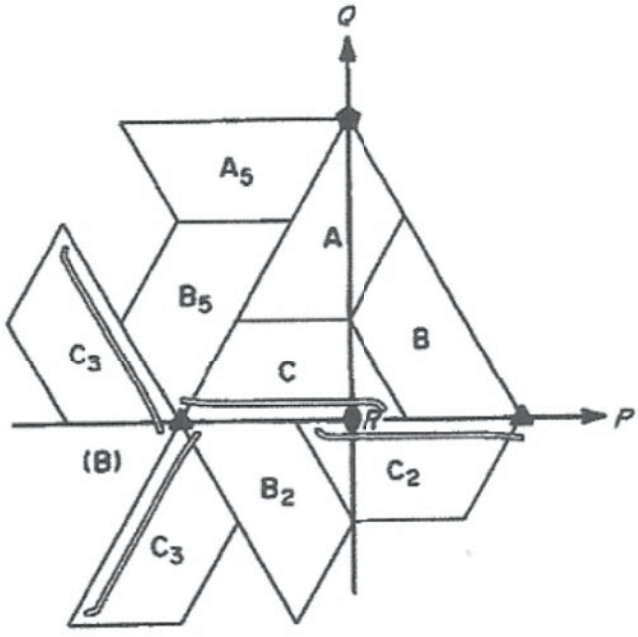

Fig. 7. Diagrammatic representation of the spatial distribution of the three quasi-equivalent subunits A, B, and C in the $T=3$ surface lattice of TBSV and SBMV. The triangle represents one icosahedral asymmetric unit. Reproduced from Hermodson *et al.* (1982).

2.2 RNA animal viruses

It took another 6 years after the determination of the first RNA plant virus structures to determine the first RNA animal virus structures. To study these viruses required the use of cell culture on which the viruses could be propagated, but with much lower yields than had been possible for the RNA plant viruses. The viruses of choice were the small RNA picornaviruses (pico = small, rna) such as the rhino- (common cold), polio, and foot and mouth disease viruses for which there was available a great deal of information. Fortunately, technology had advanced allowing the use of synchrotron radiation. This greatly speeded up and improved the quality of the data collection process. It also created a revolution in how diffraction data were collected not only for viruses but also for all crystals of biological samples by bringing the 'American Method' (shoot first, think later) from the Wild West to the sanctity of the synchrotron (Rossmann & Erickson, 1983). The high intensity of the synchrotron X-ray beam quickly damaged the crystal, making it difficult or impossible to 'set' the crystal into a known orientation relative to the camera axes by a series of short exposures before the crystal was dead. Thus, a good data 'shot' had to be taken right away without first 'setting' the crystal into a known orientation and thinking about how the crystal was set had to be done later based on the resultant diffraction pattern.

Another critically important innovation introduced in the study of rhinoviruses was the process of phase extension (see Eq. (1) above). This limited the need for an initial phase set to only low resolution. In the case of the human rhinovirus serotype 14 determination, an Au derivative was used to obtain the initial phases to 6 Å resolution. These were then extended to 3·5 Å resolution in gradual steps using icosahedral averaging at each iteration to improve and extend the phasing (Arnold *et al.* 1987; Rossmann *et al.* 1985), resulting in an easily interpretable electron

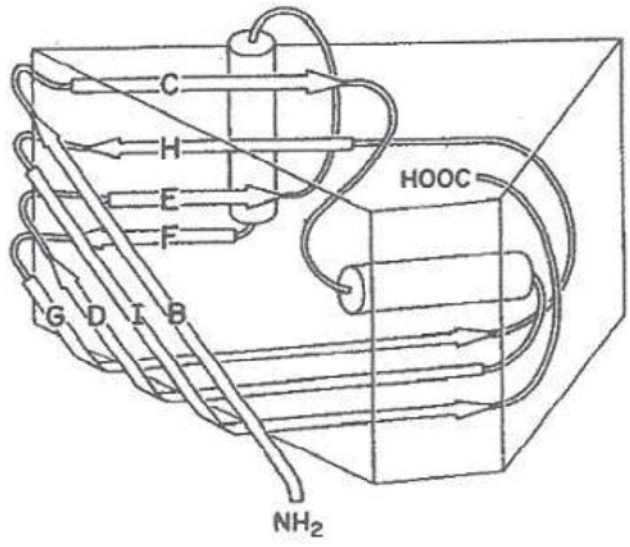

Fig. 8. The jelly-roll fold. There are two anti-parallel, four stranded β-sheets that enclose a β-barrel. If the sequence of β-strands in the polypeptide are B, C, …, I, then one sheet is composed of the BIDG strands and the other of the CHEF strands. Strand A exists only in the jelly-roll structure of satellite tobacco necrosis virus (Liljas *et al.* 1982). Reproduced with permission from Hogle *et al.* (1985).

density map. This technological success also finally demonstrated that NCS is a powerful and useful tool for phase determination, an ambition first enunciated by me in 1962 (Rossmann & Blow, 1962). Much has been written since that time about the use of NCS for phase determination (Arnold *et al.* 2012; Plevka *et al.* 2011b).

Picornaviruses are small, icosahedral animal viruses of about 300 Å diameter with a positive sense $\sim$8000 base RNA genome. The capsid consists of 60 copies of the four viral proteins VP1, VP2, VP3, and VP4. The first three of these proteins have a molecular weight of about 32 kDa, whereas the small $\sim$5 kDa VP4 is internal in the mature virus. The big surprise that came out of the rhinovirus structure determination was that the rhinovirus structure was closely similar to the RNA, $T=3$ plant virus structures of TBSV and SBMV. However, instead of having three covalently identical subunits per icosahedral asymmetric unit as in the small icosahedral RNA plant virus structures, the VP1, VP2, and VP3 rhinovirus proteins, although having no obvious amino acid sequence similarity, all had a similar jelly-roll fold such as the RNA plant viruses. Furthermore, the rhinovirus proteins were arranged with similar quaternary structure as the RNA plant virus producing a pseudo $T=3$ (or $P=3$) structure (Fig. 9). The structure of poliovirus followed a few months later (Hogle *et al.* 1985) as did mengovirus (Luo *et al.* 1987) and foot and mouth disease virus (Acharya *et al.* 1989) in the following years. These were of different picornavirus genera showing the evolutionary divergence of picornaviruses from a common ancestral virus.

Another significant observation provided by the rhinovirus structure was a deep canyon running around each pentameric vertex. Rueckert and Sherry (Sherry & Rueckert, 1985; Sherry *et al.* 1986) had studied a series of neutralizing antibodies and determined the amino acids on the virus that, when mutated, provided an escape from neutralization. These escape mutations, when plotted onto the three-dimensional structure of the rhinovirus, were all on the viral surface outside the canyon (Rossmann *et al.* 1985). It was suggested that the canyon was too narrow to

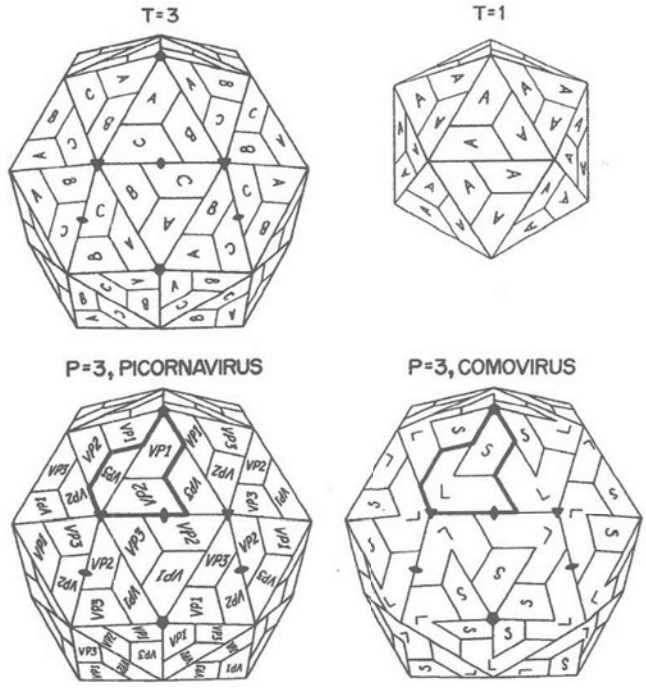

Fig. 9. Diagrammatic representation of quasi-icosahedral symmetry in different virus structures. The $T=1$ shell (top right) contains 60 subunits related by icosahedral symmetry. Each subunit is represented by a trapezoid that is the approximate shape of a jelly-roll β-barrel. All subunits in the $T=1$ capsid are identical and labeled A for comparison with the $T=3$ capsid. The asymmetric unit of the $T=3$ capsid (top left) contains subunits A, B, and C. The subunits labeled A, B, and C have the same amino acid sequence but are in slightly different environments. While similar to the triangle in the $T=1$ structure, the 3-fold axis relating A, B, and C is not exact. This quasi-3-fold axis relates the quasi-6-fold axes (left and right vertexes of the triangle) to the 5-fold axis (top vertex). Like the $T=1$ structure, the $T=3$ structures are formed by identical subunits with the jelly-roll β-barrel fold. The $P=3$ picornavirus shell (bottom left), technically a $T=1$ particle, is closely related to the $T=3$ shell. However, the three subunits in the central triangle (labeled VP1, VP2, and VP3) have different amino acid sequences. The bold outline represents the periphery of one protomer that is post translationally cleaved into its three components. In the comoviruses these cleavages are incomplete thus making the large subunit into a double jelly-roll fold that generates a pseudo-hexameric capsomer around the icosahedral 3-fold axes. Reproduced from Rossmann & Johnson (1989).

permit the binding of antibodies into the canyon and, indeed, the antibody binding sites, as mapped by the escape mutations, were all outside the canyon. In addition, the amino acids that lined the walls and floor of the canyon were more conserved among rhino- and polioviruses (Rossmann & Palmenberg, 1988). This gave rise to the 'canyon hypothesis' (Rossmann, 1989a; Rossmann *et al.* 1985) (Fig. 10) that suggested the cellular receptor would be narrow enough to allow it to bind into the canyon where the mutation rate is small, while the surface of the virus varied more quickly, thereby escaping from the host's immune system. A few years later it was found that the cellular receptor for the major group of rhinoviruses was intercellular adhesion molecule 1 (ICAM-1) (Greve *et al.* 1989; Staunton *et al.* 1989). The predicted binding site of the receptor in the canyon was verified in a cryoEM study of HRV14 incubated with the soluble component of ICAM-1 (Fig. 11) (Olson *et al.* 1993). However, the reasoning for the canyon hypothesis was later challenged when the structure of some Fab fragments bound to HRV14 were found to bind partially into the canyon (Smith *et al.* 1996), although the neutralizing

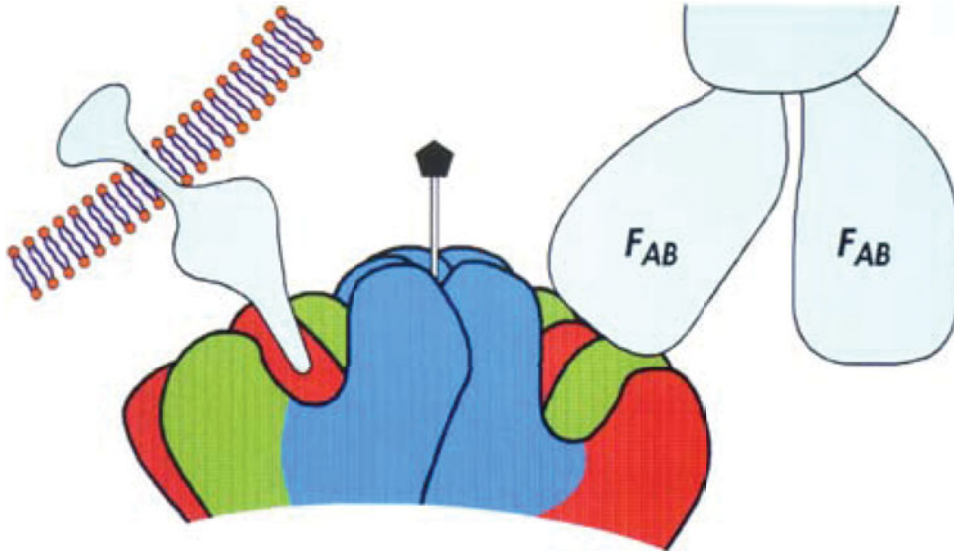

Fig. 10. The canyon hypothesis. The surface of human rhinovirus 14 was found to have a deep 'canyon' circulating around each of the 5-fold vertices. The sites for neutralizing antibodies were all outside the canyon. This, and the known size of Fab fragments, suggested that the antibodies were sterically hindered from reaching the floor of the canyon. Thus the virus could achieve protection from the host's immune surveillance and could continue binding to its usual receptor. Modified from Rossmann (1989a).

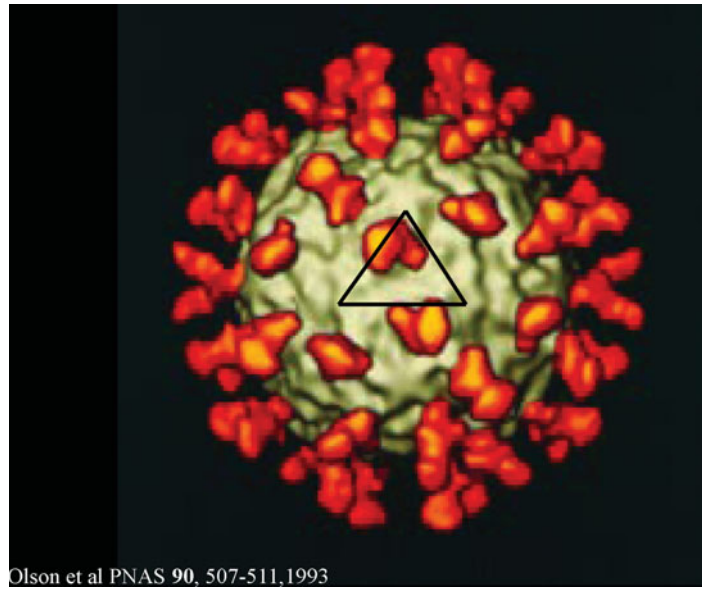

Fig. 11. Human rhinovirus 14 (gray) complexed with intercellular adhesion molecule 1 (ICAM1) (orange). The black triangle marks the boundary of the icosahedral asymmetric unit. The ICAM1 receptor molecule was found to bind to the canyon as had been predicted. Similar results were obtained for the poliovirus and coxsackie B3 virus receptors. Reproduced from Olson *et al.* (1993).

monoclonal antibody work had clearly demonstrated that the important residues involved in receptor binding were on the virus surface and not in the canyon. Furthermore, the structures of polio- and coxsackieviruses also had deep canyons and their receptor binding interface was also deep in the canyon (He *et al.* 2000, 2001; Zhang *et al.* 2008). Conversely, the recent structure of enterovirus 71 was shown to have a rather shallow canyon and the receptor binding site is probably outside the canyon (Plevka *et al.* 2012b; Wang *et al.* 2012).

The rhinovirus structure also showed that a hydrophobic pocket within VP1, underneath the canyon floor, could be occupied by anti-viral compounds that inhibited attachment and uncoating on entrance into a cell (Smith *et al.* 1986). These compounds displaced a 'pocket factor' probably a mixture of fatty acids (Filman *et al.* 1989; Oliveira *et al.* 1993; Rossmann, 1994; Smyth *et al.* 1995), stabilizing the virus by filling the hydrophobic pocket with a hydrophobic compound and altering the structure of the canyon floor that forms the receptor binding site. These compounds were developed into a potent anti-common cold virus drug (Hayden *et al.* 2003; Xiao *et al.* 2011), although they were not licensed by the FDA primarily because of significant side effects related to women taking oral contraceptives (Xiao *et al.* 2011).

Since the early work on picornavirus structure, there has been extensive work on receptor recognition (Bella *et al.* 1998, 1999; He *et al.* 2000; Kolatkar *et al.* 1999; Rossmann *et al.* 2000, 2002; Xiao *et al.* 2001; Tuthill *et al.* 2010) and, especially for poliovirus, on the series of structural changes that occur when the virus infects a cell using both crystallographic and cryoEM techniques (Bostina *et al.* 2007; He *et al.* 2000; Hogle, 2002; Zhang *et al.* 2008; Bostina *et al.* 2011).

2.3 ssDNA parvoviruses

Parvoviruses infect vertebrates, insects and crustaceans. In vertebrates, autonomous parvoviruses replicate only in dividing cells and, hence, have a predilection for rather young or unborn animals or for tissues of older animals that contain proliferating cells. Parvoviruses have an about 5000 base, ssDNA genome. They have a $T=1$ icosahedral capsid that is about 260 Å in diameter. Most mammalian parvoviruses have three types of polypeptides: VP1, VP2, and VP3. There are 60 (VP1 + VP2 + VP3) subunits in the viral capsid. VP1 and VP2 are formed by alternative splicing of the messenger RNA from the viral DNA. VP3 is formed by cleavage of 15–20 amino acids from the amino terminus of VP2. The amino terminus of VP1 forms a phospholipase (Canaan *et al.* 2004; Dorsch *et al.* 2002) that is externalized from the capsid and is required for the virus to enter and/or exit a cell.

The structures of human B19 virus-like particles (Kaufmann *et al.* 2004, 2005, 2008), canine (Simpson *et al.* 2000a; Tsao *et al.* 1991, 1992; Wu & Rossmann, 1993; Xie & Chapman, 1996), feline (Agbandje *et al.* 1993; Simpson *et al.* 2000a), porcine virus-like particles (Simpson *et al.* 2002), and murine (Agbandje-McKenna *et al.* 1998) parvoviruses have all been determined using X-ray crystallography to near atomic resolution. In addition to the autonomous parvovirus structures, extensive investigations have been made on dependo-parvoviruses such as various adeno-associated parvoviruses (AAV) using both cryoEM reconstructions (DiMattia *et al.* 2012; Hueffer *et al.* 2003; Padron *et al.* 2005; Walters *et al.* 2004) and X-ray crystallography (DiMattia *et al.* 2012; Xie *et al.* 2002). AAVs can only be grown in the presence of adenovirus. They have been extensively studied on account of their potential for use in gene therapy (Lerch *et al.* 2012; Padron *et al.* 2005).

The first parvovirus structure to be determined (Tsao *et al.* 1991) was that of canine parvovirus. The initial solution was based using a hollow sphere as a starting model to determine

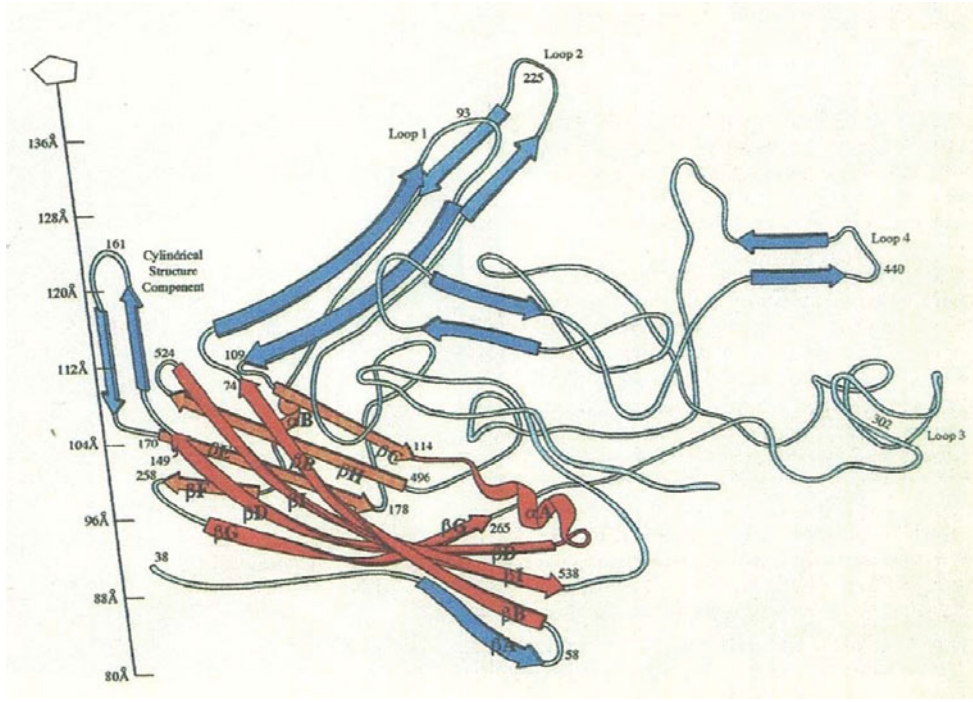

Fig. 12. The canine parvovirus major capsid protein structure. The core jelly-roll structure is in red. The large insertions are in blue. Reproduced from Tsao *et al.* (1991).

phases to 20 Å resolution. This was sufficient to find the position of heavy-atom sites of a Pt derivative. These provided 8 Å resolution single isomorphous replacement phasing to initiate phase extension by virtue of the 60-fold NCS redundancy (Chapman *et al.* 1992; Tsao *et al.* 1992).

The capsid protein of parvoviruses has a jelly-roll structure (Fig. 12) with large insertions that form much of the external surface of the virus and the contacts between neighboring subunits. The amino end of VP1 and VP2 are exposed by way of a narrow channel along the 5-fold axes of the virion. The residues in this channel are a conserved sequence of small, mainly glycine, residues (Chapman & Rossmann, 1996; Tsao *et al.* 1991). About 13% of the ssDNA genome of canine parvovirus and 33% of the genome of murine parvovirus had an icosahedrally ordered secondary structure. Partially ordered RNA genomes have also been found in bean pod mottle virus (Chen *et al.* 1989) and, in particular, in STMV (Larson *et al.* 1993).

Although human B19 VP2 has only 26% sequence identity to VP3 of AAV-2, 72% of the Cα atoms can be aligned structurally with a root-mean-square deviation of 1·8 Å. A total of 28% of the structurally aligned residues are identical between B19 and AAV-2. Both viruses require an integrin as a co-receptor and both depend on human cells for replication, suggesting a common receptor-binding region (Kaufmann *et al.* 2004). Among the available parvovirus structures the greatest similarity is between canine and feline parvoviruses. The only nucleotide changes between feline and canine parvoviruses that affect host range correspond to residues 93, 103, and 323 of VP2. These residues are on the loops, which create small spikes on the virus surface around the icosahedral 3-fold axes and are associated with the largest structural differences between these viruses. Thus, if host range is associated with receptor interaction, then the

receptor attachment site for these viruses will straddle these residues (Agbandje *et al.* 1993). This is also consistent with data for murine parvovirus (Agbandje-McKenna *et al.* 1998). Canine and feline parvoviruses can use transferrin receptor (TfR) for binding and infection (Hueffer *et al.* 2003). The site of TfR binding has been determined (Hafenstein *et al.* 2007) and is roughly consistent with the previous observations. However, the TfR molecule binds only to one or a few of the 60 equivalent sites on the virus indicating that receptor binding induces asymmetry in the virus. Asymmetric interactions of icosahedral viruses with their hosts might be a more common phenomenon than previously thought (Cherrier *et al.* 2009; Klose *et al.* 2010; Kuznetsov *et al.* 2010; Zhang *et al.* 2011b) and may have been obscured by averaging in previous crystallographic and electron microscopic structure determinations.

Many anthropod parvoviruses (densoviruses) are highly pathogenic and kill ~90% of the host larvae within days, making them potentially effective as selective pesticides. In densoviruses, there are four different initiation sites for translation of the densovirus capsid protein mRNA, giving rise to the viral proteins VP1 to VP4. A total of 60 copies of the C-terminal jelly-roll domain make up the ordered part of the icosahedral capsid. Although denso-parvoviruses have a capsid protein with a jelly-roll fold like mammalian parvoviruses, the position and extent of the large insertions are quite different (Kaufmann *et al.* 2010a, 2011; Simpson *et al.* 1998). In contrast to vertebrate parvoviruses, the N-terminal β-strand of invertebrate parvoviruses is positioned relative to the neighboring 2-fold related subunit in a 'domain-swapped' conformation.

2.4 ssDNA bacteriophage φX174

Bacteriophage φX174 is a small $T=1$ icosahedral virus (McKenna *et al.* 1992a, b) that contains a single-stranded, closed circular DNA molecule with 5386 nucleotide bases. Of the 11 gene products, the J, F, G, and H proteins exist in the mature virus. The 60 copies of the F protein create a 260 Å diameter capsid and the 60 copies of the G protein form the 12 pentameric spikes rising by about 30 Å above the capsid surface. There are 12 copies of the H 'pilot' protein that functions to guide the genome into the *Escherichia coli* cytoplasm for replication (Lei Sun, Andrei Fokine, Lindsey Young, Sergei Budko, Bentley Fane, Michael Rossmann, unpublished). The internal 60 B and external 240 D scaffolding proteins are required for assembly but are jettisoned while forming the mature virus. Unlike the majority of phages, φX174 does not contain a tail. The icosahedral structure together with the wealth of information known about this phage (Hayashi, 1978) made it an obvious choice for structural studies. With the experience gained in the study of picornaviruses, determining the structure of the mature virus (McKenna *et al.* 1992a, b) and immature procapsid (Dokland *et al.* 1997) fairly simple.

One of the surprises of these structure determinations was that the F and G proteins had jelly-roll folds (Fig. 13). The major capsid F protein had large insertions that formed much of the contacts between subunits. Nevertheless, the organization of the F proteins was much similar as in, for instance, the $T=1$ STNV structure (Liljas *et al.* 1982), suggesting that even this strange bacterial virus shared aspects of its evolution with the positive strand RNA icosahedral plant and animal viruses.

Recent results have shown that the H pilot protein forms a tail-like structure similar to the tail tube of the very common tailed dsDNA phages (Lei Sun, Andrei Fokine, Lindsey Young, Sergei Budko, Bentley Fane, Michael Rossmann, unpublished). Tailed phages use their tails as a highly efficient infection machine that usually requires only one phage to infect a cell. In contrast most animal viruses, lacking a tail, require hundreds of viruses to succeed in infecting a cell.

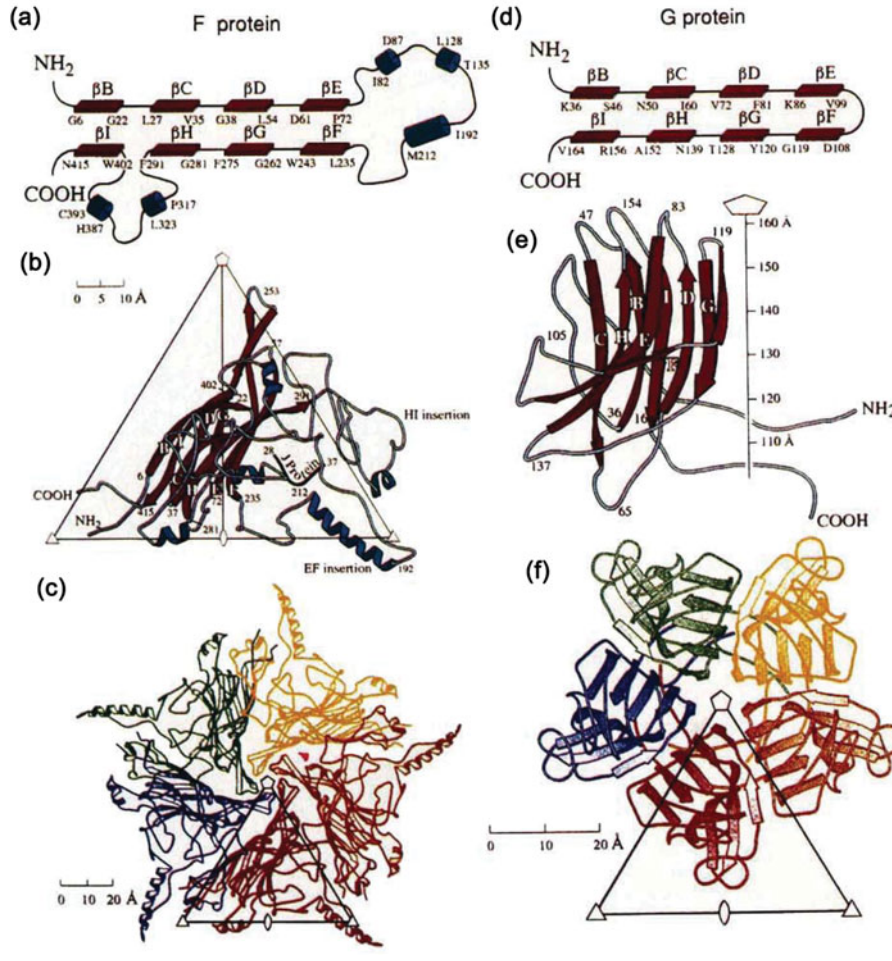

Fig. 13. The bacteriophage φX174 major capsid proteins F and G have jelly-roll structures. (*a*) and (*d*) show a topological description. (*b*) and (*e*) are ribbon diagrams of single subunits. (*c*) and (*f*) show the pentamer formation relative to the icosahedral asymmetric unit represented by a triangle. Modified from McKenna *et al.* (1992a).

Apparently, φX174 assembles a tail-like structure at the time of infection, presumably to increase its infection efficiency, whereas the most other phages have pre-assembled tails ready for use when needed.

All tailed phages that have been studied structurally to date use the 'HK97' fold to assemble their capsids, whereas φX174 uses a jelly-roll fold for the assembly of its capsid. It may be relevant that the only other known phage that assembles a tail-like structure at the time of infection is PRD1 (Cockburn *et al.* 2004; Xu *et al.* 2003) which, like φX174, uses a jelly-roll motif to assemble its capsid (Abrescia *et al.* 2004).

Structural investigations of φX174 have been particularly rewarding in the study of virus assembly aided by virally coded scaffolding proteins that are discarded when assembly is complete. The assembly of φX174 proceeds via a procapsid that includes not only the structural proteins F, G, and J, but also the internal scaffolding protein B and external scaffolding protein D. The φX174 procapsid structure is another example of a departure from strict Caspar and Klug

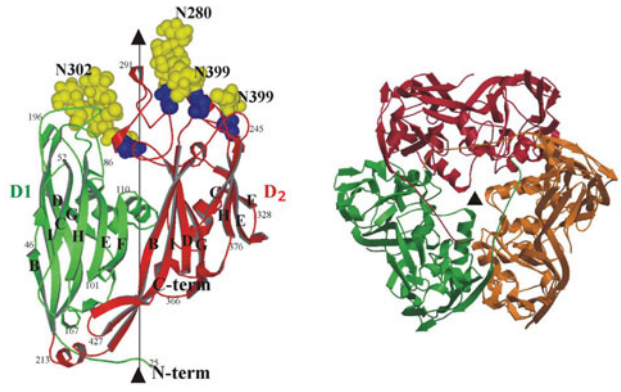

Fig. 14. The PBCV-1 double jelly-roll major capsid protein. Left: The monomer with the first jelly-roll in green and the second jelly-roll in red. Carbohydrate moieties are shown in space-filling yellow balls. Right: One capsomer consists of three monomers each with two jelly-rolls. Reproduced from Nandhagopal *et al.* (2002).

rules of quasi-symmetrical equivalence. Although the major capsid protein, F, and spike protein, G, assemble into a $T=1$ icosahedral particle, each of the four D scaffolding proteins in the icosahedral asymmetric unit have a totally different environment. The four D proteins in the icosahedral asymmetric unit have considerable structural variability, but they have a conserved core of 75 out of 152 residues for which the Cα atoms can be superimposed with a root-mean-square deviation of less than 2·5 Å between any pair of subunits. Thus, the D protein has an unusual ability to adapt itself to a different environment, which negates the requirement for each D protein to have a similar environment. The function of the D protein may be to hold together the F proteins during assembly (Dokland *et al.* 1997), but it is not clear why φX174 needs such help in assembly where this is not required in most other virus assembly processes.

3. Large icosahedral dsDNA viruses: the double jelly-roll fold

The jelly-roll motif can have different associations in different virus families. In the structures described above, each capsid protein subunit forms a jelly-roll, albeit with numerous insertions between the individual β-strands. In the structure of dsRNA viruses such as bluetongue virus (Grimes *et al.* 1998), the outer shell is composed of VP7 subunits arranged with $T=13$ quasi-symmetry. The structure of VP7 consists of an N-terminal, central, and C-terminal domain with the central domain having a jelly-roll motif. (In passing it should be mentioned that the dsRNA reoviruses have a special interest because they transcribe their genome within their capsid (Diprose *et al.* 2002; Zhang *et al.* 2003b)).

A major group of large viruses are icosahedral, have a dsDNA genome and use a 'double jelly-roll' motif in which there are two consecutive jelly-rolls along a single large polypeptide chain (Figs 14 and 15). The first such structure was determined by Roger Burnett and co-workers in their study of the adenovirus hexon (Athappilly *et al.* 1994; Roberts *et al.* 1986), a component of the capsid of an adenovirus. In these viruses, the double jelly-roll, major capsid protein forms trimeric capsomers. The six jelly-rolls in the capsomer produce pseudo-hexagonal symmetry. This is equivalent to the pseudo-hexagonal capsomers formed around the icosahedral 3-fold axes by VP2 and VP3 in pseudo $T=3$ picornaviruses. In the icosahedral dsDNA viruses the VP2 and

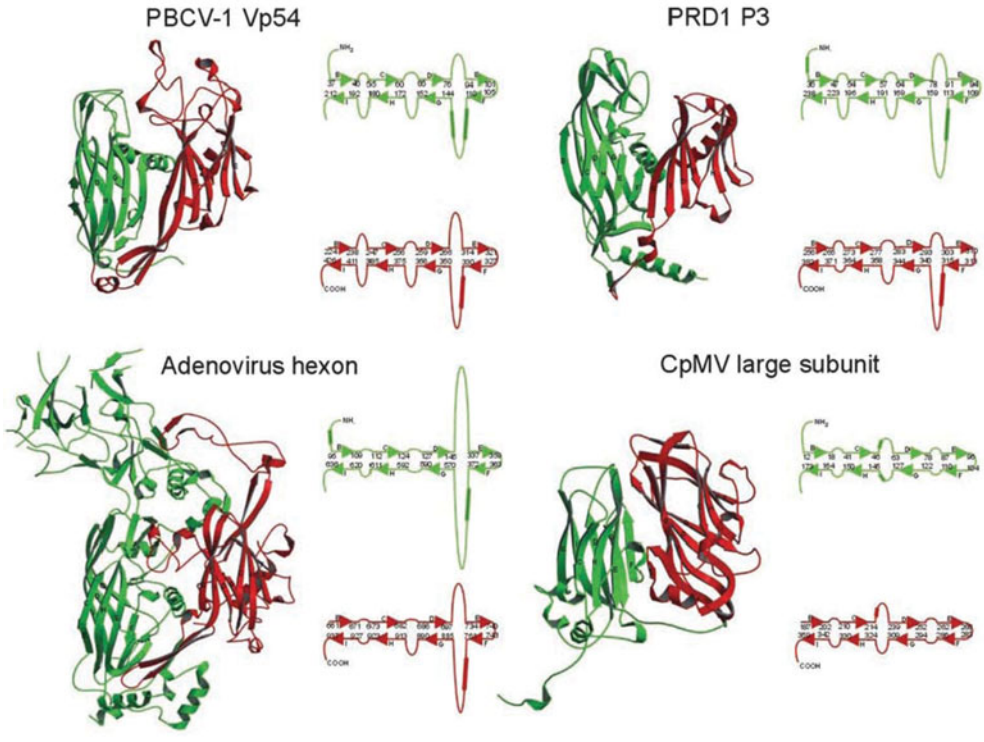

Fig. 15. Comparison of the double jelly-roll in PBCV-1 with similar structures in the bacteriophage PRD1, the human adenovirus and the plant cowpea mosaic virus. The major capsid protein of each virus is represented by a ribbon drawing color coded as in Fig. 14 and by a topological representation. Reproduced from Nandhagopal *et al.* (2002).

VP3 jelly-rolls are fused into a single polypeptide, thus creating pseudo-6-fold symmetry around the icosahedral 3-fold symmetry axes. Unlike picornaviruses, post-translational cleavage of VP2 and VP3 does not occur in cowpea mosaic virus, a 'plant picornavirus' (Fig. 9). Thus, the Casper and Klug triangulation number, *T*, represents the number of jelly-roll domains, not the number of covalently identical subunits when discussing the structures of viruses whose major capsid protein is a double jelly-roll. The implication is that the two different jelly-rolls within the major capsid protein of these viruses have similar environments in spite of having amino acid sequences that have little obvious resemblance to each other. Furthermore, with the major capsid protein having two consecutive jelly-rolls, the virus requires a separate gene to code for the penton protein containing only one jelly-roll for the formation of a pentomeric capsomer at each 5-fold vertex.

The diameter of capsomers that are assembled from six jelly-roll structures, whether they may be constructed of six covalently identical polypeptide chains as in some RNA icosahedral plant viruses, or of three double jelly-roll ploypetides, is fairly constant. The average distance between the center of one capsomer consisting of six jelly-rolls to the center of a neighboring six jelly-roll capsomer is about 75 Å. This can, therefore, act as a tentative identification of viral capsids that are composed of either six jelly-rolls in six capsid protein subunits or three double jelly-roll polypeptides in three capsid protein subunits.

Examples of viruses that have a verified double jelly-roll structure for their major capsid protein and a separate penton protein are the $T=25$ adenovirus (mentioned above) with a

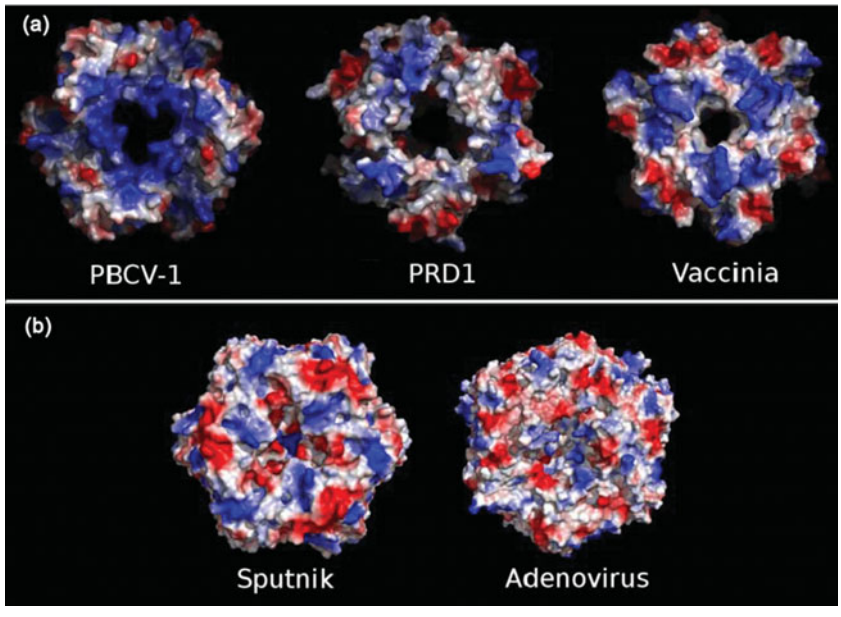

Fig. 16. Comparison of the inner surface charge distribution of dsDNA virus major capsid proteins with a double jelly-roll fold. (*a*) Viruses that have a lipid membrane envelope. (*b*) Viruses that do not have a lipid membrane envelope. Reproduced from Zhang *et al.* (2012).

diameter of about 650 Å, the $T=25$ PRD1 bacteriophage with an $\sim$700 Å diameter (Abrescia *et al.* 2004; Cockburn *et al.* 2004), the $T=25$ marine bacteriophage PM2 (Abrescia *et al.* 2008), the $T=25$ Sulfolobus turreted icosahedral virus (Fu & Johnson, 2012; Khayat *et al.* 2005), the $T=169$d *Paramecium bursaria* chlorella virus-1 (PBCV-1) with a diameter of 1900 Å (Nandhagopal *et al.* 2002; Simpson *et al.* 2003; Yan *et al.* 2000; Zhang *et al.* 2011b), and the $T=27$ Sputnik virophage (Zhang *et al.* 2012) with a diameter of about 700 Å. In addition, there are numerous viruses whose lower resolution cryoEM structures infer that the major capsid protein of the virus most probably has a double jelly-roll structure (Wrigley, 1969; Yan *et al.* 2005, 2009) including Mimivirus, the largest known virus (Kuznetsov *et al.* 2010; La Scola *et al.* 2003; Xiao *et al.* 2005, 2009). Furthermore, the large, icosahedral dsDNA virus-like *Asfarviridae*, *Phycodnaviridae*, *Iridoviridae*, *Poxviridae*, and *Mimiviridae* might all have similar icosahedral arrays of pseudo-hexameric capsomers, with each capsomere composed of three double jelly-roll major capsid proteins (Yan *et al.* 2009). Most of these large dsDNA viruses (Xiao & Rossmann, 2011) have at least one lipid membrane layer that envelopes the genome and other viral proteins. The exceptions are Adenovirus and Sputnik. The presence of a membrane is manifested not only by the presence of lipid in the virus and electron density trace of the phosphorus containing lipid membranes in cryoEM reconstructions, but also in the positive charge distribution on the capsomer surface facing the membrane (Fig. 16).

Because the capsomers made of double jelly-roll proteins that have only a pseudo-6-fold axis, there are two orientations, related by a 6-fold rotation, for every capsomer that can fit approximately into a hexagonal lattice. In many of the larger double jelly-roll icosahedral viruses, the capsomers are arranged in tri- and pentasymmetrons, within which the capsomers are all in the same orientation, but are rotated by 60° relative to the capsomers in neighboring orientations (Fig. 6). This creates cleavage planes between neighboring trisymmetrons (Fig. 6) that were first

observed by Wrigley (1969). In many of these viruses, there are 'anchor' proteins on the internal surface of the virus that tie together the tri- and pentasymmetrons (Abrescia *et al.* 2004; Yan *et al.* 2009; Zhang *et al.* 2011b). The cleavage planes are clearly the result of an ordered assembly process and also allow the virus to disassemble when called to do so during phagocytosis or other means of cell entry. However, some of these viruses also have special cell entry devices situated on a unique 5-fold vertex, such as a spike in PBCV (Cherrier *et al.* 2009; Zhang *et al.* 2011b), a 'starfish'-like feature in Mimivirus (Klose *et al.* 2010; Kuznetsov *et al.* 2010; Xiao & Rossmann, 2011; Zauberman *et al.* 2008; Zhang *et al.* 2011b) and, as mentioned above, the possibility of forming a tube connection to the host for phage PRD1 and related phages (Abrescia *et al.* 2004; Cockburn *et al.* 2004).

4. Enveloped icosahedral viruses

4.1 Alphaviruses

The next challenge, after the structure determinations of a variety of simple icosahedral viruses in the 1980s, was a structural investigation of an icosahedral viruse that contained a lipid membrane. An obvious candidate was the structural analysis of an alphavirus. Alphaviruses are a group of approximately 27 arthropod-transmitted viruses that can cause various serious, frequently fatal, diseases such as encephalitis, fever, arthritis, and rash (Kuhn, 2007). Alphaviruses can grow to high titer in cell culture, making them especially suitable for structural studies.

By the early 1990s, fairly detailed information was available on alphaviruses, such as their lipid and protein content. Low-angle X-ray scattering and electron microcroscopy (von Bonsdorff & Harrison, 1975, 1978) had established that these viruses were icosahedral with a $T=4$ quasi-symmetry surface lattice and an external diameter of about 700 Å. The introduction of cryoEM confirmed that the internal nucleocapsid was surrounded by a lipid membrane that anchored the external glycoproteins (Fuller, 1987). When the structure of the Sindbis virus capsid protein had been determined (Choi *et al.* 1991), it could be used to fit into the cryoEM density of the homologous Ross River alphavirus (Cheng *et al.* 1995) in one of the early applications of such a hybrid technique to virus structure.

The lipid membrane in alphavirus particles is derived from the host plasma membrane and envelopes one copy of a $\sim$12 kb positive-strand, genomic RNA. The genome encodes two polyproteins, both of which are post-translationally cleaved into their component proteins by viral and cellular proteinases. The structural polyproteins (capsid–E3–E2–6K–E1) are translated from a subgenomic mRNA. E1 and E2 are envelope glycoproteins and 6 K is a 6 kDa peptide of unknown function. Each alphavirus particle contains 240 copies each of the E1, E2 and capsid proteins. E1 and PE2 (precursor to the E3–E2–6K proteins before cleavage of E3, also identified as p62) are assembled as heterodimers and undergo N-linked glycosylation in the lumen of the rough endoplasmic reticulum.

Although crystals of alphaviruses had been grown, these crystals did not give detectable X-ray diffraction (Harrison *et al.* 1992), indicating that perhaps the virus particle positions but not their orientation were aligned in these crystals. An attempt was made to improve the diffracting quality of Sindbis virus crystals by mutating the potential glycosylation sites on the two external glycoproteins, E1 and E2 (Pletnev *et al.* 2001). However, this did not improve the crystal quality, but cryoEM difference maps between the wild-type and mutant viruses mapped the site of the

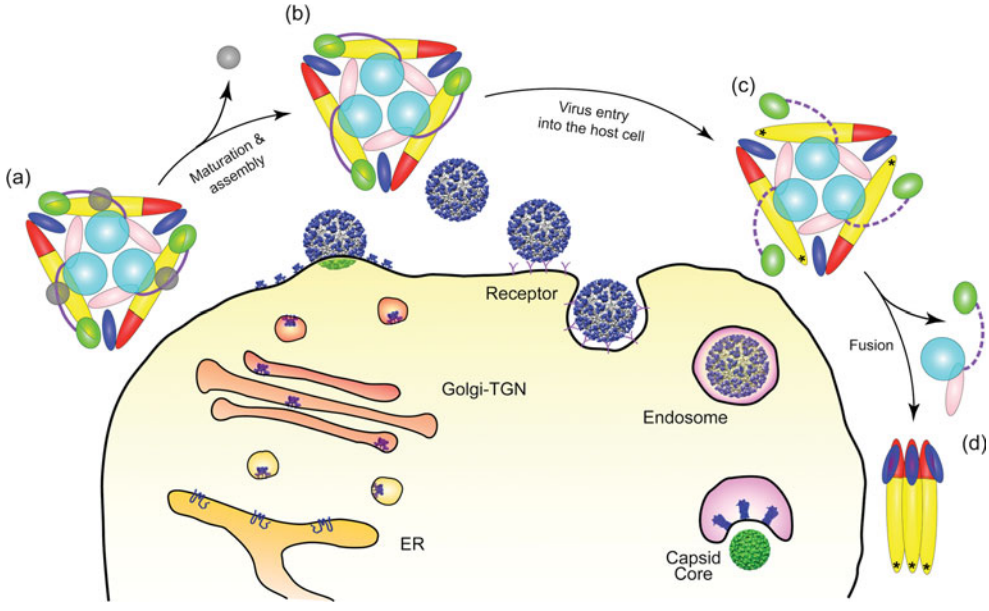

Fig. 17. Cartoon showing maturation and fusion. Color-coded as in Fig. 2. (*a*), The viral spike while being transported through the Golgi and TGN. Domain B is attached to domain II because E3 (gray) binds the β-ribbon connector to domain A and holds E2 in place at low pH. (*b*) The viral spike on the virus after E3 has been proteolytically cleaved and, in Sindbis virus, has been jettisoned. (*c*) After the virus has recognized a host cell and entered into a low pH endosomal vesicle, domain B becomes unattached from DII, exposing the fusion loop (marked with an *). The dashed line indicates the disordered β-ribbon connector. (*d*) The trimeric spike disassembles, allowing escape of E2 and formation of E1 homotrimers. Reproduced from Li *et al.* (2010).

carbohydrate sites. This information could be combined with the crystal structure of the E1 protein (Lescar *et al.* 2001) to establish the site and orientation of E1 on the virus surface (Mukhopadhyay *et al.* 2006; Pletnev *et al.* 2001; Zhang *et al.* 2002).

The E1E2 heterodimers, to which E3 may or may not remain associated depending on the virus, self-assemble into 80 trimeric spikes on the virus surface (Cheng *et al.* 1995; Zhang *et al.* 2002), 60 with quasi-3-fold symmetry and 20 coincident with the icosahedral 3-fold axes arranged with $T=4$ quasi-symmetry (Fig. 17). E1 is responsible for cell fusion that requires an acidic environment as well as cholesterol and sphingolipid membrane components (Kielian, 1995). E2 is primarily involved in receptor binding and cell entry (Gibbons & Kielian, 2002; Li *et al.* 2010; Voss *et al.* 2010). E1 and E2 both have one transmembrane helix that traverses the lipid membrane surrounding the nucleocapsid core (Cheng *et al.* 1995; Zhang *et al.* 2002).

The X-ray crystal structures of the ectodomain of the E1 glycoprotein of Semliki Forest virus (Lescar *et al.* 2001; Roussel *et al.* 2006) was found to be homologous to the flavivirus E glycoprotein (Modis *et al.* 2003; Rey *et al.* 1995; Zhang *et al.* 2004) and consists of three β-barrel domains. Domain I contains the amino terminus and is spatially located between domains II and III. The carboxy-terminus lies within domain III and the fusion loop is at the distal end of domain II. Three E1 monomers were found to be associated with each of the surface spikes.

The crystal structures of the Sindbis virus $(E2–E1)_3$ trimeric spike (Li *et al.* 2010) and of the Chikungunya virus heterodimer (E2–E1) (Voss *et al.* 2010) show that E2 is a long, thin molecule, complexed with E1, which covers the fusion loop at the top of E1. The structure of E2 consists

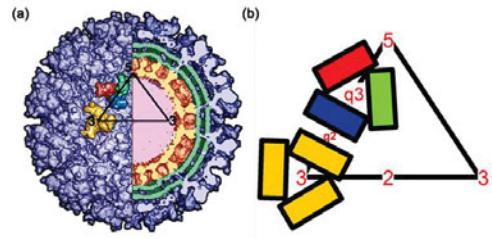

Fig. 18. The $T=4$ quasi-symmetry of SINV. (*a*) Left half: Surface-shaded view of Sindbis virus at 9 Å resolution. The four independent E1–E2 heterodimers related by $T=4$ quasi-symmetry are colored yellow, blue, green, and red. The three yellow heterodimers constitute the spikes at the icosahedral 3-fold axes. The blue, green, and red heterodimers constitute the spikes at the quasi-3-fold axes. The triangle delineates one of the 60 icosahedral asymmetric units. Right half: A cross section of the virus particle showing the overall organization of the particle. The nucleocapsid is colored red, the lipid bilayer is green, and the glycoproteins are shown in blue. (*b*) Diagrammatic representation of one asymmetric unit using the same color code. Reproduced from Mukhopadhyay *et al.* (2006).

of the N-terminal A, the middle B and the C-terminal C domains. Domain B is connected to domains A and C by long connecting linker peptides (the 'β-ribbon connector'). Although domain B, at the tip of the spike, is ordered at neutral pH, it is mostly disordered at low pH. Both domain A and domain C have the topology of an immunoglobulin fold. This is consistent with E2 functioning as a cell receptor-binding protein (Gibbons & Kielian, 2002).

Subsequent to receptor recognition, the virus initiates the formation of an endosomic vesicle that is released into the cytoplasm (Kielian & Rey, 2006; Marsh & Helenius, 1989) (Fig. 18). The low pH in the endosome causes the virion to undergo an irreversible conformational change in which the E1–E2 heterodimers associate more tightly into trimers. The fusion process in alpha- and flaviviruses (Lu *et al.* 1999), defined as 'class 2 fusion' (Lescar *et al.* 2001), requires an internal hydrophobic sequence of about a dozen moderately conserved and hydrophobic amino acids known as the fusion loop.

Fusion of the viral membrane with the cellular plasma membrane in alphaviruses requires the dissociation of the E1–E2 heterodimer and formation of E1 homotrimers at low pH in the presence of a lipid membrane. Subsequently, in the post-fusion E1 homotrimeric structure, the E1 molecules are arranged with their long direction running roughly parallel to a common 3-fold axis of rotation, exposing the fusion loops of each monomer at one end of the trimeric complex. Residues of E1 that were in contact with E2 in the pre-fusion trimeric spike are located on the surface of the post-fusion E1 homotrimer, facing away from the spike axes, and are predominantly in surface loops (Li *et al.* 2010). The conformational changes required to form a post-fusion trimer requires the removal of the E2 molecules from the center of the trimeric spike to the outside and rotation of the E1 molecules by about 180° about their long axes to face each other. Similar, huge conformational changes occur in the maturation and fusion of flaviviruses (Li *et al.* 2008; Yu *et al.* 2008).

4.2 Flaviviruses

Flaviviridae (Latin flavus meaning yellow, because of the jaundice induced by yellow fever virus) is a family of icosahedral, lipid-enveloped viruses that infect mammals. The *Flaviviridae* family has been divided into three genera, namely flaviviruses, pestiviruses, and hepatitis C virus. The Flavivirus genus itself has been further subdivided into the mosquito-borne viruses

(e.g. yellow fever, dengue, and Japanese encephalitis), the tick-borne viruses (e.g. TBEV), and those that do not use arthropod vectors (Strauss & Strauss, 2002).

Mature flaviviruses are approximately 500 Å in diameter, are composed of a single, positive-strand RNA genome packaged with multiple copies of the virus capsid protein within a host-derived lipid bilayer, surrounded by 180 copies of two glycoproteins (Lindenbach & Rice, 2001). Infection by dengue virus is usually characterized by fever and severe joint pain, but a more serious syndrome, dengue hemorrhagic fever, can occur when there has been an infection by more than one of the four dengue serotypes. Dengue hemorrhagic fever is fatal in about 10% of infected individuals, causing about 24 000 deaths each year worldwide. The taming of yellow fever virus by control of the *Aedes aegypti* mosquito vector's water habitat (Gubler, 1988) at the time of the building of the Panama Canal and, later, the development of an effective vaccine is well known. Nevertheless, at this time, there are no effective vaccines against most human flavivirus pathogens. Indeed, control of dengue through the use of vaccination has proven to be elusive (Burke & Monath, 2001) because of the pathogenic enhancement during a second infection.

The ~10·8 kb genome of flaviviruses has one long open reading frame encoding a single polyprotein. This protein is anchored in the endoplasmic reticulum (ER) by membrane-spanning amino acid sequences. The polypeptide is subsequently cleaved by cellular and virally coded proteinases, as well as glycosylated by cellular glycosyltransferases. The capsid protein, C, consists of ~100 amino acids and is involved with packaging of the viral genome and forming the nucleocapsid core. The prM (~165 amino acids) and E (~495 amino acids) glycoproteins each have two transmembrane helices. The prM protein might act as a chaperone for the folding and assembly of the E protein (Lorenz *et al.* 2002) prior to its cleavage to yield the pr peptide and M protein (~75 amino acids) during particle maturation. The E protein contains a cellular receptor-binding site and a fusion peptide (Allison *et al.* 2001; Lorenz *et al.* 2002; Pokidysheva *et al.* 2006; Stadler *et al.* 1997).

Flaviviruses enter host cells through receptor-mediated endocytosis (Lindenbach & Rice, 2001). The initial target for dengue and West Nile viruses is immature dendritic cells (Davis *et al.* 2006; Lozach *et al.* 2005; Pokidysheva *et al.* 2006) and Langmuir cells resident in skin (Wu *et al.* 2000). As in the life cycle of alphaviruses so also for flaviviruses, the acidic environment of the endosome triggers an irreversible trimerization of the E protein that results in fusion of the viral and cell membranes (Harrison, 2008; Stiasny & Heinz, 2006). After fusion has occurred, the nucleocapsid is released into the cytoplasm, the capsid protein and RNA dissociate, and replication of the RNA genome and particle assembly are initiated (Brinton, 2002; Lindenbach & Rice, 2001, 2003).

Immature particles are formed in the lumen of the ER. These particles (Zhang *et al.* 2003c, 2007), which contain E, prM, lipid membrane, and nucleocapsid, cannot induce host–cell fusion, making them non-infectious (Guirakhoo *et al.* 1991, 1992). After assembly, immature virions undergo a protease- and pH-dependent maturation during their transit from the ER through the trans-Golgi network to the cell surface (Elshuber *et al.* 2003; Li *et al.* 2008; Stadler *et al.* 1997; Yu *et al.* 2008, 2009). During maturation, the E proteins go through a positional reorganization while the prM protein is cleaved by the cellular subtilisin-like endoproteinase furin (Stadler *et al.* 1997). The proteolytic cleavage releases the glycosylated, N-terminal pr fragment from the small, membrane-anchored M portion that remains virion-associated. The pr peptide is subsequently shed from the virion during the exocytotic release of virus progeny (Yu *et al.* 2008, 2009). The removal of the previously covalently attached pr peptides from the virions primes the viral

particles, ready for the low-pH triggered fusion events during the upcoming cell entry (Harrison, 2008; Stadler *et al.* 1997).

The structure of the E protein ectodomain in mature flaviviruses is a dimer (Kuhn *et al.* 2002; Rey *et al.* 1995). Each monomer has three domains: the structurally central amino-terminal domain, the dimerization domain that contains the fusion peptide and the carboxy-terminal immunoglobulin-like domain. A total of 30 sets of three parallel dimers constitute the rather smooth surface of mature flaviviruses (Kuhn *et al.* 2002; Mukhopadhyay *et al.* 2003) (Fig. 19). This arrangement of 180 monomers lacks the conventional $T=3$ symmetry and makes the lipid bilayer largely inaccessible from the viral exterior, indicative of a major conformational change prior to fusion (Kuhn *et al.* 2002). The two transmembrane domains of the E and M proteins form anti-parallel helices (Zhang *et al.* 2003a). Neither set of transmembrane helices extends beyond the inner lipid leaflet to make contact with the nucleocapsid. The two helices in the E protein's stem region, connecting the ectodomain and the transmembrane helices, lies in the outer lipid leaflet (Zhang *et al.* 2003a). The stem region of the M protein is also helical and is partially buried in the outer lipid leaflet.

Immature flavivirus particles can be accumulated by adding NH_4Cl to the growth medium in the final steps of propagation in order to inactivate the cellular furin (Heinz *et al.* 1994). The immature particles are strikingly different to the mature virus (Zhang *et al.* 2003c, 2007). They have 60 prominent, irregular, trimeric surface spikes, making the external diameter 600 Å, somewhat larger than mature virions. Like the mature virions, these particles do not have classical $T=3$ quasi-symmetry. Each spike is a trimer of prM–E heterodimers. The three prM proteins in the spike cap the fusion peptides, which are at the external tip of the E protein. Once cleavage of the pr peptide from prM has occurred during particle maturation in the TGN, the trimer is disrupted, which then induces a major rearrangement of the E and M proteins to form the structure of the mature virus (Fig. 19). The 60 E trimers rearrange themselves into 90 E dimers, accompanied by a 27° change in the hinge angle between domains I and II in each E protein (Zhang *et al.* 2004) and a rearrangement of the E and M stem and anchor regions with respect to the icosahedral axes (Zhang *et al.* 2003c). The cleaved pr peptide is released when the pH returns to neutral in the extracellular environment (Yu *et al.* 2008, 2009) (Fig. 20).

Another major rearrangement occurs when the virus undergoes fusion with the host cell (Kaufmann & Rossmann, 2011; Kielian & Rey, 2006). The parallel E dimers of the mature virus dissociate into monomers, which then reassociate into trimers (Allison *et al.* 1995, 1999; Bressanelli *et al.* 2004; Modis *et al.* 2004; Stiasny *et al.* 1996). One hypothesis (Kuhn *et al.* 2002) is that this fusogenic form of the virus forms via an intermediate that is similar to a $T=3$ structure that had been extrapolated from small $T=1$ non-infectious particles, which can occur during normal propagation (Ferlenghi *et al.* 2001) upon receptor binding or in the endosome before forming E trimers. This $T=3$ structure is composed of trimeric units and has extensive regions of exposed membranes. Furthermore, the proposed intermediate structure has $T=3$ symmetry, as would be expected if the affinity between monomers is decreased by lowering the pH. This hypothesis has been at least partially confirmed in the structure determination of West Nile virus complexed with Fab fragment of monoclonal antibody E16 (Kaufmann *et al.* 2006) that inhibits fusion. In this study the bound Fab molecules inhibit the conformational changes that occur when the virus becomes fusogenic (Kaufmann *et al.* 2009). Further support for this hypothesis comes from the structure of dengue virus when incubated at 37°C (Zhang *et al.* 2013).

The capsid protein of flaviviruses, like many other nucleocapsid proteins, is highly basic. In many naked RNA plant viruses, as well as in the enveloped alphaviruses, the capsid protein has

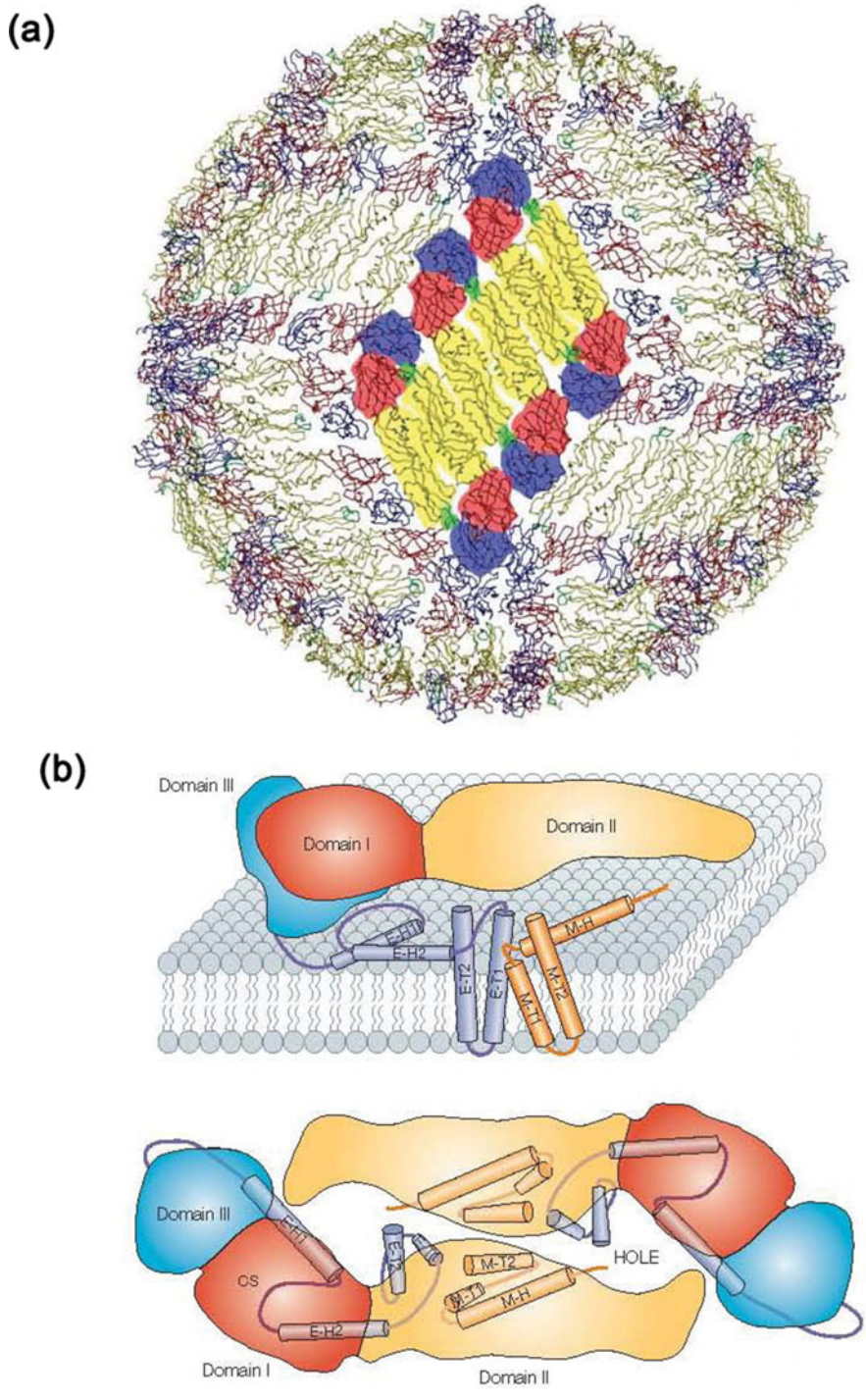

Fig. 19. Packing of the E protein in mature flavivirus virions. (*a*) One raft, consisting of three parallel dimers, is highlighted. Domains I, II, and III are in red, yellow, and blue, respectively. The fusion peptide is in green. Reprinted from (Zhang *et al.* 2004). (*b*) Diagrams of the dengue virus E protein ectodomain and transmembrane helices. Reproduced from Zhang *et al.* (2003a).

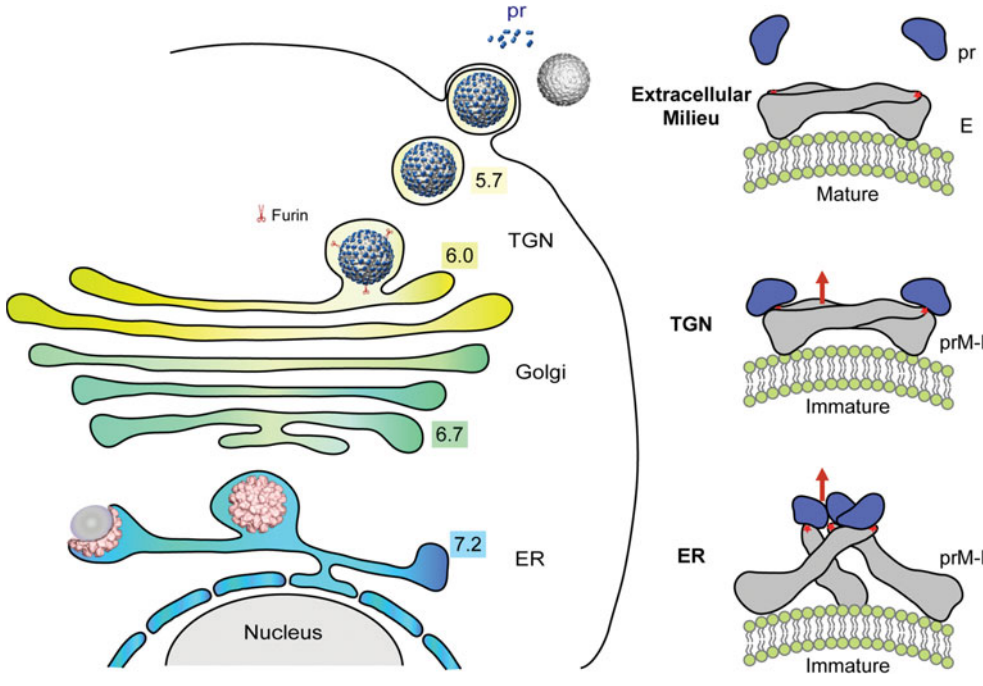

Fig. 20. Flavivirus maturation. Immature particles bud into the ER as spiky particles and are transported through the Golgi into the trans-Golgi network (TGN) where acidification induces a conformational change. Furin cleavage takes place in the TGN, but pre peptide (pr) remains associated until the virion is released into the extracellular environment where the pH is neutral. Reproduced from Yu *et al.* (2008).

an N-terminal, basic region that neutralizes the packaged genome. Thus, the small, about 100, amino acids of the flavivirus capsid protein might act in a similar manner. Nevertheless, crystallographic and NMR studies show that the flavivirus C protein has a unique structure (Dokland *et al.* 2004; Ma *et al.* 2004).

5. Pleomorphic lipid-enveloped viruses

Many of the major human viral pathogens are pleomorphic. These include HIV/AIDS (human immunodeficiency virus/acquired immunodeficiency syndrome) belonging to the *Retroviridae* family, measles and mumps belonging to the *Paramyxoviridae* family and influenza virus belonging to the *Orthomyxoviridae* family. Structural studies of these viruses have concentrated on the determination of the major structural proteins by X-ray crystallography. For instance, the structures of the two major surface glycoproteins of influenza virus, hemoglutinin (Wiley *et al.* 1981) and neuraminidase (Colman *et al.* 1983) were determined some time ago and aided in the design of vaccines and anti-viral agents, respectively. Less is known about these structural proteins in the context of the virus, because pleomorphic viruses can neither be crystallized nor their EM projections averaged to give three-dimensional information. However, with the development of cryoET in recent years, it is now possible to obtain three-dimensional images of individual virus-sized particles (Rossmann *et al.* 2011). This technique provides rather anisotropic information on virus structure at best to about 40 Å resolution. However, identification of similar structures in the same or different tomograms can sometimes improve the resolution and make the resolution

more isotropic. Here I give one recent example of the study of a pleomorphic virus using a combination of X-ray crystallography and cryoCT.

5.1 Newcastle disease virus

Newcastle disease virus (NDV) is an avian paramyxovirus. Paramyxoviruses have two transmembrane glycoproteins, the attachment protein HN (~75 kDa), G or H (depending upon the virus) and the F protein (~60 kDa), that form spikes protruding from the lipid bilayer. HN is a dual-function hemagglutinin-neuraminidase, capable of binding to cell surface sialic acids (Scheid *et al.* 1972). F is required for fusion with the host cell plasma membrane (Nagai *et al.* 1976; Yin *et al.* 2005, 2006). The nucleocapsid protein, NP (~50 kDa), together with the genomic RNA, form a helical structure that encapsidates and protects the viral RNA genome. The matrix (M) protein of many pleomorphic, membrane-enveloped viruses directs assembly and budding (Dessen *et al.* 2000; Harris *et al.* 2001; Money *et al.* 2009; Neumann *et al.* 2009; Rao *et al.* 1995).

The morphology and size of NDV virions (Battisti *et al.* 2012) varies from approximately spherical to ellipsoidal with diameters ranging from about 100–350 nm. A layer of glycoprotein spikes extends about 12–18 nm beyond the surface of the 4–5 nm thick membrane. Within the virions, the nucleocapsids are mostly long, probably helical, structures with a diameter of 20 nm and a pitch of ~7 nm. Some virions have an additional layer of density on the inner surface of the viral membrane (Fig. 21). This layer is about 4–5 nm thick and represents the matrix protein. Although this layer is visible in only about 10% of the virions, the matrix protein is in approximately the same abundance as the other structural proteins. In the NDV particles that have at least a partial matrix layer, the surface glycoproteins and internal nucleocapsids associate with the matrix array (Fig. 21).

Tomographic sections tangential to the viral membrane (Battisti *et al.* 2012) show that the matrix layer forms a grid-like array (Fig. 22). The repeating unit of matrix protein has a nearly square shape, 5 nm on edge, which forms an array with a 7 nm repeat in orthogonal directions along the diagonals of the square. Averaging the glycoprotein layer using the periodicity of the matrix protein showed the existence of an extracellular array with high densities in register with the gaps between repeating units in the matrix array on the inner side of the membrane. The volume of the gaps between the matrix protein units is ample to accommodate the cytoplasmic tails of the tetrameric HN (Yuan *et al.* 2011) or trimeric F (Yin *et al.* 2006) glycoproteins.

The structure of the NDV 364 amino-acid-long, recombinantly expressed M protein was found to be a dimer in three different crystal forms. The monomer had two similarly folded domains related by an approximate 4-fold axis coincident with the dimer 2-fold axis and has structural similarity to the matrix protein in other paramyxoviruses (Respiratory Syncytial Virus (Money *et al.* 2009), Ebola virus (Dessen *et al.* 2000) and Borna virus (Neumann *et al.* 2009)). The dimeric M protein crystal structure fits well into the subunit density in the matrix protein layer of the tomograms, verifying that the crystallographic dimer is the physiologically relevant unit. Furthermore, the side of the matrix protein facing the membrane is highly positively charged as is required to associate with the negatively charged membrane surface.

Inspection of NDV cryoEM tomograms showed that there were only a few particles that had matrix protein arrays. Presumably these particles represented viruses immediately after budding, since this event requires the matrix protein arrays for the generation of membrane curvature (Shnyrova *et al.* 2007). Particles that do not show distinct matrix protein arrays (although the virus still contained abundant matrix protein) probably represented the virus sometime after

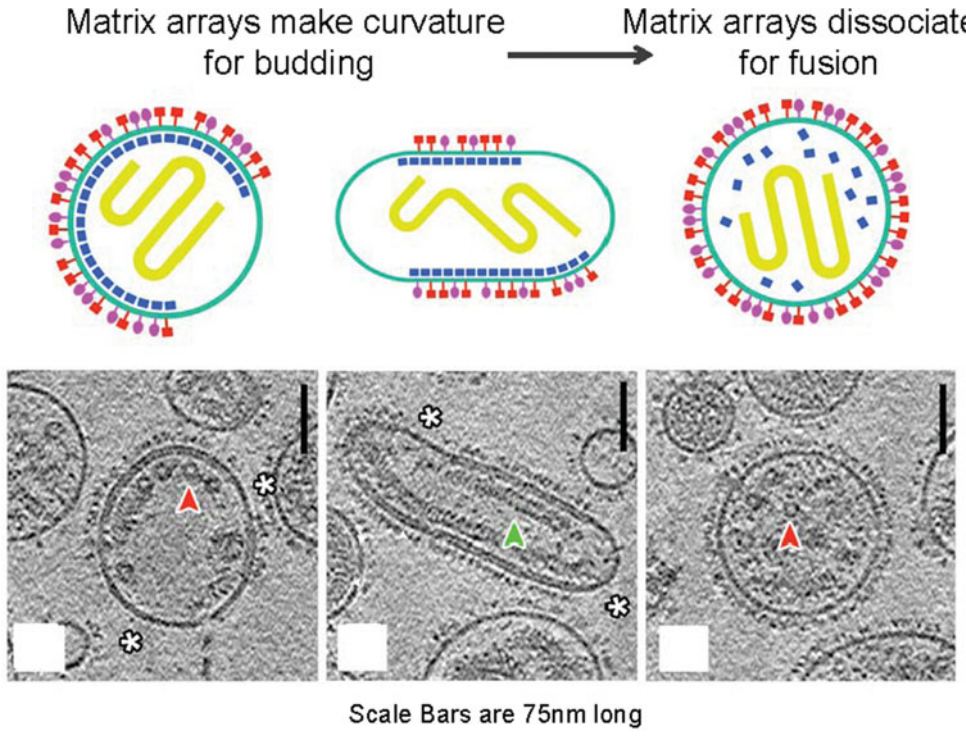

Fig. 21. Cryo-ET images and diagrammatic representations of Newcastle disease virus. The white asterisks mark places where the external glycoproteins, the internal nucleocapsids and matrix protein are absent. The red arrowheads indicate portions of the nucleocapsid viewed end-on. The green arrowhead indicates a side view of a nucleocapsid. The virions are shown schematically on the right where the two types of surface glycoproteins, HN and F, are shown in red and magenta respectively; the matrix protein is shown in blue and the nucleocapsid is shown in yellow. Reproduced from Battisti *et al.* (2012).

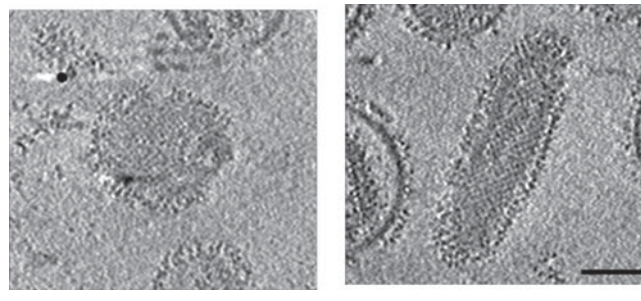

Fig. 22. Cryo-ET images of NDV showing arrays of matrix protein. The repeat distance is about 70 Å in perpendicular directions. Reproduced from Battisti *et al.* (2012).

budding. The disassembly of the matrix protein arrays is probably necessary for infection when the glycoproteins must be freed from the matrix array, allowing the F glycoprotein to undergo the conformational change required to transform from a pre- to post-fusion state (Lamb & Jardetzky, 2007). In addition, the matrix protein has to be released from the membrane for the formation of fusion pores and the release of the nucleocapsid into the host's cytoplasm (Battisti *et al.* 2012).

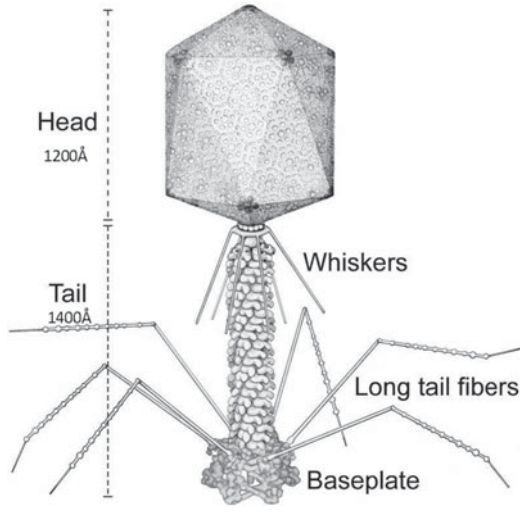

Fig. 23. Structure of bacteriophage T4. Modified from and reproduced with permission from Eiserling & Black (1994).

6. Tailed bacteriophages

Bacteriophages (Calendar, 2006) are the most abundant forms of life in the biosphere (Hatfull *et al.* 2006). They are also easier to propagate in the laboratory than most other viruses and have therefore been a major source of knowledge of virus structure, assembly and infectivity. The structures of dsDNA bacteriophages are unique among viruses in that most of them have tails that are specialized host cell attachment organelles. Unlike most animal viruses, tailed phages require on average only one encounter between virus and host to accomplish infection. This is primarily on account of the successful design of the tail that overcomes cell wall and membrane barriers in order to inject the viral genome into the host. Phages that possess a tail are collectively called *Caudovirales*, which is divided into three sub-families according to the tail morphology: *Myoviridae* (long contractile tail), *Siphoviridae* (long non-contractile tail), and *Podoviridae* (short non-contractile tail). Of these, *Myoviridae* phages have the most complex tail structures. Probably the most studied of all bacteriophages is T4 (Leiman *et al.* 2010, 2003), which belongs to the *Myoviridae* family. Thus, no account of virus structure can be complete without at least a brief mention of T4.

6.1 Bacteriophage T4

Bacteriophage T4 is a large, dsDNA virus that infects *E. coli* (Fig. 23). It has a genome that contains 274 open reading frames out of which more than 40 encode structural proteins. The mature virus consists of a 1150 Å-long and 850 Å-wide prolate head with hemi-icosahedral ends (Fokine *et al.* 2004) encapsidating the genomic DNA; a 1000 Å-long, 210 Å-diameter co-cylindrical contractile tail (Aksyuk *et al.* 2009a, 2011; Kostyuchenko *et al.* 2005), terminated with a 460 Å-diameter baseplate (Kostyuchenko *et al.* 2003) and six 1450 Å-long fibers attached to the baseplate (Aksyuk *et al.* 2009b). The head, tail, and fibers assemble via independent ordered pathways and join together to form a mature virus particle (Berget & King, 1978; Coombs & Arisaka, 1994; Kikuchi & King, 1975; Leiman *et al.* 2003; Mosig & Eiserling, 2006).

Upon infection, the phage shuts down host-specific nucleic acid and protein syntheses, thus ensuring production of only its own components in amounts sufficient to assemble up to 200 progeny virus particles per infected cell. Bacteriophage T4 head, tail, and fiber structures have been studied using X-ray crystallography to determine the various protein components and by cryoEM to establish the relative positions of the components in the virus (Leiman *et al.* 2003, 2010), combined with extensive biological, genetic, and biochemical data.

Infection of *E. coli* by phage T4 is initiated when the mature virus senses the presence of a potential host with the distal ends of the long tail fibers (gene product 37, gp37) (Bartual *et al.* 2010). The long tail fibers communicate recognition of a host to the baseplate which then unfolds the short tail fibers (gp12) (Thomassen *et al.* 2003; van Raaij *et al.* 2001) and bind to lipopolysaccharides on the surface of the *E. coli* host. Unfolding of gp12 is accompanied by a large conformational change of the dome-shaped baseplate to a hexagonal star-shaped baseplate (Crowther *et al.* 1977; Kostyuchenko *et al.* 2003; Leiman *et al.* 2004). At the same time the tail sheath, one end of which is attached to the baseplate, starts to contract (Aksyuk *et al.* 2009a, 2011). The sheath consists of 138 copies of the gp18 sheath protein arranged as a 6-ply helix attached to the baseplate at one end and to gp15 in the neck below the head at the other end (Fokine *et al.* 2012). As the sheath contracts, the tail tube is pushed through the center of the baseplate while it is rotating as a consequence of the sheath protein subunit rearrangement during contraction. The tip of the tail tube is composed of gp5 that acts as a cell puncturing device (Browning *et al.* 2012; Kanamaru *et al.* 2002), that enters the periplasm of the host bacterium and then encounters the peptidoglycan cell wall. The cell wall is digested by the lysozyme formed by the central domain of gp5. Once this obstacle has been overcome the tail tube is further extended to penetrate the inner membrane, allowing the genome to be ejected into the host's cytoplasm. The *Podoviridae* φ29, lacking the convenience of a contractile tail, similarly uses enzymes to bore its way through the host's cell wall (Cohen *et al.* 2009; Fokine *et al.* 2008; Xiang *et al.* 2008).

In all known tailed phages, the head is assembled as an empty prohead, which is subsequently packaged with the genome. Some animal viruses use the same strategy by assembling an initial empty prohead (Brown & Newcomb, 2011; Plevka *et al.* 2011a). Both DNA packaging and DNA ejection proceed via the special portal located at one of the icosahedral 5-fold vertices. This portal is modified by an assembly of 12 portal proteins that together form a tube (the 'connector') through which the DNA passes both during entry and exit (Lebedev *et al.* 2007; Olia *et al.* 2011; Simpson *et al.* 2000b). The first detailed atomic resolution structure of a phage head was determined by X-ray crystallography for the icosahedral, $T=7$, dsDNA phage HK97 (Wikoff *et al.* 2000). In phage T4 an X-ray crystallographic structure of the protein that forms 11 of the 12 vertices, $(gp24)_5$, was found to have an HK97-like structure. Furthermore, there is significant sequence similarity between gp24 and the T4 major capsid protein, gp23. Thus, the major capsid protein of phage T4 has a similar structure as the HK97 major capsid protein. Indeed, cryoEM reconstructions have shown that all dsDNA tailed phages examined so far (Jiang *et al.* 2003, 2006; Morais *et al.* 2005) have been found to use the HK97 capsid fold for their major capsid proteins.

The duplication of genes within the genome is a common event, presumably when there is a functional advantage in the separate but different evolution of the two genes. One example mentioned above is the HK97-like major capsid protein gene, gp23, in phage T4 that is followed in the genome by the gene for gp24 that provides a similar but slightly modified structure at the pentameric vertices of the phage head (Fokine *et al.* 2006). Another example of gene duplication

followed by separate evolution of the two genes is the duplication of the single jelly-roll gene is the double jelly-roll gene in numerous larger icosahedral dsDNA viruses (see above).

Packaging of the genome requires energy in order to compress the negative charge on the phosphates into the small space within the head. Thus, the tailed phages code for a DNA-packaging machine, usually consisting of a 'large terminase' and a 'small terminase'. The term 'terminase' relates to the synthesis of genomic DNA from a circular genome, thus generating a new concatenated genome. Hence, when a new head is filled with the new genome it has to be terminated when the head is full, resulting in the packaging of slightly more than one complete genome. The large terminase provides the ATPase motor and the nuclease for termination (Kanamaru *et al.* 2004; Sun *et al.* 2008, 2010). The small terminase is required for initiation of the packaging process (Büttner *et al.* 2012; Roy *et al.* 2012; Sun *et al.* 2012). Detailed analysis on the mechanisms of the phage DNA packaging motor has been the subject of many crystallographical, cyoEM and optical tweezer investigations of T4 and other tailed and tail-less phages (Aathavan *et al.* 2009; Fuller *et al.* 2007; Mancini *et al.* 2004; Moffitt *et al.* 2009; Smith *et al.* 2001; Sun *et al.* 2007, 2008; Zhang *et al.* 2011c) and has been reviewed by Sun *et al.* (2010) and by Rao & Feiss (2008). Like most biological motors that depend on an ATPase for providing energy, the structure of the motor itself is the highly conserved nucleotide-binding domain (Rossmann *et al.* 1974) that is coupled to a variety of other domains in order to transform the chemical energy stored in ATP to mechanical energy for translation of the DNA into the prohead. After a long dispute, there seems to be mostly agreement that the DNA packaging motors of φ29 and T4 are 5-stroke machines as had been suggested by us (Morais *et al.* 2008; Tao *et al.* 1998) consisting of five identical ATPases situated around the portal at the unique 5-fold vertex.

7. Conclusion

The definition of what is a virus remains in flux, especially after the discovery of Mimivirus, a virus as big as the smallest known bacteria (La Scola *et al.* 2003). One of the defining properties of a virus is its capsid that protects the genome against a hostile environment. Numerous viruses from the small STNV to the huge Mimivirus utilize the jelly-roll motif in the formation of their capsids. However, there are probably even more bacteriophages that utilize the HK97 motif for their major capsid protein. Furthermore, the HK97 motif is also used by some dsDNA mammalian viruses such as herpesvirus (Baker *et al.* 2005) in the structure of their major capsid protein. Although the jelly-roll motif and the HK97 motif are the most common capsid protein motifs, there are also other lineages such as the E glycoprotein of icosahedral enveloped alpha- and flaviviruses, the four helical bundle in TMV and hepatitis B virus, and the capsid structure found in some $T=1$ icosahedral RNA phages such as MS2 and Qβ (Plevka *et al.* 2009; Stockley *et al.* 1995; Valegård *et al.* 1990, 1994).

Viruses gather genes from their hosts as they evolve. In essence, a viral genome is a collection of useful tools that can be kept in an old tool chest. The above account shows that the structure of the capsid is evolving less rapidly than the proteins that are associated with a specific capsid motif. Thus a virus can be compared with an old tool chest whose contents can be changed more easily than the chest itself. An icosahedral viral capsid can be enlarged by increasing the T number or by adding equatorial capsomers to make a prolate head. Alternatively, the lipid envelope can be enlarged to assimilate additional components and larger genomes. Given this analogy, it is not surprising that large groups of rather different viruses evolved from a common motif for their capsid protein (Bamford *et al.* 2005; Benson *et al.* 2004; Nandhagopal *et al.* 2002) once a common

genetic code had been established. Nor is it surprising that the non-structural viral genes code proteins that are similar to proteins essential for other organisms different to that of the current host (e.g. the central domain of T4 gp5 is similar in structure to hen egg white lysozyme (Kanamaru *et al.* 2002)). The preference for the HK97 motif for prokaryotic viruses and of the jelly-roll motif for eukaryotic virus might suggest that the HK97 motif existed before the jelly-roll motif became established. Will future structural investigations find the missing link?

8. Acknowledgements

My deep appreciation and thanks go to all the many collaborators, post-docs, graduate students and others who have participated in the accomplishments over more than 50 years that form the basis of much of what is reported in this review. I wish to thank Sheryl Kelly, Shee-Mei Lok, and Siyang Sun for their help in preparing this paper. I also thank the National Science Foundation, National Institutes of Health and Purdue University for their generous financial support over many years.

9. References

Aathavan, K., Politzer, A. T., Kaplan, A., Moffitt, J. R., Chemla, Y. R., Grimes, S., Jardine, P. J., Anderson, D. L. & Bustamante, C. (2009). Substrate interactions and promiscuity in a viral DNA packaging motor. *Nature* **461**, 669–673.

Abad-Zapatero, C., Abdel-Meguid, S. S., Johnson, J. E., Leslie, A. G. W., Rayment, I., Rossmann, M. G., Suck, D. & Tsukihara, T. (1980). Structure of southern bean mosaic virus at 2·8 Å resolution. *Nature* **286**, 33–39.

Abrescia, N. G., Grimes, J. M., Kivela, H. M., Assenberg, R., Sutton, G. C., Butcher, S. J., Bamford, J. K., Bamford, D. H. & Stuart, D. I. (2008). Insights into virus evolution and membrane biogenesis from the structure of the marine lipid-containing bacteriophage PM2. *Molecular Cell* **31**, 749–761.

Abrescia, N. G. A., Cockburn, J. J. B., Grimes, J. M., Sutton, G. C., Diprose, J. M., Butcher, S. J., Fuller, S. D., San Martin, C., Burnett, R. M., Stuart, D. I., Bamford, D. H. & Bamford, J. K. H. (2004). Insights into assembly from structural analysis of bacteriophage PRD1. *Nature* **432**, 68–74.

Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D. & Brown, F. (1989). The three-dimensional structure of foot-and-mouth disease virus at 2·9 Å resolution. *Nature* **337**, 709–716.

Adrian, M., Dubochet, J., Lepault, J. & McDowall, A. W. (1984). Cryo-electron microscopy of viruses. *Nature* **308**, 32–36.

Agbandje, M., McKenna, R., Rossmann, M. G., Strassheim, M. L. & Parrish, C. R. (1993). Structure determination of feline panleukopenia virus empty particles. *Proteins* **16**, 155–171.

Agbandje-McKenna, M., Llamas-Saiz, A. L., Wang, F., Tattersall, P. & Rossmann, M. G. (1998). Functional implications of the structure of the murine parvovirus, minute virus of mice. *Structure* **6**, 1369–1381.

Åkervall, K., Strandberg, B., Rossmann, M. G., Bengtsson, U., Fridborg, K., Johannisen, H., Kannan, K. K., Lövgren, S., Petef, G., Öberg, B., Eaker, D., Hjertén, S., Rydén, L. & Moring, I. (1972). X-ray diffraction studies of the structure of satellite tobacco necrosis virus. *Cold Spring Harbor Symposia on Quantitative Biology* **36**, 469–483.

Aksyuk, A. A., Kurochkina, L. P., Fokine, A., Forouhar, F., Mesyanzhinov, V. V., Tong, L. & Rossmann, M. G. (2011). Structural conservation of the *Myoviridae* phage tail sheath protein fold. *Structure* **19**, 1885–1894.

Aksyuk, A. A., Leiman, P. G., Kurochkina, L. P., Schneider, M. M., Kostyuchenko, V. A., Mesyanzhinov, V. V. & Rossmann, M. G. (2009a). The tail sheath structure of bacteriophage T4: a molecular machine for infecting bacteria. *EMBO Journal* **28**, 821–829.

Aksyuk, A. A., Leiman, P. G., Shneider, M. M., Mesyanzhinov, V. V. & Rossmann, M. G. (2009b). The structure of gene product 6 of bacteriophage T4, the hinge-pin of the baseplate. *Structure* **17**, 800–808.

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W. & Heinz, F. X. (2001). Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. *Journal of Virology* **75**, 4268–4275.

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C. & Heinz, F. X. (1995). Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *Journal of Virology* **69**, 695–700.

Allison, S. L., Stiasny, K., Stadler, K., Mandl, C. W. & Heinz, F. X. (1999). Mapping of functional elements in

the stem-anchor region of tick-borne encephalitis virus envelope protein E. *Journal of Virology* **73**, 5605–5612.

Argos, P. & Rossmann, M. G. (1974). Determining heavy-atom positions using non-crystallographic symmetry. *Acta Crystallographica Section A* **30**, 672–677.

Arndt, U. W. & Wonacott, A. J. (1977). *The Rotation Method in Crystallography: Data Collection from Macromolecular Crystals*. Amsterdam: North-Holland.

Arnold, E., Himmel, D. M. & Rossmann, M. G. (2012). *International Tables for Crystallography, Volume F, Crystallography of Biological Macromolecules*. West Sussex, UK: John Wiley & Sons, Ltd.

Arnold, E., Vriend, G., Luo, M., Griffith, J. P., Kamer, G., Erickson, J. W., Johnson, J. E. & Rossmann, M. G. (1987). The structure determination of a common cold virus, human rhinovirus 14. *Acta Crystallographica Section A* **43**, 346–361.

Athappilly, F. K., Murali, R., Rux, J. J., Cai, Z. & Burnett, R. M. (1994). The refined crystal structure of hexon, the major coat protein of adenovirus type 2, at 2·9 Å resolution. *Journal of Molecular Biology* **242**, 430–455.

Baker, M. L., Jiang, W., Rixon, F. J. & Chiu, W. (2005). Common ancestry of herpesvirus and tailed DNA bacteriophages. *Journal of Virology* **79**, 14967–14970.

Baker, T. S., Caspar, D. L. & Murakami, W. T. (1983). Polyoma virus 'hexamer' tubes consist of paired pentamers. *Nature* **303**, 446–448.

Baker, T. S., Newcomb, W. W., Olson, N. H., Cowsert, L. M., Olson, C. & Brown, J. C. (1991). Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophysical Journal* **60**, 1445–1456.

Bamford, D. H., Grimes, J. M. & Stuart, D. I. (2005). What does structure tell us about virus evolution? *Current Opinion in Structural Biology* **15**, 655–663.

Bartual, S. G., Otero, J. M., Garcia-Doval, C., Llamas-Saiz, A. L., Kahn, R., Fox, G. C. & van Raaij, M. J. (2010). Structure of the bacteriophage T4 long tail fiber receptor-binding tip. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 20287–20292.

Battisti, A. J., Meng, G., Winkler, D. C., McGinnes, L. W., Plevka, P., Steven, A. C., Morrison, T. G. & Rossmann, M. G. (2012). Structure and assembly of a paramyxovirus matrix protein. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13996–14000.

Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M. & Rossmann, M. G. (1998). The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 4140–4145.

Bella, J., Kolatkar, P. R., Marlor, C. W., Greve, J. M. & Rossmann, M. G. (1999). The structure of the two amino-terminal domains of human intercellular adhesion molecule-1 suggests how it functions as a rhinovirus receptor. *Virus Research* **62**, 107–117.

Benson, S. D., Bamford, J. K. H., Bamford, D. H. & Burnett, R. M. (2004). Does common architecture reveal a viral lineage spanning all three domains of life? *Molecular Cell* **16**, 673–685.

Berget, P. B. & King, J. (1978). Isolation and characterization of precursors in T4 baseplate assembly. The complex of gene 10 and gene 11 products. *Journal of Molecular Biology* **124**, 469–486.

Bernal, J. D. & Fankuchen, I. (1937). Structure types of protein crystals from virus-infected plants. *Nature* **139**, 923–924.

Bernal, J. D. & Fankuchen, I. (1941). X-Ray and crystallographic studies of plant virus preparations : I. Introduction and preparation of specimens. II. Modes of aggregation of the virus particles. III. (1) The structure of the particles and (2) biological implications. *Journal of General Physiology* **25**, 111–165.

Bloomer, A. C., Champness, J. N., Bricogne, G., Staden, R. & Klug, A. (1978). Protein disk of tobacco mosaic virus at 2·8 Å resolution showing the interactions within and between subunits. *Nature* **276**, 362–368.

Bostina, M., Bubeck, D., Schwartz, C., Nicastro, D., Filman, D. J. & Hogle, J. M. (2007). Single particle cryoelectron tomography characterization of the structure and structural variability of poliovirus-receptor-membrane complex at 30 Å resolution. *Journal of Structural Biology* **160**, 200–210.

Bostina, M., Levy, H., Filman, D. J. & Hogle, J. M. (2011). Poliovirus RNA is released from the capsid near a twofold symmetry axis. *Journal of Virology* **83**, 776–783.

Bragg, W. H. & Bragg, W. L. (1913). The reflection of X-rays in crystals. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **88**, 428–438.

Bragg, W. L. (1913). The structure of some crystals as indicated by their diffraction of X-rays. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* **89**, 248–277.

Brenner, S. & Horne, R. W. (1959). A negative staining method for high resolution electron microscopy of viruses. *Biochimica et Biophysica Acta* **34**, 103–110.

Bressanelli, S., Stiasny, K., Allison, S. L., Stura, E. A., Duquerroy, S., Lescar, J., Heinz, F. X. & Rey, F. A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. *EMBO Journal* **23**, 728–738.

Brinton, M. A. (2002). The molecular biology of West Nile virus: a new invader of the western hemisphere. *Annual Review of Microbiology* **56**, 371–402.

Brown, J. C. & Newcomb, W. W. (2011). Herpesvirus capsid assembly: insights from structural analysis. *Current Opinion in Virology* **1**, 142–149.

Browning, C., Shneider, M. M., Bowman, V. D., Schwarzer, D. & Leiman, P. G. (2012). Phage pierces the host cell membrane with the iron-loaded spike. *Structure* **20**, 326–339.

Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1974). Structure determination of crystalline lobster D-glyceraldehyde-3-phosphate dehydrogenase. *Journal of Molecular Biology* **82**, 563–585.

Burke, D. S. & Monath, T. P. (2001). Flaviviruses. In *Fields Virology* (eds. D. M. Knipe & P. M. Howley), pp. 1043–1125. Philadelphia: Lippincott Williams & Wilkins.

Butler, P. J. & Klug, A. (1972). Assembly of tobacco mosaic virus *in vitro*: effect of state of polymerization of the protein component. *Proceedings of the National Academy of Sciences of the United States of America* **69**, 2950–2953.

Büttner, C. R., Chechik, M., Ortiz-Lombardia, M., Smits, C., Ebong, I. O., Chechik, V., Jeschke, G., Dykeman, E., Benini, S., Robinson, C. V., Alonso, J. C. & Antson, A. A. (2012). Structural basis for DNA recognition and loading into a viral packaging motor. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 811–816.

Calendar, R. (2006). *The Bacteriophages* (ed. R. Calendar). New York: Oxford University Press.

Canaan, S., Zadori, Z., Ghomashchi, F., Bollinger, J., Sadilek, M., Moreau, M. E., Tijssen, P. & Gelb, M. H. (2004). Interfacial enzymology of parvovirus phospholipases A2. *The Journal of Biological Chemistry* **279**, 14502–14508.

Caspar, D. L. D. (1956). Structure of bushy stunt virus. *Nature* **177**, 475–476.

Caspar, D. L. D. & Klug, A. (1962). Physical principles in the construction of regular viruses. *Cold Spring Harbor Symposia on Quantitative Biology* **27**, 1–24.

Champness, J. N., Bloomer, A. C., Bricogne, G., Butler, P. G. & Klug, A. (1976). The structure of the protein disk of tobacco mosaic virus to 5 Å resolution. *Nature* **259**, 20–24.

Chapman, M. S. & Rossmann, M. G. (1996). Structural refinement of the DNA-containing capsid of canine parvovirus using *RSRef*, a resolution-dependent stereochemically restrained real-space refinement method. *Acta Crystallographica Section D: Biological Crystallography* **52**, 129–142.

Chapman, M. S., Tsao, J. & Rossmann, M. G. (1992). *Ab initio* phase determination for spherical viruses: parameter determination for spherical-shell models. *Acta Crystallogaphica Section A* **48**, 301–312.

Chen, Z. G., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonossoff, G. & Johnson, J. E. (1989). Protein-RNA interactions in an icosahedral virus at 3·0 Å resolution. *Science* **245**, 154–159.

Cheng, R. H., Kuhn, R. J., Olson, N. H., Rossmann, M. G., Choi, H. K., Smith, T. J. & Baker, T. S. (1995). Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell* **80**, 621–630.

Cherrier, M. V., Kostyuchenko, V. A., Xiao, C., Bowman, V. D., Battisti, A. J., Yan, X., Chipman, P. R., Baker, T. S., Van Etten, J. L. & Rossmann, M. G. (2009). An icosahedral algal virus has a complex unique vertex decorated by a spike. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 11085–11089.

Choi, H. K., Tong, L., Minor, W., Dumas, P., Boege, U., Rossmann, M. G. & Wengler, G. (1991). Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. *Nature* **354**, 37–43.

Cockburn, J. J. B., Abrescia, N. G. A., Grimes, J. M., Sutton, G. C., Diprose, J. M., Benevides, J. M., Thomas Jr, G. J., Bamford, J. K. H., Bamford, D. H. & Stuart, D. I. (2004). Membrane structure and interactions with protein and DNA in bacteriophage PRD1. *Nature* **432**, 122–125.

Cohen, D. N., Sham, Y. Y., Haugstad, G. D., Xiang, Y., Rossmann, M. G., Anderson, D. L. & Popham, D. L. (2009). Shared catalysis in virus entry and bacterial cell wall depolymerization. *Journal of Molecular Biology* **387**, 607–618.

Colman, P. M., Varghese, J. N. & Laver, W. G. (1983). Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature* **303**, 41–44.

Coombs, D. H. & Arisaka, F. (1994). T4 tail structure and function. In *Molecular Biology of Bacteriophage T4* (ed. J. D. Karam), pp. 259–281. Washington, DC: American Society for Microbiology.

Crick, F. H. C. & Watson, J. D. (1956). Structure of small viruses. *Nature* **177**, 473–475.

Crick, F. H. C. & Watson, J. D. (1957). Virus structure: general principles. In *Ciba Foundation Symposium on the Nature of Viruses* (eds. G. E. W. Wolstenholme, E. C. P. Millar and Ciba Foundation), pp. 5–13. London: J. & A. Churchill.

Crowfoot, D. & Schmidt, G. M. J. (1945). X-ray crystallographic measurements on a single crystal of a tobacco necrosis virus derivative. *Nature* **155**, 504–505.

Crowther, R. A., Amos, L. A., Finch, J. T., DeRosier, D. J. & Klug, A. (1970). Three-dimensional reconstructions of spherical viruses by Fourier synthesis from electron micrographs. *Nature* **226**, 421–425.

Crowther, R. A., Lenk, E. V., Kikuchi, Y. & King, J. (1977). Molecular reorganization in the hexagon to star transition of the baseplate of bacteriophage T4. *Journal of Molecular Biology* **116**, 489–523.

Davis, C. W., Nguyen, H. Y., Hanna, S. L., Sánchez, M. D., Doms, R. W. & Pierson, T. C. (2006). West Nile virus discriminates between DC-SIGN

and DC-SIGNR for cellular attachment and infection. *Journal of Virology* **80**, 1290–1301.

Dessen, A., Volchkov, V., Dolnik, O., Klenk, H.-D. & Weissenhorn, W. (2000). Crystal structure of the matrix protein VP40 from Ebola virus. *EMBO Journal* **19**, 4228–4236.

DiMattia, M. A., Nam, H. J., Van Vliet, K., Mitchell, M., Bennett, A., Gurda, B. L., McKenna, R., Olson, N. H., Sinkovits, R. S., Potter, M., Byrne, B. J., Aslanidi, G., Zolotukhin, S., Muzyczka, N., Baker, T. S. & Agbandje-McKenna, M. (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. *Journal of Virology* **86**, 6947–6958.

Diprose, J. M., Grimes, J. M., Sutton, G. C., Burroughs, J. N., Meyer, A., Maan, S., Mertens, P. P. & Stuart, D. I. (2002). The core of bluetongue virus binds double-stranded RNA. *Journal of Virology* **76**, 9533–9536.

Dokland, T., McKenna, R., Ilag, L. L., Bowman, B. R., Incardona, N. L., Fane, B. A. & Rossmann, M. G. (1997). Structure of a viral procapsid with molecular scaffolding. *Nature* **389**, 308–313.

Dokland, T., Walsh, M., Mackenzie, J. M., Khromykh, A. A., Ee, K.-H. & Wang, S. (2004). West Nile virus core protein: tetramer structure and ribbon formation. *Structure* **12**, 1157–1163.

Dorsch, S., Liebisch, G., Kaufmann, B., von Landenberg, P., Hoffmann, J. H., Drobnik, W. & Modrow, S. (2002). The VP1 unique region of parvovirus B19 and its constituent phospholipase A2-like activity. *Journal of Virology* **76**, 2014–2018.

Dubochet, J., Adrian, M., Chang, J. J., Homo, J. C., Lepault, J., McDowall, A. W. & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics* **21**, 129–228.

Eiserling, F. A. & Black, L. W. (1994). Pathways in T4 morphogenesis. In *Molecular Biology of Bacteriophage T4* (eds. J. D. Karam, J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, C. Carlson & E. S. Miller), pp. 209–212. Washington, DC: American Society for Microbiology.

Elshuber, S., Allison, S. L., Heinz, F. X. & Mandl, C. W. (2003). Cleavage of protein prM is necessary for infection of BHK-21 cells by tick-borne encephalitis virus. *Journal of General Virology* **84**, 183–191.

Ferlenghi, I., Clarke, M., Ruttan, T., Allison, S. L., Schalich, J., Heinz, F. X., Harrison, S. C., Rey, F. A. & Fuller, S. D. (2001). Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus. *Molecular Cell* **7**, 593–602.

Filman, D. J., Syed, R., Chow, M., Macadam, A. J., Minor, P. D. & Hogle, J. M. (1989). Structural factors that control conformational transitions and serotype specificity in type 3 poliovirus. *EMBO Journal* **18**, 1567–1579.

Finch, J. T. & Klug, A. (1959). Structure of poliomyelitis virus. *Nature* **183**, 1709–1714.

Fokine, A., Battisti, A. J., Kostyuchenko, V. A., Black, L. W. & Rossmann, M. G. (2006). Cryo-EM structure of a bacteriophage T4 gp24 bypass mutant: the evolution of pentameric vertex proteins in icosahedral viruses. *Journal of Structural Biology* **154**, 255–259.

Fokine, A., Chipman, P. R., Leiman, P. G., Mesyanzhinov, V. V., Rao, V. B. & Rossmann, M. G. (2004). Molecular architecture of the prolate head of bacteriophage T4. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6003–6008.

Fokine, A., Miroshnikov, K. A., Shneider, M. M., Mesyanzhinov, V. V. & Rossmann, M. G. (2008). Structure of the bacteriophage phiKZ lytic transglycosylase, gp144. *Journal of Biological Chemistry* **283**, 7242–7250.

Fokine, A., Zhang, Z., Kanamaru, S., Bowman, V. D., Aksyuk, A. A., Arisaka, F., Rao, V. B. & Rossmann, M. G. (2013). The molecular architecture of the bacteriophage T4 neck. *Journal of Molecular Biology* **425**, 1731–1744.

Franklin, R. E. & Holmes, K. C. (1958). Tobacco mosaic virus: application of the method of isomorphous replacement to the determination of the helical parameters and radial density distribution. *Acta Crystallographica* **11**, 213–220.

Franklin, R. E., Klug, A. & Holmes, K. C. (1957). X-ray diffraction studies of the structure and morphology of tobacco mosaic virus. In *Ciba Foundation Symposium on the Nature of Viruses* (eds. G. E. W. Wolstenholme, E. C. P. Millar and Ciba Foundation), pp. 39–55. London: J. & A. Churchill.

Fu, C. Y. & Johnson, J. E. (2012). Structure and cell biology of archaeal virus STIV. *Current Opinion in Virology* **2**, 122–127.

Fuller, D. N., Raymer, D. M., Kottadiel, V. I., Rao, V. B. & Smith, D. E. (2007). Single phage T4 DNA packaging motors exhibit large force generation, high velocity, and dynamic variability. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 16868–16873.

Fuller, S. D. (1987). The $T=4$ envelope of Sindbis virus is organized by interactions with a complementary T = 3 capsid. *Cell* **48**, 923–934.

Gaykema, W. P., Volbeda, A. & Hol, W. G. (1986). Structure determination of *Panulirus interruptus* haemocyanin at 3·2 Å resolution. Successful phase extension by sixfold density averaging. *Journal of Molecular Biology* **187**, 255–275.

Gibbons, D. L. & Kielian, M. (2002). Molecular dissection of the Semliki Forest virus homotrimer reveals two functionally distinct regions of the fusion protein. *Journal of Virology* **76**, 1194–1205.

Greve, J. M., Davis, G., Meyer, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarck, M. E. &

McClelland, A. (1989). The major human rhinovirus receptor is ICAM-1. *Cell* **56**, 839–847.

Grimes, J. M., Burroughs, J. N., Gouet, P., Diprose, J. M., Malby, R., Ziéntara, S., Mertens, P. P. C. & Stuart, D. I. (1998). The atomic structure of the bluetongue virus core. *Nature* **395**, 470–478.

Gubler, D. J. (1988). Dengue. In *Epidemiology of Arthropod-Borne Viral Diseases* (ed. T. P. Monath), pp. 223–260. Boca Raton, FL: CRC Press.

Guirakhoo, F., Bolin, R. A. & Roehrig, J. T. (1992). The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein. *Virology* **191**, 921–931.

Guirakhoo, F., Heinz, F. X., Mandl, C. W., Holzmann, H. & Kunz, C. (1991). Fusion activity of flaviviruses: comparison of mature and immature (prM-containing) tick-borne encephalitis virions. *Journal of General Virology* **72**, 1323–1329.

Hafenstein, S., Bowman, V. D., Sun, T., Nelson, C. D. S., Palermo, L., Chipman, P. R., Battisti, A. J., Parrish, C. R. & Rossmann, M. G. (2009). The structural analysis of eight different FAb interacting with parvovirus capsids reveals alternative mechanisms of neutralization and binding to host range variant capsids. *Journal of Virology* **83**, 5556–5566.

Hafenstein, S., Palermo, L. M., Kostyuchenko, V. A., Xiao, C., Morais, M. C., Nelson, C. D. S., Bowman, V. D., Battisti, A. J., Chipman, P. R., Parrish, C. R. & Rossmann, M. G. (2007). Asymmetric binding of transferrin receptor to parvovirus capsids. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 6585–6589.

Harris, A., Forouhar, F., Qiu, S., Sha, B. & Luo, M. (2001). The crystal structure of the influenza matrix protein M1 at neutral pH: M1-M1 protein interfaces can rotate in the oligomeric structures of M1. *Virology* **289**, 34–44.

Harrison, S. C. (1968). A point-focusing camera for single-crystal diffraction. *Journal of Applied Crystallography* **1**, 84–90.

Harrison, S. C. (2008). Viral membrane fusion. *Nature Structural and Molecular Biology* **15**, 690–698.

Harrison, S. C. (2010). Virology. Looking inside adenovirus. *Science* **329**, 1026–1027.

Harrison, S. C., Olson, A. J., Schutt, C. E., Winkler, F. K. & Bricogne, G. (1978). Tomato bushy stunt virus at 2·9 Å resolution. *Nature* **276**, 368–373.

Harrison, S. C., Strong, R. K., Schlesinger, S. & Schlesinger, M. J. (1992). Crystallization of Sindbis virus and its nucleocapsid. *Journal of Molecular Biology* **226**, 277–280.

Hatfull, G. F., Pedulla, M. L., Jacobs-Sera, D., Cichon, P. M., Foley, A., Ford, M. E., Gonda, R. M., Houtz, J. M., Hryckowian, A. J., Kelchner, V. A., Namburi, S., Pajcini, K. V., Popovich, M. G., Schleicher, D. T., Simanek, B. Z., Smith, A. L., Zdanowicz, G. M., Kumar, V., Peebles, C. L., Jacobs Jr, W. R., Lawrence, J. G. & Hendrix, R. W. (2006). Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genetics* **2**, e92.

Hayashi, M. A. (1978). Morphogenesis of the isometric phages. In *Cold Spring Harbor Monograph: The Single-Stranded DNA Phages*, vol. 8 (eds. D. T. Denhardt, D. Dressler & D. S. Ray), pp. 531–547. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Hayden, F. G., Herrington, D. T., Coats, T. L., Kim, K. H., Cooper, E. C., Villano, S. A., Liu, S., Hudson, S., Pevear, D. C., Collett, M. & McKinlay, M. A. (2003). Efficacy and safety of oral pleconaril for treatment of picornavirus colds in adults: results of two double-blind, randomized, placebo-controlled trials. *Clinical Infectious Diseases* **36**, 1523–1532.

He, Y., Bowman, V. D., Mueller, S., Bator, C. M., Bella, J., Peng, X., Baker, T. S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2000). Interaction of the poliovirus receptor with poliovirus. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 79–84.

He, Y., Chipman, P. R., Howitt, J., Bator, C. M., Whitt, M. A., Baker, T. S., Kuhn, R. J., Anderson, C. W., Freimuth, P. & Rossmann, M. G. (2001). Interaction of coxsackievirus B3 with the full-length coxsackievirus-adenovirus receptor. *Nature Structural Biology* **8**, 874–878.

Heinz, F. X., Stiasny, K., Püschner-Auer, G., Holzmann, H., Allison, S. L., Mandl, C. W. & Kunz, C. (1994). Structural changes and functional control of the tick-borne encephalitis virus glycoprotein E by the heterodimeric association with protein prM. *Virology* **198**, 109–117.

Hermodson, M. A., Abad-Zapatero, C., Abdel-Meguid, S. S., Pundak, S., Rossmann, M. G. & Tremaine, J. H. (1982). Amino acid sequence of southern bean mosaic virus coat protein and its relation to the three-dimensional structure of the virus. *Virology* **119**, 133–149.

Hogle, J. M. (2002). Poliovirus cell entry: common structural themes in viral cell entry pathways. *Annual Review of Microbiology* **56**, 677–702.

Hogle, J. M., Chow, M. & Filman, D. J. (1985). Three-dimensional structure of poliovirus at 2·9 Å resolution. *Science* **229**, 1358–1365.

Holmes, K. C., Stubbs, G. J., Mandelkow, E. & Gallwitz, U. (1975). Structure of tobacco mosaic virus at 6·7 Å resolution. *Nature* **254**, 192–196.

Horne, R. W., Brenner, S., Waterson, A. P. & Wildy, P. (1959). The icosahedral form of an adenovirus. *Journal of Molecular Biology* **1**, 84–86.

Hueffer, K., Govindasamy, L., Agbandje-McKenna, M. & Parrish, C. R. (2003). Combinations of two capsid regions controlling canine host range determine canine transferrin receptor binding by canine and feline parvoviruses. *Journal of Virology* **77**, 10099–10105.

Huxley, H. E. & Zubay, G. (1960). The structure of the protein shell of turnip yellow mosaic virus. *Journal of Molecular Biology* **2**, 189–196.

Jiang, W., Chang, J., Jakana, J., Weigele, P., King, J. & Chiu, W. (2006). Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* **439**, 612–616.

Jiang, W., Li, Z., Zhang, Z., Baker, M. L., Prevelige Jr, P. E. & Chiu, W. (2003). Coat protein fold and maturation transition of bacteriophage P22 seen at sub-nanometer resolutions. *Nature Structural Biology* **10**, 131–135.

Kanamaru, S., Kondabagil, K., Rossmann, M. G. & Rao, V. B. (2004). The functional domains of bacteriophage T4 terminase. *Journal of Biological Chemistry* **279**, 40795–40801.

Kanamaru, S., Leiman, P. G., Kostyuchenko, V. A., Chipman, P. R., Mesyanzhinov, V. V., Arisaka, F. & Rossmann, M. G. (2002). Structure of the cell-puncturing device of bacteriophage T4. *Nature* **415**, 553–557.

Kaufmann, B., Baxa, U., Chipman, P. R., Rossmann, M. G., Modrow, S. & Seckler, R. (2005). Parvovirus B19 does not bind to membrane-associated globoside *in vitro*. *Virology* **332**, 189–198.

Kaufmann, B., Bowman, V. D., Li, Y., Szelei, J., Waddell, P. J., Tijssen, P. & Rossmann, M. G. (2010a). Structure of *Penaeus stylirostris* densovirus, a shrimp pathogen. *Journal of Virology* **84**, 11289–11296.

Kaufmann, B., Chipman, P. R., Holdaway, H. A., Johnson, S., Fremont, D. H., Kuhn, R. J., Diamond, M. S. & Rossmann, M. G. (2009). Capturing a flavivirus pre-fusion intermediate. *PLoS Pathogens* **5**, e1000672.

Kaufmann, B., Chipman, P. R., Kostyuchenko, V. A., Modrow, S. & Rossmann, M. G. (2008). Visualization of the externalized VP2 N-termini of infectious human parvovirus B19. *Journal of Virology* **82**, 7306–7312.

Kaufmann, B., El-Far, M., Plevka, P., Bowman, V. D., Li, Y., Tijssen, P. & Rossmann, M. G. (2011). Structure of *Bombyx mori* densovirus 1, a silkworm pathogen. *Journal of Virology* **85**, 4691–4697.

Kaufmann, B., Nybakken, G. E., Chipman, P. R., Zhang, W., Diamond, M. S., Fremont, D. H., Kuhn, R. J. & Rossmann, M. G. (2006). West Nile virus in complex with the Fab fragment of a neutralizing monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 12400–12404.

Kaufmann, B., Plevka, P., Kuhn, R. J. & Rossmann, M. G. (2010b). Crystallization and preliminary X-ray diffraction analysis of West Nile virus. *Acta Crystallographica Section F* **66**, 558–562.

Kaufmann, B. & Rossmann, M. G. (2011). Molecular mechanisms involved in the early steps of flavivirus cell entry. *Microbes and Infection* **13**, 1–9.

Kaufmann, B., Simpson, A. A. & Rossmann, M. G. (2004). The structure of human parvovirus B19. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 11628–11633.

Kay, L. E. (1986). W. M. Stanley's crystallization of the tobacco mosaic virus, 1930–1940. *Isis* **77**, 450–472.

Kellenberger, E. (1976). DNA viruses: cooperativity and regulation through conformational changes as features of phage assembly. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **276**, 3–13.

Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **181**, 662–666.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. & Shore, V. C. (1960). Structure of myoglobin. A three-dimensional Fourier synthesis at 2 Å resolution. *Nature* **185**, 422–427.

Khayat, R., Tang, L., Larson, E. T., Lawrence, C. M., Young, M. & Johnson, J. E. (2005). Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 18944–18949.

Kielian, M. (1995). Membrane fusion and the alphavirus life cycle. *Advances in Virus Research* **45**, 113–151.

Kielian, M. & Rey, F. A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. *Nature Reviews Microbiology* **4**, 67–76.

Kikuchi, Y. & King, J. (1975). Assembly of the tail of bacteriophage T4. *Journal of Supramolecular Structure* **3**, 24–38.

Klose, T., Kuznetsov, Y. G., Xiao, C., Sun, S., McPherson, A. & Rossmann, M. G. (2010). The three-dimensional structure of Mimivirus. *Intervirology* **53**, 268–273.

Klug, A. (1972). Interpretation of the rotation function map of satellite tobacco necrosis virus: octahedral packing of icosahedral particles. *Cold Spring Harbor Symposia on Quantitative Biology* **36**, 483–487.

Klug, A., Finch, J. T. & Franklin, R. E. (1957). The structure of turnip yellow mosaic virus: X-ray diffraction studies. *Biochimica et Biophysica Acta* **25**, 242–252.

Kolatkar, P. R., Bella, J., Olson, N. H., Bator, C. M., Baker, T. S. & Rossmann, M. G. (1999). Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. *EMBO Journal* **18**, 6249–6259.

Kostyuchenko, V. A., Chipman, P. R., Leiman, P. G., Arisaka, F., Mesyanzhinov, V. V. & Rossmann, M. G. (2005). The tail structure of bacteriophage T4 and its mechanism of contraction. *Nature Structural and Molecular Biology* **12**, 810–813.

Kostyuchenko, V. A., Leiman, P. G., Chipman, P. R., Kanamaru, S., van Raaij, M. J., Arisaka, F., Mesyanzhinov, V. V. & Rossmann, M. G. (2003).

Three-dimensional structure of bacteriophage T4 baseplate. *Nature Structural Biology* **10**, 688–693.

Kuhn, R. J. (2007). Chapter 30: *Togaviridae*: the viruses and their replication. In *Fields Virology* (eds. D. M. Knipe & P. M. Howley), pp. 1001–1021. Philadelphia, PA: Lippincott Williams & Wilkins.

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V., Corver, J., Lenches, E., Jones, C. T., Mukhopadhyay, S., Chipman, P. R., Strauss, E. G., Baker, T. S. & Strauss, J. H. (2002). Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell* **108**, 717–725.

Kuznetsov, Y. G., Xiao, C., Sun, S., Raoult, D., Rossmann, M. G. & McPherson, A. (2010). Atomic force microscopy investigation of the giant mimivirus. *Virology* **404**, 127–137.

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., Birtles, R., Claverie, J. M. & Raoult, D. (2003). A giant virus in amoebae. *Science* **299**, 2033.

Lamb, R. A. & Jardetzky, T. S. (2007). Structural basis of viral invasion: lessons from paramyxovirus F. *Current Opinion in Structural Biology* **17**, 427–436.

Larson, S. B., Koszelak, S., Day, J., Greenwood, A., Dodds, J. A. & McPherson, A. (1993). Double-helical RNA in satellite tobacco mosaic virus. *Nature* **361**, 179–182.

Lebedev, A. A., Krause, M. H., Isidro, A. L., Vagin, A. A., Orlova, E. V., Turner, J., Dodson, E. J., Tavares, P. & Antson, A. A. (2007). Structural framework for DNA translocation via the viral portal protein. *EMBO Journal* **26**, 1984–1994.

Lee, K. K. & Johnson, J. E. (2003). Complementary approaches to structure determination of icosahedral viruses. *Current Opinion in Structural Biology* **13**, 558–569.

Leiman, P. G., Arisaka, F., van Raaij, M. J., Kostyuchenko, V. A., Aksyuk, A. A., Kanamaru, S. & Rossmann, M. G. (2010). Morphogenesis of the T4 tail and tail fibers. *Virology Journal* **7**, 355.

Leiman, P. G., Chipman, P. R., Kostyuchenko, V. A., Mesyanzhinov, V. V. & Rossmann, M. G. (2004). Three-dimensional rearrangement of proteins in the tail of bacteriophage T4 on infection of its host. *Cell* **118**, 419–429.

Leiman, P. G., Kanamaru, S., Mesyanzhinov, V. V., Arisaka, F. & Rossmann, M. G. (2003). Structure and morphogenesis of bacteriophage T4. *Cellular and Molecular Life Sciences* **60**, 2356–2370.

Lerch, T. F., O'Donnell, J. K., Meyer, N. L., Xie, Q., Taylor, K. A., Stagg, S. M. & Chapman, M. S. (2012). Structure of AAV-DJ, a retargeted gene therapy vector: cryo-electron microscopy at 4·5 Å resolution. *Structure* **20**, 1310–1320.

Lescar, J., Roussel, A., Wein, M. W., Navaza, J., Fuller, S. D., Wengler, G., Wengler, G. & Rey, F. A. (2001). The fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* **105**, 137–148.

Li, L., Jose, J., Xiang, Y., Kuhn, R. J. & Rossmann, M. G. (2010). Structural changes of envelope proteins during alphavirus fusion. *Nature* **468**, 705–708.

Li, L., Lok, S.-M., Yu, I.-M., Zhang, Y., Kuhn, R. J., Chen, J. & Rossmann, M. G. (2008). The flavivirus precursor membrane-envelope protein complex: structure and maturation. *Science* **319**, 1830–1834.

Liljas, L., Unge, T., Jones, T. A., Fridborg, K., Lövgren, S., Skoglund, U. & Strandberg, B. (1982). Structure of satellite tobacco necrosis virus at 3·0 Å resolution. *Journal of Molecular Biology* **159**, 93–108.

Lindenbach, B. D. & Rice, C. M. (2001). *Flaviviridae*: the viruses and their replication. In *Fields Virology* (eds. D. M. Knipe & P. M. Howley), pp. 991–1041. Philadelphia: Lippincott Williams & Wilkins.

Lindenbach, B. D. & Rice, C. M. (2003). Molecular biology of flaviviruses. *Advances in Virus Research* **59**, 23–61.

Liu, H., Jin, L., Koh, S. B., Atanasov, I., Schein, S., Wu, L. & Zhou, Z. H. (2010). Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329**, 1038–1043.

Llamas-Saiz, A. L., Agbandje-McKenna, M., Wikoff, W. R., Bratton, J., Tattersall, P. & Rossmann, M. G. (1997). Structure determination of minute virus of mice. *Acta Crystallographica Section D: Biological Crystallography* **53**, 93–102.

Lorenz, I. C., Allison, S. L., Heinz, F. X. & Helenius, A. (2002). Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum. *Journal of Virology* **76**, 5480–5491.

Lozach, P. Y., Burleigh, L., Staropoli, I., Navarro-Sanchez, E., Harriague, J., Virelizier, J. L., Rey, F. A., Despres, P., Arenzana-Seisdedos, F. & Amara, A. (2005). Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN)-mediated enhancement of dengue virus infection is independent of DC-SIGN internalization signals. *Journal of Biological Chemistry* **280**, 23698–23708.

Lu, Y. E., Cassese, T. & Kielian, M. (1999). The cholesterol requirement for Sindbis virus entry and exit and characterization of a spike protein region involved in cholesterol dependence. *Journal of Virology* **73**, 4272–4278.

Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M. G., Boege, U., Scraba, D. G., Duke, G. M. & Palmenberg, A. C. (1987). The atomic structure of Mengo virus at 3·0 Å resolution. *Science* **235**, 182–191.

Ma, L., Jones, C. T., Groesch, T. D., Kuhn, R. J. & Post, C. B. (2004). Solution structure of dengue virus capsid protein reveals another fold. *Proceedings of the*

National Academy of Sciences of the United States of America **101**, 3414–3419.

Magdoff, B. S. (1960). Sub-units in southern bean mosaic virus. *Nature* **185**, 673–674.

Main, P. & Rossmann, M. G. (1966). Relationships among structure factors due to identical molecules in different crystallographic environments. *Acta Crystallographica* **21**, 67–72.

Mancini, E. J., Kainov, D. E., Grimes, J. M., Tuma, R., Bamford, D. H. & Stuart, D. I. (2004). Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **118**, 743–755.

Marsh, M. & Helenius, A. (1989). Virus entry into animal cells. *Advances in Virus Research* **36**, 107–151.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L., Krishnaswamy, S., Rossmann, M. G., Olson, N. H., Baker, T. S. & Incardona, N. L. (1992a). Atomic structure of single-stranded DNA bacteriophage φX174 and its functional implications. *Nature* **355**, 137–143.

McKenna, R., Xia, D., Willingmann, P., Ilag, L. L. & Rossmann, M. G. (1992b). Structure determination of the bacteriophage φX174. *Acta Crystallographica Section B* **48**, 499–511.

Modis, Y., Ogata, S., Clements, D. & Harrison, S. C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 6986–6991.

Modis, Y., Ogata, S., Clements, D. & Harrison, S. C. (2004). Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**, 313–319.

Modis, Y., Trus, B. L. & Harrison, S. C. (2002). Atomic model of the papillomavirus capsid. *EMBO Journal* **21**, 4754–4762.

Moffitt, J. R., Chemla, Y. R., Aathavan, K., Grimes, S., Jardine, P. J., Anderson, D. L. & Bustamante, C. (2009). Intersubunit coordination in a homomeric ring ATPase. *Nature* **457**, 446–450.

Money, V. A., McPhee, H. K., Mosely, J. A., Sanderson, J. M. & Yeo, R. P. (2009). Surface features of a *Mononegavirales* matrix protein indicate sites of membrane interaction. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 4441–4446.

Morais, M. C., Choi, K. H., Koti, J. S., Chipman, P. R., Anderson, D. L. & Rossmann, M. G. (2005). Conservation of the capsid structure in tailed dsDNA bacteriophages: the pseudo-atomic structure of φ29. *Molecular Cell* **18**, 149–159.

Morais, M. C., Koti, J. S., Bowman, V. D., Reyes-Aldrete, E., Anderson, D. L. & Rossmann, M. G. (2008). Defining molecular and domain boundaries in the bacteriophage φ29 DNA packaging motor. *Structure* **16**, 1267–1274.

Mosig, G. & Eiserling, F. (2006). T4 and related phages: structure and development. In *The Bacteriophages* (ed. R. Calendar), pp. 225–267. New York: Oxford University Press.

Muckelbauer, J. K., Kremer, M., Minor, I., Tong, L., Zlotnick, A., Johnson, J. E. & Rossmann, M. G. (1995). Structure determination of coxsackievirus B3 to 3·5 Å resolution. *Acta Crystallographica Section D: Biological Crystallography* **51**, 871–887.

Mukhopadhyay, S., Kim, B.-S., Chipman, P. R., Rossmann, M. G. & Kuhn, R. J. (2003). Structure of West Nile virus. *Science* **302**, 248.

Mukhopadhyay, S., Zhang, W., Gabler, S., Chipman, P. R., Strauss, E. G., Strauss, J. H., Baker, T. S., Kuhn, R. J. & Rossmann, M. G. (2006). Mapping the structure and function of the E1 and E2 glycoproteins in alphaviruses. *Structure* **14**, 63–73.

Nagai, Y., Klenk, H.-D. & Rott, R. (1976). Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of Newcastle disease virus. *Virology* **72**, 494–508.

Nandhagopal, N., Simpson, A. A., Gurnon, J. R., Yan, X., Baker, T. S., Graves, M. V., Van Etten, J. L. & Rossmann, M. G. (2002). The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14758–14763.

Neumann, P., Lieber, D., Meyer, S., Dautel, P., Kerth, A., Kraus, I., Garten, W. & Stubbs, M. T. (2009). Crystal structure of the Borna disease virus matrix protein (BDV-M) reveals ssRNA binding properties. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 3710–3715.

Nixon, H. L. & Gibbs, A. J. (1960). Electron microscope observations on the structure of turnip yellow mosaic virus. *Journal of Molecular Biology* **2**, 197–200.

Olia, A. S., Prevelige Jr, P. E., Johnson, J. E. & Cingolani, G. (2011). Three-dimensional structure of a viral genome-delivery portal vertex. *Nature Structural and Molecular Biology* **18**, 597–603.

Oliveira, M. A., Zhao, R., Lee, W., Kremer, M. J., Minor, I., Rueckert, R. R., Diana, G. D., Pevear, D. C., Dutko, F. J., McKinlay, M. A. & Rossmann, M. G. (1993). The structure of human rhinovirus 16. *Structure* **1**, 51–68.

Olson, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. (1993). Structure of a human rhinovirus complexed with its receptor molecule. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 507–511.

Padron, E., Bowman, V., Kaludov, N., Govindasamy, L., Levy, H., Nick, P., McKenna, R., Muzyczka, N., Chiorini, J. A., Baker, T. S. & Agbandje-McKenna, M. (2005). Structure of adeno-associated virus type 4. *Journal of Virology* **79**, 5047–5058.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. T. (1960). Structure of

haemoglobin. A three-dimensional Fourier synthesis at 5·5-Å resolution, obtained by X-ray analysis. *Nature* **185**, 416–422.

Pletnev, S. V., Zhang, W., Mukhopadhyay, S., Fisher, B. R., Hernandez, R., Brown, D. T., Baker, T. S., Rossmann, M. G. & Kuhn, R. J. (2001). Locations of carbohydrate sites on alphavirus glycoproteins show that E1 forms an icosahedral scaffold. *Cell* **105**, 127–136.

Plevka, P., Battisti, A. J., Winkler, D. C., Tars, K., Holdaway, H. A., Bator, C. M. & Rossmann, M. G. (2012a). Sample preparation induced artifacts in cryo-electron tomographs. *Microscopy and Microanalysis* **18**, 1043–1048.

Plevka, P., Hafenstein, S., Li, L., D'Abrgamo Jr, A., Cotmore, S. F., Rossmann, M. G. & Tattersall, P. (2011a). Structure of a packaging-defective mutant of minute virus of mice indicates that the genome is packaged via a pore at a 5-fold axis. *Journal of Virology* **85**, 4822–4827.

Plevka, P., Kaufmann, B. & Rossmann, M. G. (2011b). Analysis of phases in the structure determination of an icosahedral virus. *Acta Crystallographica Section D: Biological Crystallography* **67**, 568–577.

Plevka, P., Perera, R., Cardosa, J., Kuhn, R. J. & Rossmann, M. G. (2012b). Crystal structure of human enterovirus 71. *Science* **336**, 1274.

Plevka, P., Perera, R., Cardosa, J., Kuhn, R. J. & Rossmann, M. G. (2012c). Structure determination of enterovirus 71. *Acta Crystallographica Section D: Biological Crystallography* **68**, 1217–1222.

Plevka, P., Tars, K. & Liljas, L. (2009). Structure and stability of icosahedral particles of a covalent coat protein dimer of bacteriophage MS2. *Protein Science* **18**, 1653–61.

Pokidysheva, E., Zhang, Y., Battisti, A. J., Bator-Kelly, C. M., Chipman, P. R., Xiao, C., Gregorio, G. G., Hendrickson, W. A., Kuhn, R. J. & Rossmann, M. G. (2006). Cryo-EM reconstruction of dengue virus in complex with the carbohydrate recognition domain of DC-SIGN. *Cell* **124**, 485–493.

Rao, V. B. & Feiss, M. (2008). The bacteriophage DNA packaging motor. *Annual Review of Genetics* **42**, 647–681.

Rao, Z., Belyaev, A. S., Fry, E., Roy, P., Jones, I. M. & Stuart, D. I. (1995). Crystal structure of SIV matrix antigen and implications for virus assembly. *Nature* **378**, 743–747.

Rayment, I., Baker, T. S., Caspar, D. L. & Murakami, W. T. (1982). Polyoma virus capsid structure at 22·5 Å resolution. *Nature* **295**, 110–115 (issue cover).

Reddy, V. S., Natchiar, S. K., Stewart, P. L. & Nemerow, G. R. (2010). Crystal structure of human adenovirus at 3·5 Å resolution. *Science* **329**, 1071–1075.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C. & Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature* **375**, 291–298.

Roberts, M. M., White, J. L., Grütter, M. G. & Burnett, R. M. (1986). Three-dimensional structure of the adenovirus major coat protein hexon. *Science* **232**, 1148–1151.

Rossmann, M. G. (1972). *The Molecular Replacement Method: A Collection of Papers on the Use of Non-crystallographic Symmetry*. New York: Gordon & Breach.

Rossmann, M. G. (1979). Processing oscillation diffraction data for very large unit cells with an automatic convolution technique and profile fitting. *Journal of Applied Crystallography* **12**, 225–238.

Rossmann, M. G. (1984). Synchrotron radiation studies of large proteins and supramolecular structures. In *Biological Systems: Structure and Analysis* (ed. G. P. G. Diakun, C. D.), pp. 28–40. Daresbury: Science and Engineering Research Council.

Rossmann, M. G. (1989a). The canyon hypothesis. Hiding the host cell receptor attachment site on a viral surface from immune surveillance. *Journal of Biological Chemistry* **264**, 14587–14590.

Rossmann, M. G. (1989b). Determination of virus structures by the use of molecular replacement density averaging. In *Improving Protein Phases. Proceedings of the Study Weekend held at Daresbury, 1988* (eds. S. Bailey, E. Dodson & S. Phillips), pp. 49–56. Daresbury: Science and Engineering Research Council.

Rossmann, M. G. (1994). Viral cell recognition and entry. *Protein Science* **3**, 1712–1725.

Rossmann, M. G. (1999). Synchrotron radiation as a tool for investigating virus structures. *Journal of Synchrotron Radiation* **6**, 816–821.

Rossmann, M. G. (2001). Molecular replacement – historical background. *Acta Crystallographica Section D: Biological Crystallography* **57**, 1360–1366.

Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B. & Vriend, G. (1985). Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature* **317**, 145–153.

Rossmann, M. G., Battisti, A. J. & Plevka, P. (2011). Future prospects. In *Advances in Protein Chemistry and Structural Biology, Volume 82: Recent Advances in Electron Cryomicroscopy, Part B* (eds. S. J. Ludtke & B. V. V. Prasad), pp. 101–121. San Diego, CA: Academic Press.

Rossmann, M. G., Bella, J., Kolatkar, P. R., He, Y., Wimmer, E., Kuhn, R. J. & Baker, T. S. (2000). Minireview: cell recognition and entry by rhino- and enteroviruses. *Virology* **269**, 239–247.

Rossmann, M. G. & Blow, D. M. (1962). The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallographica* **15**, 24–31.

Rossmann, M. G. & Erickson, J. W. (1983). Oscillation photography of radiation-sensitive crystals using a synchrotron source. *Journal of Applied Crystallography* **16**, 629–636.

Rossmann, M. G., Ford, G. C., Watson, H. C. & Banaszak, L. J. (1972). Molecular symmetry of glyceraldehyde-3-phosphate dehydrogenase. *Journal of Molecular Biology* **64**, 237–249.

Rossmann, M. G., He, Y. & Kuhn, R. J. (2002). Picornavirus-receptor interactions. *Trends in Microbiology* **10**, 324–331.

Rossmann, M. G. & Johnson, J. E. (1989). Icosahedral RNA virus structure. *Annual Review of Biochemistry* **58**, 533–573.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). Processing and post-refinement of oscillation camera data. *Journal of Applied Crystallography* **12**, 570–581.

Rossmann, M. G., McKenna, R., Tong, L., Xia, D., Dai, J., Wu, H., Choi, H. K., Marinescu, D. & Lynch, R. E. (1992). Molecular replacement real space averaging. In *Molecular Replacement. Proceedings of the CCP4 Study Weekend, 31 January – 1 February 1992* (eds. E. Dodson, S. Gover & W. Wolf), pp. 33–48. Daresbury: Science and Engineering Research Council.

Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and biological evolution of a nucleotide-binding protein. *Nature* **250**, 194–199.

Rossmann, M. G. & Palmenberg, A. C. (1988). Conservation of the putative receptor attachment site in picornaviruses. *Virology* **164**, 373–382.

Roussel, A., Lescar, J., Vaney, M. C., Wengler, G. & Rey, F. A. (2006). Structure and interactions at the viral surface of the envelope protein E1 of Semliki Forest virus. *Structure* **14**, 75–86.

Roy, A., Bhardwaj, A., Datta, P., Lander, G. C. & Cingolani, G. (2012). Small terminase couples viral DNA binding to genome-packaging ATPase activity. *Structure* **20**, 1403–1413.

Scheid, A., Caliguiri, L. A., Compans, R. W. & Choppin, P. W. (1972). Isolation of paramyxovirus glycoproteins. Association of both hemagglutinating and neuraminidase activities with the larger SV5 glycoprotein. *Virology* **50**, 640–652.

Sherry, B., Mosser, A. G., Colonno, R. J. & Rueckert, R. R. (1986). Use of monoclonal antibodies to identify four neutralization immunogens on a common cold picornavirus, human rhinovirus 14. *Journal of Virology* **57**, 246–257.

Sherry, B. & Rueckert, R. (1985). Evidence for at least two dominant neutralization antigens on human rhinovirus 14. *Journal of Virology* **53**, 137–143.

Shnyrova, A. V., Ayllon, J., Mikhalyov, I. I., Villar, E., Zimmerberg, J. & Frolov, V. A. (2007). Vesicle formation by self-assembly of membrane-bound matrix proteins into a fluidlike budding domain. *Journal of Cell Biology* **179**, 627–633.

Simpson, A. A., Chandrasekar, V., Hébert, B., Sullivan, G. M., Rossmann, M. G. & Parrish, C. R. (2000a). Host range and variability of calcium binding by surface loops in the capsids of canine and feline parvoviruses. *Journal of Molecular Biology* **300**, 597–610.

Simpson, A. A., Chipman, P. R., Baker, T. S., Tijssen, P. & Rossmann, M. G. (1998). The structure of an insect parvovirus (*Galleria mellonella* densovirus) at 3·7 Å resolution. *Structure* **6**, 1355–1367.

Simpson, A. A., Hebert, B., Sullivan, G. M., Parrish, C. R., Zadori, Z., Tijssen, P. & Rossmann, M. G. (2002). The structure of porcine parvovirus: comparison with related viruses. *Journal of Molecular Biology* **315**, 1189–1198.

Simpson, A. A., Nandhagopal, N., Van Etten, J. L. & Rossmann, M. G. (2003). Structural analyses of *Phycodnaviridae* and *Iridoviridae*. *Acta Crystallographica Section D: Biological Crystallography* **59**, 2053–2059.

Simpson, A. A., Tao, Y., Leiman, P. G., Badasso, M. O., He, Y., Jardine, P. J., Olson, N. H., Morais, M. C., Grimes, S., Anderson, D. L., Baker, T. S. & Rossmann, M. G. (2000b). Structure of the bacteriophage φ29 DNA packaging motor. *Nature* **408**, 745–750.

Smith, D. E., Tans, S. J., Smith, S. B., Grimes, S., Anderson, D. L. & Bustamante, C. (2001). The bacteriophage φ29 portal motor can package DNA against a large internal force. *Nature* **413**, 748–752.

Smith, T. J., Chase, E. S., Schmidt, T. J., Olson, N. H. & Baker, T. S. (1996). Neutralizing antibody to human rhinovirus 14 penetrates the receptor-binding canyon. *Nature* **383**, 350–354.

Smith, T. J., Kremer, M. J., Luo, M., Vriend, G., Arnold, E., Kamer, G., Rossmann, M. G., McKinlay, M. A., Diana, G. D. & Otto, M. J. (1986). The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science* **233**, 1286–1293.

Smyth, M., Tate, J., Hoey, E., Lyons, C., Martin, S. & Stuart, D. (1995). Implications for viral uncoating from the structure of bovine enterovirus. *Nature Structural Biology* **2**, 224–231.

Stadler, K., Allison, S. L., Schalich, J. & Heinz, F. X. (1997). Proteolytic activation of tick-borne encephalitis virus by furin. *Journal of Virology* **71**, 8475–8481.

Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D. & Springer, T. A. (1989). A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell* **56**, 849–853.

Stehle, T., Gamblin, S. J., Yan, Y. & Harrison, S. C. (1996). The structure of simian virus 40 refined at 3·1 Å resolution. *Structure* **4**, 165–182.

Stiasny, K., Allison, S. L., Marchler-Bauer, A., Kunz, C. & Heinz, F. X. (1996). Structural requirements for low-pH-induced rearrangements in the envelope

glycoprotein of tick-borne encephalitis virus. *Journal of Virology* **70**, 8142–8147.

Stiasny, K. & Heinz, F. X. (2006). Flavivirus membrane fusion. *Journal of General Virology* **87**, 2755–2766.

Stockley, P. G., Stonehouse, N. J., Murray, J. B., Goodman, S. T., Talbot, S. J., Adams, C. J., Liljas, L. & Valegard, K. (1995). Probing sequence-specific RNA recognition by the bacteriophage MS2 coat protein. *Nucleic Acids Research* **23**, 2512–2518.

Strauss, J. H. & Strauss, E. G. (2002). *Viruses and Human Disease*. San Diego: Academic Press.

Stubbs, G., Warren, S. & Holmes, K. (1977). Structure of RNA and RNA binding site in tobacco mosaic virus from 4 Å map calculated from X-ray fibre diagrams. *Nature* **267**, 216–221.

Sun, S., Gao, S., Kondabagil, K., Xiang, Y., Rossmann, M. G. & Rao, V. B. (2012). Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 817–822.

Sun, S., Kondabagil, K., Draper, B., Alam, T. I., Bowman, V. D., Zhang, Z., Hegde, S., Fokine, A., Rossmann, M. G. & Rao, V. B. (2008). The structure of the phage T4 DNA packaging motor suggests a mechanism dependent on electrostatic forces. *Cell* **135**, 1251–1262.

Sun, S., Kondabagil, K., Gentz, P. M., Rossmann, M. G. & Rao, V. B. (2007). The structure of the ATPase that powers DNA packaging into bacteriophage T4 procapsids. *Molecular Cell* **25**, 943–949.

Sun, S., Rao, V. B. & Rossmann, M. G. (2010). Genome packaging in viruses. *Current Opinion in Structural Biology* **20**, 114–120.

Sun, S., Xiang, Y., Wataru, A., Holdaway, H., Pal, P., Zhang, X., Diamond, M. S., Nabel, G. J. & Rossmann, M. G. (2013). Structural analyses at pseudo atomic resolution of Chikungunya virus and antibodies show mechanisms of neutralization. *eLIFE*, in press. doi: 10.7554/eLife.00435.

Tang, J., Jose, J., Chipman, P., Zhang, W., Kuhn, R. J. & Baker, T. S. (2011). Molecular links between the E2 envelope glycoprotein and nucleocapsid core in Sindbis virus. *Journal of Molecular Biology* **414**, 442–459.

Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G. & Baker, T. S. (1998). Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95**, 431–437.

Thomassen, E., Gielen, G., Schutz, M., Schoehn, G., Abrahams, J. P., Miller, S. & van Raaij, M. J. (2003). The structure of the receptor-binding domain of the bacteriophage T4 short tail fibre reveals a knitted trimeric metal-binding fold. *Journal of Molecular Biology* **331**, 361–373.

Tong, L. & Rossmann, M. G. (1990). The locked rotation function. *Acta Crystallographica Section A* **46**, 783–792.

Tsao, J., Chapman, M. S., Agbandje, M., Keller, W., Smith, K., Wu, H., Luo, M., Smith, T. J., Rossmann, M. G., Compans, R. W. & Parrish, C. R. (1991). The three-dimensional structure of canine parvovirus and its functional implications. *Science* **251**, 1456–1464.

Tsao, J., Chapman, M. S., Wu, H., Agbandje, M., Keller, W. & Rossmann, M. G. (1992). Structure determination of monoclinic canine parvovirus. *Acta Crystallographica Section B* **48**, 75–88.

Tuthill, T. J., Groppelli, E., Hogle, J. M. & Rowlands, D. J. (2010). Picornaviruses. *Current Topical Microbiology Immunology* **343**, 43–89.

Valegård, K., Liljas, L., Fridborg, K. & Unge, T. (1990). The three-dimensional structure of the bacterial virus MS2. *Nature* **345**, 36–41.

Valegård, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J. & Liljas, L. (1994). Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature* **371**, 623–626.

van Raaij, M. J., Schoehn, G., Burda, M. R. & Miller, S. (2001). Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *Journal of Molecular Biology* **314**, 1137–1146.

von Bonsdorff, C. H. & Harrison, S. C. (1975). Sindbis virus glycoproteins form a regular icosahedral surface lattice. *Journal of Virology* **16**, 141–145.

von Bonsdorff, C. H. & Harrison, S. C. (1978). Hexagonal glycoprotein arrays from Sindbis virus membranes. *Journal of Virology* **28**, 578–583.

Voss, J. E., Vaney, M.-C., Duquerroy, S., Vonrhein, C., Girard-Blanc, C., Crublet, E., Thompson, A., Bricogne, G. & Rey, F. A. (2010). Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature* **468**, 709–712.

Walters, R. W., Agbandje-McKenna, M., Bowman, V. D., Moninger, T. O., Olson, N. H., Seiler, M., Chiorini, J. A., Baker, T. S. & Zabner, J. (2004). Structure of adeno-associated virus serotype 5. *Journal of Virology* **78**, 3361–3371.

Wang, X., Peng, W., Ren, J., Hu, Z., Xu, J., Lou, Z., Li, X., Yin, W., Shen, X., Porta, C., Walter, T. S., Evans, G., Axford, D., Owen, R., Rowlands, D. J., Wang, J., Stuart, D. I., Fry, E. E. & Rao, Z. (2012). A sensor-adaptor mechanism for enterovirus uncoating from structures of EV71. *Nature Structural and Molecular Biology* **19**, 424–429.

Watson, J. D. & Crick, F. H. C. (1953a). Genetical implications of the structure of deoxyribonucleic acid. *Nature* **171**, 964–967.

Watson, J. D. & Crick, F. H. C. (1953b). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738.

Wikoff, W. R., Liljas, L., Duda, R. L., Tsuruta, H., Hendrix, R. W. & Johnson, J. E. (2000). Topologically

linked protein rings in the bacteriophage HK97 capsid. *Science* **289**, 2129–2133.

Wiley, D. C., Wilson, I. A. & Skehel, J. J. (1981). Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* **289**, 373–378.

Williams, R. C. & Smith, K. M. (1958). The polyhedral form of the Tipula iridescent virus. *Biochimica et Biophysica Acta* **28**, 464–469.

Wilson, H. R. (1966). *Diffraction of X-rays by Proteins, Nucleic Acids and Viruses*. London: Edward Arnold Publishers Ltd.

Wrigley, N. G. (1969). An electron microscope study of the structure of Sericesthis iridescent virus. *Journal of General Virology* **5**, 123–134.

Wu, H. & Rossmann, M. G. (1993). The canine parvovirus empty capsid structure. *Journal of Molecular Biology* **233**, 231–244.

Wu, S. J. L., Grouard-Vogel, G., Sun, W., Mascola, J. R., Brachtel, E., Putvatana, R., Louder, M. K., Filgueira, L., Marovich, M. A., Wong, H. K., Blauvelt, A., Murphy, G. S., Robb, M. L., Innes, B. L., Birx, D. L., Hayes, C. G. & Frankel, S. S. (2000). Human skin Langerhans cells are targets of dengue virus infection. *Nature Medicine* **6**, 816–820.

Xiang, Y., Morais, M. C., Cohen, D. N., Bowman, V. D., Anderson, D. L. & Rossmann, M. G. (2008). Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage φ29 tail. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 9552–9557.

Xiao, C., Bator, C. M., Bowman, V. D., Rieder, E., He, Y., Hébert, B., Bella, J., Baker, T. S., Wimmer, E., Kuhn, R. J. & Rossmann, M. G. (2001). Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1. *Journal of Virology* **75**, 2444–2451.

Xiao, C., Chipman, P. R., Battisti, A. J., Bowman, V. D., Renesto, P., Raoult, D. & Rossmann, M. G. (2005). Cryo-electron microscopy of the giant Mimivirus. *Journal of Molecular Biology* **353**, 493–496.

Xiao, C., Kuznetsov, Y. G., Sun, S., Hafenstein, S. L., Kostyuchenko, V. A., Chipman, P. R., Suzan-Monti, M., Raoult, D., McPherson, A. & Rossmann, M. G. (2009). Structural studies of the giant Mimivirus. *PLoS Biology* **7**, e1000092.

Xiao, C., McKinlay, M. A. & Rossmann, M. G. (2011). Design of capsid-binding antiviral agents against human rhinoviruses. In *RSC Biomolecular Sciences Series, No. 21, Structural Virology* (eds. M. Agbandje-McKenna & R. McKenna), pp. 321–339. London, England: Royal Society of Chemistry.

Xiao, C. & Rossmann, M. G. (2011). Structures of giant icosahedral eukaryotic dsDNA viruses. *Current Opinion in Virology* **1**, 101–109.

Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A. & Chapman, M. S. (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 10405–10410.

Xie, Q. & Chapman, M. S. (1996). Canine parvovirus capsid structure, analyzed at 2·9 Å resolution. *Journal of Molecular Biology* **264**, 497–520.

Xu, L., Benson, S. D., Butcher, S. J., Bamford, D. H. & Burnett, R. M. (2003). The receptor binding protein P2 of PRD1, a virus targeting antibiotic-resistant bacteria, has a novel fold suggesting multiple functions. *Structure* **11**, 309–322.

Yan, X., Chipman, P. R., Castberg, T., Bratbak, G. & Baker, T. S. (2005). The marine algal virus PpV01 has an icosahedral capsid with $T=219$ quasi-symmetry. *Journal of Virology* **79**, 9236–9243.

Yan, X., Olson, N. H., Van Etten, J. L., Bergoin, M., Rossmann, M. G. & Baker, T. S. (2000). Structure and assembly of large lipid-containing dsDNA viruses. *Nature Structural Biology* **7**, 101–103.

Yan, X., Yu, Z., Zhang, P., Battisti, A. J., Holdaway, H. A., Chipman, P. R., Bajaj, C., Bergoin, M., Rossmann, M. G. & Baker, T. S. (2009). The capsid proteins of a large, icosahedral dsDNA virus. *Journal of Molecular Biology* **385**, 1287–1299.

Yin, H. S., Paterson, R. G., Wen, X., Lamb, R. A. & Jardetzky, T. S. (2005). Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9288–9293.

Yin, H. S., Wen, X., Paterson, R. G., Lamb, R. A. & Jardetzky, T. S. (2006). Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* **439**, 38–44.

Yu, I.-M., Holdaway, H. A., Chipman, P. R., Kuhn, R. J., Rossmann, M. G. & Chen, J. (2009). Association of the pr peptides with dengue virus at acidic pH blocks membrane fusion. *Journal of Virology* **83**, 12101–12107.

Yu, I.-M., Zhang, W., Holdaway, H. A., Li, L., Kostyuchenko, V. A., Chipman, P. R., Kuhn, R. J., Rossmann, M. G. & Chen, J. (2008). Structure of the immature dengue virus at low pH primes proteolytic maturation. *Science* **319**, 1834–1837.

Yuan, P., Swanson, K. A., Leser, G. P., Paterson, R. G., Lamb, R. A. & Jardetzky, T. S. (2011). Structure of the Newcastle disease virus hemagglutinin-neuraminidase (HN) ectodomain reveals a four-helix bundle stalk. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 14920–14925.

Zauberman, N., Mutsafi, Y., Halevy, D. B., Shimoni, E., Klein, E., Xiao, C., Sun, S. & Minsky, A. (2008). Distinct DNA exit and packaging portals in the virus *Acanthamoeba polyphaga* mimivirus. *PLoS Biology* **6**, e114.

Zhang, P., Mueller, S., Morais, M. C., Bator, C. M., Bowman, V. D., Hafenstein, S., Wimmer, E. & Rossmann, M. G. (2008). Crystal structure of CD155 and electron microscopic studies of its complexes with

polioviruses. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 18284–18289.

Zhang, R., Hryc, C. F., Cong, Y., Liu, X., Jakana, J., Gorchakov, R., Baker, M. L., Weaver, S. C. & Chiu, W. (2011a). 4·4 Å cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. *EMBO Journal* **30**, 3854–3863.

Zhang, W., Chipman, P. R., Corver, J., Johnson, P. R., Zhang, Y., Mukhopadhyay, S., Baker, T. S., Strauss, J. H., Rossmann, M. G. & Kuhn, R. J. (2003a). Visualization of membrane protein domains by cryo-electron microscopy of dengue virus. *Nature Structural Biology* **10**, 907–912.

Zhang, W., Mukhopadhyay, S., Pletnev, S. V., Baker, T. S., Kuhn, R. J. & Rossmann, M. G. (2002). Placement of the structural proteins in Sindbis virus. *Journal of Virology* **76**, 11645–11658.

Zhang, X., Sheng, J., Plevka, P., Kuhn, R. J., Diamond, M. S. & Rossmann, M. G. (2013). Dengue structure differs at the temperatures of its human and mosquito hosts. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 6795–6799.

Zhang, X., Sun, S., Xiang, Y., Wong, J., Klose, T., Raoult, D. & Rossmann, M. G. (2012). Structure of Sputnik, a virophage, at 3·5 Å resolution. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 18431–18436.

Zhang, X., Walker, S. B., Chipman, P. R., Nibert, M. L. & Baker, T. S. (2003b). Reovirus polymerase l3 localized by cryo-electron microscopy of virions at a resolution of 7·6 Å. *Nature Structural Biology* **10**, 1011–1018.

Zhang, X., Xiang, Y., Dunigan, D. D., Klose, T., Chipman, P. R., Van Etten, J. L. & Rossmann, M. G. (2011b). Three-dimensional structure and function of the *Paramecium bursaria* chlorella virus capsid. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 14837–14842.

Zhang, Y., Corver, J., Chipman, P. R., Zhang, W., Pletnev, S. V., Sedlak, D., Baker, T. S., Strauss, J. H., Kuhn, R. J. & Rossmann, M. G. (2003c). Structures of immature flavivirus particles. *EMBO Journal* **22**, 2604–2613.

Zhang, Y., Kaufmann, B., Chipman, P. R., Kuhn, R. J. & Rossmann, M. G. (2007). Structure of immature West Nile virus particles. *Journal of Virology* **81**, 6141–6145.

Zhang, Y., Zhang, W., Ogata, S., Clements, D., Strauss, J. H., Baker, T. S., Kuhn, R. J. & Rossmann, M. G. (2004). Conformational changes of the flavivirus E glycoprotein. *Structure* **12**, 1607–1618.

Zhang, Z., Kottadiel, V. I., Vafabakhsh, R., Dai, L., Chemla, Y. R., Ha, T. & Rao, V. B. (2011c). A promiscuous DNA packaging machine from bacteriophage T4. *PLoS Biology* **9**, e1000592.