Structure and Assembly of Membrane-Containing dsDNA Bacteriophages

JUHA T. HUISKONEN

Institute of Biotechnology and
Department of Biological and Environmental Sciences
Division of Genetics
Faculty of Biosciences
University of Helsinki

and

National Graduate School in Informational and Structural Biology

ACADEMIC DISSERTATION

To be presented with the permission of the Faculty of Biosciences of the University of Helsinki for public examination in the auditorium 1041 of Biocenter 2, Viikinkaari 5, Helsinki, on December 16th, 2005, at 12 noon

HELSINKI 2005

Supervisor

Docent Sarah J. Butcher
Institute of Biotechnology
University of Helsinki

Reviewers

Professor Mark S. Johnson
Department of Biochemistry and Pharmacy
Åbo Akademi University

Doctor Jonathan M. Grimes
Division of Structural Biology
Wellcome Trust Centre for Human Genetics
University of Oxford

Opponent

Doctor R. Anthony Crowther
Division of Structural Studies
MRC Laboratory of Molecular Biology
Cambridge

ISBN 952-10-2803-3 (paperback)
ISBN 952-10-2804-1 (PDF, http://ethesis.helsinki.fi/)

ISSN 1795-7079

Yliopistopaino, Helsinki University Printing House
Helsinki 2005

To Niina

Original publications

This thesis is based on the following articles, which are referred to in the text by their Roman numerals.

I San Martín, C., Huiskonen, J.T., Bamford, J.K., Butcher, S.J., Fuller, S.D., Bamford, D.H. and Burnett, R.M. (2002) Minor proteins, mobile arms and membrane-capsid interactions in the bacteriophage PRD1 capsid. *Nat Struct Biol*, **9**, 756–763.

II Huiskonen, J.T., Kivelä, H.M., Bamford, D.H., Butcher, S.J. (2004) The PM2 virion has a novel organization with an internal membrane and pentameric receptor binding spikes. *Nat Struct Mol Biol*, **11**, 850–856.

III Merckel, M.C., Huiskonen, J.T., Bamford, D.H., Goldman, A. and Tuma, R. (2005) The structure of the bacteriophage PRD1 spike sheds light on the evolution of viral capsid architecture. *Mol Cell*, **18**, 161–170.

IV Laurinmäki, P.A.*, Huiskonen, J.T.*, Bamford, D.H., and Butcher, S.J. (2005) Membrane proteins modulate the bilayer curvature in the bacterial virus Bam35. *Structure*, **13**, in press.

* These authors contributed equally.

Also unpublished data will be presented.

Abbreviations

σ	standard deviation
2D	two-dimensional
3D	three-dimensional
ATPase	adenosine triphosphatase
bp	base pairs
BTV	Bluetongue virus
C-	carboxy-
CCD	charge-coupled device
CIV	Chilo iridescent virus
CTF	contrast transfer function
DNA	deoxyribonucleic acid
ds	double-stranded
EM	electron microscopy
FEG	field emission gun
FSC	Fourier shell correlation
gp	gene product
HK97	bacteriophage Hong Kong 97
HSV-1	herpes simplex virus type 1
LC	lipid core (of PM2)
MS	mass spectrometry
N-	amino-
NMR	nuclear magnetic resonance
PBCV-1	paramecium bursaria chlorella virus type 1
PFT	polar Fourier transform
PpV01	*Phaeocystic puchetii* virus
RNA	ribonucleic acid
SFV	Semliki Forest virus
SNR	signal-to-noise ratio
ss	single-stranded
SV40	Simian virus 40
T	triangulation number
TNF	tumour necrosis factor
TMV	tobacco mosaic virus
wt	wild type

Summary

Viruses present a paradigm for assembly and function of large macromolecular complexes. Bacterial viruses PRD1, Bam35 and PM2 studied in this work contain double stranded DNA genomes, internal lipid-bilayers and icosahedrally symmetric protein capsids. Their structures were studied using X-ray crystallography and cryo-electron microscopy combined with three-dimensional image reconstruction. Cryo-electron microscopy, established nationally at the Institute of Biotechnology in 2001, provided a valuable method to study the structure of these large assemblies at sub-nanometre resolution. The structures shed light on the assembly principles and evolution of these viruses. Receptor binding spike structures provided information on virus–host cell interactions. The analysis of the structures of the membrane components revealed that membrane proteins modulate the bilayer curvature and thickness.

Table of contents

A. INTRODUCTION

Viruses are obligatory parasites that require the cellular machinery of their host organism. Their sole meaning is to replicate—to transfer their genome into the host cell for the production of new virions. Outside the host cell, the viral particle does not have any metabolism. However, inside the cell, the virus recruits the cellular machinery for its own functions. Viruses infecting humans are one of the major causes of human suffering (WHO, 1996). However, viruses are by no means limited to only infecting humans, but they infect probably all organisms from all domains of life: Eukarya, Archea and Bacteria. Viruses are ubiquitous and they are the most abundant organisms on the Earth. Their number exceeds the number of their host cells by at least a factor of ten. There are millions of bacterial viruses in one millilitre of seawater and it has been estimated that there are 10^{30} viruses in the biosphere (Bergh *et al*., 1989; Wommack and Colwell, 2000; Hendrix, 2002). Thus, viruses may have a huge impact on ecosystems.

The viral genome can be either deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) and either double-stranded (ds) or single stranded (ss). In addition, the genome can be either circular or linear. Some viruses, such as the retroviruses, are reverse-transcribing, copying their RNA genome into DNA in the process of host-genome integration. Genome type and particle morphology are considered as the ultimate criteria in the classification of viruses into different families and higher taxa. This classification reflects their phylogenies, or evolutionary relationships (Murphy and Kingsbury, 1990). The virus taxonomy is maintained by the International Committee on Taxonomy of Viruses (Fauquet *et al*., 2004).

Viruses that infect bacteria are called bacteriophages (eaters of bacteria) or simply phages. Phages were discovered independently by Frederick Twort in 1915 and by Félix d'Hérelle in 1917. Phages have been reported for more than 140 bacterial genera (Ackermann, 2001). All bacteria can be divided into two subgroups, based on their outer cell wall structure. These groups are referred to as Gram-positive and Gram-negative. Belonging into a certain group depends on how the bacterium behaves in a staining developed by the Danish bacteriologist Hans Christian Gram in 1882.

Two types of phages exist, lytic and temperate. Lytic phages infect the host bacterium, replicate and break up the host cell wall to release the progeny phages. Temperate phages have an additional option to this lytic pathway, a lysogenic pathway. This pathway enables normal host cell growth, and only the phage genome endures in the host cell. Such a phage is designated as a prophage. A prophage can integrate into the host genome or exist as extra-chromosomal DNA in the form of a circular or linear plasmid. Environmental signals can cause the activation of the prophage leading to lytic replication (Casjens, 2003).

Viruses are considered as model systems in molecular biology and biophysics and have been used to study many fundamental aspects of biology. Bacteriophages are especially easy to grow and purify for experimental studies. Phage research has lead to many important discoveries (for a review see Cairns *et al*., 1966). For example, research on bacteriophage T2 contributed to the identification of DNA as genetic material (Hershey and Chase, 1952). Eukaryotic viruses are used as model systems to study cellular events, such as cellular signalling, cytoskeleton reorganization and membrane trafficking (Pelkmans, 2005). Viruses are also used in several applications, for example in gene therapy (Kootstra and Verma, 2003), phage therapy (Summers, 2001), phage display (O'Neil and Hoess, 1995) and as self-assembling materials for nanotechnology (Wang *et al*., 2002).

1. Virus structure and assembly

1.1. Virion morphology and symmetry

The function of the virion is to act as a container for the nucleic acid, shielding it from the environment outside the host cell. In infection, the virion recognizes a new host and transfers the nucleic acid into it. The container is generally formed by a protein layer called the capsid. In enveloped viruses, an additional membrane layer also exists. A thick layer called the tegument consisting of additional proteins may reside between the capsid and the membrane envelope (Harrison, 1990). The virion is not simply an inert particle, but contains enzymatically-active proteins. Many bacteriophage particles, for example, contain lysozymes which are required for host cell infection (Rossmann *et al*., 2004).

Viruses assemble into virions with different shapes and sizes. The encoding capacity of the genetic material limits the number of possible capsid architectures, since the genome has to be sufficiently large to be able to encode all of the viral functions, including the capsid protein, but it must also fit into the volume enclosed by the capsid. Crick and Watson (1956) first realized that viruses must utilize multiple copies of the capsid protein instead of using just one or two large proteins to form the spherical protein shell. In electron microscopy (EM) images of virions, the individual capsid protein subunits are unresolved, but repeating clusters of capsid proteins that appear as morphologically distinct units, designated as capsomers, are visible.

Capsomers assemble into rod-shaped, spherical or pleomorphic capsids (Casjens, 1997). Rod-shaped capsids, such as the capsid of tobacco mosaic virus (TMV), exhibit helical symmetry (Namba and Stubbs, 1986). In these capsids, proteins assemble around the nucleic acid and all of the interactions between the capsomers are identical. The sizes of helical capsids vary, depending on the number of capsid proteins and their organization in the helix. Spherical viruses also exploit symmetry (Caspar and Klug, 1962). Of all symmetries capable of generating a closed spherical capsid, the icosahedral symmetry is the highest and the only one observed in such capsids (Casjens, 1997)(see also **Section 1.2**). In capsids with icosahedral symmetry, the largest ratio between the capsid volume and the size of the capsomer is reached, and maximal encoding capacity is retained for functions other than encoding the capsid. Some closed capsids, such as the cone-shaped cores of human immunodeficiency virus, do not exploit icosahedral symmetry (Briggs *et al*., 2003). The capsid proteins of these pleomorphic viruses assemble into more complex patterns, and their structure is less understood.

Different appendages may be attached to viral capsids adding extra functionality. Most dsDNA bacteriophages have a tail structure attached to one of the five-fold vertices. The tail is used to translocate the genome into a host cell (Rossmann *et al*., 2004). Many viruses such as adenoviruses and reoviruses have elongated fibre-like structures that they use to bind to the host cell surface receptor (van Raaij *et al*., 1999; Chappell *et al*., 2002). The membrane of many enveloped viruses is fully covered with protruding spike proteins, which assist in the fusion of the viral and host membranes. Examples include the hemagglutinin of influenza virus (Wilson *et al*., 1981; Ruigrok *et al*., 1984) and the glycoproteins of alpha- and flaviviruses (Rey *et al*., 1995; Mancini *et al*., 2000; Lescar *et al*., 2001; Kuhn *et al*., 2002).

1.2. Icosahedrally symmetric capsids and quasi-equivalence

Spherical virus capsids exploit icosahedral symmetry, denoted 532-symmetry, since it defines five-fold, three-fold and two-fold symmetry axes. Icosahedral symmetry is present in two Platonic solids, the regular icosahedron and the regular dodecahedron. The regular icosahedron consists of 12 pentagonal vertices, 20 equilateral triangular faces and 30 edges (**Figure 1a**). Six five-fold axes pass through the icosahedral vertices, ten three-fold axes through the icosahedral faces (also called 'facets') and fifteen two-fold axes through the icosahedral edges. The regular dodecahedron, in turn, is composed of 20 triangular vertices, 12 equilateral pentagonal faces and 30 edges (**Figure 1b**). The shapes of icosahedrally symmetric virus capsids lie somewhere between the ideal icosahedron and the ideal dodecahedron (Baker *et al.*, 1999).

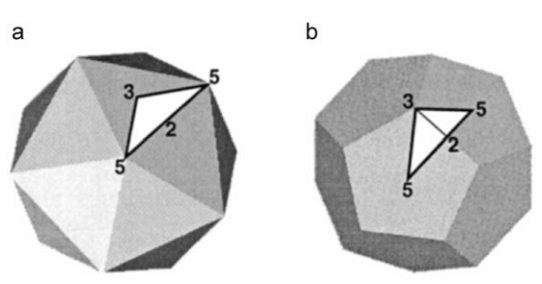

Figure 1. Platonic solids with icosahedral symmetry. (**a**) A regular icosahedron. (**b**) A regular dodecahedron. The numbers (2, 3 and 5) indicate the positions of some of the icosahedral symmetry axes that traverse through the indicated position and the centre of the solid. The white triangle outlines the asymmetric unit. With permission from Baker *et al*. (1999).

Viral capsids with icosahedral symmetry consist of 60 asymmetric units (**Figure 1**). In the simplest capsids, each asymmetric unit is composed of one single protein and all of the interactions between capsid proteins are equivalent. In more complex viruses, the viral genome requires more encoding capacity for additional functions, which renders the genome too large to fit into such a simple capsid. In such cases, asymmetric units are assembled using more than one copy of the capsid protein in order to increase the size of the capsid. The theory of quasi-equivalency proposed by Caspar and Klug (1962) predicts that only certain types of capsid architectures would lead to these larger capsids. Abstractly thinking, a three-dimensional (3D) icosahedral lattice can be constructed from a two-dimensional (2D) hexagonal net by replacing twelve appropriate hexagons with pentagons (**Figure 2**). This leads to a set of possible lattice types, which can be described with a set of triangulation (*T*) numbers. $T = h^2 + h \cdot k + k^2$, where h and k are positive integers indicating the coordinates where a hexagon is replaced with a pentagon. When $h \neq 0$, $k \neq 0$ and $h \neq k$, the lattice is said to be handed or skewed. For such a lattice, two different enantiomers (mirror lattices), called *laevo* and *dextro*, exist (**Figure 3**). These two enantiomers refer only to the lattice, not to the biomolecules themselves within the capsid.

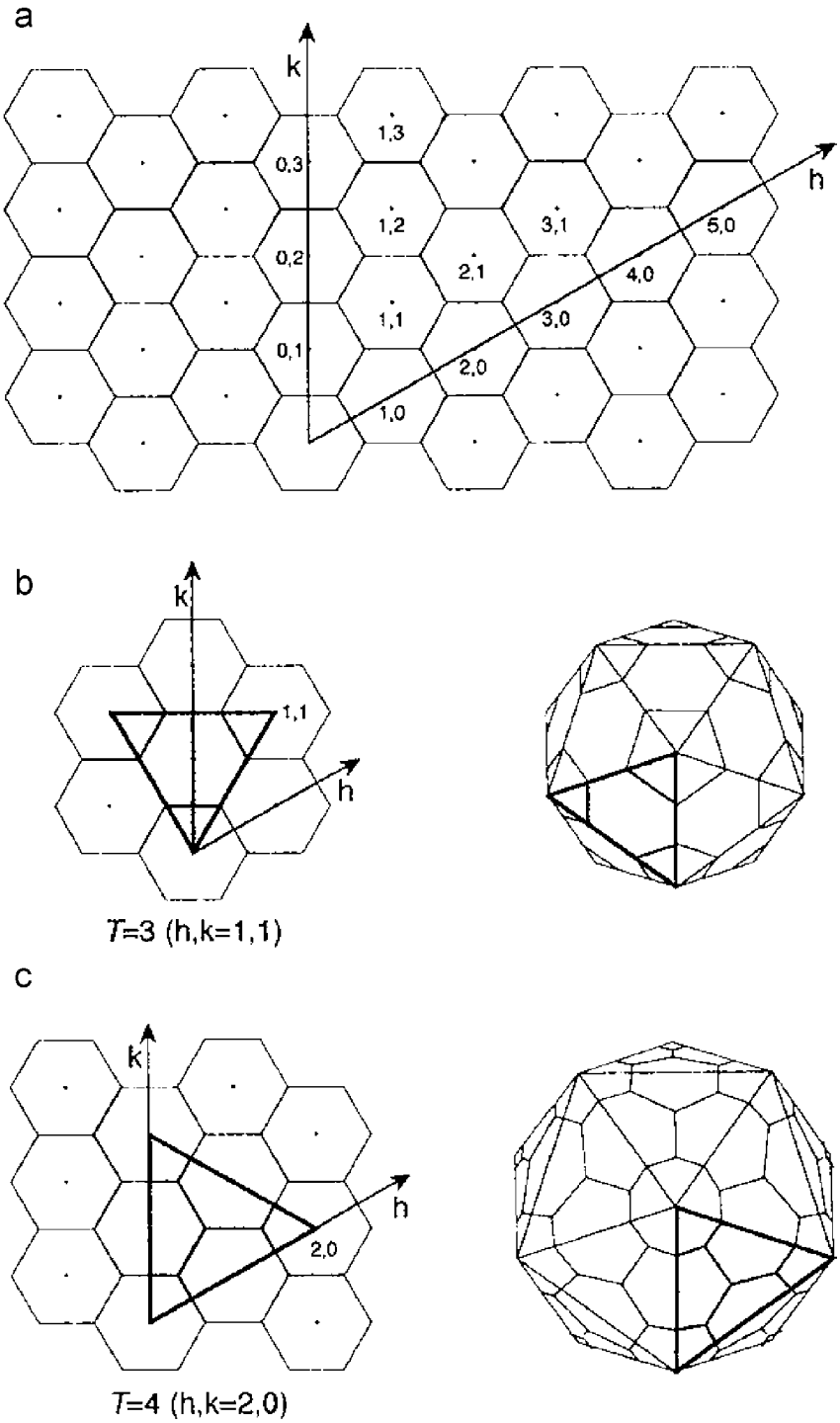

Figure 2. Geometric principles for constructing icosahedral lattices of defined *T* (triangulation) number. (**a**) A 2D hexagonal lattice with lattice indices h and k. A closed shell can be generated by replacing the hexagon at twelve appropriate lattice positions (h, k) with a pentagon. (**b**) Constructing a *T*=3 icosahedron. (**c**) Constructing a *T*=4 icosahedron. One facet is highlighted in *b* and *c*. Modified from Baker *et al*. (1999) with permission.

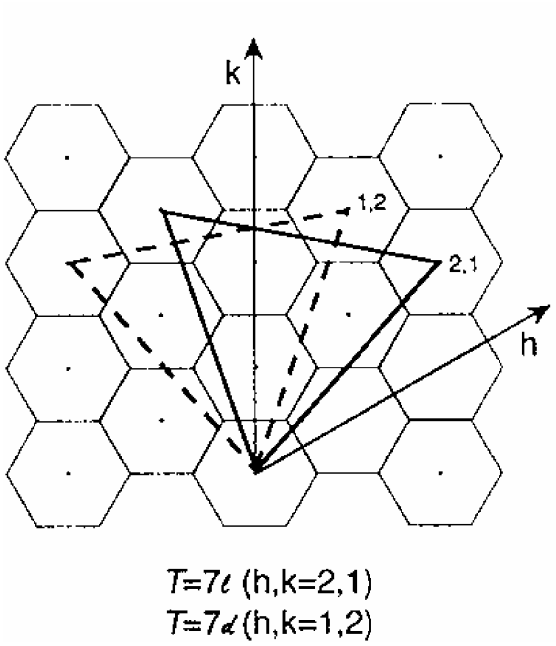

Figure 3. Mirror lattices. When h≠0, k≠0 and h≠k, as in the example of the *T*=7 lattice, two possible lattices exist, laevo (l) and dextro (d). These two are related by mirror symmetry. Modified from Baker *et al.* (1999) with permission.

In capsids with icosahedral symmetry, the geometrical lattices predicted by Caspar–Klug theory usually define the positions of the capsomers. Six capsid protein monomers contribute to each hexavalent position and five monomers to each pentavalent position. Furthermore, *T* usually equals the number of capsid proteins in the asymmetric unit. In the simplest capsid, h=1 and k=0. Thus, *T*=1, and there is only one protein in the asymmetric unit. Bigger *T*-numbers (3, 4, 7, 9, 12, 13, 16, 19, 21, 25 *etc.*) lead to larger capsids with a total of 60*T* capsid proteins (Baker *et al.*, 1999).

Triangulation is an abstract mathematical concept, and biological systems show many exceptions to it. Firstly, L-A virus (Naitow *et al.*, 2002), members of the *Reoviridae* (Grimes *et al.*, 1998; Hill *et al.*, 1999; Nakagawa *et al.*, 2003) and bacteriophage ϕ6 (Huiskonen, J.T., de Haas, F., Bubeck, D., Bamford, D.H., Fuller, S.D. and Butcher, S.J., manuscript) have capsid shells with 120 capsid proteins. In these capsid shells, two capsid proteins occupy each *T*=1 lattice position. Because of this, these lattices have been described as '*T*=2', a disallowed *T*-number (Grimes *et al.*, 1998). Secondly, chemically different units may occupy different quasi-equivalent positions. For example, the *T*=3 capsids of picornaviruses are composed of three different, although structurally similar, proteins occupying the three quasi-equivalent positions (Rossmann *et al.*, 1985). Moreover, a single protein with two similar domains can be thought to consist of two subunits. For example, the hexon of adenovirus consists of three capsid proteins, each with a double β-barrel fold (Roberts *et al.*, 1986)(see **Section 2.2**). Consequently, there are six similar subunits in each trimer. These pseudo-hexameric trimers occupy the hexavalent positions in the lattice. Five hexons encircle the icosahedral five-fold axes of symmetry leaving a hole at the vertex. This hole is filled by a separate penton protein, which is itself a pentamer (Stewart *et al.*, 1991). As a result, 24 domains from four hexons and 1 domain of the penton protein contribute to one asymmetric unit, as expected for a *T*=25 capsid. These types of lattices are described with a pseudo *T*-number to distinguish them from capsids that have *T* identical lattice proteins per each asymmetric unit. The triangulation of picornaviruses is said to be 'pseudo *T*=3' and the triangulation of adenoviruses is 'pseudo *T*=25'. Thirdly, some capsids are built entirely of pentameric capsomers. For example, polyomavirus (Rayment *et al.*, 1982) and the highly similar simian virus 40 (SV40)(Baker *et al.*, 1989; Liddington *et al.*, 1991)(see **Section 2.1**) capsids consist of 72 pentameric capsomers located at both the pentavalent and the hexavalent positions of a *T*=7 lattice. A generalization of the Caspar–Klug theory has been recently proposed to explain these capsid architectures that deviate from the classical theory (Twarock, 2004).

Capsid proteins must form a stable capsid. In capsids where *T*>1, all the subunits (*T* copies) occupy symmetrically

non-equivalent positions in the asymmetric unit. A consequence of this is that the interactions between these capsid proteins cannot be strictly equivalent. The Caspar–Klug theory predicts that the same bonding patterns occur between all of the subunits but the interactions are slightly deformed. These interactions are said to be quasi-equivalent (Caspar and Klug, 1962). It has been observed that most of the capsid protein structure is nearly identical at all quasi-equivalent positions. Only the protein termini, amino- (N) or carboxy- (C) termini or both, are flexible and have different conformations. This allows them to bridge the neighbouring capsomers together (Harrison *et al*., 1978; Liddington *et al*., 1991; Wikoff *et al*., 2000; Abrescia *et al*., 2004). In addition, minor capsid proteins, called glue or cementing proteins, may also play a role in making a stable capsid. A cementing protein has been observed, for example, in the adenovirus capsid (Furcinitti *et al*., 1989; Saban *et al*., 2005).

1.3. Assembly of viral particles and precursors

All individual virions of an icosahedrally symmetric virus are practically identical in structure (Crick and Watson, 1956). How is it that the different building blocks of a virus precisely self-assemble into a functional virion? Reconstitution of TMV was the first demonstration of macromolecular self-assembly *in vitro*. Fraenkel-Conrat and Williams (1955) managed to assemble infectious TMV virions from separate capsid protein and RNA components. Following this historical experiment on TMV, the assembly of more complex viruses has been studied in several well-characterized systems, such as bacteriophages P22 (King and Casjens, 1974) and ϕ6 (Poranen *et al*., 2001).

A general theme in virus assembly is that the newly synthesized subunits assemble together in an orderly fashion and the process is split into multiple, controlled steps (Casjens, 1997). This process is called the assembly pathway. The pathway is not always linear but can also be branched. For example, the head, tail and the fibres of bacteriophage T4 are first assembled individually following their separate pathways and finally combined together to form an infectious virion (for a review see Mosig and Eiserling, 1988).

A general assembly pathway for a dsDNA bacteriophage capsid is outlined in **Figure 4**. The first step in the pathway is the proper folding of the capsid proteins, usually controlled by host chaperones, and possible multimerisation into larger subunits. A nucleation event initiates the assembly of the capsid. Capsid proteins, or their multimers, then bind to the growing assembly, until a spherical closed structure is obtained. This may involve conformational changes in the capsid protein subunits that change them from a form that cannot associate to a form that can (Caspar, 1980; Mosig and Eiserling, 1988). These conformational changes, designated as conformational switches, have been observed, for example, in the assembly of TMV (Namba and Stubbs, 1986).

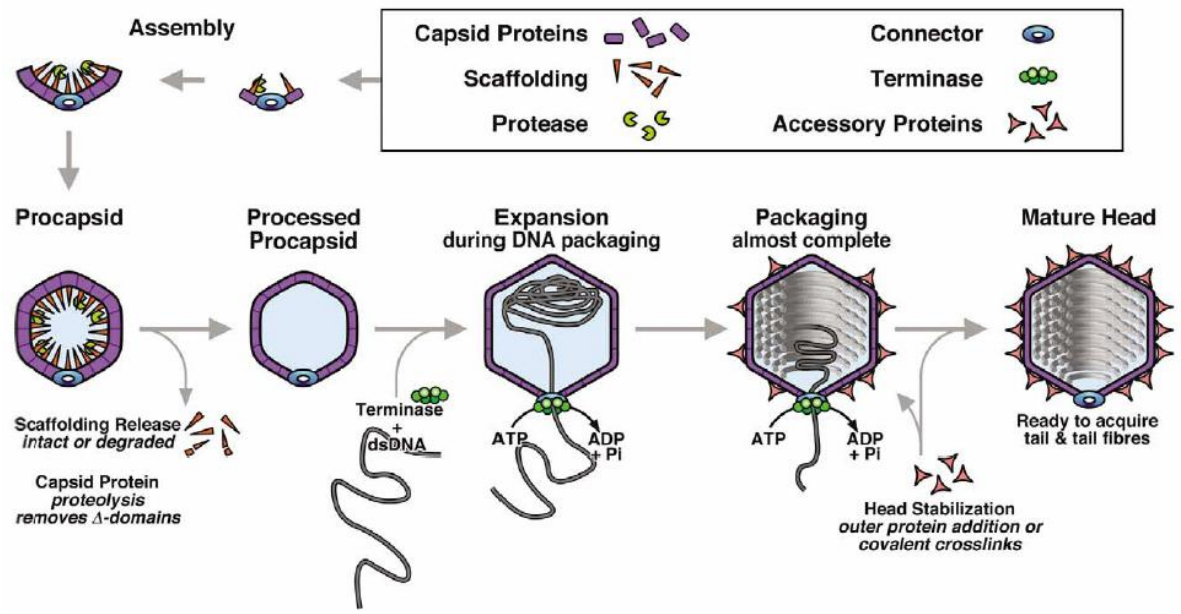

Figure 4. A generic dsDNA bacteriophage capsid assembly pathway. The connector acts as an assembly initiator. The procapsid is first assembled under the guidance of the scaffolding protein. The scaffolding is then removed. Proteolysis of a capsid protein domain (Δ) may be required for subsequent steps. DNA is packaged into the processed procapsid. During the packaging an expansion may occur. In most dsDNA viruses, the DNA is probably organized as a co-axial spool (Booy *et al*., 1991; Cerritelli *et al*., 1997). After packaging, the capsid may be stabilized and additional appendages may be attached. With permission from Steven *et al*. (2005).

Even in some complex viruses, the capsid proteins are capable of self assembling into spherical shells without any additional factors. For example, infectious nucleocapsids of bacteriophage ϕ6 can be assembled *in vitro* from purified capsid proteins and genomic dsRNA segments (Poranen *et al*., 2001). However, specific accessory proteins are often required for the proper continuation of capsid assembly after initiation. King and Casjens (1974) found in bacteriophage P22 a protein "that specifies the correct dimensions of a virus, catalytically, by temporarily complexing with the major coat protein to form a transient precursor shell of the virus" and designated it as a 'scaffolding protein'. Scaffolding proteins have three characteristics (for a review see Fane and Prevelige, 2003). Firstly, hundreds of copies of the scaffolding protein are required for each capsid. Secondly, the scaffolding protein temporarily becomes an integral part of the assembly intermediate and the release of it is a triggered reaction, usually due to a proteolytic cleavage. Thirdly, the scaffolding protein is not generally part of the mature capsid. Two functions of scaffolding proteins are crucial in virus assembly: (i) they lower the nucleation barrier by bringing a sufficient number of capsid protein molecules together for nucleation to occur (ii) they determine the proper form of the assembly. Without scaffolding proteins, the initiation of assembly may be statistically improbable (Erickson and Pantaloni, 1981). The lowering of the nucleation barrier by a scaffolding protein also presents a control step in the initiation of capsid assembly, which may be used to regulate the assembly process. Although the local binding properties of the capsid proteins may enable them to cluster together into curved assemblies, the absence of the scaffolding protein may lead to assemblies of incorrect shape and size. This has been observed, for example, in bacteriophage P22, where

assemblies with incorrect curvature accumulated in cells infected with a scaffolding protein defective mutant (Earnshaw and King, 1978).

The scaffolding function should not be strictly limited to the above described features. Other kinds of form-determining functions exist, too. In some cases, the scaffold may be present in the final particle. For example, the inner '*T*=2' protein shell present in many dsRNA viruses, such as in the Blue tongue virus (BTV), is thought to serve a scaffold-like function in the assembly of the outer *T*=13 shell (Grimes *et al*., 1998). Similarly, the inner membranes of some viruses such as PRD1 may act as a template for capsid protein assembly (Rydman *et al*., 2001)(**Section 3**). Some virus capsids assemble on precondensed nucleic acid, which may also serve a form-determining function (Fane and Prevelige, 2003).

Many viruses package their genome into an empty viral particle, designated as a procapsid. This may present another control step, since only correctly formed procapsids are packaged. Packaging in dsDNA viruses occurs through a specific portal/connector protein present at one of the five-fold vertices (Valpuesta and Carrascosa, 1994). During or immediately after packaging, the virus particle matures into its infectious form. Large conformational changes, which may involve protein cleavage, are seen in many viral capsids during their maturation (Steven *et al*., 2005). In many dsDNA bacteriophages and in the bacteriophage ϕ6, maturation involves a substantial expansion of the procapsids (Poranen and Tuma, 2004; Steven *et al*., 2005). Maturation may also include the addition of other protein layers, appendages such as spikes or tails, or a membrane envelope.

1.4. Host cell infection

In the assembly process, viruses have to be able to form stable virions that survive in the harsh environment outside the cell. However, this poses a dilemma, since the virion needs to be at least partially disassembled before subsequent infection. Thus, the virion structure is not completely rock-solid, but some parts contain hidden functionality. These parts are in a metasable state and capable of undergoing conformational changes or disassembly once a new host cell is encountered.

Most viruses recognize their host cell with a special surface protein, which binds to the host cell surface receptor (Haywood, 1994). After recognition, the virus may enter the host cell as a viral or subviral particle. Viral membranes are exploited in fusion with host cell membranes. Fusion can be assisted by fusion proteins, as in influenza virus (Bullough *et al*., 1994). Alternatively, the virus may just transfer its genome into the host, as seen in many dsDNA bacteriophages, such as T4 (Rossmann *et al*., 2004). In dsRNA viruses, whose genome is alien to the host cell machinery, the polymerase complex provides the necessary function for genome transcription and replication and must be transferred into the host (Poranen and Tuma, 2004).

2. Structures of viruses and their capsid proteins

The structures of many spherical, icosahedrally symmetric viruses have been determined using cryo-EM combined with 3D image reconstruction (Baker *et al*., 1999)(see **Section 7**). In addition, X-ray crystallography has been used to solve the atomic structures of virions, subviral particles and their proteins (Lee and Johnson, 2003). Most structures of the viral capsid proteins have revealed an anti-

parallel, eight-stranded β-barrel fold with a jellyroll topology (reviewed in Chapman and Liljas, 2003). The barrel is divided in two β-sheets and usually there is no hydrogen bonding between the end-strands of the sheets. This viral β-barrel fold is seen in the capsids of positive-sense ssRNA vertebrate viruses, such as human rhinovirus (*Picornaviridae*)(Rossmann *et al*., 1985) and Norwalk virus (*Caliciviridae*)(Prasad *et al*., 1999), several positive-sense ssRNA plant viruses and their satellite viruses (Harrison *et al*., 1978; Abad-Zapatero *et al*., 1980; Liljas *et al*., 1982; Chen *et al*., 1989; Speir *et al*., 1995; Canady *et al*., 1996) and ssDNA bacteriophages, such as ϕX174 (McKenna *et al*., 1992). In the capsids of these viruses, the β-barrels lie roughly tangentially to the viral capsid (**Figure 5**).

The major capsid protein of dsDNA virus SV40 (*Polyomaviridae*) also has a viral β-barrel fold, but the barrels are oriented radially, not tangentially as in positive-sense, ssRNA viruses (Liddington *et al*., 1991)(**Section 2.1**). In adenoviruses (*Adenoviridae*)(**Section 2.2**) and bacteriophage PRD1 (*Tectiviridae*)(**Section 3**) the capsid protein has two of these folds within one polypeptide chain (Roberts *et al*., 1986; Benson *et al*., 1999) and the β-barrels lie radially in the viral capsid (Stewart *et al*., 1991; Butcher *et al*., 1995). Some example structures of icosahedral dsDNA viruses and their capsid proteins, the topic of this thesis, are reviewed in more detail below.

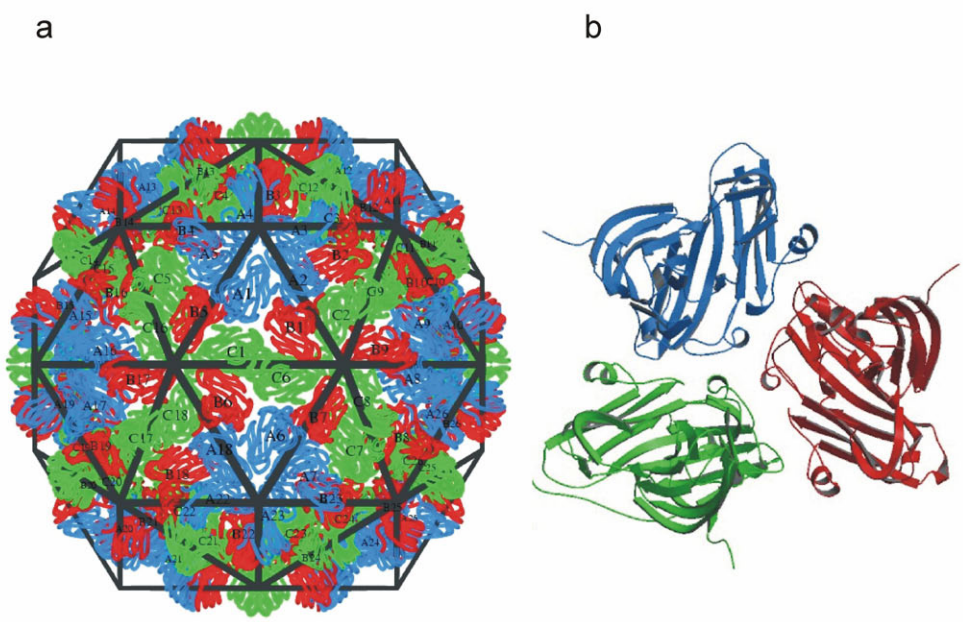

Figure 5. The structure of tomato bushy stunt virus. (**a**) The three quasi-equivalent capsid proteins of the *T*=3 lattice are coloured in blue, red and green. (**b**) A close-up of one asymmetric unit revealing the β-barrel fold of the capsid protein, and their tangential orientation in the capsid. The figure was rendered in Virus particle explorer (Reddy *et al*., 2001) from the coordinates of 2BTV (Harrison *et al*., 1978) obtained from the RCSB Protein Data Bank.

2.1. SV40

SV40 (*Polyomaviridae*) is a relatively small dsDNA virus, ~50 nm in diameter. The capsid consists of 72 capsomers organized on a *T*=7 lattice. All of the capsomers are major capsid protein pentamers (Baker *et al*., 1989). Thus, the structure breaks the rules of quasi-equivalence (**Section 1.2**) and there are 360 capsid proteins instead of 420 predicted by the Caspar–Klug theory (Caspar and Klug, 1962). The capsid protein is an eight-stranded β-barrel and lies perpendicular to the capsid surface (Liddington *et al*., 1991). The core of the protein is nearly identical in all quasi-equivalent positions. The flexible C-terminal part is responsible for the quasi-equivalent interactions between the pentamers (Liddington *et al*., 1991).

2.2. Adenovirus

Adenoviruses (*Adenoviridae*) infect birds and mammals. Human adenoviruses cause respiratory infections. Adenovirus was one of the first viruses visualized using the cryo-EM technique (Adrian *et al*., 1984) and nowadays it is one of the best structurally characterized virus (Rux and Burnett, 2004). The 3D structure of the virion has been solved from cryo-electron micrographs using image reconstruction (Stewart *et al*., 1991; Fabry *et al*., 2005; Saban *et al*., 2005)(**Figure 6**). The capsid structure (~750 Å in diameter) displays 240 pseudo-hexameric capsomers, organized on a pseudo *T*=25 lattice. Nine capsomers cluster together to form an assembly called a 'group of nine' (**Figure 6**). A minor component, protein IX cements the hexons together in these assemblies (Furcinitti *et al*., 1989).

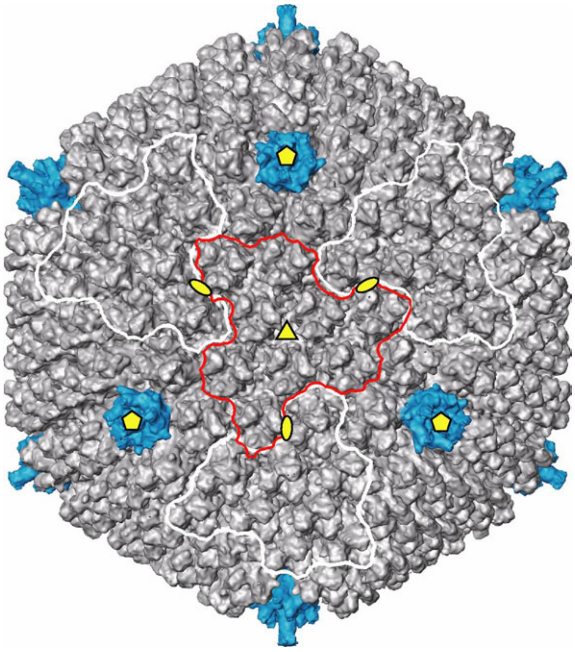

Figure 6. The 3D reconstruction of adenovirus at 10-Å resolution seen along an icosahedral three-fold axis of symmetry. Some of the icosahedral two-fold, three-fold and five-fold axes are denoted with ellipses, triangles and pentagons, respectively. The pentons are coloured in blue and four 'groups of nine' are outlined in red and white. With permission from Fabry *et al*. (2005).

The X-ray structure of the capsomer, the hexon, has been solved (Roberts *et al*., 1986). The hexon protein is a trimer. The monomer consists of two viral β-barrels with a jellyroll topology. The three-fold axis of the trimer is perpendicular to the capsid surface. Long surface loops rise above the core of the protein and are partially disordered in the X-ray structure. These loops are seen as prominent protrusions in the virion cryo-EM structure (Stewart *et al*., 1991). Another type of capsomer, the penton base, occupies the vertices. More recently, the X-ray structure of the penton base protein has been solved, too (Zubieta *et al*., 2005). It has the same β-barrel fold as that present in the hexon. Two long insertions form an additional distal domain, which has the cellular integrin binding RGD-sequence.

Attached to the penton bases are trimers of the fibre protein that extend above the capsid. Adenovirus uses the fibre to

attach to its cell surface receptor. The trimeric fibre consists of a knob domain and a shaft. The fold of the receptor binding knob domain is an eight-stranded, anti-parallel β-barrel with a Greek-key topology (Xia *et al*., 1994). The structure of the shaft has revealed a novel fold, a triple-β-spiral, which is repeated 22 times to make an elongated protein (van Raaij *et al*., 1999).

The fitting and subsequent subtraction of the hexon density from the virion reconstruction enabled the analysis of the minor components in the capsid (Stewart *et al*., 1991; Stewart *et al*., 1993). More recently, fitting of both the hexon and the penton structures into a refined 3D reconstruction at 10-Å resolution yielded a quasi-atomic model of the adenovirus capsid (Fabry *et al*., 2005). Another group refined the structure to 9-Å resolution and had simultaneously followed a similar approach (Saban *et al*., 2005). These studies shed light on the position of proteins IIIa, VI, VIII and IX and their role in stabilizing the capsid. The interaction of proteins IIIa and IX with the hexons occurs mainly above the hexon β-barrel domains (Fabry *et al*., 2005). Twelve molecules of protein IX form a stable assembly with the nine hexons of the 'group of nine' (Furcinitti *et al*., 1989). Protein VI is thought to reside under the peripentonal hexons (Stewart *et al*., 1993; Fabry *et al*., 2005). Protein VIII is suggested to locate under the capsid shell organized on a '*T*=2' lattice. Thin unassigned densities are present at the two-fold axis of symmetry (Fabry *et al*., 2005).

2.3. Large membrane-containing dsDNA viruses

Paramecium bursaria Chlorella virus type 1 (PBCV-1)(Yan *et al*., 2000), Chilo iridescent virus (CIV)(Yan *et al*., 2000), *Phaeocystic puchetii* virus (PpV01)(Yan *et al*., 2005) and a giant mimivirus (Xiao *et al*., 2005) belong to a group of large dsDNA viruses. Similarly to PRD1 (**Section 3**), they all seem to possess a lipid-bilayer under the protein capsid and have pseudo-hexameric capsomers which are trimers of their corresponding major capsid proteins. The twelve vertices are occupied with a pentameric capsomer of a different protein.

PBCV-1 (*Phycodnaviridae*) infects chlorella-like unicellular green algae. The genome is 331,000 base pairs (bp) long, and has 387 putative genes. The 190-nm capsid is composed of 1,680 hexavalent capsomers organized on a pseudo *T*=169 *dextro* lattice (**Figure 7**)(Yan *et al*., 2000). The X-ray structure of its major capsid protein revealed the viral double β-barrel (Nandhagopal *et al*., 2002).

CIV (*Iridoviridae*) is an insect virus with a 185-nm diameter capsid. The capsid consists of 1,460 hexavalent capsomers organized on a pseudo *T*=147 lattice (Yan *et al*., 2000)(**Figure 7**). A second capsid shell seems to be sandwiched between the outer capsid and the lipid-bilayer. Such a shell was not observed in PBCV-1.

PpV01 is a marine algal virus. It is not yet taxonomically assigned, but shares a number of common properties with members of the *Phycodnaviridae*, such as PBCV-1. The diameter of the capsid is 220 nm. The capsid consists of 2,180 hexavalent capsomers organized on a *T*=219 *dextro* lattice (Yan *et al*., 2005)(**Figure 7**). One hexavalent capsomer is different from all the others in the asymmetric unit, since a 6-nm long fibre extends from it.

A giant mimivirus (*Mimiviridae*) has been isolated from amoebae. It is the largest virus known (diameter 500 nm) and its genome size (1.2 Mbp) exceeds those of some small bacteria (La Scola *et al*., 2003). Its structure has been solved recently (Xiao *et al*., 2005). Based on the capsid dimensions and assuming a similar size for the capsomer as in PBCV-1, it was estimated that the capsid has a *T*=1,179 triangulation.

The large icosahedral virus capsid structures can be thought to consist of groups of capsomers called trisymmetrons

and pentasymmetrons (**Figure 7**). Twenty trisymmetrons and twelve pentasymmetrons are centered at the three-fold and five-fold axes of symmetry, respectively. These sub-structures may be true capsid assembly intermediates, as suggested for Sericesthis iridescent virus in the late 1960s (Wrigley, 1969). In PBCV-1, CIV and PpV01, the pentasymmetrons consist of 1 pentamer and 30 trimers. However, the number of trimers in the trisymmetrons is different. The trisymmetrons have 66, 55 and 91 trimers in PBCV-1, CIV and PpV01, respectively (**Figure 7**). This explains the different capsid sizes and triangulation numbers of these viruses (Yan *et al.*, 2005). Analogously, capsid sub-structures called 'groups of nine' have been observed for adenovirus (**Section 2.2**). However, it is still an open question whether the tri- and pentasymmetrons are general assembly intermediates in dsDNA viruses and what factors limit their size during assembly.

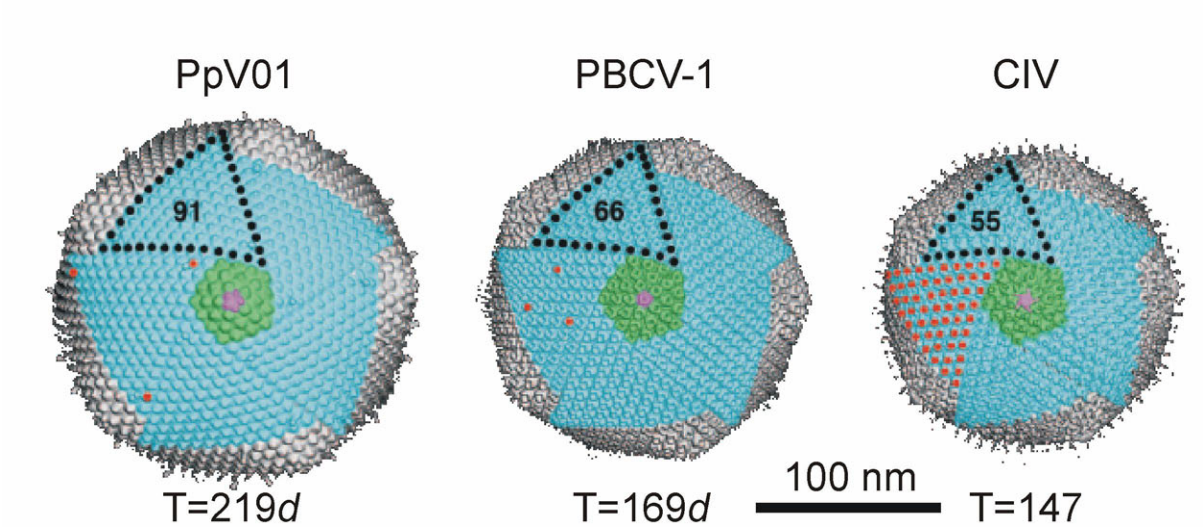

Figure 7. 3D reconstructions of PpV01, PBCV-1 and CIV. The pentasymmetrons are shown in green and the trisymmetrons in blue. Some pentameric (magenta) and hexameric (red) capsomers are highlighted. The number of capsomers in each trisymmetron is indicated. Modified from Yan *et al.* (2005) with permission.

2.4. Herpes simplex virus

Herpes simplex virus type 1 (HSV-1, *Herpesviridae*) is a large membrane-containing dsDNA virus, ~2250 Å in diameter. It is a human pathogen causing cold sores and, more rarely, severe encephalitis. The membrane surrounds an icosahedrally symmetric nucleocapsid (~1250 Å in diameter). The 3D structure of the nucleocapsid has been solved using icosahedral reconstruction (Schrag *et al.*, 1989; Newcomb and Brown, 1991; Newcomb *et al.*, 1993; Zhou *et al.*, 2000). These structures reveal a capsid consisting of 150 hexons and 12 pentons of the same protein, organized on a *T*=16 lattice. In contrast to adenovirus (**Section 2.2**), the hexamers are truly six-fold symmetric and the fold is different from the classic viral β-barrel (Bowman *et al.*, 2003). Smaller subunits called triplexes connect these hexameric and pentameric capsomers in groups of three. The densely packed DNA is ordered as a liquid crystal (Booy *et al.*, 1991). More recently, cryo-electron tomography revealed the structure of the whole virion, including a tegument layer and an outer membrane covered by glycoprotein spikes (Grünewald *et al.*, 2003).

The procapsid matures into a mature capsid through conformational changes.

These changes were observed by solving the structures of different intermediates (Trus *et al*., 1996; Heymann *et al*., 2003). DNA enters the empty, preformed virus particle through a complex called the portal (Newcomb *et al*., 2001). The portal exists in the place of one pentamer at the five-fold vertex of the capsid. Single particle analysis has solved the portal structure (Trus *et al*., 2004). It is a multimer of 12 copies of the same protein. These results suggest that in many aspects, herpesviruses are functionally very similar to dsDNA bacteriophages that package their genome into preformed capsids through a portal.

3. Bacteriophage PRD1

PRD1 is a membrane-containing bacteriophage and a member of the *Tectiviridae* family. The family name originates from the word *tectus* meaning covered, which reflects the fact that the viruses of this family all have a protein shell fully covering the membrane. Tectiviruses are divided into two subgroups based on their host organism. They either infect Gram-negative (PRD1 and its close relatives PR3, PR4, PR5, PR772 and L17) or Gram-positive bacteria (*e.g*. Bam35, see **Section 4**)(Fauquet *et al*., 2004). Those infecting Gram-negative bacteria can infect the host cell only when a multiple drug resistance conjugative plasmid (IncP-, IncN- or IncW-type) is present (Stanisich, 1974; Bradley and Rutherford, 1975). Thus, they are called donor-specific phages, since this plasmid enables the cell to act as a genetic donor and to transfer genes to a recipient cell. The host range of PRD1 and its relatives is broad, including many human pathogens, such as *Escherichia coli*, *Salmonella enterica* and *Pseudomonas aeruginosa*. The genome sequences of PRD1 and its relatives are 91.9–99.8 % identical, based on pair-wise comparisons (Saren *et al*., 2005).

PRD1, the type member of the *Tectiviridae*, is well characterized both biologically and structurally (reviewed in Caldentey *et al*., 1990; Bamford *et al*., 1995; Grahn *et al*., 2005). It can be grown in the well-known *E. coli* host and purified in large quantities. PRD1 is thus a model system amenable for functional and structural studies and is of general interest for several reasons. Firstly, the virion possesses a lipid-bilayer, allowing PRD1 to be used to study membrane structure and morphogenesis. Secondly, PRD1 offers the possibility to study protein-primed replication of linear DNA. Thirdly, and more recently as information on the structure has started to accumulate, it has become evident that PRD1 shares similarities with adenoviruses, infecting birds and mammals, including humans. This has led to a proposal that these viruses share a common ancestor and form a ‘viral lineage’. How broad this lineage extends, is currently under study (Bamford *et al*., 2002a; Bamford, 2003; Benson *et al*., 2004).

3.1. Virion architecture

Cryo-EM and 3D reconstruction at 25-Å resolution have revealed the overall architecture of the PRD1 virion (Butcher *et al*., 1995). The linear dsDNA genome is enclosed by a membrane vesicle, which closely lines the inner surface of the capsid. The capsid surface is punctuated with protrusions of the major capsid protein P3. P3 assembles as trimers (240 copies) into an isometric capsid with pseudo *T*=25 icosahedral symmetry. Penton protein P31 occupies the pentavalent positions of the capsid (Rydman *et al*., 1999). The same pseudo *T*=25 organization is seen in

adenovirus (**Sections 1.2**, **2.2**). A minor capsid component, protein P30, is required for particle assembly, as infection with P30^{-} virus resulted in particles, which possessed only 5–10% of P3. P30 is thought to act as a glue protein stabilizing the capsid, analogously to adenovirus glue proteins (Rydman *et al*., 2001).

The X-ray structure of P3 has revealed its fold (Benson *et al*., 1999). It is composed of two β-barrels with a jelly-roll topology. This fold has also been observed in adenovirus (Athappilly *et al*., 1994)(**Section 2.2**). The larger β-barrel rises higher on the virion surface and appears as a protruding tower. The P3 trimer–trimer contacts have been defined by fitting the X-ray structure into the 3D cryo-EM reconstruction (San Martín *et al*., 2001).

P31 pentamers occupy the holes at the twelve vertices of the P3-shell. They form the base of the vertex complex, composed of the spike protein P5 and the receptor binding protein P2. The presence of P2 is dependent on P5. Both P2 and P5 are missing in mutants lacking the penton protein P31 (Rydman *et al*., 1999). Although the spikes are seen in the electron micrographs, they are unresolved in the icosahedrally averaged reconstructions (Butcher *et al*., 1995; Rydman *et al*., 1999). Thus, the spatial arrangement of the proteins remains unknown. An X-ray structure has been determined for P2 (Xu *et al*., 2003). The P2 structure reveals an elongated molecule with the shape of a sea horse—a β-strand tail and a head with a pseudo β-propeller fold. The head part is proposed to be distal to the virus and responsible for receptor binding. The other component of the spike complex, P5, can be isolated from the virion as a trimer. Small angle X-ray scattering has revealed its elongated shape (Sokolova *et al*., 2001). P5 has at least two separate domains, an N- and a C-terminal domain. The N-terminal domain attaches P5 to the virion (Bamford and Bamford, 2000), while the distal C-terminal domain is required for trimerization (Caldentey *et al*., 2000). The two domains are separated by a collagen-like motif and a hinge of eight glycines. The deletion of these regions and the incorporation of mutations into the viral genome have defined the phenotypes of such mutations (Huiskonen *et al*., 2003). The glycine-hinge deletion slowed infection while the deletion of the collagen-like region reduced the fraction of receptor-bound virus particles to less than half.

3.2. Replication cycle

PRD1 uses protein P2 to recognize the host cell surface receptor, which is encoded by a conjugative plasmid (Mindich *et al*., 1982b). The vertex structure is metastable and able to disassemble during genome injection (Rydman *et al*., 1999). The genome is released into the host cell, while the empty capsid remains outside (Lundström *et al*., 1979). The analysis of suppressor sensitive mutants has revealed proteins essential for infectivity. While the spike complex proteins (P2, P5, and P31) are required for absorption, proteins P11, P14, P18, and P32 are involved in DNA delivery (Mindich *et al*., 1982a; Mindich *et al*., 1982b; Rydman *et al*., 1999; Grahn *et al*., 2002a). Electrochemical measurements of cells infected with mutant viruses suggested that the first one involved in DNA delivery is protein P11 (Grahn *et al*., 2002a). The missing of any of the other proteins (P14, P18 or P32) abolishes the ability of the round lipid membrane to transform into a tube-like structure (Grahn *et al*., 2002b; Grahn *et al*., 2002a). This membrane tube is suggested to act as a DNA injection device, analogously to the tails of many dsDNA phages (Lundström *et al*., 1979; Bamford and Mindich, 1982).

The linear dsDNA genome replication mechanism has been studied in a minimal *in vitro* replication system using purified components (Savilahti *et al*., 1991). Phage encoded DNA polymerase carries out

the genome replication reactions. Covalently linked terminal proteins prime the replication similarly to priming in adenovirus and ϕ29-type phages (reviewed in Salas, 1991).

In the assembly of progeny virions, the major capsid protein P3 and spike complex proteins P2, P5 and P31 are expressed in soluble form in the host cytoplasm, whereas the membrane-proteins (*e.g.* P11, P14, P18 and P32) are targeted to the host cell cytoplasmic membrane. Particles consisting of the protein capsid and the membrane with phage membrane proteins, but devoid of DNA, are assembled first. These preformed procapsids are then packaged with the genome. Packaging adenosine triphosphatase (ATPase) P9, a structural component of the virion, then translocates the viral genome into these preformed procapsids (Strömsten *et al.*, 2005), similarly to HSV-1 (see **Section 2.4**). The packaging occurs through a portal structure located at a single vertex (Gowen *et al.*, 2003). The degradation of the host cell wall by two phage-encoded lytic enzymes and a helper protein results in the release of the progeny virions (Mindich *et al.*, 1982a; Caldentey *et al.*, 1994; Rydman and Bamford, 2002).

4. Bacteriophage Bam35

Bam35, a Tectivirus morphologically similar to PRD1, infects Gram-positive bacteria including *Bacillus thuringiensis* serovar alesti (Ackermann *et al.*, 1978; Ravantti *et al.*, 2003). Four other tectiviruses infecting Gram-positive bacteria have been identified: bacteriophage AP50 infecting *B. anthracis* (Nagy, 1974), ϕNS11 infecting *B. acidocaldarius* (Sakaki *et al.*, 1977), GIL01 infecting *B. thuringiensis* serovar *israelensis* (Verheust *et al.*, 2003) and GIL16 infecting *B. thuringiensis* strain B16 (Verheust *et al.*, 2005). Some of these tectiviral phages can exist as a prophage. For example, GIL01 was first identified as a linear plasmid pGIL01. Its sequence is almost identical to the Bam35 genome, with only 11 mismatches and one single nucleotide frame-shift in both of the inverted terminal repeats. A prophage state was also demonstrated for Bam35 (Strömsten *et al.*, 2003). The linear plasmid pBClin15 of *B. cereus* also has a genome sequence similar to that of Bam35. pBClin15 is suggested to be a degenerate prophage that has lost its ability to assemble virions (Ivanova *et al.*, 2003; Strömsten *et al.*, 2003; Verheust *et al.*, 2005).

4.1. Virion architecture

Bam35 has an isometric protein capsid surrounding a membrane (Ackermann *et al.*, 1978; Laurinavicius *et al.*, 2004). The membrane encloses the linear dsDNA genome. The genome has been sequenced. It has a similar genome organization and size as the PRD1 genome (~15 kbp), but no significant sequence similarity (Ravantti *et al.*, 2003). Thirty-two open reading frames were predicted to encode for proteins. Ten of these have predicted transmembrane helices. N-terminal sequencing has identified eight structural proteins, including the major capsid protein (gene product [gp] 18)(Ravantti *et al.*, 2003). Its fold is predicted to be similar to that of the PRD1 protein P3 (Benson *et al.*, 1999; Ravantti *et al.*, 2003; Strömsten *et al.*, 2003; Benson *et al.*, 2004). No Bam35 X-ray structures are currently available.

4.2. Replication cycle

Bam35 displays membrane tubes similar to those of PRD1 (Ackermann *et al*., 1978). In PRD1, these tail-like structures are thought to penetrate the cell outer membrane and deliver the genome into the host (Grahn *et al*., 2002a). The Bam35 membrane might also be involved in delivering the viral genome into the host cell, although the structure of the cell wall is different in Gram-positive bacteria (Gaidelyte *et al*., 2005). Two lytic enzymes have been reported for both Bam35 and Gil01 (Ravantti *et al*., 2003; Verheust *et al*., 2004). These enzymes are possibly responsible for degrading the bacterial peptidoglycan layer. In the lysogenic pathway, Bam35 late functions were suppressed (Gaidelyte *et al*., 2005). The lytic pathway of Bam35 leads to assembly of progeny virions. The viral lipids are derived from the host plasma membrane. This process is selective, enriching some lipid species (Laurinavicius *et al*., 2004). Progeny virions are released upon host cell lysis.

5. Bacteriophage PM2

PM2 is the only characterized member of the *Corticoviridae* family. It is a dsDNA virus with a circular super-coiled genome. It infects the marine, Gram-negative bacteria *Pseudoalteromonas espejiana* BAL-31 and *Pseudoalteromonas* sp. ER72M2 (Espejo and Canelo, 1968; Kivelä *et al*., 1999). PM2 was the first bacteriophage that was shown to contain lipids and was actively studied as a membrane morphogenesis model in the 1970s. The PM2 system has been reviewed recently (Bamford and Bamford, 2005).

5.1. Virion architecture

The virion is spherical and 60 nm in diameter. The composition is 72% protein, 14% lipid and 14% nucleic acid (Franklin *et al*., 1976). The virion contains at least nine viral proteins (P1–P9)(Kivelä *et al*., 1999; Männistö *et al*., 1999; Kivelä *et al*., 2002). Protein P2 is the major component of the capsid. Protein P1, a minor component, is responsible for host cell attachment (Kivelä *et al*., 2002). Freeze-thawing or chelation of calcium ions release these proteins from the virions resulting in the formation of lipid core (LC) particles. P2 proteins are released as trimers and P1 proteins as monomers. The rest of the structural proteins (P3–P9), however, stay associated with the LCs (Kivelä *et al*., 2002). The highly super-coiled genome resides inside of the inner LC particle.

5.2. Replication cycle

P1 is required for the recognition of a host cell surface receptor (Kivelä *et al*., 2002). Viral replication occurs in proximity to the cytoplasmic membrane (Brewer, 1978). The PM2 genome is replicated using the rolling circle mechanism (Espejo *et al*., 1971; Männistö *et al*., 1999). The 21 putative PM2 genes are arranged in three operons (Männistö *et al*., 2003). A host RNA polymerase carries out the transcription of these operons (Zimmer and Millette, 1975), although viral transcription factors control the promoters (Männistö *et al*., 2003). Studies using temperature sensitive PM2 mutants identified an assembly intermediate similar to the LC

(Brewer, 1980). Viral membrane morphogenesis results in an uneven phospholipid distribution; phosphatidyl-glycerol is enriched in the outer leaflet, and phosphatidylethanolamine in the inner leaflet (Schäfer *et al*., 1974). A putative ATPase (P9) may package the viral DNA (Männistö *et al*., 1999; Kivelä *et al*., 2002). Upon lysis, approximately 300 virions are released from each single host cell (Kivelä *et al*., 1999).

6. Evolution of viruses and viral proteins

Biological organisms are generally classified hierarchically based on their morphology. This taxonomic approach was founded by the Swedish natural scientist Carl von Linné (1707–1778). Due to developments of genome sequencing and structural biology, it is possible to study the possible homologies at the sequence and protein structural level. Viruses that share homologous key components are thought to share a common ancestor. Since viruses do not have any fossil record, the study of homologous structures and sequences in the currently available viruses is the only approach to construct a scenario of virus evolution. However, horizontal gene transfer and high mutation rates hamper such classification (Hendrix, 2003).

While sequence based methods are suitable for studying closely related viruses, morphology or structure based classification would be more suitable for classification of more distantly related viruses, since structure is preserved much longer than the nucleic acid sequence in the course of evolution. Structures of whole viruses and their components determined using X-ray crystallography and cryo-EM are accumulating rapidly. These structures offer a wealth of information for structure-based classification. This has led to the hypothesis of evolutionary relatedness within groups of viruses or 'viral lineages' (Bamford *et al*., 2002a; Bamford, 2003; Benson *et al*., 2004). These lineages include PRD1-type, bacteriophage Hong Kong 97 (HK97) -type and BTV-type viruses, following a recently established nomenclature (Bamford *et al*., 2005). These lineages have members infecting different domains of life. For example, the members of BTV-type infect eukaryotes (BTV)(Grimes *et al*., 1998) and bacteria (ϕ6)(Huiskonen, J.T., de Haas, F., Bubeck, D., Bamford, D.H., Fuller, S.D., Butcher, S.J., manuscript). The group of PRD1-type viruses is discussed in more detail below.

6.1. PRD1-type viruses

The X-ray structure of the PRD1 major capsid protein revealed a double β-barrel with jelly-roll topology, a fold also observed in the adenovirus hexon (Benson *et al*., 1999)(**Figure 8**). In addition, PRD1 and adenovirus share the same pseudo *T*=25 capsid architecture (Stewart *et al*., 1991; Butcher *et al*., 1995) and a genome type, a linear dsDNA molecule with terminal proteins. These findings lead to a hypothesis that PRD1 (infecting Gram-negative bacteria) and adenovirus (infecting birds and mammals, including humans) share a common ancestor (Benson *et al*., 1999). This ancestor would have existed before the domains of life separated.

More recently determined structures strongly support this hypothesis and reveal similarities between viruses infecting all domains of life—Bacteria, Archea and Eukarya. PBCV-1 infects chlorella-like unicellular green algae (**Section 2.3**). The X-ray structure of its major capsid protein revealed the same fold as found in

adenovirus and PRD1 (Nandhagopal *et al*., 2002)(**Figure 8**). PBCV-1 belongs to a group of large dsDNA viruses which possess a membrane under the icosahedral protein shell. Other members with a similar morphology, as seen from their cryo-EM reconstructions, are CIV (Yan *et al*., 2000), PpV01 (Yan *et al*., 2005) and the giant mimivirus (Xiao *et al*., 2005)(**Section 2.3**). The cryo-EM structure of an archeal virus, *Sulfobulos* turreted icosahedral virus, revealed capsomers similar to those of PRD1 (Rice *et al*., 2004). Sequence analysis and morphology of Bam35, a bacteriophage infecting Gram-positive bacteria, suggested another candidate in the group of PRD1-type viruses (Ravantti *et al*., 2003; Benson *et al*., 2004).

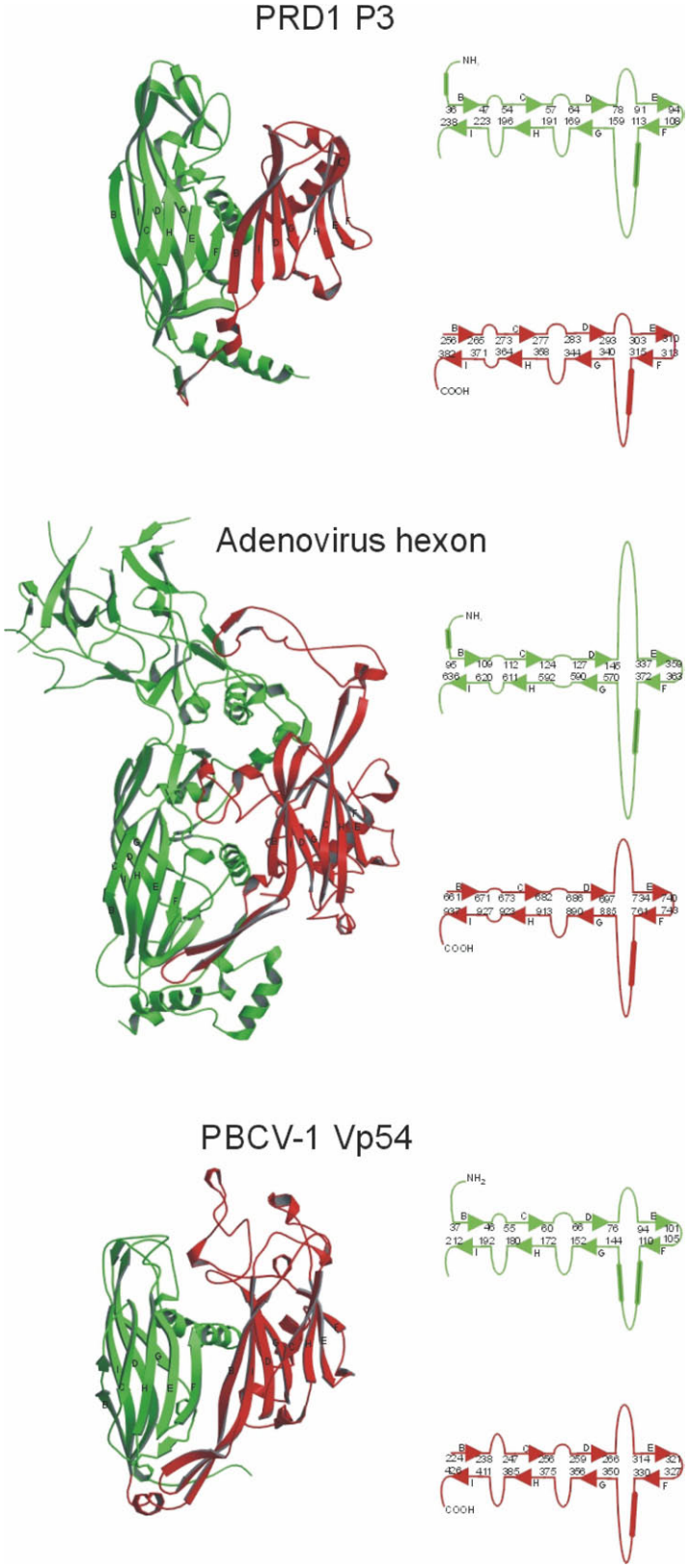

Figure 8. Comparison of capsid protein monomer structures from adenovirus (the hexon), bacteriophage PRD1 (P3) and PBCV-1 (Vp54). The structures reveal the same double β-barrel fold with a jelly-roll topology. The two β-barrel domains are coloured in green and red. The schematic ribbon diagrams show the arrangement of β-strands (arrows), helices (barrels) and loops (lines). Modified with permission from Nandhagopal *et al*. (2002).

7. Cryo-electron microscopy and image processing

7.1. Electron microscopy of biological specimens

Different microscopy techniques are suitable for imaging objects of different sizes and at different resolutions. While light microscopy is suitable for imaging objects as small as the size of a bacterium (~1 μm), the resolution is limited by the wavelength of visible light (400–700 nm). Thus, to image objects smaller than this, and at higher resolutions, the much shorter wavelength of electrons can be exploited. This makes transmission EM a suitable technique to image biological macromolecule structures and their complexes such as viruses (typical diameter 10–100 nm).

Biological specimens contain large amounts of water that would evaporate in the vacuum of the electron microscope. Removal of the water by drying the specimen leads to collapse of biological structures, such as membranes. One solution to the problem is the negative staining technique. However, in this method only a metal replica of the object can be imaged and information about its internal structure is lost. In addition, the structure may be distorted due to staining artefacts as exemplified by the structure of hemocyanin (Boisset *et al*., 1990). An alternative approach to conventional negative staining is the technique of cryo-EM developed by Jacques Dubochet and his colleagues at the European Molecular Biology Laboratory in Heidelberg in the 1980s (Adrian *et al*., 1984; Dubochet *et al*., 1988). The idea is to embed the specimen in vitreous ice, an amorphous state of water, which is a non-crystalline aqueous medium that preserves the specimen in the vacuum of the microscope (**Figure 9**).

7.2. Preparation of vitrified specimens

Since the thickness of the specimen limits the vitrification process, suitable samples for vitrification are aliquots of aqueous suspensions and small pieces of tissue. In the former case, a small amount of the suspension (~3 μl) is pipetted onto an EM-copper grid with holey carbon film. Excess suspension is removed by blotting the grid with a piece of filter paper leaving only a very thin film of suspension (preferably <200 nm). The vitreous state is achieved by rapidly cooling the sample to very low temperature (–140 ^{o}C). In the plunge-freezing technique this is achieved by plunging the sample into liquid ethane cooled by liquid nitrogen (Adrian *et al*., 1984; Dubochet *et al*., 1988). In the case of thicker samples, such as blocks of tissue (~1 mm^3), high-pressure freezing is required to achieve rapid cooling. Due to the rapid cooling rate (> 1000 ^{o}C/s), water molecules do not form ice crystals. This is beneficial for two reasons. First, ordinary freezing of water leads to expansion, which may break biological structures. Second, hexagonal ice in the sample would produce unwanted electron diffraction, hindering the imaging of the specimen itself (**Figure 9a**). In addition, the low temperature leads to ‘cryo-protection’, reducing the radiation damage caused by the electron beam during imaging in the EM (Roos and Morgan, 1990).

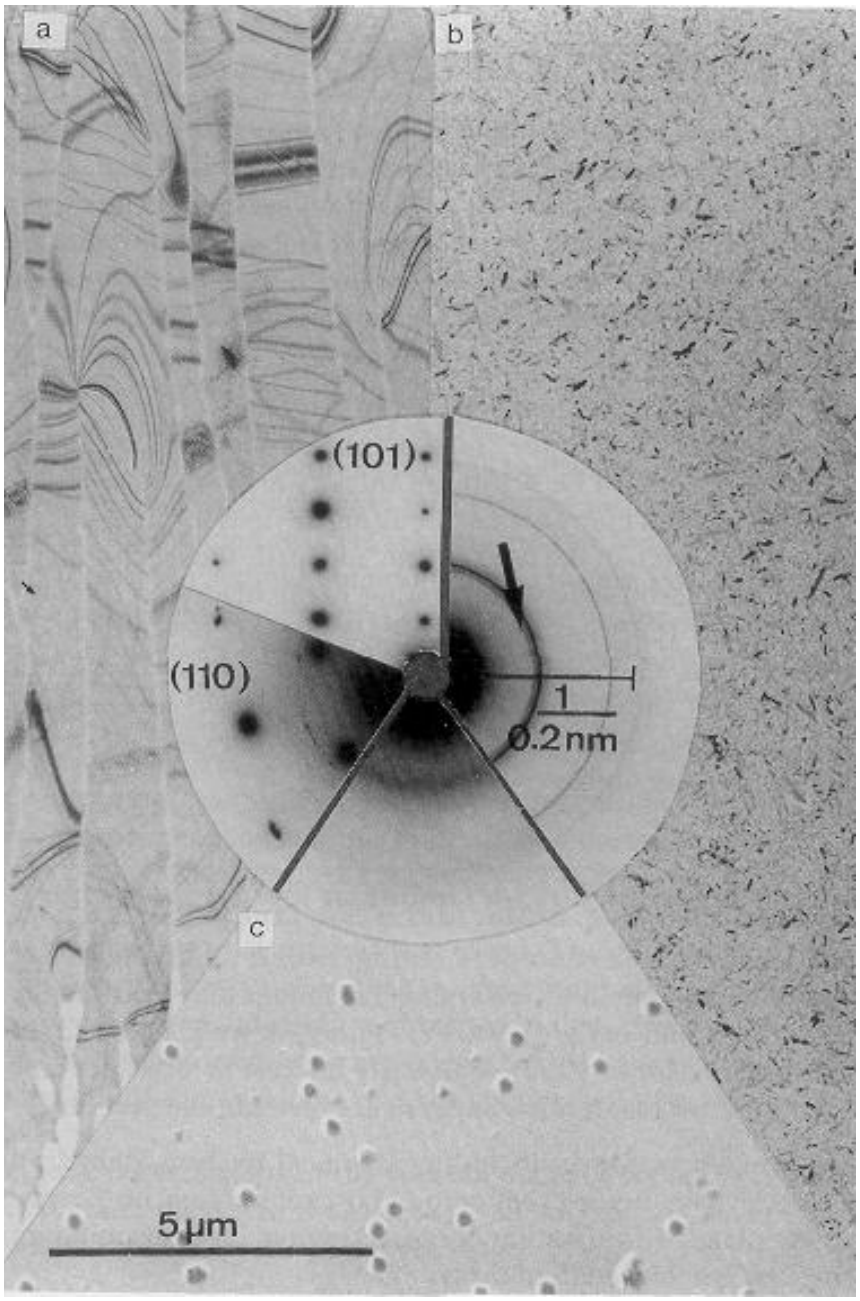

Figure 9. Typical images and electron diffraction patterns of three forms of ice. (**a**) Hexagonal ice. The layer is 50–80 nm thick. The reflection spots in the diffractograms taken from another sample of hexagonal ice reveal the crystalline order. Two lattice planes of the ice crystal are indicated with (101) and (110). (**b**) Cubic ice, obtained by warming a layer of vitrified ice (see *c*). The arrow indicates a small contribution of hexagonal ice in the layer. The layer is ~70 nm thick. (**c**) Vitrified ice. The layer is ~70 nm thick. The polystyrene beads added as focusing aid are clearly visible. The images are all presented at the same magnification 7,500×. With permission from Dubochet *et al.* (1982).

7.3. The electron microscope

In the electron microscope, electrons pass information about the object to a detector. The design of an electron microscope is illustrated in **Figure 10**. Two types of electron sources exist: filaments and field emission guns (FEGs). A filament is usually made of a tungsten or lanthanum hexaboride (LaB_6) wire, from which the electrons are emitted. The emitted electrons are accelerated across a large potential difference between the anode and cathode (on the order of 100,000 V). FEGs provide a highly coherent and bright electron beam, especially important for imaging biological molecules at high resolution (Siegel, 1971; Orlova and Saibil, 2004).

Lenses consisting of magnetic coils are used to focus the electron beam. The condenser lens system focuses the electron beam onto the specimen plane creating good imaging conditions. The objective lens and the projector lens system create the magnified image of the object. The objective lens determines the resolution of the image and the condenser lenses make the final magnified image. The most critical lens in high resolution imaging is the objective lens, since the lens aberrations are magnified in the subsequent lenses. Apertures in each lens are used to restrict the width of the beam. Different aperture sizes can be used to limit the number of electrons hitting the specimen (Chescoe and Goodhew, 1990).

Since electrons cannot pass through air, a complicated and effective series of pumps is required to maintain a vacuum in the microscope column. Energy filters can be used to filter electrons that have scattered inelastically from the sample. This filtering decreases the noise component in the image (Yonekura *et al.*, 2002)(see **Section 7.5**).

Different detectors can be attached under the specimen for imaging. Conventionally, film is used to collect the images. The developed films are then digitized with a scanner for computer processing. A more recent alternative to film are charged-coupled devices (CCDs). They are able to form a digital image directly. With a CCD, the electron image is first transformed into a light image in a scintillator. The light image is then passed to the CCD and recorded.

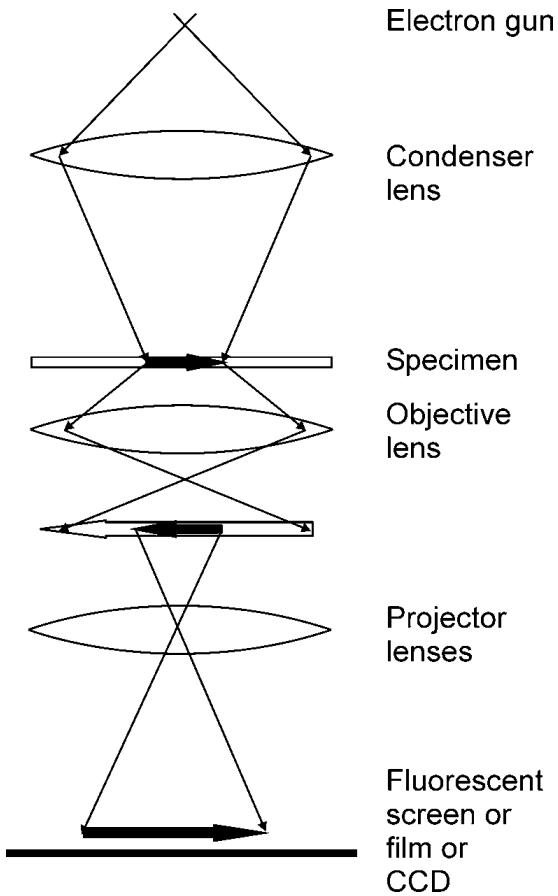

Figure 10. Schematic illustration of a transmission EM. Modified from Chescoe and Goodhew (1990).

7.4. Imaging of vitrified specimens

After preparation, the sample has to be kept at the temperature of liquid nitrogen in order to prevent the conversion of vitrified water into crystalline water. This poses challenges for EM-imaging. A special cryo-transfer station is required to move the sample grid from the storage container into a cryo-specimen holder which is then inserted into the microscope. During imaging, the holder and the specimen therein are cooled by liquid nitrogen or helium (Roos and Morgan, 1990).

The radiation dose is kept minimal using a low-dose imaging mode in order to prevent radiation damage caused by inelastic electron scattering events (see **Section 7.5**). In this mode, the grid is first scanned at low magnification and dose (0.01 $e^-/Å^2$) in order to locate suitable areas for imaging (*e.g.* areas with proper ice thickness). Focusing and the correction of astigmatism require high magnification and sufficient intensity. For these operations, the beam is shifted or deflected next to the area being imaged. The actual area is exposed to the high intensity beam only once. Thus, the image can be recorded using the maximal dose (10 $e^-/Å^2$), which directly leads to maximal signal-to-noise ratio (SNR) in the recorded images (also called electron micrographs)(Roos and Morgan, 1990; Frank, 1996). If conventional film is used for recording, it is underexposed in low dose EM-imaging. For compensation, the push processing technique is used, where the film is over-developed during the processing phase. This allows the silver grains in the film to grow larger than usual, and the signal is enhanced.

7.5. Image formation in the electron microscope and image contrast

In order to interpret the electron micrographs and relate these 2D images to the actual 3D structure of the object, one has to understand the EM image formation process: How does the electron beam interact with the specimen when passing through it? Electron scattering occurs when beam electrons interact with the electrostatic field of nuclei or the electrons in the electron shells surrounding the nuclei. Beam electrons may also pass outside the range of these fields and remain unscattered. During the scattering event, the electron may lose energy to the specimen (inelastic scattering), or its energy may be unaltered (elastic scattering). Beam electrons that hit the electrons of the specimen scatter inelastically. This scattering occurs mainly at small angles and is associated both with an energy loss and a change in wavelength. The lost energy is deposited in the specimen and may damage it. Beam electrons that pass close to the atom nuclei scatter elastically. They exhibit a change in direction.

Contrast is a quality of the image required to see the object separate from its background. Contrast in the EM image is an outcome from two different sources, historically called ‘amplitude’ and ‘phase’ effects (Erickson and Klug, 1971). Electrons that scatter at large angles do not contribute to the image but are blocked by the objective aperture. This gives raise to amplitude contrast (also called scattering contrast), since some areas are more electron opaque than others and thus appear less intense in the micrograph. Phase contrast (also called interference contrast), in turn, is a result of interference of the elastically scattered and the unscattered parts of the electron beam. Since biological samples consist of very light atoms, scattering is generally weak and the images show very little amplitude contrast unless a heavy metal stain is used. Cryo-EM specimens considered here are generally unstained and the main source of contrast is phase contrast. A defocused electron beam enhances the phase contrast, since the path lengths of the scattered electrons become relatively longer

than those of the unscattered electrons. Normally, the beam is focused under the sample plane, instead of over-focusing, in order to avoid a cancelling of 'amplitude' and 'phase' effects, which have opposite signs at low spatial frequencies at over-focus.

Due to the described contrast effects, image formation in EM is not straightforward. In addition, the electromagnetic lenses are usually not perfect, but have spherical aberration. This poses a theoretical resolution limit (*d*) to EM imaging (Wade, 1992), which can be defined as

$$d = (C_s \cdot \lambda^3)^{1/4} \qquad (\mathbf{1})$$

where C_s is the spherical aberration coefficient of the objective lens and λ is the electron wavelength, which depends on the electron acceleration voltage of the microscope.

The relation of the object and its projection image is best described mathematically as 2D Fourier transforms (Siegel, 1971; Wade, 1992; Frank, 1996; Baker *et al*., 1999). The image transform T_i is not directly related to the object transform T_o but is affected by a contrast transfer function (CTF) so that

$$T_i(\nu)=T_o(\nu) \cdot CTF(\nu), \qquad (\mathbf{2})$$

where ν denotes the spatial frequency (reciprocal resolution). ν equals α/λ, where α is the scattering angle in the microscope. The CTF thus describes the relative contrast in the image as a function of spatial frequency and is characteristic of the microscope and the imaging conditions used. The function includes phase- and amplitude-contrast components, which are independent of each other. CTF is defined as

$$CTF(\nu) = -[(1-F^2_{amp})^{1/2} \cdot \sin \chi(\nu) + F_{amp} \cdot \cos \chi(\nu)] \cdot E(\nu), \qquad (\mathbf{3})$$

where $\sin\chi(\nu)$ is the phase contrast, $\cos\chi(\nu)$ is the amplitude contrast and $E(\nu)$ is a spatial-frequency and electron source-dependent contrast attenuation function, also called the envelope function. F_{amp} is the fractional amplitude contrast and can be considered as a constant (Toyoshima and Unwin, 1988). Since in unstained cryo-EM specimens F_{amp} is approximately 0.07, from **Equation 3** it is clear that the phase component is the major source of contrast. $\chi(\nu)$ is a phase shift function describing the phase distortion due to aberrations, and is defined as

$$\chi(\nu)=\pi \cdot \lambda \, (\Delta f \cdot \nu^2 - {}^1/_2 \, C_s \cdot \lambda^2 \cdot \nu^4), \qquad (\mathbf{4})$$

where Δf is the beam defocus (positive for underfocus). Thus, the image contrast can be increased at low spatial frequencies by defocusing the electron beam, since defocusing shifts the first contrast maximum to lower spatial frequencies (**Figure 11**). At perfect focus, the contrast arises only from spherical aberration, and is very low. The contrast is maximal only at certain spatial frequencies, where $\sin \chi(\nu) = \pm 1$. At other frequencies the absolute contrast is lower, or completely absent. Notice that at some frequencies the contrast is reversed ($CTF(\nu) < 0$). Over a limited spatial frequency range, the spherical aberration can be balanced with a certain defocus value, called Scherzer defocus $\Delta f_s = (C_s \cdot \lambda)^{1/2}$, and the phase stays at $-\pi/2$ (Wade, 1992). For a 200-kV microscope with spherical aberration of 2 mm, Δf_s=71 nm. Thus it is too close to the focus to be practical for imaging of low contrast cryo-EM specimens.

The contrast is also attenuated at higher spatial frequencies by the envelope function $E(\nu)=e^{-(\delta \cdot \nu)^2}$ (**Equation 3**), which depends on the beam coherence and specimen drift. In addition, inelastic scattering has a similar contrast-dampening effect. This attenuation limits the SNR at high spatial frequencies and thus the resolution of EM images, too.

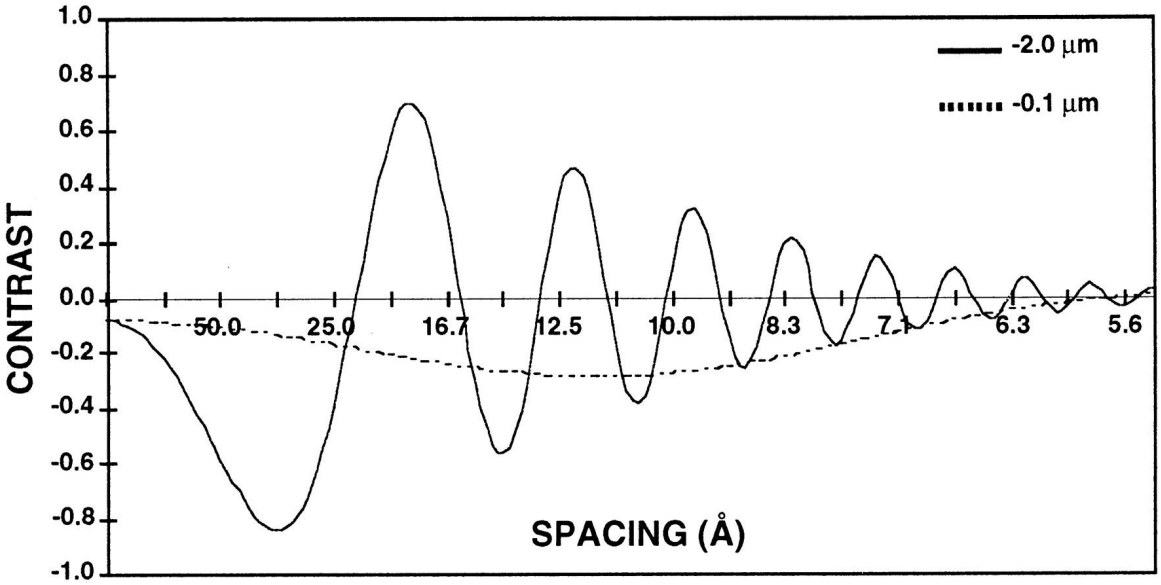

Figure 11. Simulated contrast transfer function. CTF (**Equation 3**) is plotted as a function of spatial resolution (ν) for two different defoci (Δf). The electron voltage is 200 kV, which is equal to a 0.02527-Å electron wavelength (λ). Spherical aberration (C_s) is 2.0 mm. With permission from Baker *et al*. (1999).

7.6. Analysis and pre-processing of the electron micrographs

Many factors affect the quality of the collected electron micrographs. Careless handling of the vitrified sample may cause a rise in the temperature above the vitrification temperature, and formation of cubic or hexagonal ice (**Figure 9**). Astigmatism of the lenses causes distortion of details. Movement of the specimen during exposure causes blurring of details along the direction of the movement. This specimen drift can be beam-induced or due to the instability of the specimen holder. The electron beam can charge the specimen grid causing movement. The specimen holder reacts to changes in temperature by contracting or expanding, which causes movement of the specimen in the direction parallel to the specimen holder.

The quality of the micrographs is most easily evaluated from their Fourier transforms (Erickson and Klug, 1971). This can be achieved using either an optical or a computational transform of the micrograph. In the transform, the cubic and hexagonal forms of ice are visible as sharp rings and diffraction spots, respectively (**Figure 9**). The contrast reversals are visible as Thon-rings (Thon, 1971). The ellipticity of the Thon rings is a sign of astigmatism in the image. Uneven thickness of the Thon rings in certain direction indicates specimen drift in the direction parallel to this. Weak or non-existent Thon rings indicate that high-resolution information is not present in the image. This amplitude decay may be due to excessive inelastic scattering caused by too thick specimen, beam induced specimen drift or damage, incoherent beam or other factors (Saad *et al*., 2001).

Only images with strong and even Thon rings and a suitable number of objects of interest ('particles') are selected for further processing, since the addition of poor-quality data in the analysis leads only to poor-quality results. The selection also depends on the target resolution of the final model. In the area of the whole micrograph, smaller sub-images of the particles are located and extracted for further processing. This particle picking can be manual or

computational (Kivioja *et al.*, 2000). The picked particles can be evaluated visually to discard any abnormalities, such as misidentified or broken particles.

7.7. Correction for the contrast transfer function

Due to the contrast reversals caused by the CTF, image details appear white at some spatial frequencies, and black at some other frequencies, which causes serious artefacts in the image. If the image is to be interpreted beyond the first contrast reversal, this artefact has to be corrected for. This requires calculation of the value of the CTF from the imaging parameters. While all of the other parameters are known (spherical aberration, electron wavelength and fractional amplitude contrast, which can be estimated), the exact defocus value (Δf, **Equation 4**) is unknown and has to be determined from the EM image (Mindell and Grigorieff, 2003). A CTF correction, where the phases are 'flipped' at appropriate spatial resolution, can then be performed:

$$T_o(\nu)=\begin{cases} -T_i(\nu), \text{ for } CTF(\nu) < 0 \\ T_i(\nu), \text{ for } CTF(\nu) \geq 0 \end{cases} \qquad (\mathbf{5})$$

This simple correction, however, leaves the amplitudes unaffected. In order to correct both the phases and the amplitudes, the contrast can be divided by the CTF in Fourier space. Instead of direct division by the CTF, a Wiener filter, such as the one given in **Equation 6** (Marinescu *et al.*, 2001), can be used to avoid division by zero.

$$T_o(\nu)=T_i(\nu) / [CTF(\nu) + w\cdot(1-|CTF(\nu)|)], \qquad (\mathbf{6})$$

where w is a Wiener factor (a small value, such as 0.1). The spatial frequencies at which the amplitude of the image transform is zero carry no information. Thus, one has to collect multiple images of the specimen at various different defoci and combine this information to cover the whole spatial resolution range.

7.8. Three dimensional reconstruction

The thickness of the specimen is normally much smaller than the depth of field of the microscope. Thus, all levels in the specimen are practically at the same focus level and the EM image can be considered as a 2D projection of the electron scattering potential of the 3D structure in the specimen (De Rosier and Klug, 1968). To calculate a 3D representation of the structure from the projection images, different views (projections in different angles) of the structure are required. However, since 2D projections are used, they do not directly reveal information about the absolute handedness of the biological structure. Only a 3D model of the structure and image pairs taken at different angles can determine the absolute handedness (Klug and Finch, 1968).

According to the projection theorem, Fourier coefficients calculated from a 2D projection of the object correspond to a section through the origin of the 3D Fourier transform of the original object. Thus, the 3D volume of an object can be reconstructed by 'filling' the 3D Fourier space with 2D Fourier transforms and calculating an inverse transform. The number of views required to evenly fill the Fourier space depends on the size of the object and on the desired resolution. Before the reconstruction can be calculated, one has to determine the orientation of the object that gave raise to each particular projection (De Rosier and Klug, 1968).

The type of the specimen and whether or not tilting of the specimen in the microscope is required to produce more than one view of the structure determines the

applicable technique for orientation determination and 3D reconstruction (**Table 1**). If the object is unique, or if all of its copies orient so the direction of view is always along the symmetry axis, tilting is required to produce different views. Tilting techniques include tomography (Hoppe *et al*., 1974) and the random conical tilt method (Radermacher *et al*., 1987). Large tilt angles imposed in these techniques result in a non-uniform focus in the specimen plane. This is challenging to take into account in the correction for the CTF, and thus the resolution may be limited to the first contrast reversal of the CTF in these techniques. If the object exists as a 2D ordered crystal, the structure can be solved by combining the phase-information from the images and amplitude information from electron diffraction data (Matricardi *et al*., 1972). Tilting of the crystal is required to resolve the structure in the direction orthogonal to the plane of the crystal (Glaeser, 1999).

Multiple copies of non-crystalline macromolecular assemblies are referred to as 'single particles' (van Heel *et al*., 2000; Orlova and Saibil, 2004). If the object behaves as single particles, identical copies of the object orient randomly in the vitrified water layer. Thus, already one zero-tilt EM-image provides different views naturally (Crowther, 1971). Moreover, if the object has helical symmetry, one object provides multiple views of the asymmetric unit around the helical axis. However, the technique for helical reconstruction conventionally exploits the crystalline nature of the helical assembly and thus differs from the zero-tilt single-particle methods (De Rosier and Klug, 1968).

Zero-tilt techniques can be divided into common lines based methods and projection matching methods. The Fourier space common lines approach is based on an implication of the projection theorem: Fourier transforms of 2D projections from the same 3D object share a common line, on which all the values of the transform are identical (Crowther, 1971). These pairs of common lines can be then used to determine the relative orientation between three or more projections. Angular reconstitution relies on determining common line projections (sinogram common lines) in the 2D projections of the 3D object in real space (van Heel, 1987). Class averages of projections rather than noisy single 2D images are an integral part of this approach. Both the Fourier space and real-space common lines methods are model-free, although they can also be used together with model projections. In the projection matching approach, an earlier 3D model is always required (Baker and Cheng, 1996). This model is projected in all possible different orientations with a chosen angular sampling. These model projections are then compared with the observed 2D projection of the object and the best correlating model projection defines the most likely orientation. Polar Fourier transforms (PFTs) are generally used in projection matching (see **Section 7.8.1**).

Table 1. Methods for 3D reconstruction

Method	Reference
1) Tilting methods	
a) Random conical tilt	Radermacher *et al*., 1987
b) Tomography	Hoppe *et al*., 1974
c) 2D crystallography	Matricardi *et al*., 1972
2) Zero-tilt methods	
a) Common lines methods	
i) Fourier space	Crowther *et al*., 1970b; Crowther, 1971
ii) Real space (angular reconstitution)	van Heel, 1987
b) Projection matching	
i) PFT	Baker and Cheng, 1996
c) Helical reconstruction	De Rosier and Klug, 1968

7.8.1. Icosahedrally symmetric particles

The first single particle studied by 3D reconstruction was an icosahedral virus, tomato bushy stunt virus (Crowther *et al*., 1970a). Out of all single particles, icosahedrally symmetric particles offer the highest pointgroup symmetry (532). The high symmetry of the icosahedron can be exploited in various image processing steps during the 3D reconstruction of icosahedrally symmetric objects such as many viruses (**Figure 12**)(reviewed in Baker *et al*., 1999). First of all, the symmetry can be exploited in the determination of the three angles defining the orientation of the particle (θ, ϕ and ω; **Figure 12**, inset). Due to the high symmetry, 37 pairs of common lines are present within one single projection image. These include two pairs from each of the 6 five-fold axes, one pair from each of the 10 three-fold axes and one pair from each of the 15 two-fold axes. Between two projection images, 60 pairs of cross-common lines exist. The position of these lines determines the orientation of the particle (Crowther, 1971; Fuller *et al*., 1996). The symmetry also provides an internal control. The agreement between the pairs of common lines, the so called phase residual, can be used to assess the consistency within the data set and to monitor the orientation refinement process (Thuman-Commike and Chiu, 1997). In the alternative projection matching approach, PFTs of the projections are commonly used instead of projections for computational efficiency (Baker and Cheng, 1996). The high symmetry limits the orientation search space to a relatively small asymmetric unit, speeding up the computation.

Finally, after the particle orientations are determined, a 3D reconstruction can be calculated. Due to the icosahedral symmetry, each projection images contains 60 copies of the asymmetric unit. This increases the amount of data 60-fold. Thus it is possible, although at very low resolution, to calculate an icosahedral 3D reconstruction from one single image for an icosahedrally symmetric object.

In the ‘icosahedrally symmetric viruses’, it is only the capsid that obeys this symmetry, whereas, for example, the viral genome is not ordered according to this symmetry. In addition, some capsid proteins, especially those associated with the viral vertices, may not obey the icosahedral symmetry. These symmetry-mismatched structures are unresolved using conventional structure determination exploiting icosahedral symmetry. Different non-conventional approaches have been used to study these structures (de Haas *et al*., 1999; Morais *et al*., 2001; Bubeck *et al*., 2005). I and my colleagues have recently developed a method that combines conventional icosahedral reconstruction and single-

particle analysis to resolve symmetry-mismatched structures *in situ* with their inherent symmetry (Briggs *et al*., 2005); Huiskonen, J.T., Jäälinoja, H.T., Briggs, J.A.G, Fuller, S.D, Butcher, S.J., manuscript).

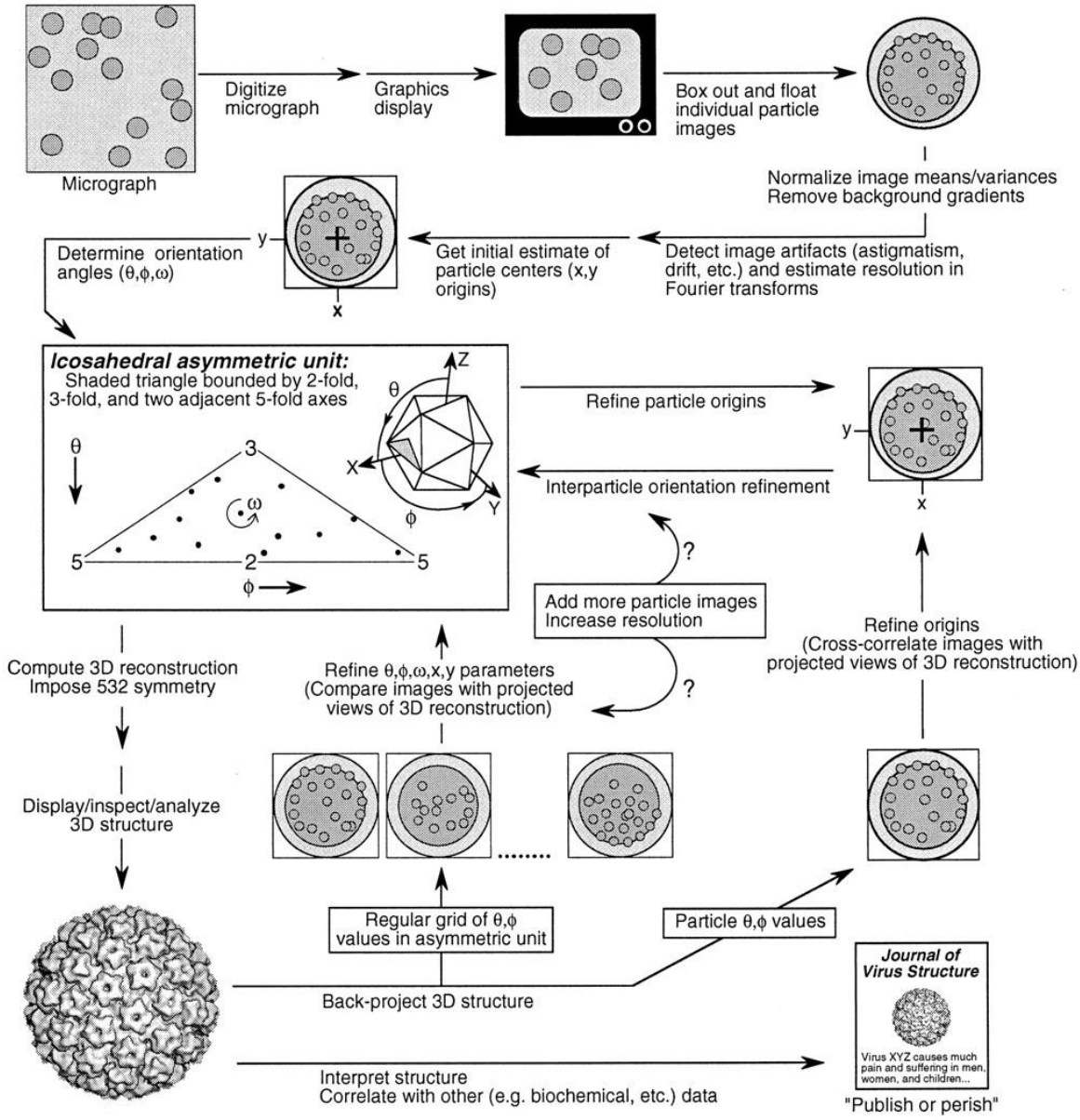

Figure 12. Schematic diagram for icosahedral reconstruction process. The process is iterative in nature and involves cyclic refinement of the origin (x, y) and the orientation (θ, ϕ, ω) parameters defined in the inset. With permission from Baker *et al*. (1999).

7.9. Analysis and interpretation of cryo-electron microscopy based density maps

A 3D reconstruction presents a density distribution of all proteins and other biomolecules (*e.g.* lipid and nucleic acid) in the object. To address any biological hypothesis, one has to interpret the density map. Specifically, one has to assign different parts of the map to different components, and further to interpret this spatial information in terms of other biological knowledge about the system. The raw EM density map itself, however, provides only a little information on how the density is divided between the different molecular components. In addition, the resolution usually limits the detection of molecular boundaries and features.

7.9.1. Resolution assessment of reconstructions

Fourier shell correlation (FSC) is a routinely-used technique to assess the resolution in the EM density maps (Harauz and van Heel, 1986). In the resolution assessment, the original data set is split into two independent halves. From these two sets, two models are reconstructed separately. These models are then correlated in the spatial frequency domain. The frequency at which the correlation drops below a certain threshold value determines the resolution. The two commonly used threshold values are 0.5 (Böttcher *et al.*, 1997; Conway *et al.*, 1997) and 3σ (standard deviations) above the noise level (Orlova *et al.*, 1997). For icosahedral reconstructions these two thresholds are roughly equal (van Heel *et al.*, 2000). The reconstruction calculated from a full data set can also be compared to an atomic model obtained with X-ray crystallography. A threshold value equivalent to this comparison has also been derived for the case, where two half maps are used (Rosenthal and Henderson, 2003). Whatever the criteria used, the amount and type of details visible in the map should reflect the estimated resolution.

7.9.2. Difference imaging

Several difference imaging approaches can be exploited in the interpretation. Firstly, if a structure of some subcomponent (*e.g.* viral major capsid protein) is already known, one can assign density corresponding to this component by searching for its position in the whole density map. In order to analyse any minor, unknown components, one can then subtract the known structure from the density map and analyse the residual densities separately. If the known structure is an atomic model, a continuous density has to be generated from it before subtraction (Belnap *et al.*, 1999). This approach was used for the first time to study the minor components in the adenovirus capsid (Stewart *et al.*, 1991; Stewart *et al.*, 1993). Secondly, if treatments, such as limited proteolysis or the use of detergents specifically affect only certain parts of the complex, then these parts can be assigned to the corresponding density by biochemical analysis and difference imaging between the native and treated complexes. Similarly, some proteins can be removed or shortened using mutagenesis. For example, bacteriophage PRD1 has been subjected to SDS-treatment and mutagenesis of structural proteins prior to structural analysis using cryo-EM (Butcher *et al.*, 1995; Rydman *et al.*, 1999). Thirdly, if the object can be, for instance, studied with and without a bound antibody, difference imaging can define the position to which the antibody binds (Smith *et al.*, 1993).

7.9.3. Protein mass estimates from cryo-EM density maps

The density of proteins is relatively constant (1.35 g/ml)(Harpaz *et al*., 1994). If the level of the molecular surface can be scaled using the volume of a known component, it is possible to determine the molecular mass from segmented parts of the density map or from residual densities in a difference map. These mass estimates can lead to the assignment of a density to a known protein of the complex, or at least rule out some candidates.

7.10. The potential of cryo-electron microscopy in structural biology

During the last ten years, cryo-EM combined with single particle analysis and 3D reconstruction has yielded structures of biological macromolecule complexes at sub-nanometre resolution (Henderson, 2004). In fact, modern electron microscopes can attain a resolving power better than 2 Å. In practice, however, the low contrast of the cryo-EM images and the sensitivity of biological materials to electron radiation limit the resolution to a much lower level. The limited resolution poses limits for structure interpretation. Only when the reconstruction approaches 7-Å resolution, does the internal protein structure, the α-helices and β-sheets, become evident. Only in a small number of cases has resolution as high as 4-Å been reached. At this resolution, the protein backbone and large side chains become resolved and atomic models can be built. Cryo-EM based atomic models from 2D crystals include, for example, those of bacteriorhodopsin (Henderson *et al*., 1990), αβ-tubulin heterodimers (Nogales *et al*., 1998) and aquaporin (Murata *et al*., 2000). Atomic models from helically ordered assemblies include, for example, the bacterial flagellum (Yonekura *et al*., 2003) and the nicotinic acetylcholine receptor (Unwin, 2005).

The well-established techniques of X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy readily yield atomic models of molecules. However, the cryo-EM technique has certain benefits, which make it a complementary technique to X-ray crystallography and NMR (Grimes *et al*., 1999). First of all, whereas X-ray crystallography measures the amplitudes of the 3D diffraction data, in cryo-EM one observes projection images of the object. Thus, the phase component of the data is maintained, and the phase-problem of X-ray crystallography does not exist. Secondly, structure determination using X-ray crystallography requires a crystalline form of the molecule or complex, and finding suitable conditions for crystallization can be very time-consuming and frequently unsuccessful. In addition, contacts between molecules in the crystal lattice may restrict the molecule to a certain conformation or flexibility can render a certain part of the molecule unresolved. For cryo-EM, a purified sample in aqueous solution is sufficient and cryo-EM based reconstructions are not restricted to a certain crystalline conformation, but can sample a population of conformations. Thirdly, cryo-EM is suitable for solving structures of very large assemblies, not generally accessible by X-ray or NMR.

Generally, a combination of different techniques is needed to study large macromolecular assemblies such as viruses. For example, atomic models of individual subunits can be fitted into an EM-density map in order to create a ‘quasi-atomic’ model of the whole structure, as was first shown for the adenovirus capsid (Stewart *et al*., 1993). An initial 3D model from cryo-EM can also provide phase information for X-ray crystallographic structure determination at higher resolution (Baker *et al*., 1999; Chapman and Liljas, 2003).

B. AIMS OF THE PRESENT STUDY

For this study, we chose the type members for groups of membrane-containing, icosahedral dsDNA bacteriophages that have been isolated so far. Two of these, PRD1 and Bam35, have been grouped into the same family, *Tectiviridae*, although they infect Gram-negative and Gram-positive bacteria, respectively. PM2, infecting Gram-negative bacteria, is the only member of the *Corticoviridae*. As the overall architecture is similar between all three viruses, we were intrigued to see how far the similarities would extend, and whether it was possible to detect any new virus assembly principles. The architectures of these phages should provide useful information about general architecture and assembly mechanisms as well as the possible evolutionary relatedness between these phages and certain eukaryotic viruses.

The aim was to use cryo-EM and 3D icosahedral reconstruction to solve the structures of all three phages. When the cryo-EM technique alone was not sufficient we also wanted to combine it with atomic models obtained from X-ray crystallography. The specific aims were:

1. To refine the existing low resolution cryo-EM structure of PRD1 to detect any minor components (**Study I**). Comparison of the minor components between PRD1 and adenovirus should shed light on their evolutionary relationship. The structure of protein P11^{-} deficient PRD1 mutant should reveal its position in the virion and provide further insight into the infection process.

2. To calculate a 3D reconstruction of PM2 (**Study II**). Although PM2 has been studied extensively using biochemical and biophysical approaches, no previous 3D structures existed.

3. To use X-ray crystallography to study the detailed structure of the PRD1 spike protein P5, since it is not resolved using conventional 3D reconstruction methods exploiting the icosahedral symmetry (**Study III**).

4. To solve the first 3D structure of Bam35 virion for comparison with PRD1 (**Study IV**). Earlier, electron micrographs had revealed a similar morphology between PRD1 and Bam35, but no 3D structure existed for Bam35.

We realize that a raw 3D reconstruction of a virus provides little information about how the virus carries out its functions. The goal was to interpret such reconstructions in the light of additional biochemical and genetic knowledge. One general aim was to develop new data analysis tools whenever applicable to aid such interpretation. For example, the 3D density maps obtained for the different membrane viruses provide a 3D structure for the membrane and possibly reveal some associated membrane proteins as well. We exploited this information to study how membrane proteins affect membrane curvature and thickness in these model membranes (**Study IV**). One specific aim was to develop a visualization tool to analyze these membrane parameters.

This work includes the first published results from the national cryo-EM facility, established in the Institute of Biotechnology at the University of Helsinki at the beginning of 2001. Thus, the work is an integral part of the installation and early application of this emerging technique in Finland.

C. MATERIALS AND METHODS

For cryo-EM of PRD1 (**Study I**), an amber mutant (*sus*607) was produced in addition to the wild type (wt) virions. This mutant was devoid of P11 protein. For analysis of PRD1 major capsid protein–membrane interaction, a set of mutations were generated in the major capsid protein. These are described in **Study I**, Table 1. For cryo-EM of PM2, in addition to the native virion, two proteolytically treated virion specimens were prepared, using proteinase-K or bromelain in the digestion as described in **Study II**. For cryo-EM of Bam35, the clear-plaque derivative (Bam35c) was used. The virions were grown and purified as described in the original publications (**Studies I**, **II**, **IV**). For X-ray analysis of the PRD1 spike protein P5, the wt protein was expressed and purified as described in **Study III**. The construction of the mutant lacking the glycine-hinge, using standard molecular biology methods, has been described elsewhere (Huiskonen *et al*., 2003).

Methods used in this study are listed in **Table 2** and described in the original publications. PRD1 data were collected on an FEI Tecnai F20 and a Philips CM200 FEG microscopes, both operating at 200 kV and with a spherical aberration of 2 mm and a nominal magnification of 50,000×. PRD1 reconstructions were calculated as described elsewhere (San Martín *et al*., 2001) using Fourier–Bessel inversion in cylindrical polar coordinates (Crowther *et al*., 1970a; Crowther, 1971). PM2 and Bam35 data were collected similarly on an FEI Tecnai F20 and reconstructions were calculated as described in **Studies II** and **IV**, respectively. Calculations were performed on a parallel Linux cluster. The 3D reconstructions were calculated using Fourier inversion in Cartesian coordinates (Crowther *et al*., 1970a; Marinescu *et al*., 2001). A full CTF-correction for PM2 and Bam35 images was performed using the Wiener filter described in **Equation 6** and a Wiener factor of 0.1.

Table 2. Methods used in this study

Method	Used in			
Computational analysis of protein sequences		II	III	IV
Cryo-electron microscopy and 3D reconstruction	I	II		IV
Difference imaging	I	II		IV
Dissociation of phage particles		II		
Fitting of X-ray structures into EM density maps	I			IV
Membrane curvature and thickness analysis				IV
Molecular cloning techniques	I		III	
Mass spectrometry		II		
Protease treatment of phage particles		II		
Protein concentration measurement	I	II	III	IV
Protein over-expression			III	
Protein purification			III	
SDS polyacrylamide gel electrophoresis		II	III	
Site directed mutagenesis	I			
Three-dimensional reconstruction	I	II		IV
Virus growth and purification	I	II		IV
X-ray crystallography			III	

D. RESULTS AND DISCUSSION

We used cryo-EM and 3D image reconstruction to solve the structures of three different membrane-containing bacteriophages, PRD1 (**Section A.3**; **Study I**), Bam35 (**Section A.4**; **Study IV**) and PM2 (**Section A.5**; **Study II**). When studying icosahedrally symmetric capsid structure, copies of the particles of one virus can be assumed to be identical. The icosahedrally symmetric spherical viruses studied here oriented nearly randomly in the vitrified water layer. Thus we chose a 'single particle' approach combined with zero-tilt imaging (see **Section A.7.8**). The 3D reconstructions calculated are listed in **Table 3**. Advances in computation (Marinescu *et al*., 2001; Ji *et al*., 2003) and larger data sets allowed the refinement of Bam35 and PM2 to higher resolution than the earlier PRD1 reconstructions. The isosurface representations of the reconstructions are shown in **Figure 13** for comparison. While the PRD1 and Bam35 are highly similar in appearance and size, the PM2 capsid is smaller and has prominent spike structures. The spike component of PRD1, which does not obey icosahedral symmetry, was studied separately using X-ray crystallography (**Study III**).

Table 3. Three-dimensional reconstructions

Reconstruction	No. of particles	No. of micrographs	Underfocus (μm)	Nominal pixel size (Å/pixel)	Resolution (Å)*
PRD1					
wt	852	13	0.9–4.1	2.8	13
sus607	569	21	1.0–4.7	2.8	13
Bam35					
native	4,474	69	0.6–1.9	1.4	7.3
empty	379	32	0.6–1.7	2.8	18
PM2					
native-a	5,284	72	0.7–4.4	2.8	12
native-b	13,887	111	0.6–2.9	1.4	8.4
proteinase K	752	8	1.1–2.2	2.8	13
bromelain	665	15	1.1–2.0	2.8	13

* FSC was used in resolution assessment, with a threshold of 0.5. In PRD1, the FSC was calculated between the quasi-atomic models of the capsids and the corresponding reconstructions. In the other cases, the FSC was calculated between two maps calculated from independent halves of the data (see **Section A.7.9.1**).

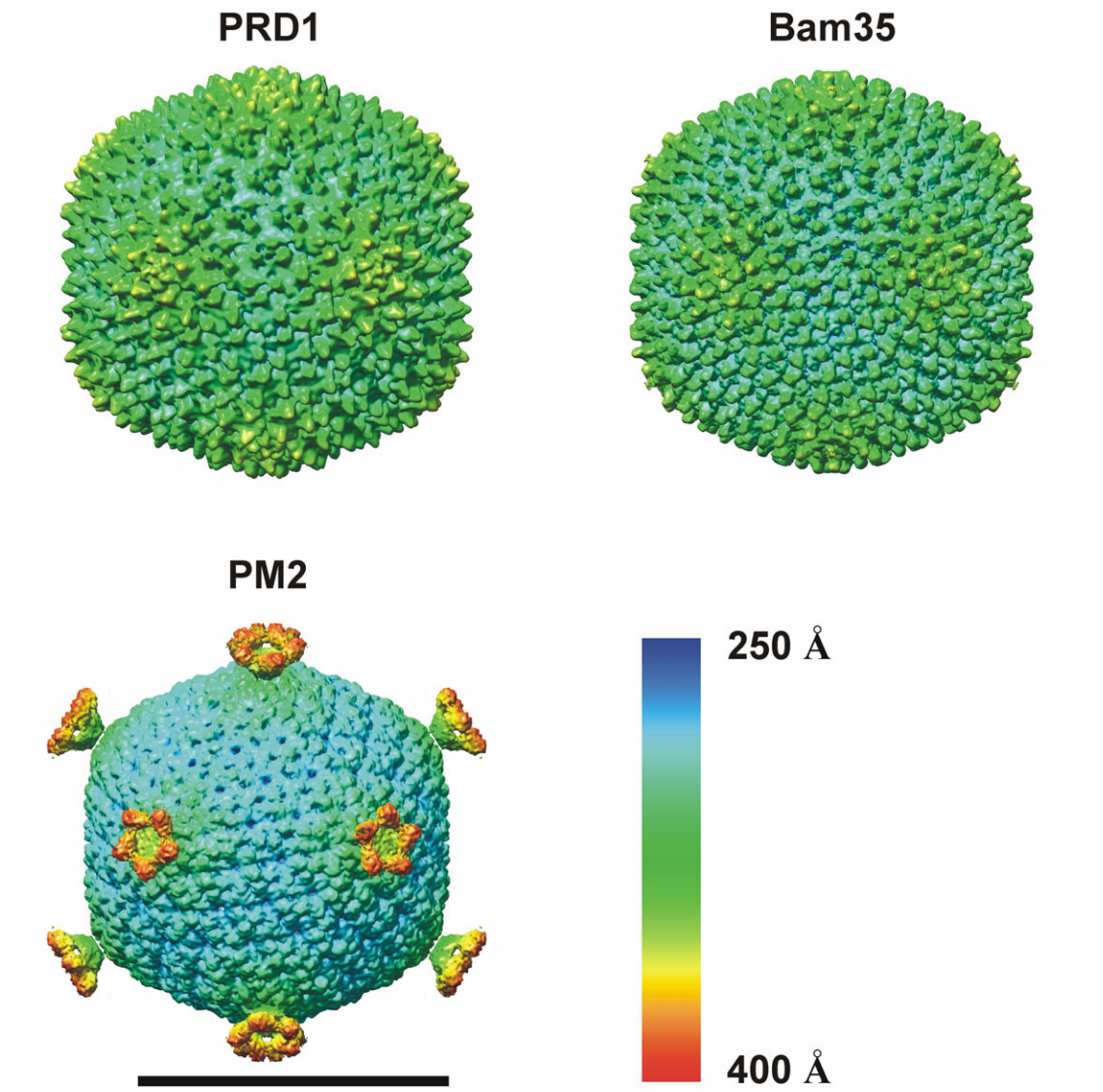

Figure 13. Isosurface renderings of PRD1wt (PRD1), Bam35 native (Bam35) and PM2 native-a (PM2) reconstructions seen along an icosahedral two-fold axis of symmetry. The renderings were radially depth-cued from blue (at radius of 250 Å) to red (at radius of 400 Å) as shown by the colorbar. The isosurfaces were drawn at 1.5σ above the mean density. The scale bar represents 500 Å. The figure was produced using Chimera (Huang *et al*., 1996).

1. PRD1 capsid architecture

1.1. Reconstructions and quasi-atomic models

Cryo-EM and 3D image reconstruction provided two new PRD1 models (**Table 3**). We refined the PRD1 virion structure (diameter ~640 Å between opposite faces) to 13-Å resolution and calculated a new reconstruction to the same resolution for a mutant lacking the membrane aggregation protein P11 (*sus607*)(**Study I**). Since an earlier PRD1 wt reconstruction existed (Butcher *et al*., 1995), it was used as a model in a model-based orientation search (see **Section A.7.8.1**). The

fitting of the atomic model of the major capsid protein P3 into these reconstructions provided quasi-atomic models of the capsids. These reconstructions and models complemented two previously published reconstructions and their quasi-atomic models (Butcher *et al*., 1995; San Martín *et al*., 2001). The first reconstruction was of a mutant devoid of protein P9 (*sus1*) that does not package DNA. The second reconstruction was of an SDS-treated shell, where the DNA, membrane, spike complexes and peripentonal trimers have been removed (Butcher *et al*., 1995; San Martín *et al*., 2001). In this shell, in addition to P3, a minor protein P30 is present, too (Rydman *et al*., 2001).

1.2. EM/X-ray difference imaging revealed minor components in the capsid

The quasi-atomic models contained all of the residues that were ordered in the P3 X-ray crystallographic electron density maps. Subtraction of these P3-shell models from the four reconstructions yielded difference maps (see **Section A.7.9.2**), which defined the non-P3 components in the capsid, and those P3 residues that were disordered in the crystal structure of the isolated protein but ordered in the reconstruction (**Study I**; Figures 1,2). These components included minor proteins and the N- and C-termini of P3. Analysis of these components provided new insights into PRD1 capsid assembly.

The biggest capsid density in the difference maps was the vertex complex (**Study I**, Figure 3). The wt and *sus1* difference maps defined the resolved part of the spike complex, the P31 pentamer. The mass estimate from the difference map agreed with the known mass (see **Section A.7.9.3**). A novel P3–membrane interaction was revealed under the trimers 3 and 4 in the DNA filled particles (wt and *sus607*) but not in the empty particles (*sus1* and SDS-treated shell)(**Study I**; Figures 2b, 4d,e). We saw strong peaks of density at the positively charged central cavity of P3. These peaks may correspond to negatively-charged lipids abundant in the membrane, or some membrane proteins. At the capsid edges, we observed significant non-P3 density (**Study I**; Figures 2b–d, 6). Since this density was not only present in the wt reconstruction but also in the SDS-treated shell, we could assign it to protein P30, the only other component in the SDS-treated shell. We suggested that P30 is a glue protein connecting the facets together to form a closed capsid. Another key finding was that the termini of the major capsid protein act as molecular switches. The state of the switch reflects the quasi-equivalent position of the capsomer in the capsid (**Study I**; Figures 2, 4a–c, 5). The N-termini link the capsomers to the membrane at certain positions, whereas at other positions these connections are not present. The C-termini bridge the capsomer together in the middle of the facet, but at the edge of the facet they possibly interact with the minor protein P30 (**Study I**, Figure 7).

Mutagenesis of the N-terminal residues of the major capsid protein provided insight into the capsid–membrane interaction (**Study I**, Table 1). Three glutamine-residues on one side of the N-terminal helix seen in the X-ray structure were predicted to interact with the membrane (Benson *et al*., 1999). These three glutamines are preceded by another strip of four glutamines, which are disordered in the crystal structure. The change of these four residues into alanines partially decreased the virus titer in a complementation assay. Changes in residues closer to the N-terminus were gradually more severe than those further away (**Study I**, Table 1). These results confirmed that the N-terminal helix plays a key role in PRD1 capsid assembly.

Since protein P11 is involved in the first step of the entry process, it may be

associated with the vertex. In addition, it is a membrane-associated protein and a major component of the virion. Thus, we hypothesized that P11 would occupy the unassigned density (described as 'clasping arms') under the penton protein (**Study I**, Figure 3). However, our P11⁻ reconstruction turned out to be similar to the wt, ruling out this assumption (**Study I**). The distribution of P11 most probably does not follow icosahedral symmetry.

The recent success in crystallizing the whole PRD1 virion and its X-ray analysis has resulted in an atomic model for the viral capsid including four different proteins, P3, P16, P30 and P31 (Bamford *et al*., 2002b; Cockburn *et al*., 2003; Abrescia *et al*., 2004; Cockburn *et al*., 2004). The PRD1 quasi-atomic model was used to phase the X-ray diffraction data (San Martín *et al*., 2001). This structure confirmed most of our results and provided a description of the capsid at 4-Å resolution. Protein P30, as we suggested in **Study I**, locates between two adjacent facets. The X-ray crystallographic electron density map revealed it as an extended polypeptide chain running from one vertex to the two-fold position. It was hypothesized to act as a 'tape-measure protein', determining the size of the facet during assembly, analogously to the bacteriophage tail tape-measure proteins (reviewed in Mosig and Eiserling, 1988). The P3 termini were seen to act as molecular switches. Two different conformations were observed for the N-terminus and four conformations for the C-terminus. The P3–membrane interaction visible in our reconstructions (**Study I**) was not observed in the X-ray electron density map, although the membrane is otherwise resolved (Cockburn *et al*., 2004). The X-ray structure showed that proteins P16 and P31 constitute the 'clasping arm' densities (Abrescia *et al*., 2004; Jaatinen *et al*., 2004) but the location of P11 remained unresolved.

2. The structure of the PRD1 spike

The PRD1 spike complex, consisting of proteins P2 and P5, was unresolved in our cryo-EM based reconstruction (**Study I**). In the more recent X-ray crystallographic analysis of the whole virion, a mutant virus lacking protein P2 was used for crystallization (Bamford *et al*., 2002b; Cockburn *et al*., 2003). However, P5 was also unresolved in that analysis (Abrescia *et al*., 2004). These failures to resolve the spike complex are probably due to the mismatch between the symmetries of P2 (a monomer with no symmetry) and P5 (a trimer with three-fold symmetry) and the icosahedral symmetry of the viral capsid imposed in the structure determination. In addition, P5 has been suggested to contain a flexible hinge (Caldentey *et al*., 2000; Huiskonen *et al*., 2003), which would further impede the structure determination.

We decided to study purified P5 protein using X-ray crystallography to solve its atomic structure (**Study III**). The C-terminal fragment of P5 (residues 136–340) obtained with collagenase digestion was successfully crystallized. The structure of this fragment revealed a globular trimeric knob domain (residues 197–340), the rest of the chain being unresolved (**Study III**, Figure 1). This was due to a flexible hinge region of eight glycines (residues 191–198) just before the C-terminal knob. I had constructed a mutant protein lacking the hinge earlier (Huiskonen *et al*., 2003). The removal of the hinge ordered the rest of the P5 C-terminal fragment in the crystals and we were able to trace the polypeptide chain for an additional shaft domain (residues 141–190; **Study III**, Figure 2).

The knob domain consists of a ten-stranded β-barrel fold with a jelly-roll topology (**Study III**, Figure 3) and it is strikingly similar to the tumour necrosis factor (TNF) protein. TNF-α fragment

superposed well on the knob structure (root mean square distance 2.6 Å between the C_α atoms). Significant similarities were also present at the quaternary structure level of the trimer. Two similar patches of hydrophobic residues and one patch of polar residues form the trimer subunit interface in both proteins. Interestingly, a member of the TNF-superfamily, sTALL-1 assembles into virus-like particles (Liu *et al*., 2002). These findings suggest that the TNF fold is suitable for different protein-protein interactions and it may exemplify a building block of an ancestral viral capsid.

The fold of the resolved part of the P5 shaft (**Study III**, Figure 2) resembles the triple-β-spiral fold observed in the fibres of adenovirus (van Raaij *et al*., 1999) and reovirus (Chappell *et al*., 2002). Only two shaft repeats are present in the P5 structure. A collagen-like sequence motif (residues 124–141) precedes the triple-β-spiral shaft (Caldentey *et al*., 2000), illustrating the fact that PRD1 uses two different structural motifs to make an elongated protein.

The N-terminal domain of P5 (residues 1–123) is homologous to the penton protein P31 (38 % sequence identity). Thus the P31 X-ray structure (Abrescia *et al*., 2004) could be used as a template to model the P5 N-terminal domain. The model of whole P5, together with the earlier structure of P2, could be used to build an atomic model of the entire PRD1 spike complex. We are currently using a method developed recently (Briggs *et al*., 2005); Huiskonen, J.T., Jäälinoja, H.T., Briggs, J.A.G, Fuller, S.D, Butcher, S.J., manuscript) to reconstruct a low resolution model of the spike complex, into which the atomic models of P2 and P5 could be fitted.

3. Difference imaging in the absence of an atomic model

A difference map created by subtracting the atomic model of PRD1 P3 protein from the 3D reconstruction of PRD1 was valuable in the interpretation of the structure (**Study I**). For Bam35 and PM2, however, no atomic model of the major capsid protein existed. Thus, we developed an approach to carry out the difference imaging using the EM density alone (see also **Studies II**, **IV**). The approach exploits the averaging of the non-symmetry related copies of the capsid protein in the lattice. This non-icosahedral averaging has been also used to further increase the signal in addition to icosahedral averaging, for example to study the Semliki Forest virus (SFV) capsid protein (Mancini *et al*., 2000). Our approach is outlined here:

1. The position of the repeating target density (*e.g*. major capsid protein) that is to be subtracted is first identified manually for one copy.

2. A search model is created by manually extracting a density at the defined position. Alternatively, a ‘pseudo-atomic’ model created from such density can also act as a search model. The search model is used in an exhaustive search to define the accurate positions and orientations of all copies of the density.

3. Once the positions and orientations of the target density copies are defined, the densities can be averaged together. Further averaging can be imposed, if the density has a local symmetry axis.

4. The averaging enhances the molecular boundary of the target density, since its surroundings are averaged out. The target density can then be extracted using interactive volume segmentation. Density values below a certain threshold (*e.g*.

one standard deviation above the mean) are set to zero to remove any contribution from features that are not the same in all copies of the density.

5. The averaged density is subtracted at the known positions and orientations to yield a difference map.

Generally in difference imaging approaches, the differences between the two models have to be first minimized (Baker *et al.*, 1999; San Martín *et al.*, 2001). This includes minimizing the differences in size, intensity levels and amplitudes. For difference imaging between a density generated from an atomic model and an EM density (or between two EM densities generated from data collected with different microscopes), the pixel size of the EM-density has to be carefully calibrated, since true magnifications of microscopes can deviate from the nominal magnifications. In addition, the amplitude component of the EM-density has to be scaled to match that of the X-ray model, since amplitudes are inaccurately measured in EM images due to noise.

Our difference imaging approach outlined above is suitable only for certain applications, such as subtracting a major repeating component from the rest of the density. The advantage is that since the density to be subtracted is generated from the same density map, any artefacts arising due to inappropriate scaling are avoided. The success of this approach depends on step 4. If the resolution of the map is too low, one cannot reliably define the molecular boundary. However, in our hands the approach was useful for interpreting reconstructions at sub-nanometre resolution (**Studies II**, **IV**).

4. Bam35 structure revealed its relatedness to PRD1

Our 3D reconstructions of Bam35 provided the first structural model of the virion (**Table 3**; **Study IV**, Figure 1). The PRD1 reconstruction (**Study I**) provided a starting model for a model-based orientation search (see **Section A.7.8.1**). The Bam35 capsid was reconstructed to 7.3-Å resolution, which allowed a detailed analysis of the capsid protein structures, while the reconstruction from empty particles defined the extent of the inner DNA component (**Study IV**). The major capsid proteins were organized on a pseudo *T*=25 lattice, similarly to PRD1 (**Study IV**, Figure 1d). In addition, the size of the Bam35 capsid (diameter ~650 Å between opposite faces) turned out to be similar to that of PRD1 (**Figure 13**). These similarities and the recent atomic model of the PRD1 capsid (Abrescia *et al.*, 2004) made possible a comparison between the PRD1 protein structures and the Bam35 density (**Study IV**).

4.1. The major capsid protein has a double β-barrel fold

The Bam35 major capsid protein was seen as a double β-barrel as predicted earlier (**Study IV**)(Ravantti *et al.*, 2003). Averaging of the four trimer densities further enhanced the capsid protein signal and allowed its extraction from the rest of the capsid density (see **Section 3**). At 7-Å resolution, α-helices and the separation between the two β-sheets in the β-barrels were clearly resolved (**Study IV**, Figure 2a–d). The PRD1 major capsid protein trimer structure was fitted into the averaged Bam35 capsid protein trimer density for comparison (Benson *et al.*, 2002). This confirmed the

earlier sequence alignment between the two proteins (Strömsten *et al*., 2003). The fitting revealed only minor differences, mainly in the loops (**Study IV**, Figure 2a–d). Those loops in the PRD1 structure that were predicted to be missing in Bam35 based on sequence analysis protruded out from the density. The rest of the structure aligned remarkably well given the fact that there is no significant sequence similarity between the two proteins (Ravantti *et al*., 2003).

4.2. EM-based difference imaging revealed minor components in the capsid

In the difference map where the averaged major capsid protein has been subtracted (see **Section 3**), all of the minor components of the capsid are still present (**Study IV**). Furthermore, the parts of the major capsid protein which have different conformations in different copies of the molecule, are present. The difference map helped us to analyse such components and compare them to PRD1.

Differences were observed at the termini of the quasi-equivalent major capsid protein trimers. The N-terminus of the major capsid protein is shorter in Bam35 and does not exhibit the same conformational freedom as in PRD1 (**Study IV**)(Strömsten *et al*., 2003; Abrescia *et al*., 2004). Depending on the position, the N-terminus either reaches down to the membrane or not (**Study IV**; Figure 2e,f). Different possible conformations of the C-terminus were observed (**Study IV**; Figure 2g,h). In some positions, the C-termini were seen to bridge the capsomers, similar to the situation with PRD1. In some other positions, these bridges were not present (**Study IV**, Figure 6).

The most dominant capsid component features in the Bam35 difference map are the penton bases located at the vertices (**Study IV**, Figure 3a–d). Based on the appearance of this density and the good fit of the PRD1 penton protein P31 structure to the Bam35 difference density, an equivalent protein in Bam35 is likely to share the same β-barrel architecture. A thicker knob of density is present on the top of the Bam35 penton than in PRD1. This knob probable corresponds to a resolved part of the spike complex. Threading analysis of all the Bam35 open reading frames against all the solved PRD1 protein structures revealed a PRD1 receptor binding protein homologue in Bam35 (gp29). This is consistent with the finding that in micrographs spikes were visible (**Study IV**, Figure 1). However, a Bam35 protein homologous to the PRD1 spike protein was not detected in this analysis. Thus, the Bam35 spike complex may have a different organization than in PRD1, which would be expected, since these two viruses infect very different hosts (Gram-positive *vs*. Gram-negative bacteria).

Thin densities running from each five-fold symmetric position towards the two-fold positions are visible in the difference map (**Study IV**, Figure 3e). These densities overlay with the atomic model of PRD1 protein P30, a suggested tape-measure protein (Abrescia *et al*., 2004). This observation strongly suggested that Bam35 also has an analogous protein. Thus, the tape-measuring function suggested for PRD1 could be a general mechanism for size-determination in PRD1-type capsids. However, such a protein has not yet been identified for adenovirus, although there are other minor components, some of which are located at the capsid edges (**Section A.2.2**)(Fabry *et al*., 2005). Since some spherical membrane-containing dsDNA viruses, such as PBCV-1 and CIV, have much larger capsids (**Section A.2.3**), this would suggest that they also have significantly longer proteins that could act as tape-measures (Abrescia *et al*., 2004).

5. Structure of PM2 revealed unique features

We calculated a 3D icosahedral reconstruction of the PM2 virion (**Table 3**; **Study II**, Figure 2). Common lines -approaches were used to reconstruct an initial model, which was then used in model based refinement (see **Section A.7.8.1**). The virion structure refined to 8.4-Å resolution, as measured from the major capsid protein shell, which is the most ordered part in the virion. The DNA and the membrane were resolved to a lower resolution (**Study II**, Figure 3). The capsid is ~570 Å in diameter as measured between opposite faces. Thus, it is smaller than the capsids of PRD1 and Bam35. Prominent spikes are present at the icosahedral vertices (**Figure 13**).

In addition to the native virion structure, we reconstructed models for proteinase K and bromelain treated virions (**Table 3**; **Study II**, Figure 2). In these protease treated particles, virus infectivity was abolished (Kivelä *et al*., 2002). Biochemical characterization of these particles revealed that it is the receptor binding protein P1 that is cleaved. In the case of proteinase K treatment, a larger N-terminal fragment of protein P1 stayed associated with the viral capsid than in the case of bromelain treatment which cleaved most of the protein leaving only a small N-terminal fragment.

5.1. PM2 capsid revealed a novel capsid organization

PM2 possesses a capsid lattice not observed in any other virus, described as pseudo *T*=21 (**Study II**, Figure 4). The *T*=21 capsid is handed, with two possible enantiomers, *laevo* and *dextro*. Tilting experiments defined the hand to be *dextro* (Belnap *et al*., 1997). HK97 proheads, which have a handed *T*=7 *laevo* lattice, were used as an internal control in the PM2 cryo-EM samples. The PM2 capsid consists of 200 trimers of protein P2 (**Study II**, Figure 4). The trimers have a pseudo-hexagonal shape similar to that of the PRD1 major capsid protein (Butcher *et al*., 1995; Benson *et al*., 2002)(**Study I**). Four pseudo-hexagonal trimers contribute to each asymmetric unit. One of them locates at the icosahedral three-fold axis. Thus, there are three trimers and one-third of the three-fold trimer in each asymmetric unit, giving ten monomers. Since each of the monomers of the pseudo hexameric trimer contains two structural units, the trimers contribute 20 structural units. The base of the pentameric spike contributes one structural unit to each asymmetric unit around it. The total number of structural is thus 21 as expected for a *T*=21 capsid.

According to the theory of quasi-equivalence, adjustments are needed in the capsomers to compensate for the different environments in the lattice (Caspar and Klug, 1962)(**Section A.1.2**). Similarly to PRD1 (**Study I**) and Bam35 (**Study IV**), we observed variation in the PM2 capsid protein trimers depending on the lattice position (**Study II**; Figure 4e,f). Under the trimer located at the icosahedral three-fold axis, three α-helices were observed. Under the other trimers, however, only one α-helix was observed at the equivalent position. We suggested that the other, unobserved helices would be either disordered, or in some other conformation we could not detect. In the trimer adjacent to the pentameric spike, the unresolved helix could interact with the spike base. In the two other non-threefold trimers, the unresolved helices could be part of the connection down to the LC (**Study II**; Figure 4e,g).

5.2. The inner lipid core is anchored to the capsid

The inner lipid core particles that have lost the outer capsid appear round in the cryo-EM images (**Study II**). We determined the diameter (~480 Å) of the LC from averages of 2D LC images. Attempts to reconstruct a 3D model from isolated LC images did not refine into a consistent model. This was probably due to lack of icosahedral symmetry in the LC. However, the subtraction of the contribution of the P2 protein in the reconstruction yielded an accurate definition for the LC particle (**Study II**, Figure 2). The LC is connected to the capsid at both sides of the icosahedral two-fold axes, via 60 membrane protein anchors (**Study II**, Figure 4g). The protein composition of the anchors remains unknown. However, we identified a previously unknown component of the LC, the putative membrane protein P10. P10 is the third largest protein in the LC and might well correspond to the observed anchor density. Under the membrane resides the circular super-coiled DNA. Although it is circular, in contrast to the linear DNA in PRD1, its packaging density, spacing of the first two layers, and the close proximity of the first layer to the membrane were similar to PRD1 (**Study II**, Figure 3b).

5.3. Three-domain structure of the pentameric receptor binding spike

Pentameric spikes locate at the icosahedral five-fold vertices and protrude above the capsid. Difference imaging of the native reconstruction and reconstructions calculated from protease treated particles revealed that the spike density was diminished in the latter. Thus, we could assign the spike density to protein P1 (**Study II**, Figure 5). The proteolysis and difference imaging revealed three spike domains. Five globular distal tips attach to the central domain. The distal tip domains are removed in the proteinase K digestion, while both the distal and central domains were cleaved in bromelain digestion. The proximal domain is embedded in the major capsid protein shell and is thus protected from proteases.

A combination of MALDI-TOF mass-spectrometry (MS) and liquid chromatography MS/MS sequence analysis defined the exact cleavage sites in the P1 sequence. These correspond to the N-terminal proximal domain (residues 1–60), a central domain (residues 61–158) and the distal C-terminal domain (residues 159–335). Since removal of the distal domain already abolished infectivity, we proposed that the distal domain is responsible for receptor binding. Recently, the distal domain sequence was cloned, the protein was over-expressed and crystallized for structure determination using X-ray crystallography (N. Abrescia and D. Stuart, personal communication).

5.4. Structural view of the PM2 replication cycle

Based on the structure solved in this work and on several earlier results, a hypothetical PM2 replication scheme can be envisaged (**Figure 14**; see also **Section A.5.2**). During the infection of a host cell, the spike protein distal domain is responsible for receptor binding (**Study II**). We propose that binding of P1 to the host triggers uncoating of the virion, similarly to effects of freeze-thawing and the chelation of calcium ions (**Figure 14a**,**b**). This would allow the fusion-active LC to interact with the host outer membrane (Kivelä *et al.*, 2004). The virion-associated lytic enzymes can then degrade the peptidoglycan (Ravantti *et al.*, 2003), eventually leading to delivery of the genome, DNA replication, transcription and protein synthesis (**Figure 14c**,**d**).

How is it then that the PM2 particle is assembled and the circular supercoiled genome packaged? In PM2, there is no evidence so far for an empty procapsid as in PRD1 (Bamford *et al*., 1995). Furthermore, we observed neither a nucleocapsid inside the vesicle nor a specific packaging vertex (**Study II**). However, empty LC-sized vesicles have been observed in cells infected with a temperature sensitive mutant. These mutant viruses mature into infectious virions when shifted from a restrictive to a permissive temperature (Brewer, 1980). This suggests that empty vesicles that bud internally are packaged (**Figure 14e–g**). The membrane vesicle could wrap around the circular supercoiled genome in the host cytoplasm. The budding could also be combined with genome encapsidation. The capsid assembly could be nucleated by interaction with the LC anchor proteins, possibly made of protein P10, since these are the major contacts that take place between the capsid and membrane (**Figure 14h**, **Study II**). Lateral interactions between the major capsid proteins and pentamerisation and assembly of the spike proteins would then finalize the particle (**Figure 14i**). The replication cycle ends in host cell lysis (**Figure 14j**). Mutagenesis of different PM2 proteins and the analysis of the mutant phenotypes are required to test this hypothesis.

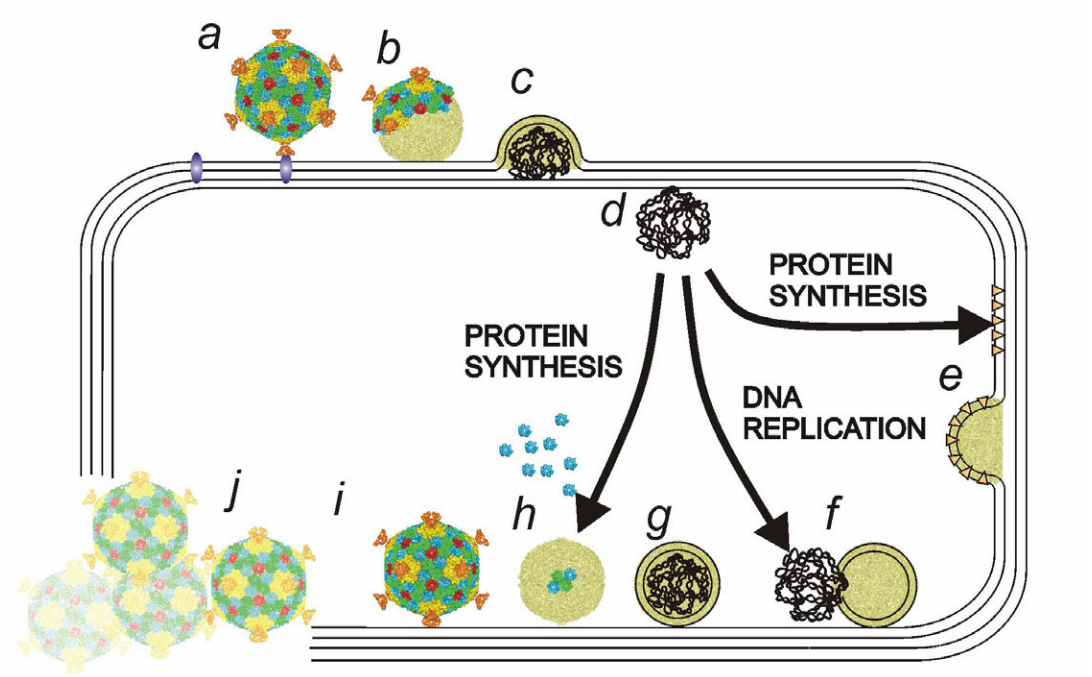

Figure 14. Proposed replication cycle of PM2. (**a**) PM2 virion binds to host cell receptor with the spike protein distal domain. (**b**) Receptor binding triggers uncoating of the virion. (**c**) Lipid core fuses with the host outer membrane leading to genome delivery. (**d**) DNA replication and protein synthesis is associated with the host cytoplasmic membrane. (**e**) Lipid vesicle buds internally with the help of viral membrane proteins. (**f**) The empty vesicle associates with the circular, supercoiled genome. (**g**) The vesicle wraps around the genome. (**h**) Assembly of the capsid nucleates from the P2 trimers interacting with the lipid core anchor proteins. (**i**) Lateral interactions between P2 trimers mediate the assembly of the rest of the capsid. Monomeric P1 proteins associate into pentameric spikes. (**j**) Host cell lysis.

6. Determination of bilayer curvature and thickness

Membrane lipids have a role in controlling protein function via lipid–protein interactions. The mechanisms involved, curvature stress and hydrophobic matching are reviewed by Jensen and Mouritsen (2004). Membrane curvature can, for instance, activate a membrane protein. In addition, increased curvature may render the membrane more susceptible to fusion. Hydrophobic matching refers to the tendency of lipids with certain acyl side chain lengths to cluster around membrane proteins with similar-sized transmembrane regions (or *vice versa*). This has been observed indirectly in simple model systems and molecular dynamics simulations (Mouritsen and Bloom, 1984; Jensen and Mouritsen, 2004). Membrane virus structures provide sufficient information to study the membrane thickness and curvature directly. In this work, we developed a tool for quantitative analysis and visualisation of the membrane thickness and curvature (**Study IV**).

6.1. Membrane proteins modulate the bilayer curvature

The membrane thickness and curvature analysis from **Study IV** was extended here, and five different membrane virus structures, PRD1 (**Study I**), Bam35 (**Study IV**), PM2 (**Study II**), Dengue virus (Zhang *et al*., 2003) and SFV (Mancini *et al*., 2000) were analysed. The membrane curvature analysis is shown in **Figure 15**. The membranes of PRD1, Bam35 and PM2 are all highly curved at the icosahedral five folds ($\theta=90^o$, $\phi= \pm31.7^o$). A transmembrane segment of protein P16 is present in this position as seen in the PRD1 X-ray structure (Abrescia *et al*., 2004). The membranes have less curvature between the two-fold ($\theta=90^o$, $\phi=0^o$) and three-fold ($\theta=69.1^o$, $\phi=0^o$) symmetry axes. The largest difference between PRD1 and Bam35 is the locally higher curvature in Bam35 at the position of transmembrane complex. This complex clearly induces the observed curvature, since PRD1 and Bam35 capsids are otherwise highly similar (**Study IV**). The α-helices present in the bases of the major capsid proteins in all three viruses may contact the membrane causing some of the observed curvature. In the Dengue virus, the E and M transmembrane glycoproteins are clustered in the area with the highest curvature (**Figure 15**, Dengue). The other areas are of low curvature or they are flat. This is inline with the conclusion the authors made from the structure: "…the envelope is markedly polygonal rather than spherical, with constrictions where proteins cross the membrane forming the vertices of the polygon" (Zhang *et al*., 2003). Also in SFV, the transmembrane segments of proteins E1 and E2 correlate with positions, where the curvature changes (**Figure 15**, SFV). They cluster around the two-fold ($\theta=90^o$, $\phi=0^o$) and five-fold positions ($\theta=90^o$, $\phi= \pm31.7^o$), which appear flat. Again, this agrees with the conclusion made in the original publication: "The bilayer of SFV is not round. The membrane appears flattened between the positions of penetration by the transmembrane regions of spikes" (Mancini *et al*., 2000). In conclusion, the tool developed in **Study IV** provided a quantitative and visually descriptive method to analyse membrane curvature.

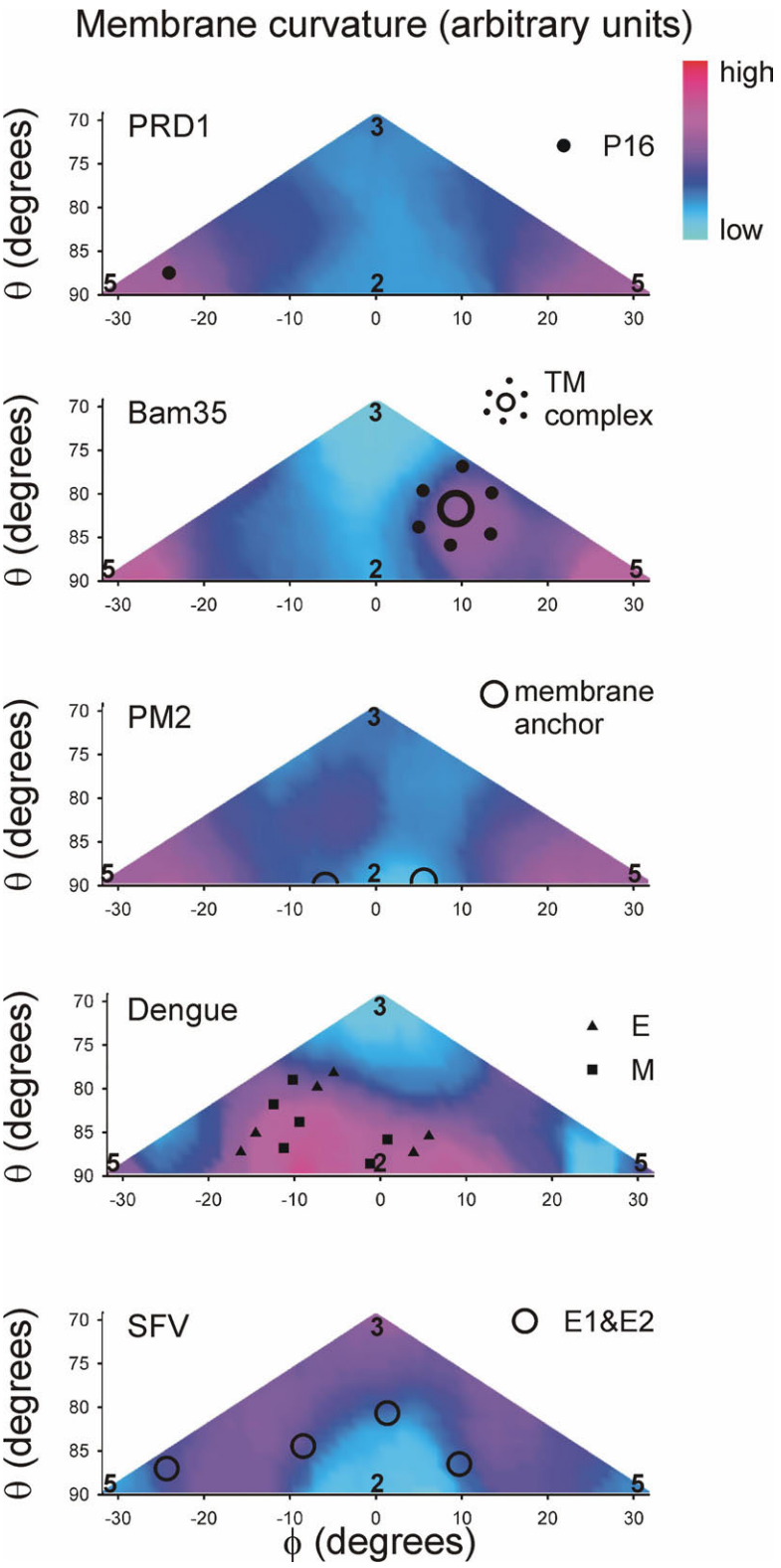

Figure 15. Quantitative membrane curvature analysis of several membrane-containing viruses. The membrane curvature (arbitrary units, see **Study IV**) is plotted as a function of θ and ϕ over the icosahedral asymmetric unit (for the angle definitions, see **Figure 12**, inset). Indicated symbols refer to positions of transmembrane proteins in different viruses. The PM2 membrane anchor crosses the triangular asymmetric unit and two halves are displayed. The positions of the icosahedral two- (2), three- (3) and five-fold (5) axes are indicated.

E. CONCLUSIONS

The relatively high resolution of the 3D reconstructions from cryo-EM images and difference imaging approaches presented in this thesis allowed a detailed analysis of the capsid structures of three different dsDNA membrane-containing bacteriophages, PRD1, Bam35 and PM2 (**Studies I**, **II**, **IV**). The capsids of all three phages are composed of multiple copies of the major capsid protein, which have very similar structures in different quasi-equivalent positions. However, α-helices in their bases, in addition to the C-termini in PRD1 and Bam35, make quasi-equivalent connections to the other capsid proteins and to the membrane. A tape-measure protein, suggested to limit the extent of PRD1 facets (Abrescia *et al*., 2004), was also observed in Bam35. It will be interesting to see if the tape measure protein turns out to be a general assembly mechanism in PRD1-type phages. Despite the fact that the PRD1 capsid is now understood in atomic detail (Abrescia *et al*., 2004), the organization and the structures of the spike proteins still remain unknown. We managed to solve the X-ray structure of one missing piece in the puzzle, the C-terminal part of the spike protein P5 (**Study III**). However, further work will be required to determine the spatial relationships of the spike components. In addition to the capsid architectures, the membrane components were also carefully analysed (**Study IV**, **Figure 15**). The quantitative analysis of the membrane structures showed that protein transmembrane segments modulate the bilayer curvature and thickness.

Although Bam35 infects Gram-positive bacteria, it is clearly very closely related to PRD1 infecting Gram-negative bacteria, since their capsid structures are highly similar (**Study IV**). The two lineages of the host bacteria have diverged more than a billion years ago. However, the last common ancestor of PRD1 and Bam35 might be of much later origin. Conjugative plasmids, which encode for a PRD1-receptor, can transfer from Gram-negative bacteria to Gram-positives (Trieu-Cuot *et al*., 1990). Thus, the phage receptor may have transferred horizontally from a Gram-negative host to a Gram-positive host, or *vice versa*. This might then have allowed the phage to infect a new type of a host, utilizing the same receptor as in the original host. The fact that a PRD1-receptor binding protein homologue was detected in Bam35 (**Study IV**) supports this hypothesis. This hypothesis also predicts that Gram-positive bacteria have cell surface receptors similar to those of PRD1 present in Gram-negative bacteria.

The PM2 structure revealed similarities to PRD1-type viruses, but also unique differences, such as the capsid organization and the receptor binding spikes (**Study II**). Empty capsid shells outside the host cells are observed in infections of PRD1-type viruses, such as PRD1 and PBCV-1 (Meints *et al*., 1984; Bamford *et al*., 1995), but not in PM2 infections. Thus, the infection mechanism of PM2 is likely to be different and may involve the disassembly of the capsid (**Figure 14**). However, the major capsid protein may share the same fold as the PRD1 major capsid protein, since they both are pseudo-hexagonal trimers of the same width and they pack similarly to form a pseudo-type lattice. X-ray crystallographic analysis of PM2 virion is currently underway and the whole virion has been crystallized. Our cryo-EM density map (**Study IV**) was used to solve the phase problem of the X-ray structure determination. The X-ray structure is currently approaching 7-Å resolution (N. Abrescia and D. Stuart, personal communication).

This comparative study of three different dsDNA membrane-containing bacteriophages shed light on multiple different aspects of their capsid and membrane structure, assembly and

evolution. Even though X-ray crystallography can provide atomic models for complete viral capsids, as demonstrated for PRD1, due to its easier applicability cryo-EM combined with 3D image reconstruction continues to be the method of choice in studying large and complex macromolecular assemblies such as viruses.

F. ACKNOWLEDGEMENTS

I carried out this work in the Finnish Academy Centre of Excellence 2000–2005, Programme on Structural Virology, Institute of Biotechnology and Department of Biological and Environmental Sciences, Division of Genetics, at the University of Helsinki.

I deeply appreciate Academy professor Dennis Bamford, the head of the Centre. He took me as a trainee into his research group at a very early age and his support and encouragement have guided me ever since. Although the Bamford laboratory has diversified enormously during the years, it has always offered a pleasant and highly motivating working atmosphere.

My biggest gratitude goes to my supervisor and group leader Docent Sarah Butcher. She introduced me to cryo-EM, so undoubtedly, this thesis would have not been possible without her. Although involved in a zillion of projects (more recently only one that matters), Sarah has always had time for a cup of tea and constructive and supportive discussion, whether work related or not.

I thank Academy professor Mårten Wikström, the director of the Structural Biology and Biophysics research programme, for creating a stimulating working environment, Academy professor Esko Ukkonen (Department of Computer Science) for showing interest in my work, Professor Arto Annila for acting as my thesis advisory board member, and Professor Mark Johnson and Dr. Jonathan Grimes for carefully reviewing the thesis manuscript.

I thank my thesis paper coauthors Professors Stephen Fuller (University of Oxford), Roger Burnett (The Wistar Institute, Philadelphia) and Adrian Goldman, and Drs. Jaana Bamford, Carmen San Martín (The Wistar Institute, Philadelphia), Roman Tuma, Michael Merckel and Hanna Kivelä for good collaboration.

I warmly thank all the past and present members of the Bamford laboratory. Now being separated either in the EM unit basement or in our office in another building, I really miss the lively and groovy atmosphere of the ‘wet lab’ that I experienced when I was still doing my pipetting there. However, the cheerful laughter of Minni Koivunen has reached even these more distant and silent places. Special thanks also go to Jíří Lísal, Harri Jäälinoja, Pasi Laurinmäki, Benita Koli and Jani Seitsonen, for all the good time in work and during free time. I thank Risto Tetri and Teppo Kankaanpää for computer support and Riitta Tarkiainen for skillful technical assistance.

The staff and students of the Electron microscopy unit and the head of the unit Dr. Eija Jokitalo are thanked for their warm attitude towards our cryo-business.

The head of the National Graduate School in Structural and Informational Biology (ISB), Professor Mark Johnson, and the former school coordinator Kaija Söderlund are warmly thanked for creating such a splendid graduate school. The financial support from ISB and Kuopion Luonnon Ystäväin Yhdistys ry (Kuopio Naturalists’ Society) is greatly appreciated.

I am grateful to my relatives and friends for all the time we have spent together. My warmest appreciation goes to my parents, Pertti and Helka Huiskonen and to my sister Johanna Huiskonen.

I express my deepest gratitude to my wife Niina Hautamäki for her love and support.

Juha Huiskonen
Helsinki, December 2005

G. REFERENCES

Abad-Zapatero, C., Abdel-Meguid, S.S., Johnson, J.E., Leslie, A.G.W., Rayment, I., Rossmann, M.G., Suck, D. and Tsukihara, T. (1980) Structure of southern bean mosaic virus at 2.8 Å resolution. *Nature*, **286**, 33-39.

Abrescia, N.G., Cockburn, J.J., 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. and Bamford, J.K. (2004) Insights into assembly from structural analysis of bacteriophage PRD1. *Nature*, **432**, 68-74.

Ackermann, H.-W. (2001) Frequency of morphological phage descriptions in the year 2000. Brief review. *Arch Virol*, **146**, 843-857.

Ackermann, H.-W., Roy, R., Martin, M., Murthy, M.R. and Smirnoff, W.A. (1978) Partial characterization of a cubic Bacillus phage. *Can J Microbiol*, **24**, 986-993.

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

Athappilly, F.K., Murali, R., Rux, J.J., Cai, Z. and Burnett, R.M. (1994) The refined crystal structure of hexon, the major coat protein of adenovirus type 2, at 2.9 Å resolution. *J Mol Biol*, **242**, 430-455.

Baker, T.S. and Cheng, R.H. (1996) A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J Struct Biol*, **116**, 120-130.

Baker, T.S., Drak, J. and Bina, M. (1989) The capsid of small papova viruses contains 72 pentameric capsomeres: direct evidence from cryo-electron-microscopy of simian virus 40. *Biophys J*, **55**, 243-253.

Baker, T.S., Olson, N.H. and Fuller, S.D. (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol Mol Biol Rev*, **63**, 862-922.

Bamford, D. and Mindich, L. (1982) Structure of the lipid-containing bacteriophage PRD1: disruption of wild-type and nonsense mutant phage particles with guanidine hydrochloride. *J Virol*, **44**, 1031-1038.

Bamford, D.H. (2003) Do viruses form lineages across different domains of life? *Res Microbiol*, **154**, 231-236.

Bamford, D.H. and Bamford, J.K.H. (2005) Lipid-containing bacteriophage PM2, the type organism of Corticoviridae. In Calendar, R.L. (ed.), *The Bacteriophages*. Oxford University Press, pp. 170-173.

Bamford, D.H., Burnett, R.M. and Stuart, D.I. (2002a) Evolution of viral structure. *Theor Popul Biol*, **61**, 461-470.

Bamford, D.H., Caldentey, J. and Bamford, J.K. (1995) Bacteriophage PRD1: a broad host range DSDNA tectivirus with an internal membrane. *Adv Virus Res*, **45**, 281-319.

Bamford, D.H., Grimes, J.M. and Stuart, D.I. (2005) What does structure tell us about virus evolution? *Curr Opin Struct Biol*, **15**, 655-663.

Bamford, J.K. and Bamford, D.H. (2000) A new mutant class, made by targeted mutagenesis, of phage PRD1 reveals that protein P5 connects the receptor binding protein to the vertex. *J Virol*, **74**, 7781-7786.

Bamford, J.K., Cockburn, J.J., Diprose, J., Grimes, J.M., Sutton, G., Stuart, D.I. and Bamford, D.H. (2002b) Diffraction quality crystals of PRD1, a 66-MDa dsDNA virus with an internal membrane. *J Struct Biol*, **139**, 103-112.

Belnap, D.M., Kumar, A., Folk, J.T., Smith, T.J. and Baker, T.S. (1999) Low-resolution density maps from atomic models: how stepping "back" can be a step "forward". *J Struct Biol*, **125**, 166-175.

Belnap, D.M., Olson, N.H. and Baker, T.S. (1997) A method for establishing the handedness of biological macromolecules. *J. Struct. Biol.*, **120**, 44-51.

Benson, S.D., Bamford, J.K., Bamford, D.H. and Burnett, R.M. (1999) Viral evolution revealed by bacteriophage PRD1 and human adenovirus coat protein structures. *Cell*, **98**, 825-833.

Benson, S.D., Bamford, J.K., Bamford, D.H. and Burnett, R.M. (2002) The X-ray crystal structure of P3, the major coat protein of the lipid-containing bacteriophage PRD1, at 1.65 Å resolution. *Acta Crystallogr D Biol Crystallogr*, **58**, 39-59.

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

Bergh, O., Borsheim, K.Y., Bratbak, G. and Heldal, M. (1989) High abundance of viruses found in aquatic environments. *Nature*, **340**, 467-468.

Boisset, N., Taveau, J.C., Lamy, J., Wagenknecht, T., Radermacher, M. and Frank, J. (1990) Three-dimensional reconstruction of native *Androctonus australis* hemocyanin. *J Mol Biol*, **216**, 743-760.

Booy, F.P., Newcomb, W.W., Trus, B.L., Brown, J.C., Baker, T.S. and Steven, A.C. (1991) Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell*, **64**, 1007-1015.

Bowman, B.R., Baker, M.L., Rixon, F.J., Chiu, W. and Quiocho, F.A. (2003) Structure of the herpesvirus major capsid protein. *EMBO J*, **22**, 757-765.

Bradley, D.E. and Rutherford, E.L. (1975) Basic characterization of a lipid-containing bacteriophage specific for plasmids of the P, N, and W compatibility groups. *Can J Microbiol*, **21**, 152-163.

Brewer, G.J. (1978) Membrane-localized replication of bacteriophage PM2. *Virology*, **84**, 242-245.

Brewer, G.J. (1980) Control of membrane morphogenesis in bacteriophage. *Int Rev Cytol*, **68**, 53-96.

Briggs, J.A., Huiskonen, J.T., Fernando, K.V., Gilbert, R.J., Scotti, P., Butcher, S.J. and Fuller, S.D. (2005) Classification and three-dimensional reconstruction of unevenly distributed or symmetry mismatched features of icosahedral particles. *J Struct Biol*, **150**, 332-339.

Briggs, J.A., Wilk, T., Welker, R., Krausslich, H.G. and Fuller, S.D. (2003) Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J*, **22**, 1707-1715.

Bubeck, D., Filman, D.J. and Hogle, J.M. (2005) Cryo-electron microscopy reconstruction of a poliovirus-receptor-membrane complex. *Nat Struct Mol Biol*, **12**, 615-618.

Bullough, P.A., Hughson, F.M., Skehel, J.J. and Wiley, D.C. (1994) Structure of influenza haemagglutinin at the pH of membrane fusion. *Nature*, **371**, 37-43.

Butcher, S.J., Bamford, D.H. and Fuller, S.D. (1995) DNA packaging orders the membrane of bacteriophage PRD1. *EMBO J.*, **14**, 6078-6086.

Böttcher, B., Wynne, S.A. and Crowther, R.A. (1997) Determination of the fold of the core protein of hepatitis B virus by electron cryomicroscopy. *Nature*, **386**, 88-91.

Cairns, J., Stent, G.S. and Watson, J.D. (eds.). (1966) *Phage and The Origins of Molecular Biology*. Cold Spring Harbor Laboratory of Quantitative Biology, New York.

Caldentey, J., Bamford, J.K. and Bamford, D.H. (1990) Structure and assembly of bacteriophage PRD1, and *Escherichia coli* virus with a membrane. *J Struct Biol*, **104**, 44-51.

Caldentey, J., Hänninen, A.L. and Bamford, D.H. (1994) Gene XV of bacteriophage PRD1 encodes a lytic enzyme with muramidase activity. *Eur J Biochem*, **225**, 341-346.

Caldentey, J., Tuma, R. and Bamford, D.H. (2000) Assembly of bacteriophage PRD1 spike complex: role of the multidomain protein P5. *Biochemistry*, **39**, 10566-10573.

Canady, M.A., Larson, S.B., Day, J. and McPherson, A. (1996) Crystal structure of turnip yellow mosaic virus. *Nat Struct Biol*, **3**, 771-781.

Casjens, S. (1997) Principles of virion structure, function, and assembly. In Chiu, W., Burnett, R.M. and Garcea, R. (eds.), *Structural biology of viruses*. Oxford Univeristy Press, New York, pp. 3-37.

Casjens, S. (2003) Prophages and bacterial genomics: what have we learned so far? *Mol Microbiol*, **49**, 277-300.

Caspar, D.L. (1980) Movement and self-control in protein assemblies. Quasi-equivalence revisited. *Biophys J*, **32**, 103-138.

Caspar, D.L.D. and Klug, A. (1962) Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol*, **27**, 1-24.

Cerritelli, M.E., Cheng, N., Rosenberg, A.H., McPherson, C.E., Booy, F.P. and Steven, A.C. (1997) Encapsidated conformation of bacteriophage T7 DNA. *Cell*, **91**, 271-280.

Chapman, M.S. and Liljas, L. (2003) Structural folds of viral proteins. In Chiu, W. and Johnson, J.E. (eds.), *Virus Structure*. Academic press, Vol. 64, pp. 125-196.

Chappell, J.D., Prota, A.E., Dermody, T.S. and Stehle, T. (2002) Crystal structure of reovirus attachment protein sigma1 reveals evolutionary relationship to adenovirus fiber. *EMBO J*, **21**, 1-11.

Chen, Z.G., Stauffacher, C., Li, Y., Schmidt, T., Bomu, W., Kamer, G., Shanks, M., Lomonossoff, G. and Johnson, J.E. (1989) Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution. *Science*, **245**, 154-159.

Chescoe, D. and Goodhew, P.J. (1990) *The operation of transmission and scanning electron microscopes*. Oxford University Press, New York.

Cockburn, J.J., Bamford, J.K., Grimes, J.M., Bamford, D.H. and Stuart, D.I. (2003) Crystallization of the membrane-containing bacteriophage PRD1 in quartz capillaries by vapour diffusion. *Acta Crystallogr D Biol Crystallogr*, **59**, 538-540.

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

Conway, J.F., Cheng, N., Zlotnick, A., Wingfield, P.T., Stahl, S.J. and Steven, A.C. (1997) Visualization of a 4-helix bundle in the hepatitis B virus capsid by cryo-electron microscopy. *Nature*, **386**, 91-94.

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

Crowther, R.A. (1971) Procedures for three-dimensional reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Phil Trans Roy Soc Lond B*, **261**, 221-230.

Crowther, R.A., Amos, L.A., Finch, J.T., De Rosier, D.J. and Klug, A. (1970a) Three dimensional reconstructions of spherical viruses by fourier synthesis from electron micrographs. *Nature*, **226**, 421-425.

Crowther, R.A., De Rosier, D.J. and Klug, A. (1970b) The reconstruction of a three-dimensional structure from projections and its application to electron microscopy. *Phil Trans Roy Soc Lond A*, **317**, 319-340.

de Haas, F., Paatero, A.O., Mindich, L., Bamford, D.H. and Fuller, S.D. (1999) A symmetry mismatch at the site of RNA packaging in the polymerase complex of dsRNA bacteriophage ϕ6. *J Mol Biol*, **294**, 357-372.

De Rosier, D.J. and Klug, A. (1968) Reconstruction of three dimensional structures from electron micrographs. *Nature*, **217**, 130-134.

Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowall, A.W. and Schultz, P. (1988) Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys*, **21**, 129-228.

Dubochet, J., Groom, M. and Mueller-Neuteboom, S. (1982) The mounting of macromolecules for electron microscopy with particular reference to surface phenomena and the treatment of support films by glow discharge. In Barer, R. and Cosslett, V.E. (eds.), *Advances in optical and electron microscopy*. Academic Press, London, Vol. 8, pp. 107-135.

Earnshaw, W. and King, J. (1978) Structure of phage P22 coat protein aggregates formed in the absence of the scaffolding protein. *J Mol Biol*, **126**, 721-747.

Erickson, H.P. and Klug, A. (1971) Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. *Phil Trans Roy Soc Lond B*, **261**, 105-118.

Erickson, H.P. and Pantaloni, D. (1981) The role of subunit entropy in cooperative assembly. Nucleation of microtubules and other two-dimensional polymers. *Biophys J*, **34**, 293-309.

Espejo, R.T. and Canelo, E.S. (1968) Properties and characterization of the host bacterium of bacteriophage PM2. *J Bacteriol*, **95**, 1887-1891.

Espejo, R.T., Canelo, E.S. and Sinsheimer, R.L. (1971) Replication of bacteriophage PM2 deoxyribonucleic acid: a closed circular double-stranded molecule. *J. Mol. Biol.*, **56**, 597-621.

Fabry, C.M., Rosa-Calatrava, M., Conway, J.F., Zubieta, C., Cusack, S., Ruigrok, R.W. and Schoehn, G. (2005) A quasi-atomic model of human adenovirus type 5 capsid. *EMBO J*, **24**, 1645-1654.

Fane, B.A. and Prevelige, P.E. (2003) Mechanism of scaffolding-assisted viral assembly. In Chiu, W. and Johnson, J.E. (eds.), *Virus Structure*. Academic press, Vol. 64, pp. 259-299.

Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U. and Bell, L.A. (eds.). (2004) *Virus Taxonomy, The Eight Report of the ICTV*. Academic Press, London.

Fraenkel-Conrat, H. and Williams, R.C. (1955) Reconstitution of active tobacco mosaic virus from its inactive protein and nucleic acid components. *PNAS*, **41**, 690-698.

Frank, J. (1996) *Three-dimensional electron microscopy of macromoleculer assemblies*. Academic Press, San Diego, California.

Franklin, R.M., Hinnen, R., Schäfer, R. and Tsukagoshi, N. (1976) Structure and assembly of lipid-containing viruses, with special reference to bacteriophage PM2 as one type of model system. *Phil Trans Roy Soc Lond B*, **276**, 63-80.

Fuller, S.D., Butcher, S.J., Cheng, R.H. and Baker, T.S. (1996) Three-dimensional reconstruction of icosahedral particles--the uncommon line. *J Struct Biol*, **116**, 48-55.

Furcinitti, P.S., van Oostrum, J. and Burnett, R.M. (1989) Adenovirus polypeptide IX revealed as capsid cement by difference images from electron microscopy and crystallography. *EMBO J*, **8**, 3563-3570.

Gaidelyte, A., Jaatinen, S.T., Daugelavicius, R., Bamford, J.K. and Bamford, D.H. (2005) The linear double-stranded DNA of phage Bam35 enters lysogenic host cells, but the late phage functions are suppressed. *J Bacteriol*, **187**, 3521-3527.

Glaeser, R.M. (1999) Review: electron crystallography: present excitement, a nod to the past, anticipating the future. *J Struct Biol*, **128**, 3-14.

Gowen, B., Bamford, J.K., Bamford, D.H. and Fuller, S.D. (2003) The tailless icosahedral membrane virus PRD1 localizes the proteins involved in genome packaging and injection at a unique vertex. *J Virol*, **77**, 7863-7871.

Grahn, A.M., Butcher, S.J., Bamford, J.K.H. and Bamford, D.H. (2005) PRD1 - dissecting the genome, structure and entry. In Calendar, R.L. (ed.), *The Bacteriophages*. Oxford University Press, pp. 160-169.

Grahn, A.M., Daugelavicius, R. and Bamford, D.H. (2002a) Sequential model of phage PRD1 DNA delivery: active involvement of the viral membrane. *Mol Microbiol*, **46**, 1199-1209.

Grahn, A.M., Daugelavicius, R. and Bamford, D.H. (2002b) The small viral membrane-associated protein P32 is involved in bacteriophage PRD1 DNA entry. *J Virol*, **76**, 4866-4872.

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

Grimes, J.M., Fuller, S.D. and Stuart, D.I. (1999) Complementing crystallography: the role of cryo-electron microscopy in structural biology. *Acta Crystallogr D Biol Crystallogr*, **55 (Pt 10)**, 1742-1749.

Grünewald, K., Desai, P., Winkler, D.C., Heymann, J.B., Belnap, D.M., Baumeister, W. and Steven, A.C. (2003) Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science*, **302**, 1396-1398.

Harauz, G. and van Heel, M. (1986) Exact filters for general geometry three dimensional reconstruction. *Optik*, **73**, 146-156.

Harpaz, Y., Gerstein, M. and Chothia, C. (1994) Volume changes on protein folding. *Structure*, **2**, 641-649.

Harrison, S.C. (1990) Principles of virus structure. In Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., Monath, T.P. and Roizman, B. (eds.), *Virology*. Raven Press, New York, pp. 37-61.

Harrison, S.C., Olson, A.J., Schutt, C.E., Winkler, F.K. and Bricogne, G. (1978)

Tomato bushy stunt virus at 2.9 Å resolution. *Nature*, **276**, 368-373.

Haywood, A.M. (1994) Virus receptors: binding, adhesion strengthening, and changes in viral structure. *J Virol*, **68**, 1-5.

Henderson, R. (2004) Realizing the potential of electron cryo-microscopy. *Q Rev Biophysics*, **37**, 3-13.

Henderson, R., Baldwin, J.M., Ceska, T.A., Zemlin, F., Beckmann, E. and Downing, K.H. (1990) Model for the structure of bacteriorhodopsin based on high-resolution electron cryo-microscopy. *J Mol Biol*, **213**, 899-929.

Hendrix, R.W. (2002) Bacteriophages: evolution of the majority. *Theor Popul Biol*, **61**, 471-480.

Hendrix, R.W. (2003) Bacteriophage genomics. *Curr Opin Microbiol*, **6**, 506-511.

Hershey, A.D. and Chase, M. (1952) Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J Gen Physiol*, **36**, 39-56.

Heymann, J.B., Cheng, N., Newcomb, W.W., Trus, B.L., Brown, J.C. and Steven, A.C. (2003) Dynamics of herpes simplex virus capsid maturation visualized by time-lapse cryo-electron microscopy. *Nat Struct Biol*, **10**, 334-341.

Hill, C.L., Booth, T.F., Prasad, B.V., Grimes, J.M., Mertens, P.P., Sutton, G.C. and Stuart, D.I. (1999) The structure of a cypovirus and the functional organization of dsRNA viruses. *Nat Struct Biol*, **6**, 565-568.

Hoppe, W., Gassmann, J., Hunsmann, N., Schramm, H.J. and Sturm, M. (1974) Three-dimensional reconstruction of individual negatively stained yeast fatty-acid synthetase molecules from tilt series in the electron microscope. *Hoppe Seylers Z Physiol Chem*, **355**, 1483-1487.

Huang, C.C., Couch, G.S., Pettersen, E.F. and Ferrin, T.E. (1996) Chimera: an extensible molecular modeling application constructed using standard components. *Pacific Symposium on Biocomputing*, **1**.

Huiskonen, J.T., Laakkonen, L., Toropainen, M., Sarvas, M., Bamford, D.H. and Bamford, J.K. (2003) Probing the ability of the coat and vertex protein of the membrane-containing bacteriophage PRD1 to display a meningococcal epitope. *Virology*, **310**, 267-279.

Ivanova, N., Sorokin, A., Anderson, I., Galleron, N., Candelon, B., Kapatral, V., Bhattacharyya, A., Reznik, G., Mikhailova, N., Lapidus, A., Chu, L., Mazur, M., Goltsman, E., Larsen, N., D'Souza, M., Walunas, T., Grechkin, Y., Pusch, G., Haselkorn, R., Fonstein, M., Ehrlich, S.D., Overbeek, R. and Kyrpides, N. (2003) Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. *Nature*, **423**, 87-91.

Jaatinen, S.T., Viitanen, S.J., Bamford, D.H. and Bamford, J.K.H. (2004) Integral membrane protein P16 of bacteriophage PRD1 stabilizes the adsorption vertex structure. *J Virol*, **78**, 9790-9797.

Jensen, M.O. and Mouritsen, O.G. (2004) Lipids do influence protein function-the hydrophobic matching hypothesis revisited. *Biochim Biophys Acta*, **1666**, 205-226.

Ji, Y., Marinescu, D.C., Zhang, W. and Baker, T.S. (2003) Orientation refinement of virus structures with unknown symmetry. *Proc. 17th Ann. Int'l Parallel & Distrib. Process. Symp*. IEEE Press, Nice, France.

King, J. and Casjens, S. (1974) Catalytic head assembling protein in virus morphogenesis. *Nature*, **251**, 112-119.

Kivelä, H.M., Daugelavicius, R., Hankkio, R.H., Bamford, J.K. and Bamford, D.H. (2004) Penetration of membrane-containing double-stranded-DNA bacteriophage PM2 into *Pseudoalteromonas* hosts. *J Bacteriol*, **186**, 5342-5354.

Kivelä, H.M., Kalkkinen, N. and Bamford, D.H. (2002) Bacteriophage PM2 has a protein capsid surrounding a spherical proteinaceous lipid core. *J Virol*, **76**, 8169-8178.

Kivelä, H.M., Männistö, R.H., Kalkkinen, N. and Bamford, D.H. (1999) Purification and protein composition of PM2, the first lipid-containing bacterial virus to be isolated. *Virology*, **262**, 364-374.

Kivioja, T., Ravantti, J., Verkhovsky, A., Ukkonen, E. and Bamford, D. (2000) Local average intensity-based method for identifying spherical particles in

electron micrographs. *J Struct Biol*, **131**, 126-134.

Klug, A. and Finch, J.T. (1968) Structure of viruses of the papilloma-polyoma type. IV. Analysis of tilting experiments in the electron microscope. *J Mol Biol*, **31**, 1-12.

Kootstra, N.A. and Verma, I.M. (2003) Gene therapy with viral vectors. *Annu Rev Pharmacol Toxicol*, **43**, 413-439.

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. and Strauss, J.H. (2002) Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. *Cell*, **108**, 717-725.

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

Laurinavicius, S., Käkelä, R., Somerharju, P. and Bamford, D.H. (2004) Phospholipid molecular species profiles of tectiviruses infecting Gram-negative and Gram-positive hosts. *Virology*, **322**, 328-336.

Lee, K.K. and Johnson, J.E. (2003) Complementary approaches to structure determination of icosahedral viruses. *Curr Opin Struct Biol*, **13**, 558-569.

Lescar, J., Roussel, A., Wien, M.W., Navaza, J., Fuller, S.D., Wengler, G. and 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.

Liddington, R.C., Yan, Y., Moulai, J., Sahli, R., Benjamin, T.L. and Harrison, S.C. (1991) Structure of simian virus 40 at 3.8-Å resolution. *Nature*, **354**, 278-284.

Liljas, L., Unge, T., Jones, T.A., Fridborg, K., Lovgren, S., Skoglund, U. and Strandberg, B. (1982) Structure of satellite tobacco necrosis virus at 3.0 Å resolution. *J Mol Biol*, **159**, 93-108.

Liu, Y., Xu, L., Opalka, N., Kappler, J., Shu, H.B. and Zhang, G. (2002) Crystal structure of sTALL-1 reveals a virus-like assembly of TNF family ligands. *Cell*, **108**, 383-394.

Lundström, K.H., Bamford, D.H., Palva, E.T. and Lounatmaa, K. (1979) Lipid-containing bacteriophage PR4: structure and life cycle. *J Gen Virol*, **43**, 583-592.

Mancini, E.J., Clarke, M., Gowen, B.E., Rutten, T. and Fuller, S.D. (2000) Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol Cell*, **5**, 255-266.

Marinescu, D.C., Ji, Y. and Lynch, R.E. (2001) Space-time tradeoffs for parallel 3D reconstruction algorithms for atomic virus structure determination. *Concurrency Computat: Pract Exper*, **13**, 1083-1106.

Matricardi, V.R., Moretz, R.C. and Parsons, D.F. (1972) Electron diffraction of wet proteins: catalase. *Science*, **177**, 268-270.

McKenna, R., Xia, D., Willingmann, P., Ilag, L.L. and Rossmann, M.G. (1992) Structure determination of the bacteriophage ϕX174. *Acta Crystallogr B*, **48 (Pt 4)**, 499-511.

Meints, R.H., Lee, K., Burbank, D.E. and Van Etten, J.L. (1984) Infection of a Chlorella-like alga with the virus, PBCV-1: ultrastructural studies. *Virology*, **138**, 341-346.

Mindell, J.A. and Grigorieff, N. (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol*, **142**, 334-347.

Mindich, L., Bamford, D., Goldthwaite, C., Laverty, M. and Mackenzie, G. (1982a) Isolation of nonsense mutants of lipid-containing bacteriophage PRD1. *J Virol*, **44**, 1013-1020.

Mindich, L., Bamford, D., McGraw, T. and Mackenzie, G. (1982b) Assembly of bacteriophage PRD1: particle formation with wild-type and mutant viruses. *J Virol*, **44**, 1021-1030.

Morais, M.C., Tao, Y., Olson, N.H., Grimes, S., Jardine, P.J., Anderson, D.L., Baker, T.S. and Rossmann, M.G. (2001) Cryoelectron-microscopy image reconstruction of symmetry mismatches in bacteriophage ϕ29. *J Struct Biol*, **135**, 38-46.

Mosig, G. and Eiserling, F. (1988) Phage T4 structure and metabolism. In Calendar, R. (ed.), *The Bacteriophages*. Plenum Press, New York, Vol. 2, pp. 521-606.

Mouritsen, O.G. and Bloom, M. (1984) Mattress model of lipid-protein interactions in membranes. *Biophys J*, **46**, 141-153.

Murata, K., Mitsuoka, K., Hirai, T., Walz, T., Agre, P., Heymann, J.B., Engel, A. and Fujiyoshi, Y. (2000) Structural determinants of water permeation through aquaporin-1. *Nature*, **407**, 599-605.

Murphy, F.A. and Kingsbury, D.W. (1990) Virus Taxonomy. In Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L., Monath, T.P. and Roizman, B. (eds.), *Virology*. Raven Press, New York, pp. 9-35.

Männistö, R.H., Grahn, A.M., Bamford, D.H. and Bamford, J.K.H. (2003) Transcription of bacteriophage PM2 involves phage-encoded regulators of heterologous origin. *J Bact*, **185**, 3278-3287.

Männistö, R.H., Kivelä, H.M., Paulin, L., Bamford, D.H. and Bamford, J.K.H. (1999) The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. *Virology*, **262**, 355-363.

Nagy, E. (1974) A highly specific phage attacking *Bacillus anthracis* strain Sterne. *Acta Microbiol Acad Sci Hung*, **21**, 257-263.

Naitow, H., Tang, J., Canady, M., Wickner, R.B. and Johnson, J.E. (2002) L-A virus at 3.4 Å resolution reveals particle architecture and mRNA decapping mechanism. *Nat Struct Biol*, **9**, 725-728.

Nakagawa, A., Miyazaki, N., Taka, J., Naitow, H., Ogawa, A., Fujimoto, Z., Mizuno, H., Higashi, T., Watanabe, Y., Omura, T., Cheng, R.H. and Tsukihara, T. (2003) The atomic structure of rice dwarf virus reveals the self-assembly mechanism of component proteins. *Structure*, **11**, 1227-1238.

Namba, K. and Stubbs, G. (1986) Structure of tobacco mosaic virus at 3.6 Å resolution: implications for assembly. *Science*, **231**, 1401-1406.

Nandhagopal, N., Simpson, A.A., Gurnon, J.R., Yan, X., Baker, T.S., Graves, M.V., Van Etten, J.L. and Rossmann, M.G. (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proc Natl Acad Sci U S A*, **99**, 14758-14763.

Newcomb, W.W. and Brown, J.C. (1991) Structure of the herpes simplex virus capsid: effects of extraction with guanidine hydrochloride and partial reconstitution of extracted capsids. *J Virol*, **65**, 613-620.

Newcomb, W.W., Juhas, R.M., Thomsen, D.R., Homa, F.L., Burch, A.D., Weller, S.K. and Brown, J.C. (2001) The UL6 gene product forms the portal for entry of DNA into the herpes simplex virus capsid. *J Virol*, **75**, 10923-10932.

Newcomb, W.W., Trus, B.L., Booy, F.P., Steven, A.C., Wall, J.S. and Brown, J.C. (1993) Structure of the herpes simplex virus capsid. Molecular composition of the pentons and the triplexes. *J Mol Biol*, **232**, 499-511.

Nogales, E., Wolf, S.G. and Downing, K.H. (1998) Structure of the alpha beta tubulin dimer by electron crystallography. *Nature*, **391**, 199-203.

O'Neil, K.T. and Hoess, R.H. (1995) Phage display: protein engineering by directed evolution. *Curr Opin Struct Biol*, **5**, 443-449.

Orlova, E.V., Dube, P., Harris, J.R., Beckman, E., Zemlin, F., Markl, J. and van Heel, M. (1997) Structure of keyhole limpet hemocyanin type 1 (KLH1) at 15 Å resolution by electron cryomicroscopy and angular reconstitution. *J Mol Biol*, **271**, 417-437.

Orlova, E.V. and Saibil, H.R. (2004) Structure determination of macromolecular assemblies by single-particle analysis of cryo-electron micrographs. *Curr Opin Struct Biol*, **14**, 584-590.

Pelkmans, L. (2005) Viruses as probes for systems analysis of cellular signalling, cytoskeleton reorganization and endocytosis. *Curr Opin Microbiol*, **8**, 331-337.

Poranen, M.M., Paatero, A.O., Tuma, R. and Bamford, D.H. (2001) Self-assembly of a viral molecular machine from purified protein and RNA constituents. *Mol Cell*, **7**, 845-854.

Poranen, M.M. and Tuma, R. (2004) Self-assembly of double-stranded RNA bacteriophages. *Virus Res*, **101**, 93-100.

Prasad, B.V., Hardy, M.E., Dokland, T., Bella, J., Rossmann, M.G. and Estes, M.K. (1999) X-ray crystallographic structure

of the Norwalk virus capsid. *Science*, **286**, 287-290.

Radermacher, M., Wagenknecht, T., Verschoor, A. and Frank, J. (1987) Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc*, **146 (Pt 2)**, 113-136.

Ravantti, J.J., Gaidelyte, A., Bamford, D.H. and Bamford, J.K.H. (2003) Comparative analysis of bacterial viruses Bam35, infecting a gram-positive host, and PRD1, infecting gram-negative hosts, demonstrates a viral lineage. *Virology*, **313**, 401-414.

Rayment, I., Baker, T.S., Caspar, D.L. and Murakami, W.T. (1982) Polyoma virus capsid structure at 22.5 Å resolution. *Nature*, **295**, 110-115.

Reddy, V.S., Natarajan, P., Okerberg, B., Li, K., Damodaran, K.V., Morton, R.T., Brooks, C.L., 3rd and Johnson, J.E. (2001) Virus Particle Explorer (VIPER), a website for virus capsid structures and their computational analyses. *J Virol*, **75**, 11943-11947.

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

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T. and Young, M. (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *Proc Natl Acad Sci U S A*, **101**, 7716-7720.

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

Roos, N. and Morgan, J.M. (1990) *Cryopreparation of thin biological specimens for electron microscopy: methods and applications*. Oxford University Press.

Rosenthal, P.B. and Henderson, R. (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol*, **333**, 721-745.

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. and Vriend, G. (1985) Structure of human cold virus and functional relationship to other picornaviruses. *Nature*, **317**, 145-153.

Rossmann, M.G., Mesyanzhinov, V.V., Arisaka, F. and Leiman, P.G. (2004) The bacteriophage T4 DNA injection machine. *Curr Opin Struct Biol*, **14**, 171-180.

Ruigrok, R.W., Andree, P.J., Hooft van Huysduynen, R.A. and Mellema, J.E. (1984) Characterization of three highly purified influenza virus strains by electron microscopy. *J Gen Virol*, **65 (Pt 4)**, 799-802.

Rux, J.J. and Burnett, R.M. (2004) Adenovirus structure. *Hum Gene Ther*, **15**, 1167-1176.

Rydman, P.S. and Bamford, D.H. (2002) The lytic enzyme of bacteriophage PRD1 is associated with the viral membrane. *J Bacteriol*, **184**, 104-110.

Rydman, P.S., Bamford, J.K.H. and Bamford, D.H. (2001) A minor capsid protein P30 is essential for bacteriophage PRD1 capsid assembly. *J Mol Biol*, **313**, 785-795.

Rydman, P.S., Caldentey, J., Butcher, S.J., Fuller, S.D., Rutten, T. and Bamford, D.H. (1999) Bacteriophage PRD1 contains a labile receptor-binding structure at each vertex. *J Mol Biol*, **291**, 575-587.

Saad, A., Ludtke, S.J., Jakana, J., Rixon, F.J., Tsuruta, H. and Chiu, W. (2001) Fourier amplitude decay of electron cryomicroscopic images of single particles and effects on structure determination. *J Struct Biol*, **133**, 32-42.

Saban, S.D., Nepomuceno, R.R., Gritton, L.D., Nemerow, G.R. and Stewart, P.L. (2005) CryoEM structure at 9 Å resolution of an adenovirus vector targeted to hematopoietic cells. *J Mol Biol*, **349**, 526-537.

Sakaki, Y., Oshima, M., Yamada, K. and Oshima, T. (1977) Bacteriophage ϕNS11: a lipid-containing phage of acidophilic thermophilic bacteria. III. Characterization of viral components. *J Biochem (Tokyo)*, **82**, 1457-1461.

Salas, M. (1991) Protein-priming of DNA replication. *Annu. Review Biochem.*, **60**, 39-71.

San Martín, C., Burnett, R.M., de Haas, F., Heinkel, R., Rutten, T., Fuller, S.D., Butcher, S.J. and Bamford, D.H. (2001) Combined EM/X-ray imaging yields a quasi-atomic model of the adenovirus-related bacteriophage PRD1 and shows key capsid and membrane interactions. *Structure*, **9**, 917-930.

Saren, A.M., Ravantti, J.J., Benson, S.D., Burnett, R.M., Paulin, L., Bamford, D.H. and Bamford, J.K. (2005) A snapshot of viral evolution from genome analysis of the tectiviridae family. *J Mol Biol*, **350**, 427-440.

Savilahti, H., Caldentey, J., Lundström, K., Syväoja, J.E. and Bamford, D.H. (1991) Overexpression, purification, and characterization of *Escherichia coli* bacteriophage PRD1 DNA polymerase. In vitro synthesis of full-length PRD1 DNA with purified proteins. *J Biol Chem*, **266**, 18737-18744.

Schrag, J.D., Prasad, B.V., Rixon, F.J. and Chiu, W. (1989) Three-dimensional structure of the HSV1 nucleocapsid. *Cell*, **56**, 651-660.

Schäfer, R., Hinnen, R. and Franklin, R.M. (1974) Structure and synthesis of a lipid-containing bacteriophage. Properties of the structural proteins and distribution of the phospholipid. *Eur J Biochem*, **50**, 15-27.

Siegel, B.M. (1971) Current and future prospects in electron microscopy for observations in biomolecular structure. *Phil Trans Roy Soc Lond B*, **261**, 5-14.

Smith, T.J., Olson, N.H., Cheng, R.H., Liu, H., Chase, E.S., Lee, W.M., Leippe, D.M., Mosser, A.G., Rueckert, R.R. and Baker, T.S. (1993) Structure of human rhinovirus complexed with Fab fragments from a neutralizing antibody. *J Virol*, **67**, 1148-1158.

Sokolova, A., Malfois, M., Caldentey, J., Svergun, D.I., Koch, M.H., Bamford, D.H. and Tuma, R. (2001) Solution structure of bacteriophage PRD1 vertex complex. *J Biol Chem*, **276**, 46187-46195.

Speir, J.A., Munshi, S., Wang, G., Baker, T.S. and Johnson, J.E. (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*, **3**, 63-78.

Stanisich, V.A. (1974) The properties and host range of male-specific bacteriophages of *Pseudomonas aeruginosa*. *J Gen Microbiol*, **84**, 332-342.

Stewart, P.L., Burnett, R.M., Cyrklaff, M. and Fuller, S.D. (1991) Image reconstruction reveals the complex molecular organization of adenovirus. *Cell*, **67**, 145-154.

Stewart, P.L., Fuller, S.D. and Burnett, R.M. (1993) Difference imaging of adenovirus: Bridging the resolution gap between X-ray crystallography and electron microscopy. *EMBO J.*, **12**, 2589-2599.

Steven, A.C., Heymann, J.B., Cheng, N., Trus, B.L. and Conway, J.F. (2005) Virus maturation: dynamics and mechanism of a stabilizing structural transition that leads to infectivity. *Curr Opin Struct Biol*, **15**, 227-236.

Strömsten, N.J., Bamford, D.H. and Bamford, J.K. (2005) In vitro DNA packaging of PRD1: a common mechanism for internal-membrane viruses. *J Mol Biol*, **348**, 617-629.

Strömsten, N.J., Benson, S.D., Burnett, R.M., Bamford, D.H. and Bamford, J.K. (2003) The *Bacillus thuringiensis* linear double-stranded DNA phage Bam35, which is highly similar to the *Bacillus cereus* linear plasmid pBClin15, has a prophage state. *J Bacteriol*, **185**, 6985-6989.

Summers, W.C. (2001) Bacteriophage therapy. *Annu Rev Microbiol*, **55**, 437-451.

Thon, F. (1971) Phase contrast electron microscopy. In Valdré, U. (ed.), *Electron microscopy in material science*. Academic Press, New York, pp. 570-625.

Thuman-Commike, P.A. and Chiu, W. (1997) Improved common line-based icosahedral particle image orientation estimation algorithms. *Ultramicroscopy*, **68**, 231-255.

Toyoshima, C. and Unwin, N. (1988) Contrast transfer for frozen-hydrated specimens: determination from pairs of defocused images. *Ultramicroscopy*, **25**, 279-291.

Trieu-Cuot, P., Carlier, C., Poyart-Salmeron, C. and Courvalin, P. (1990) A pair of

mobilizable shuttle vectors conferring resistance to spectinomycin for molecular cloning in *Escherichia coli* and in gram-positive bacteria. *Nucleic Acids Res*, **18**, 4296.

Trus, B.L., Booy, F.P., Newcomb, W.W., Brown, J.C., Homa, F.L., Thomsen, D.R. and Steven, A.C. (1996) The herpes simplex virus procapsid: structure, conformational changes upon maturation, and roles of the triplex proteins VP19c and VP23 in assembly. *J Mol Biol*, **263**, 447-462.

Trus, B.L., Cheng, N., Newcomb, W.W., Homa, F.L., Brown, J.C. and Steven, A.C. (2004) Structure and polymorphism of the UL6 portal protein of herpes simplex virus type 1. *J Virol*, **78**, 12668-12671.

Twarock, R. (2004) A tiling approach to virus capsid assembly explaining a structural puzzle in virology. *J Theor Biol*, **226**, 477-482.

Unwin, N. (2005) Refined structure of the nicotinic acetylcholine receptor at 4A resolution. *J Mol Biol*, **346**, 967-989.

Wade, R.H. (1992) A brief look at imaging and contrast transfer. *Ultramicroscopy*, **46**, 145-156.

Valpuesta, J.M. and Carrascosa, J.L. (1994) Structure of viral connectors and their function in bacteriophage assembly and DNA packaging. *Q Rev Biophys*, **27**, 107-155.

van Heel, M. (1987) Angular reconstitution: *a posteriori* assignment of projection directions for 3D reconstruction. *Ultramicroscopy*, **21**, 111-123.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. and Patwardhan, A. (2000) Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys*, **33**, 307-369.

van Raaij, M.J., Mitraki, A., Lavigne, G. and Cusack, S. (1999) A triple beta-spiral in the adenovirus fibre shaft reveals a new structural motif for a fibrous protein. *Nature*, **401**, 935-938.

Wang, Q., Kaltgrad, E., Lin, T., Johnson, J.E. and Finn, M.G. (2002) Natural supramolecular building blocks. Wild-type cowpea mosaic virus. *Chem Biol*, **9**, 805-811.

Verheust, C., Fornelos, N. and Mahillon, J. (2004) The *Bacillus thuringiensis* phage GIL01 encodes two enzymes with peptidoglycan hydrolase activity. *FEMS Microbiol Lett*, **237**, 289-295.

Verheust, C., Fornelos, N. and Mahillon, J. (2005) GIL16, a new gram-positive tectiviral phage related to the *Bacillus thuringiensis* GIL01 and the *Bacillus cereus* pBClin15 elements. *J Bacteriol*, **187**, 1966-1973.

Verheust, C., Jensen, G. and Mahillon, J. (2003) pGIL01, a linear tectiviral plasmid prophage originating from *Bacillus thuringiensis* serovar israelensis. *Microbiology*, **149**, 2083-2092.

WHO. (1996) The world health report: fighting disease, fostering development. World Health Organization, Geneva.

Wikoff, W.R., Liljas, L., Duda, R.L., Tsuruta, H., Hendrix, R.W. and Johnson, J.E. (2000) Topologically linked protein rings in the bacteriophage HK97 capsid. *Science*, **289**, 2129-2133.

Wilson, I.A., Skehel, J.J. and Wiley, D.C. (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature*, **289**, 366-373.

Wommack, K.E. and Colwell, R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev*, **64**, 69-114.

Wrigley, N.G. (1969) An electron microscope study of the structure of Sericesthis iridescent virus. *J Gen Virol*, **5**, 123-134.

Xia, D., Henry, L.J., Gerard, R.D. and Deisenhofer, J. (1994) Crystal structure of the receptor-binding domain of adenovirus type 5 fiber protein at 1.7 Å resolution. *Structure*, **2**, 1259-1270.

Xiao, C., Chipman, P.R., Battisti, A.J., Bowman, V.D., Renesto, P., Raoult, D. and Rossmann, M.G. (2005) Cryo-electron microscopy of the giant mimivirus. *J Mol Biol.*

Xu, L., Benson, S.D., Butcher, S.J., Bamford, D.H. and 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. and Baker, T.S. (2005) The marine algal virus PpV01 has an icosahedral capsid with T=219 quasisymmetry. *J Virol*, **79**, 9236-9243.

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

Yonekura, K., Maki-Yonekura, S. and Namba, K. (2002) Quantitative comparison of zero-loss and conventional electron diffraction from two-dimensional and thin three-dimensional protein crystals. *Biophys J*, **82**, 2784-2797.

Yonekura, K., Maki-Yonekura, S. and Namba, K. (2003) Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature*, **424**, 643-650.

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

Zhou, Z.H., Dougherty, M., Jakana, J., He, J., Rixon, F.J. and Chiu, W. (2000) Seeing the herpesvirus capsid at 8.5 Å. *Science*, **288**, 877-880.

Zimmer, S.G. and Millette, R.L. (1975) DNA-dependent RNA polymerase from *Pseudomonas* BAL-31. II. Transcription of the allomorphic forms of bacteriophage PM2 DNA. *Biochemistry*, 14, 300-307.

Zubieta, C., Schoehn, G., Chroboczek, J. and Cusack, S. (2005) The structure of the human adenovirus 2 penton. Mol Cell, 17, 121-135.