Packaging Motors of Cystoviruses

Denis E. Kainov

Department of Biological and Environmental Sciences,

Division of Genetics,

Institute of Biotechnology,

Graduate School in Informational and Structural Biology,

University of Helsinki

ACADEMIC DISSERTATION

To be presented for public criticism

with the permission of the Faculty of Science,

University of Helsinki, in the auditorium 1041, Viikki Biocenter 2

on November 4, 2005, at 12 noon

Helsinki 2005

Supervised by

Docent **Roman Tuma**
University of Helsinki, Finland

Reviewed by

Docent **Mikko Frilander**
University of Helsinki, Finland

Docent **Alexander Plyusnin**
University of Helsinki, Finland

Opponent

Professor **John Walker**
Medical Research Council,
Dunn Human Nutrition Unit,
Cambridge, United Kingdom

Chairman

Professor **Tapio Palva**
University of Helsinki, Finland

© Denis Kainov 2005
ISBN 952-10-2631-6 (nid) or (paperback)
ISBN 952-10-2632-4 (PDF, http://ethesis.helsinki.fi/)
ISSN 1795-7079
Yliopistopaino, University Press
Helsinki, Finland 2005

To my family

CONTENTS page

LIST OF ORIGINAL PUBLICATIONS 5

ABBREVIATIONS 6

SUMMARY 7

A. INTRODUCTION 8

A.1. GENOME ENCAPSIDATION IN VIRUSES 8

A.1.1. Capsid assembly around viral genome 8

A.1.2. Genome packaging into preformed capsid: Φ29 as a model system 11

A.2. GENOME PACKAGING OF BACTERIOPHAGE PHI6 14

A.2.1. Phi6 structure and life cycle 14

A.2.2. Structure and function of PC constituents 16

A.2.3. Model of sequential RNA packaging 19

A.3. FROM VIRUSES TO MOLEULAR MOTORS 20

A.4. HELICASE FAMILIES 21

A.4.1. Monomeric helicases 24

A.4.2. Oligomeric helicases 26

A.4.3. How NTP binding and hydrolysis effect DNA/RNA translocation in hexameric helicases 29

A.4.4. Nucleotide hydrolysis coordination within hexameric ring 29

A.4.5. Mechanism of DNA/RNA unwinding and translocation 30

B. AIMS OF THE STUDY 32

C. MATERIALS AND METHODS 33

C.1. BACTERIAL STRAINS AND PLASMIDS 33

C.2. EXPERIMENTAL METHODS 34

D. RESULTS AND DISCUSSION 35

D.1. ISOLATION AND CHARACTERISATION OF P4 PROTEINS 35

D.2. STRUCTURE OF P4 PROTEINS 35

D.3. NUCLEOTIDE BINDING AND SPECIFICITY 37

D.4. MECHANISM OF COUPLING OF ATP HYDROLYSIS TO RNA TRANSLOCATION 38

D.5. COORDINATION OF THE CATALYSIS BETWEEN SUBUNITS 40

D.6. RNA LOADING MECHANISM 42

D.7. REGULATION OF P4 ACTIVITY WITHIN THE VIRAL CORE 42

D.8. IMPLICATIONS FOR THE VIRAL LIFE CYCLE 43

E. CONCLUDING REMARKS 45

F. ACKNOWLEDGEMENTS 46

G. REFERENCES 47

REPRINTS OF ORIGINAL PUBLICATIONS 61

LIST OF ORIGINAL PUBLICATIONS

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

I Kainov, D.E., Butcher, S.J., Bamford, D.H., Tuma, R. (2003). Conserved intermediates on the assembly pathway of double-stranded RNA bacteriophages. *J Mol Biol* **328**, 791-804

II Kainov, D.E., Pirttimaa, M., Tuma, R., Butcher, S.J., Thomas, G.J. Jr., Bamford, D.H., Makeyev, E.V. (2003). RNA packaging device of double-stranded RNA bacteriophages, possibly as simple as hexamer of P4 protein. *J Biol Chem* **278**, 48084-48091

III Kainov, D.E., Lisal, J., Bamford, D.H., Tuma, R. (2004). Packaging motor from double-stranded RNA bacteriophage phi12 acts as an obligatory passive conduit during transcription. *Nucl Acid Res* **32**, 3515-3521

IV Mancini, E.J., Kainov, D.E., Grimes, J.M., Tuma, R., Bamford, D.H., Stuart, D.I. (2004). Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **118**, 743-55

V Lisal, J., Lam, T.T., Kainov, D.E., Emmett, M.J., Marshall, A.G., Tuma, R. (2005). Functional visualization of viral molecular motor by hydrogen-deuterium exchange reveals transient states. *Nat Struct Mol Biol* **12**, 460-466

VI Kainov, D.E., Mancini, E.J., Lisal, J., Grimes, J.M., Bamford, D.H., Stuart, D.I., Tuma, R. (2005) Structural basis of mechano-chemical coupling in hexameric packaging motors. *Manuscript*.

Some unpublished results are also presented.

ABBREVIATIONS

aa	amino acid
AAA	ATPases associated with various cellular functions
ATP	adenosine triphosphate
bp	base pair
BTV	bluetongue virus
CTP	cytidine triphosphate
d	deoxy
DLS	dimer linkage site
DNA	deoxyribonucleic acid
ds	double-stranded
EDTA	ethylenediaminetetraacetic acid
EGTA	ethyleneglycol-bis (β-aminoethyl ether)-N,N,N',N'-tetraacetic acid
EM	electron microscopy
gp	gene product
GTP	guanosine triphosphate
HCV	hepatitis C virus
HSV	herpes simplex virus
IPTG	isopropyl-β-D-thiogalactopyranoside
kb	kilobase
K_D	dissociation constant
kDa	kilodalton
MoMuLV	Moloney murine leukemia virus
NC	nucleocapsid
NDP	nucleoside diphosphate
NMP	nucleoside monophosphate
nt	nucleotide
NTP	nucleoside triphosphate
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PC	procapsid
RdRP	RNA-dependent RNA polymerase
RNA	ribonucleic acid
RNase	ribonuclease
RT	reverse transcriptase (RNA-dependent DNA polymerase)
PX	polymerase complex
SANS	small angle neutron scattering
SDS	sodium dodecyl sulfate
SF	superfamily
SL	stem loop
SLP	single-layered particle
ss	single-stranded
T	triangulation number
TLP	triple-layered particle
TMV	tobacco mosaic virus
UTP	uridine triphosphate
wt	wild type

SUMMARY

Viruses are molecular machines capable of selectively infecting their hosts, replicating and self-assembling new progeny. Although virus architecture, replication and assembly can be rather complex, several viruses have been developed as model systems to provide insight into all aspects of their life cycles. In particular, such systems can shed light on virus evolution, self-assembly of nanostructures, regulation of activities in cooperative multienzyme molecular machines and molecular motors, and facilitate technological application of viral-based molecular devices.

We use enveloped dsRNA bacterial viruses from the *Cystoviridae* family (phi6, phi8, phi12 and phi13) as models. In this work we investigated *in vitro* assembly of polymerase complexes and mechanism of RNA packaging. In particular, we mapped the assembly pathway of bacteriophage phi8 (I). Then we continued our studies by showing that genome packaging in *Cystoviridae* is catalyzed by a hexameric ATPase, P4 (II). Following a thorough biochemical characterization of this protein (II, III), we crystallized P4 ATPase of phi12 and, in collaboration with others, determined atomic structures of the holoenzyme and its several ligand-bound forms (IV). This work together with mutational analysis of key residues of phi12 packaging motor (VI) and characterization of the conformational dynamics of phi12 and related phi8 motors by hydrogen-deuterium exchange (V) provided a molecular mechanism for mechano-chemical coupling. This mechanism is applicable to other molecular motors that are involved in a number of biological processes, such as genome replication, repair and recombination.

A. INTRODUCTION

A.1. GENOME ENCAPSIDATION IN VIRUSES

Viruses are biological entities that can reproduce only within host cells. A viral genome contains all the information for its multiplication in the cell and for directing synthesis of its structural components. A viral genome can be either RNA or DNA, double-stranded or single-stranded. Viruses with RNA genomes are grouped into four main categories: (+) ssRNA, (-) ssRNA, dsRNA and reverse-transcribing viruses. The RNA genome may consist of a single RNA molecule or several segments (partitions). In RNA viruses, the strand that can serve as the messenger RNA in protein synthesis is known as the (+) strand and the complementary strand as the (-)-strand. The (+)-strand RNA genomes, upon introduction to the host cell, can serve directly as messengers, whereas the (-)-strand and dsRNA genomes must be transcribed into (+)-strands in the cell.

Viruses possess two basic structural symmetries with numerous variations, as the capsid can be helical or icosahedral. The size of the icosahedral capsid sets the limit to the amount of packaged nucleic acid. In contrast, the size of the genome determines the length of a helical virus capsid.

Genome packaging or encapsidation into viral capsids protects genome from nuclease degradation within the host cell as well as outside. Encapsidation also targets the genome to a new host cell or organism. Two main strategies for packaging of viral genomes exist: (1) capsid assembly around the viral nucleic acid; (2) filling of preformed capsid structures with a previously synthesized nucleic acid or a nucleic acid that is being synthesized during packaging. Our understanding of the genome packaging events has been based on a few model systems, but information from a great variety of viruses is rapidly accumulating. The following chapters describe various examples of viral packaging systems for which detailed information about the mechanism is available.

A.1.1. Capsid assembly around viral genome

Simple RNA viruses use co-condensation strategy to encapsidate their genomes. The assembly involves high affinity binding between the capsid protein and a packaging site within the viral genomic RNA. The RNA-coat protein complex nucleates capsid assembly. The energy required for genome encapsidation is provided by the interactions of the assembly components.

Detailed information of the co-condensation mechanism is available for both helical and icosahedral ssRNA viruses. Tobacco mosaic virus (TMV) is a classic case of such encapsidation mechanism for helical viruses, and it is well described elsewhere (Buck, 1999; Fritsch et al., 1973; Jonard et al., 1975).

Genome packaging in retroviruses - Moloney murine leukaemia virus (MoMuLV)

The MoMuLV is a prototypical retrovirus. It is widely used in human gene therapy and is extensively studied as a model for retrovirus assembly and genome encapsidation (Berkowitz et al., 1996). As with all retroviruses, MoMuLV packages dimeric genomic RNA (Mann et al., 1983). The genomic RNA is present in the virus particle in two identical copies joined by a limited number of Watson-Crick base pairs. The redundancy of the information provides insurance against breaks or other damage that would be fatal if only a single copy was present (Hu and Temin, 1990). In essence, retroviruses could be considered the simplest diploid organisms.

Electron microscopy studies on partially denatured RNAs from the mature particles revealed that dimeric genomic RNA were joined near their 5' ends. This region is called the dimer linkage site (DLS) (Oroudjev et al., 1999). A common feature of DLS regions is the presence of stem-loops (SLs) with four- or six-base palindromic sequences in the loops ('kissing loops') (Kim and Tinoco, 2000).

The assembly of a retrovirus particle is mediated by a nucleocapsid domain (NC) of the retroviral Gag polyprotein. Processing of Gag polyprotein generates the mature NC. NC stabilizes the genomic RNA dimer. Such stabilization or maturation of the genomic RNA dimer has also been described for RSV, HIV-1 and the retrotransposon Ty1 (Feng et al., 1995; Fu and Rein, 1993; Oertle and Spahr, 1990; Stewart et al., 1990). In the MoMuLV, NC normally incorporates the genomic RNA into the nascent particle by recognizing a *cis*-acting signal, termed ψ, present in the 5' end of the viral RNA (Prats et al., 1990). The fact that ψ and the DLS are in the same region of the viral RNA suggests that dimerization of the RNA is a prerequisite for packaging (Paillart et al., 2004).

MoMuLV ψ contains three stem-loops, designated SL-B, SL-C and SL-D (Mougel and Barklis, 1997). The bases in the SL-B stem shift register upon RNA dimerization. This shift exposes the palindromic sequence, AGCU, which becomes the loop in the dimeric form; this conformation enables the two monomers to form the kissing complex. In addition, dimerization-induced register shifts in base pairing

within the ψ region expose conserved UCUG elements that bind NC with high affinity (nM K_D). Dimerization may promote exposure of additional downstream UCUG elements to enhance the specific genome packaging (D'Souza et al., 2004; D'Souza et al., 2001).

The structure of the complex between MoMuLV NC and the 101-nucleotide 'core encapsidation' segment of the ψ site was determined (D'Souza and Summers, 2004). Structure revealed a network of interactions that promote sequence- and structure-specific binding by a single NC Cys-Cys-His-Cys zinc "knuckle". These findings support a structural RNA switch mechanism for genome encapsidation, in which the protein binding sites are sequestered by base pairing in the monomeric RNA and only become exposed upon dimerization to promote packaging of the diploid genome. Thus, a conformational switch between monomers and dimers, exposing a crucial sequence only in the dimeric RNA, may eventually be recognized as a general mechanism for the packaging of dimeric RNAs in retroviruses.

Examples of retroviral RNA dimerization also could be drawn from HIV-1-based research (Berkhout and van Wamel, 1996; Clever and Parslow, 1997; Haddrick et al., 1996; Laughrea et al., 1997; Paillart et al., 1996). In HIV-1, RNA dimerization is initiated by the base pairing of a self-complementary sequence that is located in the loop of an RNA hairpin motif. This initial loose dimer matures into a more stable (tight) dimer. Importantly, the HIV-1 dimerization initiation site identified *in vitro* is not the only sequence participating in RNA dimerization *in vivo*. Similarly to MoMuLV, in cells, HIV-1 RNA dimerization is linked to the trafficking of the Gag precursor protein and the RNA dimer can be used as a scaffold during viral assembly. HIV-1 RNA dimers undergo conformational changes during virion maturation, but the structure of the *in vivo* matured dimer remains speculative. Importantly, RNA dimerization is crucial for reverse transcription and recombination (Balakrishnan et al., 2001; Balakrishnan et al., 2003; Temin, 1991). Drugs targeting RNA dimerization could potentially limit the emergence of multidrug-resistant viruses (Jung et al., 2002; Moutouh et al., 1996; Yusa et al., 1997).

Bacteriophages MS2 and R17

Nucleation of capsid assembly around a viral nucleic acid is typical for some (+)-strand RNA viruses. Small icosahedral bacteriophages, such as MS2 and R17 are models for studying genome encapsidation (Beckett et al., 1988; Pickett and Peabody,

1993; Valegard et al., 1994). These viruses are built of one type of coat protein that exists as a stable dimer (Ni et al., 1995). Later in the infection, when the coat protein concentration in the cell increases, the coat protein binds specifically to an asymmetric RNA hairpin loop (21-nucleotides long), which is located in the translational initiation region of the replicase gene (Grahn et al., 2001; Grahn et al., 1999). This binding leads to translational repression of the replicase synthesis and stimulates the binding of additional coat-protein dimers to non-specific RNA sequences. This results in a highly cooperative assembly of the coat protein, forming a shell that simultaneously encloses the viral genome (Helgstrand et al., 2002; Horn et al., 2004; Stockley et al., 1995; Stockley et al., 1994; Valegard et al., 1994).

Yeast L-A virus

The yeast L-A virus is the only dsRNA virus of eukaryotic host which appears to nucleate the capsid around a (+)-strand transcript (Fujimura et al., 1990; Naitow et al., 2002; Sommer and Wickner, 1982). The L-A genome encodes for a major capsid protein (Gag) and a fusion protein that has the Gag upstream of the C-terminal RdRP domain (Gag-Pol) (Fujimura et al., 1992). The packaging starts with a formation of the initiation complex in which two Gag-Pol proteins dimerize and bind one (+)-strand RNA molecule (Fujimura and Esteban, 2000; Ribas and Wickner, 1998). The Gag domain of Gag-Pol primes the assembly of additional Gag proteins to form an icosahedral shell while condensing the RNA inside. Following packaging, the (-)-strand synthesis takes place inside the particle. The particle size is optimised for a single L-A genome copy. It is worth mentioning that the possibility of genome packaging into preformed capsid has not been completely ruled out for the L-A virus.

A.1.2. Genome packaging into preformed capsids - Φ29 as a model system

In many cases genomes of complex viruses are packaged into preformed empty capsids (procapsids, proheads). Examples include herpesviruses and adenoviruses, as well as tailed bacteriophages, such as Φ29, λ, P22, and T4 (Serwer, 2005). Molecular motors that translocate nucleic acid against concentration gradient at the expense of NTP hydrolysis perform the packaging reaction.

Genome packaging is best understood for the dsDNA bacteriophage Φ29, whose portal complex serves as the packaging motor (Guo, 2002; Guo, 2005). The complex contains three components: 1) the head-tail connector, a dodecamer of

protein p10, attached to a unique 5-fold prohead vertex; 2) a multimeric ring of structural RNAs; and 3) several copies of ATPase protein p16 (Grimes et al., 2002; Simpson et al., 2000).

In the model based on biochemical and structural data, one subunit of the portal dodecamer interacts with the DNA loaded in the central channel. ATP hydrolysis then drives a 12° rotation of the narrow end of the connector, resulting in a lengthwise expansion of the connector via a slight change in the angle of the long helices. In a subsequent step, the wide end of the connector follows the narrow end, allowing the structure to relax and contract while translating two base pairs into the capsid (Simpson et al., 2000).

The high efficiency of the Φ29 *in vitro* packaging system has recently made possible to measure directly the forces involved in DNA packaging (Smith et al., 2001). In these experiments, the biotinylated DNA tail protruding from stalled, partially packaged, complexes was attached to a polystyrene bead captured in an optical trap, and the capsid was tethered to another bead coated with anti-capsid antibodies. Upon the addition of ATP, the two beads moved closer together indicating that packaging was occurring. Detailed analysis of the packaging mode revealed that the motors are highly efficient: 95% of the complexes studied displayed movements of several micrometres. Packaging the entire 6.6 μm, 19-kilobase, genome was found to require about five minutes with short pauses. During these pauses, the DNA did not slip, indicating that the motor remained engaged. However, occasional slippages (averaging 44 base pairs) were observed, after which the motor immediately re-engaged and continued packaging.

Initially, the packaging rate was ~100 base pairs per second but dropped to zero as the capsid filled up and the motor stalled. A marked transition in the packaging rate was found to occur when the capsid was approximately 50% full. These data suggest that the rate decreases because of a build up pressure inside the capsid. Stalling force measurements suggest that the motor can drive DNA into the head with an approximate 30% efficiency until the internal force builds to ~50 pN. This makes the portal complex one of the strongest molecular motors studied to date, with strength approximately eight times that of kinesin and twice that of RNA polymerase.

Portal complexes of many other DNA bacteriophages are organised similarly to Φ29 (Meijer et al., 2001). DNA translocation in these systems is often accompanied

by the processing of genomic DNA concatemers (terminase activity). Since the p16 of Φ29 and terminases of others bacteriophages have been shown to catalyse NTP hydrolysis *in vitro*, these proteins are thought to generate the energy for the packaging reaction (Ibarra et al., 2001). It has been suggested that the energy is somehow transmitted from the transiently associated terminase via the RNA ring to the connector, which in turn effects unidirectional DNA translocation (Guo, 2002; Simpson et al., 2000a). However, the complexity of the portal structures in DNA viruses has made it difficult to verify experimentally the proposed packaging mechanism.

In vitro packaging of dsRNA viruses has been studied for bacteriophages phi6 and phi8 of the *Cystoviridae* family (Mindich, 1999; Mindich, 2004; Sun et al., 2003). The assembly of phi6 and phi8 bacteriophages proceed via formation of procapsids followed by packaging, like in Φ29 and other dsDNA bacteriophages (Butcher et al., 1997; Olkkonen et al., 1990; Poranen et al., 2001; Poranen and Tuma, 2004). The implications of the cystoviral model for the mechanism of packaging in other dsRNA viruses are not obvious. As we already discussed, L-A virus utilises more concerted mechanism, in which the capsid is formed around the RNA.

The mechanism of RNA packaging in rotaviruses still remains elusive, despite the structure of the putative packaging motor NSP2 had been determined (Jayaram et al., 2002). NSP2 is an octameric non-structural NTPase (Taraporewala et al., 1999). The NSP2 monomer has two distinct domains. The amino-terminal domain has a new fold. The carboxy-terminal domain resembles the ubiquitous cellular histidine triad (HIT) group of nucleotidyl hydrolases (Lima et al., 1997). This structural similarity suggests that the nucleotide-binding site is located inside the cleft between the two domains. Prominent grooves that run diagonally across the doughnut-shaped octamer are probable locations for RNA binding. Several RNA binding sites, resulting from the quaternary organization of NSP2 monomers, may be required for the helix destabilizing activity of NSP2 and its function during genome replication and packaging. NSP2 is not structurally related to the cystoviral packaging motor P4 (Jayaram et al., 2004; Taraporewala and Patton, 2004).

Since the *Cystoviridae* assembly and packaging is the main topic of this thesis an overview is presented in the following chapter. Some comprehensive description of the *in vitro* assembly (Poranen and Tuma, 2004) and packaging has been given recently (Mindich, 2004).

A.2. GENOME PACKAGING OF BACTERIOPHAGE PHI6

A.2.1. Phi6 structure and life cycle

Current knowledge of genome packaging in dsRNA bacteriophages is derived mostly from the phi6 system (Gottlieb et al., 1991; Gottlieb et al., 1990; Qiao et al., 1995; Qiao et al., 1997). Phi6 is the prototype virus of the family *Cystoviridae*, which includes eight additional members (phi7 to phi14) (Mindich, 1999; Mindich, 2004). Phi7, phi9, phi10, phi11, phi13 and phi14 exhibit clear sequence similarity to phi6, while phi8 and phi12 are more distantly related (Gottlieb et al., 2002; Hoogstraten et al., 2000; Mindich et al., 1999; Qiao et al., 2000). Phi6 infects the plant pathogenic bacterium *Pseudomonas syringae* (Semancik et al., 1973). Phi6 contains a ~13 kb genome comprising three segments (large - L, middle – M, and small - S) enclosed in an icosahedral polymerase complex which is further coated by a nucleocapsid (NC) protein shell and a lipid envelope (Emori et al., 1980; Emori et al., 1982). The envelope of phi6 contains phospholipids originating from the host plasma membrane (Laurinavicius et al., 2004; Sands and Lowlicht, 1976; Sands et al., 1975), and four virus-encoded integral membrane proteins, P6, P9, P10 and P13 (Bamford and Palva, 1980; Sinclair et al., 1975). P6 anchor the receptor binding spike protein P3 (Fig. 1).

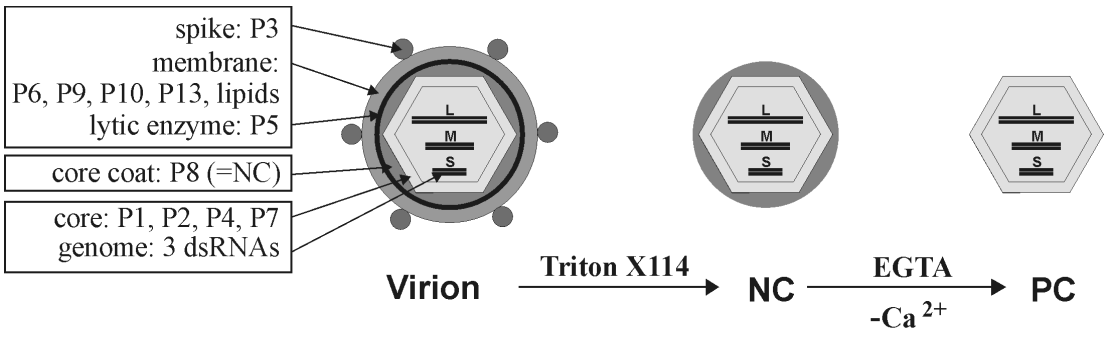


Figure 1. Organization of bacteriophage phi6 virion. The phi6 virion is composed of a icosahedral procapsid (PC) containing proteins P1, P2, P4, and P7. PC contains three dsRNA genomic segments L, M and S, and this RNA-containing particle is called the core. Core is covered by a T=13 shell of protein P8 to form a nucleocapsid (NC). The nucleocapsid is enveloped in a lipid-containing membrane composed of phospholipids and proteins P9, P10, P13, P3, and P6. P3, the receptor-binding spike, specifies the host range of the virus. P5 is a lytic endopeptidase that is associated with the surface of the nucleocapsid. Modified from (Mindich et al., 1999).

The envelope lacks icosahedral symmetry. Lipid envelope can be removed by detergent treatment (Bamford et al., 1976). The resulting NC is composed of the surface shell and the underlying polymerase complex. The surface shell is composed of P8 trimers that are arranged onto an incomplete T=13 lattice (five-fold vertices are occupied by proteins protruding from the polymerase complex) (Butcher et al., 1997; de Haas et al., 1999; Ktistakis et al., 1988). A similar arrangement has also been found in the intermediate layer of reoviruses (Grimes et al., 1998; Reinisch et al., 2000). The P8 lattice can be dissociated by calcium chelating agents, such as EGTA (Olkkonen et al., 1990). The resulting polymerase complex is composed of 120 copies of major structural protein P1 (Ktistakis and Lang, 1987; Steely and Lang, 1984), 12 monomers of RNA-dependent RNA polymerase P2 (Makeyev and Grimes, 2004), 12 hexamers of packaging motor P4 (Gottlieb et al., 1992; Juuti et al., 1998), and 30 dimers of assembly factor P7 (Juuti and Bamford, 1997). 120 copies of P1 are arranged as 60 dimers on a T=1 icosahedral lattice (Butcher et al., 1997; de Haas et al., 1999). The organisation with two molecules of the same protein in two different conformations in an icosahedral asymmetric unit is also referred to as the "T=2" structure (Grimes et al., 1998). Bluetongue virus VP3 layer as well as the rotavirus VP2 layer are organized similarly (Mertens and Diprose, 2004). This arrangement has been observed only in the core particles of the dsRNA viruses. Interesting hypothesis for the emergence of cystoviral/ reoviral "T=2" and T=13 layers was proposed by Coulibaly and co-workers and is diagrammed in Fig. 2 (Coulibaly et al., 2005).

During the virus entry the structural layers of phi6 virion are sequentially removed and the transcriptionally active core particle is delivered into the cytoplasm (Fig. 3; Bamford et al., 1976). Core plays a central role in the viral RNA metabolism (Bamford, 2000; Kakitani et al., 1980). Core is capable of synthesising (+)-RNA that is extruded from the particle (Ewen and Revel, 1988). The (+)-RNA serves as a template for viral protein synthesis and constitutes the precursor for packaging into empty PC (Frilander and Bamford, 1995). The packaged (+)-RNA is then replicated inside the PC to yield genomic dsRNA (Frilander et al., 1992; Gottlieb et al., 1991). Proteins P2 and P4 constitute the enzymatic machinery for genome transcription, replication and packaging (Butcher et al., 2001; Juuti et al., 1998).

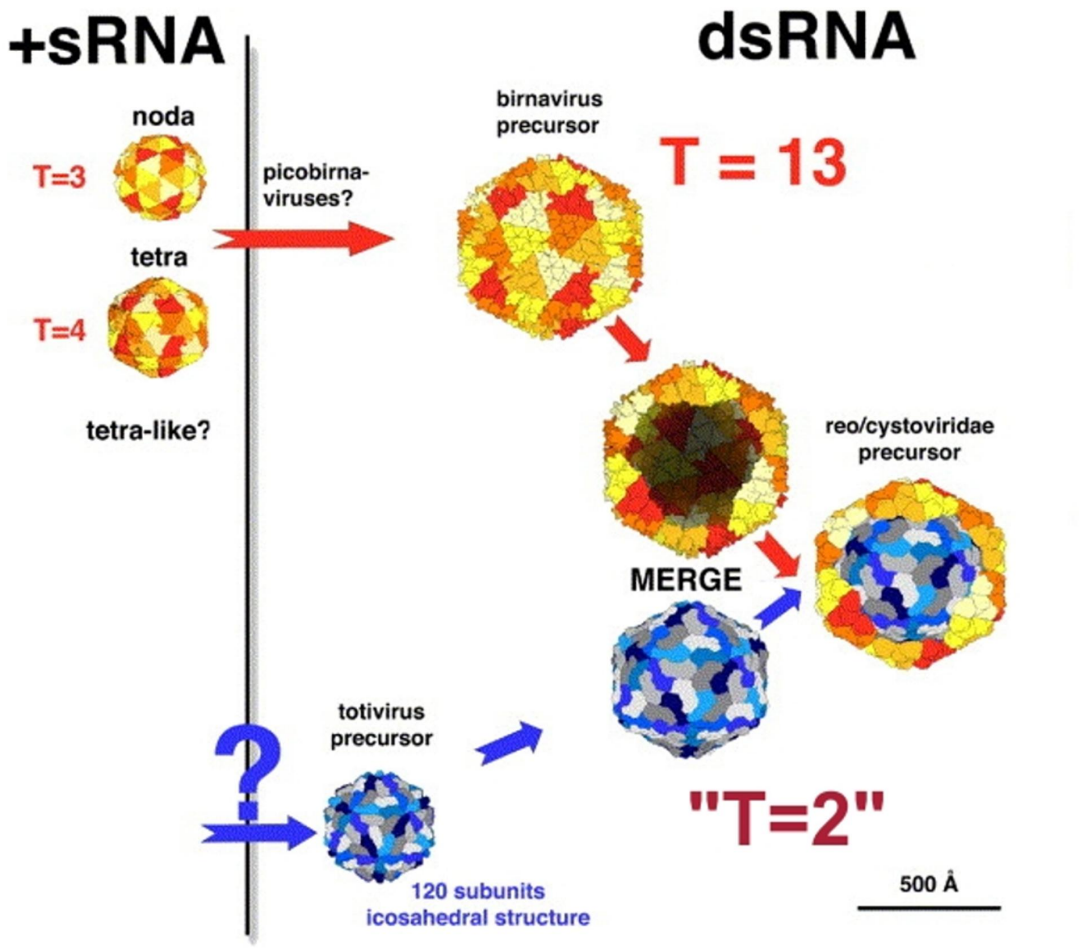


Figure 2. Putative emergence of the *Reo/Cystoviridae* by merging of genome segments from two simpler dsRNA viruses. Yellow/red colors represent *Birnaviridae*-related structures and blue/gray colors the *Totiviridae*-related counterparts. Modified from (Coulibaly et al., 2005).

A.2.2. Structure and function of PC constituents

RdRP P2 is central to genome replication and transcription

P2 is a minor constituent of PC. Twelve P2 subunits are located under the fivefold vertices of the icosahedron (Ikonen et al., 2003). It was shown that the purified recombinant P2 acts both as a replicase and a transcriptase using ssRNA and dsRNA substrates, respectively (Fig. 4) (Makeyev and Bamford, 2000a; Makeyev and Bamford, 2000b). Isolated P2 is a relatively unspecific, primer-independent, polymerase which can accept a range of heterologous templates. However, when incubated with the phi6 dsRNA segments L, M and S, phi6 P2 synthesizes predominantly (+) sense copies, M and S being more efficient templates than L.

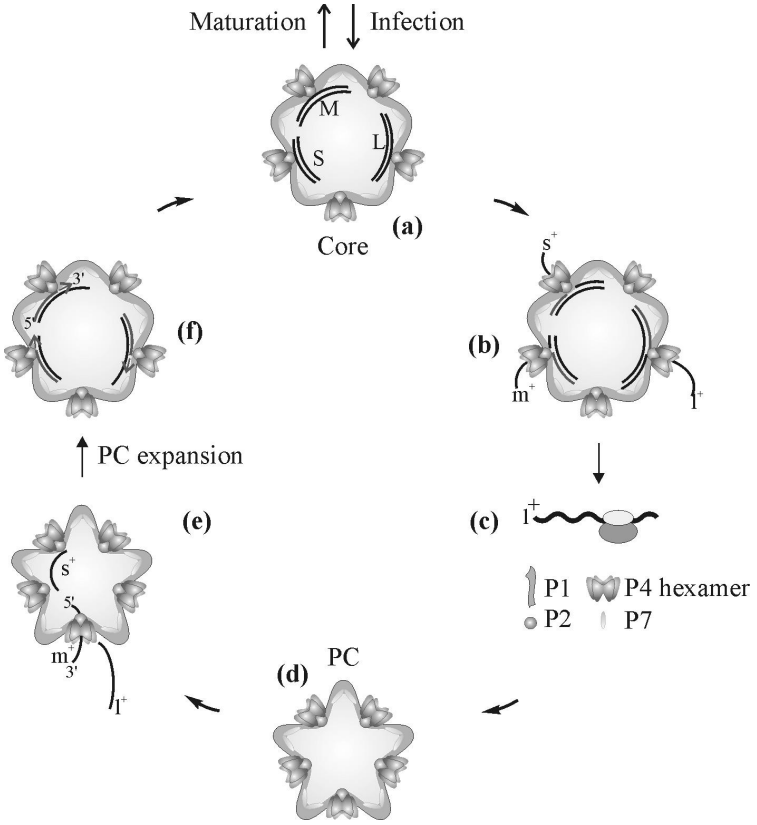


Figure 3. Simplified scheme of the cystoviral life cycle. The three dsRNA genomic segments of a cystovirus are brought into the host cell inside the viral core (a). Upon cell entry, the core catalyses semi-conservative dsRNA transcription and (+)-sense ssRNA transcripts (l^+, m^+ and s^+) are extruded into the cytoplasm (b). Ribosomes translate l^+ RNA (c) giving rise to proteins P1, P2, P4 and P7. The newly produced proteins assemble into empty polymerase complexes (PC) (d), which are capable of packaging specifically one copy of each l^+, m^+ and s^+ segments (e). Upon packaging PC expands and replication is initiated (f). The dsRNA-filled PC (core) can enter additional rounds of transcription or mature into infectious virions. The latter pathway uses proteins produced by the translation of m^+ and s^+ ssRNA segments, which is followed by the acquisition of the rest of the viral structural proteins together with the lipid membrane (not shown). The mature virus particles are released by lysis of the host cell.

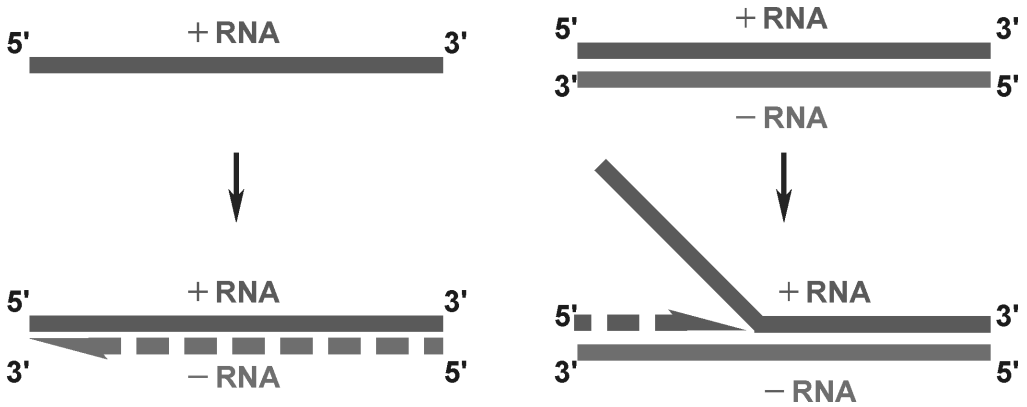


Figure 4. Diagram showing the two reactions catalysed by cystoviral RdRP: replication (left) and transcription (right). Modified from (Makeyev and Grimes, 2004).

This distribution of the test tube reaction products faithfully mimics the phi6 (+) strand synthesis *in vivo*, thus suggesting a 'polymerase-centric' model for phi6 RNA metabolism regulation (Makeyev and Grimes, 2004).

The structure of phi6 P2 was solved by X-ray crystallography (Butcher et al., 2001). It is highly similar to that of Hepatitis C virus (HCV) polymerase subunit and the caliciviral polymerase (Ago et al., 1999; Bressanelli et al., 2002; Bressanelli et al., 1999; Lesburg et al., 1999). Unlike the open-hand appearance of other polymerase structures, these three RdRPs resemble a cupped right hand with fingers and the thumb interconnected by several loops protruding from the fingers. In addition, they contain a C-terminal extension, which is important for the primer-independent initiation. This similarity suggests a possible evolutionary link between *Cystoviridae* and certain (+)-strand ssRNA viruses. Furthermore, crystal structure of the phi6 P2 complexes with oligonucleotide templates and/or NTP substrates yielded a general mechanism for initiating primer-independent RNA polymerisation (Butcher et al., 2001; Salgado et al., 2004).

Packaging NTPase P4

Protein P4 has been identified as the packaging NTPase, since particles devoid of P4 were completely inactive in ssRNA packaging assays (Ewen and Revel, 1990; Juuti and Bamford, 1995; Ojala et al., 1993). P4 plays additional role during transcription when the newly synthesised mRNA molecules are extruded from core particles (Pirttimaa et al., 2002). Furthermore, although P4 null particles can be expressed and assembled in *E.coli*, under defined *in vitro* conditions the particle assembly is dependent on hexameric P4 (Poranen et al., 2001).

P4 sequence contains characteristic NTPase motifs, which are conserved among the *Cystoviridae*, although overall sequence identity is low. The purified phi6 P4 is an unspecific NTPase, hydrolysing ribo-, deoxyribo-, and dideoxyribonucleoside triphosphates (Gottlieb et al., 1992; Paatero et al., 1995). The phi6 P4 forms doughnut-shaped hexamers in the presence of divalent cations and ATP or ADP, and the NTPase activity is associated only with the multimeric form. The NTPase activity is weakly stimulated by ssRNA (Juuti et al., 1998). Cryo-EM reconstruction of phi6 PC has localised P4 hexamers to the 5-fold vertices of PC (de Haas et al., 1999). This constitutes a symmetry mismatch. Recent EM of phi8 subviral particles showed that the localization of P4 is likely to be similar for all cystoviruses (Yang et al., 2003).

The symmetry mismatch is typical for packaging motors of tailed bacteriophages, and it has been proposed to facilitate rotation of the portal during nucleic acid translocation (Hendrix, 1978).

Although P4 NTPase activity is essential for RNA packaging, the details of the mechano-chemical coupling during RNA packaging in dsRNA viruses were poorly understood, a situation similar to the genome packaging in dsDNA viruses.

P1 and P7

Protein P1 is the major PC component. P1 provides the structural framework for assembly of the other proteins (Ktistakis and Lang, 1987; Olkkonen and Bamford, 1987). P2, P4 and P7 can associate independently with the respective deficient particles (Poranen et al., 2001). P1 participates in the initial binding and recognition of the genomic ssRNAs as well as in the organisation of dsRNA genome inside the capsid (Qiao et al., 2003a; Qiao et al., 2003b).

P7 is a minor component of the PC. Purified P7 exists as dimer which is elongated in shape (Juuti and Bamford, 1997; Kainov et al., 2004). The function of P7 is not clear, but the analysis of particles missing P7 have revealed that P7 is needed for transcription and packaging (Juuti and Bamford, 1995; Juuti and Bamford, 1997). In addition P7 accelerates the *in vitro* assembly process (Poranen et al., 2001). The exact location of P7 in PC is not known. Using SANS P7 was estimated to reside at a distance of 160 Å from the procapsid center indicating localization on the inner surface of P1 framework (Ikonen et al., 2003).

A.2.3. Model of sequential RNA packaging

Packaging is specific for phi6 RNA and packaging signals determine this specificity (Gottlieb et al., 1991). The packaging signals encompass about 210-280 nucleotide-long stretches at the 5’ non-coding ends of segments (Gottlieb et al., 1994). The first 18 nucleotides at the 5’-end are common to all segments (Pirttimaa and Bamford, 2000). In each segment there is also a 10-12 nucleotides long homologous sequence within the first 100 nucleotides from the 5’-terminus. This sequence forms a hairpin structure with a very stable tetra-loop UACG, which is a phi6-specific signal in the packaging site. In the segment s+, there is an extra copy of the tetra-loop.

Packaging of phi6 RNA is sequential (Qiao et al., 1995). The PC recognizes the three ssRNA genomic precursor molecules in a segment-specific manner and translocates the RNA molecules into the particle interior so that s+ segment is packaged first, followed by m+ and then l+ (Frilander and Bamford, 1995; Qiao et al., 1997). The PC particle undergoes considerable conformational changes during packaging (Butcher et al., 1997). Estimates of the internal volume for the empty and filled particles showed that the PC particle has to expand several times during ssRNA packaging in order to have enough space for all three segments. This may explain the sequential packaging: different segment-specific, high-affinity binding sites are exposed on the particle surface upon each particle expansion event (Frilander and Bamford, 1995).

Energy for the packaging is provided by NTP hydrolysis carried out by the P4 hexamers (Gottlieb et al., 1992). One of twelve hexamers is distinct, being tightly attached to the particle (Ewen and Revel, 1990). Mutant or detergent-treated particles containing reduced amounts of P4 (equivalent to occupancy of one vertex only) have been shown to be efficient in RNA packaging and (-) strand synthesis, but (+) strand synthesis in these particles was completely abolished. Pirttimaa et al. proposed the "special vertex" model (Pirttimaa et al., 2002). It remains to be seen whether the members of *Reoviridae* family adopt a similar scheme for their RNA packaging.

A.3. FROM VIRUSES TO MOLECULAR MOTORS

Molecular motors convert chemical energy (e.g. ATP hydrolysis) into mechanical work, which is manifested by directional motion. Two motors in cystoviruses are associated with the PC, namely the packaging motor P4, which threads ssRNA into the empty PC, and the P2 polymerase, which translocates RNA within and out of the PC (Bamford, 2000). On the basis of hexameric nature of the packaging ATPase P4 from phi6, ability to hydrolyse nucleotides and localization of P4 within PC, it was proposed that P4 might represent a new class of molecular motors, which couples ATP hydrolysis to RNA translocation (Juuti et al., 1998).

When compared to the portal complexes of dsDNA bacteriophages the portal of cystoviruses looks very simple. The hexameric P4 from phi6 and other *Cystoviridae* members have a ring-like morphology with a central channel, while dsDNA portals are usually dodecameric (Moore and Prevelige, 2002). P4 is a structural component of PC and mature capsids, while terminases are only transiently

associated with procapsids and are not present in mature dsDNA virions (Catalano et al., 1995; Kanamaru et al., 2004).

It is important to mention here, that P4 motor contains RecA-like ATPase domain, which is essential for mechano-chemical coupling. Terminase subunits from dsDNA phages and herpesvirus (U_L15 in HSV-1, gp17 in bacteriophage T4, gp16 in Φ29, gpP in P2, gpA in λ) contain similar domain (Bain et al., 2001; Catalano, 2000; Goetzinger and Rao, 2003; Guo, 2005; Linderoth et al., 1991; Przech et al., 2003; White et al., 2003). These terminases, however, are not well characterised due to their complex nature. Certain insight into the RecA-like ATPase mechanism was obtained for helicases (Wang, 2004). In the next chapter, we describe the molecular basis of helicase mechanisms in order to make a comparison with P4 mechanism.

A.4. HELICASE FAMILIES

Helicases are molecular motors that use the energy of NTP hydrolysis to translocate along nucleic acid strand and catalyze various reactions such as DNA unwinding. Approximately 1% of both eukaryotic and prokaryotic genomes codes for helicases (Gorbalenya and Koonin, 1993). They are known to play essential roles in nearly all aspects of nucleic acid metabolism, such as DNA replication, repair, recombination, and transcription (Bennett and Keck, 2004; Egelman, 1998; Matson et al., 1994). These enzymes usually act in concert with other proteins (von Hippel and Delagoutte, 2001; von Hippel and Delagoutte, 2003). All helicases share at least three common biochemical properties: (a) nucleic acid binding; (b) NTP/dNTP binding and hydrolysis; (c) NTP/dNTP hydrolysis-dependent translocation on duplex or single-stranded DNA/RNA. All helicases could be divided into classes using the following criteria: (1) by directionality, i.e. 5' ->3' or 3' ->5' helicases, with respect to the strand they bind to and move along, (2) by the type of the nucleic acid substrate they act upon, i.e. DNA or RNA helicases, (3) by quaternary or oligomeric structures.

Helicases could also be classified into superfamilies (SF), based on the extent of sequence similarity and organisation of conserved helicase motifs (Fig. 5; Gorbalenya and Koonin, 1993).

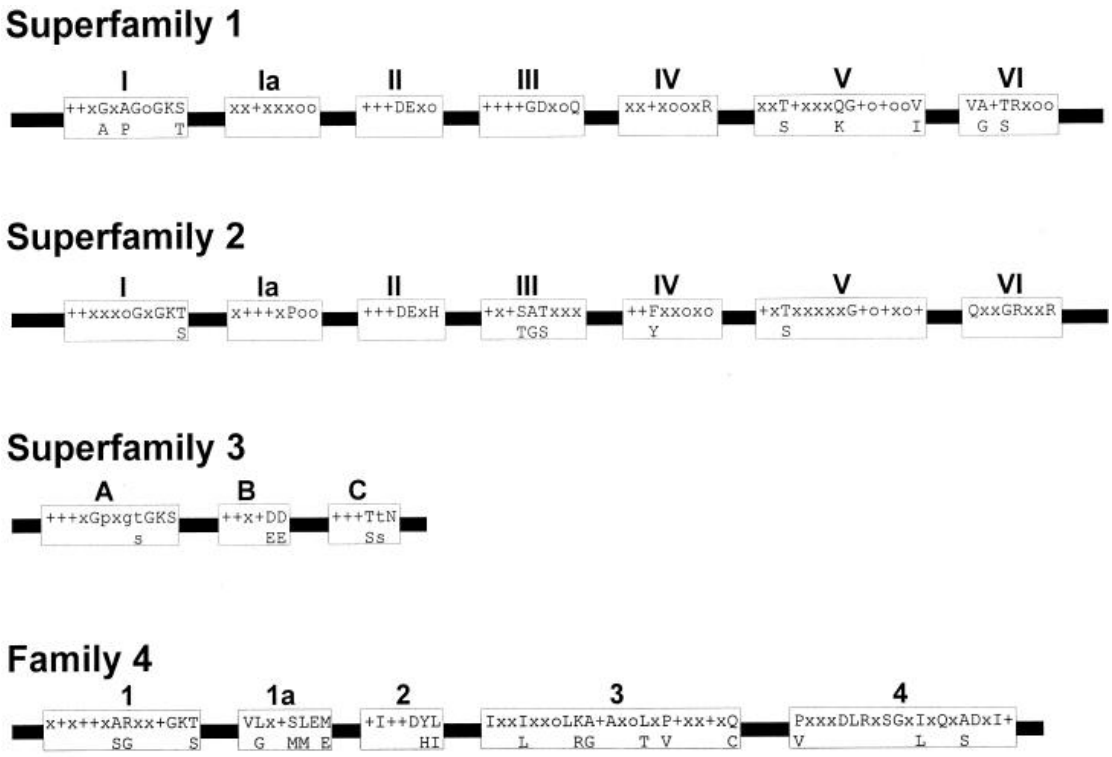


Figure 5. Schematic diagrams representing proteins from four helicase families (modified from Hall and Matson, 1999). Open boxes represent the conserved helicase motifs, and letters inside the boxes are the consensus amino acid sequences of each motif. Labels above the open boxes are the names assigned to the motifs. The relative positions of motifs and spacing between motifs are arbitrary. The consensus amino acid sequences of SF1 and SF2 were taken from (Gorbalenya and Koonin, 1993). Single-letter amino acid abbreviations indicate the presence of an amino acid in more than 75% of the family members. The consensus sequences for SF3 motifs were taken from (Gorbalenya et al., 1989). Upper-case amino acid abbreviations for SF3 indicate the presence of the residue in more than 50% of all family members and lower-case abbreviations indicate the presence of the residue in more than 50% of either the DNA viral or RNA viral family members. The consensus sequences of F4 motifs were derived from (Ilyina et al., 1992). Single-letter amino acid abbreviations were used when a residue was present in at least six of the seven aligned family members or when no more than two different residues occupied a particular position in the seven aligned sequences. In all family consensus sequences, a '+' represents a hydrophobic residue, an 'o' represents a hydrophilic residue and an 'x' represents a residue that is not restricted to being hydrophobic or hydrophilic.

SF1 and SF2 are the largest and most closely related groups, containing more than 50 and 100 members, respectively, from viral, prokaryotic and eukaryotic organisms (Gorbalenya and Koonin, 1993). In general, SF1 members (Rep, UvrD and PcrA) translocate along single stranded RNA or DNA, whereas SF2 members (RecG, PriA and HSV NS3) are double-stranded DNA translocases. They each contain seven to nine conserved amino acid motifs whose sequences, arrangements and predicted secondary structures are, in general, very similar (Gorbalenya et al., 1989).

Superfamily 3 (SF3) is much smaller and members of this group contain only three conserved motifs (Gorbalenya and Koonin, 1993). They are found generally in

RNA and DNA viruses, and many members must be considered putative helicases at this time because unwinding activity has not been demonstrated.

A fourth family consists of helicases that are related in sequence to the *E. coli* DnaB protein (Caruthers and McKay, 2002); these proteins have five motifs, unwind DNA in the 5' to 3' direction and generally form hexameric ring structures. These helicases are found in bacterial and bacteriophage systems. They are invariably associated physically with DNA primases, and thus function in DNA replication (Ilyina et al., 1992). Finally, an additional group, exemplified by the transcription termination factor Rho, was recognized as a family with sequence similarity to the β subunit of proton-translocating ATPases (Caruthers and McKay, 2002).

Conserved helicase motifs

The conserved motifs can be envisioned as the engine of a helicase, generating energy by the consumption of fuel (NTPs) and using the energy to do work (translocation along DNA or RNA and unwinding of duplex DNA or RNA) (Hall and Matson, 1999). Helicase motifs generally present themselves either at the interface between domains or at the interface with the oligonucleotide. The spatial segregation of the NTP- and oligonucleotide-binding sites suggests the partitioning of the function of the motifs between (1) those whose primary function is to bind MgNTP/MgNDP; (2) those that function solely or primarily in oligonucleotide binding; (3) those that participate in a molecular mechanism of coupling the NTPase cycle to the intramolecular conformational changes that drive the duplex unwinding and processive strand displacement or translocation activities. Third group of motifs would logically be localized to the middle region between the NTP and oligonucleotide binding sites (Caruthers and McKay, 2002).

The fingerprints that retain the greatest extent of conservation across all helicases are the Walker A and B motifs (motifs I and II, respectively). The Walker A motif (phosphate-binding loop or 'P-loop'), which was classically defined as a GxxxxGKT consensus, minimally requires the three final residues GK(T/S); the amino group of the lysine interacts with the phosphates of MgNTP/MgNDP and the hydroxyl of the threonine or serine ligates the Mg2+ ion (Gorbalenya and Koonin, 1993; Walker et al., 1982). The Walker B motif, originally defined as a single aspartic acid residue, takes the general form DExx across the SF1 and SF2 superfamilies. The carboxyl of the aspartic acid coordinates the Mg2+ ion of MgNTP/MgNDP through

outer sphere interactions, whereas the glutamic acid is suggested to act as a catalytic base in NTP hydrolysis (Gorbalenya and Koonin, 1993; Walker et al., 1982).

The remaining motifs are not common among SFs. Motifs III and V show substantial divergence in both length and amino acid sequence between SF1 and SF2. In some instances, these two motifs participate in a complex network of interactions that include ligation of MgNTP/MgNDP, formation of specific salt bridges or hydrogen bonds between domains 1 and 2, and contribute to binding oligonucleotides. Many of the interactions seen in crystal structures are specific to particular helicases, correlating with the family-to-family variations in the sequences of these motifs. The consensus sequence of motif VI is the unique feature of the SF1 and SF2 helicases. Shared by both superfamilies is an arginine in the middle of the motif, which forms a salt bridge with the γ-phosphate of the nucleotide (Velankar et al., 1999). Motifs Ia and IV contribute specific interactions with oligonucleotides (Gorbalenya and Koonin, 1993).

DnaB-like helicases contain five conserved motifs (H1, H1a, H2, H3, H4) (Gorbalenya and Koonin, 1993; Ilyina et al., 1992). The conserved H1 and H2 motifs contain the Walker A and B sequences (Walker et al., 1982). The exact role of H3 is unclear. Because of its location in the T7 gp4 helicase domain structure, it has been proposed to play a role in energy transduction (Sawaya et al., 1999). H4 may be involved in DNA binding (Washington et al., 1996). There are several residues beyond H4 that show some sequence conservation, and they are involved in nucleotide binding and hydrolysis (Patel and Picha, 2000).

A.4.1. Monomeric helicases

The first crystal structure of a helicase was that of PcrA (SF1, 3'->5' helicases) (Subramanya et al., 1996). X-ray structures of PcrA, crystallised with or without ADP, were very similar. In both cases, PcrA monomer consists of two parallel domains (1 and 2) with a deep cleft running between them. Each domain contains two additional subdomains (A and B). The subdomain 1A (which carries the Walker A and B motifs) and the subdomain 2A both structurally resemble the central region of RecA. The ADP moiety is located at the bottom of the cleft between the subdomains 1A and 2A. Only 1A can bind ADP. This cleft is lined with conserved helicase motifs. No site-bound Mg^{2+} has been found in the two crystal structures.

PcrA catalysed helicase activity comprises two structurally distinct but coupled activities. Firstly, the protein contains a single stranded DNA motor (above), hydrolysing one ATP per base translocated. Secondly, the protein undergoes ATP dependent conformational changes, which actively distort the duplex DNA ahead of the progressing ssDNA tracking motor, presumably aiding in its progress (Soultanas et al., 2000).

The crystal structure of UvrB helicase has been elucidated (Theis et al., 1999). UvrB structure was used as a template for the development of a structural model for the XPD helicase (helicase of general transcription factor IIH; TFIIH). XPD and UvrB helicases are functionally related, however, they share extremely low sequence identity (<15%). The XPD structural model has been used to understand the molecular mechanism of XPD-linked human disease and mutations (Bienstock et al., 2003).

Recently the crystal structure of the RecG helicase (3'->5', SF2) from the thermophilic bacterium *Thermotoga maritima* in a complex with its stalled replication fork substrate (three-way DNA junction, the preferred physiological substrate) has been solved (Singleton et al., 2001a). RecG protein is involved in the processing of stalled replication forks, and acts by reversing the fork past the damage to create a four-way junction that allows template switching and lesion bypass (Higgins et al., 1976; McGlynn and Lloyd, 2000). RecG protein differs from other helicases analysed at atomic resolution in that it mediates strand separation via translocation on ds- rather than ss-DNA. DNA is bound mainly to the large N-terminal domain of RecG, a domain that is not found in other DNA helicases. This region of the protein not only clamps onto and splits open the junction, but also stabilizes unwinding of the fork. In the structure, the junction has already begun to unwind, catching the complex in the initial stages of fork reversal. The template arm of the DNA (i.e., the region that would precede the moving replication fork) is bound across the interface between the N- and C-terminal domains, suggesting a novel mechanism for DNA unwinding (Singleton et al., 2001b).

The crystal structure of full-length eIF-4A from yeast has been reported to be a ‘dumbbell’ structure consisting of two compact domains connected by an extended linker (Caruthers et al., 2000). The eukaryotic translation initiation factor 4A (eIF4A) is a member of the DEA(D/H)-box RNA helicase family. By using the structures of other helicases as a template, compact structures was modeled for eIF4A that suggested (i) helicase motif IV binds RNA; (ii) Arg298, which is conserved in the

DEA(D/H)-box RNA helicase family but is absent from many other helicases, also binds RNA; and (iii) motifs V and VI "link" the carboxyl-terminal domain to the amino-terminal domain through interactions with ATP and the DEA(D/H) motif, providing a mechanism for coupling ATP binding and hydrolysis with conformational changes that modulate RNA binding.

The crystal structure of RuvB helicase together with its partner RuvA was determined, and structural basis for the Holliday junction migrating motor machinery was specified (Yamada et al., 2001). The RuvB motor protein is classified as an AAA^+ family (Neuwald et al., 1999). Like most AAA^+ family members, the oligomeric state of RuvB depends on the concentration of protein, and ions and cofactors, such as Mg^{2+}, nucleotides, and DNA (Mitchell and West, 1994). Crystallographic studies showed that RuvB has a crescent-like architecture consisting of three consecutive domains (N, M, and C) in which the first two domains (N and M) are involved in ATP hydrolysis. The overall architectures of these domains are well conserved in the AAA^+ family.

Recently, two structures of *E. coli* RecQ have been determined: a structure of the catalytic core in its unbound form at 1.8 Å and another structure of the core bound to the ATP analogue ATPγS at 2.5 Å resolution (Bernstein et al., 2003). The RecQ core comprises four conserved subdomains; two of these combine to form its helicase region, while the others form unexpected Zn^{2+} binding and winged-helix motifs. The structures reveal the molecular basis of missense mutations that cause Bloom's syndrome, a human RecQ-associated disease.

The *E. coli* Rep helicase was crystallized with ssDNA, giving the first glimpse of how the protein interacts with nucleic acids (Korolev et al., 1997). The ssDNA binding site involves the helicase motifs Ia, III, and V, whereas the ADP binding site involves helicase motifs I and IV. Residues in motifs II and VI may function to transduce the allosteric effects of nucleotides on DNA binding. These structures represent the first view of a DNA helicase bound to DNA.

A.4.2. Oligomeric helicases

Table 1 lists the various oligomeric helicases and shows that most are involved in DNA replication, recombination, repair and transcription.

Table 1. Oligomeric helicases (modified from www.helicase.net/m_struct.htm)

Protein	Directionality	DNA/RNA	PDB / ID	Structural details and references	Role
Bacteriophages and viruses					
E.coli phage T7 gp4	5'-3'	DNA	1CR0 1CR1 1CR2 1CR4 1E0J 1E0K 1Q57	3.45 Å crystal structure of the T7 primase-helicase (Toth et al., 2003). Crystal structures of the helicase domain (Sawaya et al., 1999; Singleton et al., 2000)	Helicase-primase gp4 is part of T7 replisome together with DNA pol gp5, processivity factor (thioredoxin), and ssDNA-binding protein gp2.5.
E.coli phage T4 gp41	5'-3'	DNA		27 Å EM reconstruction for T4 gp41complexed with ssDNA in the presents of MgATPγS (Norcum et al., 2005)	Helicase gp41 is part of gp41-gp61 primosome which is part of T4 replisome.
B.subtilis phage SPP1 gene 40	5'-3'	DNA		~23 Å EM reconstruction (Barcena et al., 1998)	G40P is part of SPP1 replisome
Simian virus SV40 Large Tumor Antigen	5'-3'	DNA	1N25 1SVM 1SVL 1SVO	2.8 Å crystal structure of a LTag fragment that assembles into a hexamer with helicase activity (Li et al., 2003). 1.9 Å structures of LTag hexamers in distinct nucleotide binding states (Gai et al., 2004)	Essential for initiation and elongation of viral DNA replication
Human papilloma-virus HPV E1	3'-5'	DNA	1F08 1KSX 1KSY 1TUE	Crystal structure of the DNA binding domain (Enemark et al., 2000) The X-ray structure of the E1 in complex with its matchmaker E2 (Abbate et al., 2004).	Replication initiation of papillomavirus genome
Bacterial					
E.coli DnaB	5'-3'	DNA	1B79 1JWE	Crystal structure of the N-terminal domain of the DnaB (Fass et al., 1999). NMR structure of the N-terminal domain of E. coli DnaB (Weigelt et al., 1999)	DnaB unwinds the DNA duplex at the Escherichia coli chromosome replication fork
RuvB	5'-3'	DNA	1IXS 1IXR	1.6 Å crystal structure of *Thermotoga maritima* RuvB (Putnam et al., 2001). Crystal Structure of the RuvA-RuvB Complex from *Thermus thermophilus* (Yamada et al., 2001).	Branch migration of Holliday junctions
E.coli Rho	5'-3'	DNA/RNA	1A62 1A63 1A8V 1PVO	Crystal structure of the RNA-binding domain (Allison et al., 1998; Bogden et al., 1999) Mechanism of mRNA recognition and helicase loading (Skordalakes and Berger, 2003)	Rho-dependent transcription termination
E. coli Rep protein	5'-3'	DNA	1UAA	Crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP (Korolev et al., 1997)	Replicative helicase

Archaea					
M. thermoauto-trophicum MCM	3'-5'	DNA	1LTL	Crystal structure of the N-terminal domain at 3.0 Å resolution (Fletcher et al., 2003). 23-Å 3D reconstruction of the full-length MtMCM from negatively stained particles (Pape et al., 2003).	DNA replication
Plasmid					
Plasmid encoded RSF1010 RepA	5'-3'	DNA	1OLO 1G8Y 1NLF	Crystal tructures at 2.4 and 1.95 Å resolution (Niedenzu et al., 2001).	Replication initiation
Eukaryotic					
Human Mcm4/6/7	3'-5'	DNA		EM (Sato et al., 2000).	Function at the origin of replication in all eukaryotes
Human Bloom's syndrome helicase	3'-5'	DNA		EM (Karow et al., 1999).	Involved in processing nascent DNA at blocked replication forks prior to the resumption of DNA synthesis.

Atomic structures have now been reported for several hexameric helicases, such as bacteriophage T7 gene 4 helicase domain (T7 helicase), archaeal minichromosome maintenance protein complex (MCM), Rho transcription terminator, replicative helicase (LTag) of the SV40 virus large tumor antigen, and plasmid RSF1010 replicative helicase RepA (Fletcher et al., 2003; Li et al., 2003; Niedenzu et al., 2001; Singleton et al., 2000; Skordalakes and Berger, 2003). In the case of Rho helicase, the structure has been solved in the presence and absence of AMP-PNP, but revealed the same conformation. Most importantly, no hexameric helicase has been visualized in a conformation with bound Mg^{2+}, which is strictly required for NTP hydrolysis. The structures of T7 gp4 and RepA revealed a single-domain protomer with a parallel α/β core topology that includes the RecA core fragment, as well as four β-strands extending the β-sheet on the carboxy-proximal side, three of which are similar to strands of the RecA protein (Niedenzu et al., 2001; Sawaya et al., 1999; Singleton et al., 2000).

Structural information has come also from Cryo EM studies. Low-resolution structural information is available for papillomavirus E1, bacteriophage T4 gp41, and others (Enemark et al., 2000; Enemark et al., 2002; Norcum et al., 2005). The above helicases act on DNA, except Rho, a transcription terminator protein that acts to separate RNA–DNA hybrids (Skordalakes and Berger, 2003).

A.4.3. How NTP binding and hydrolysis effect DNA/RNA translocation in hexameric helicases?

Recent structural and biochemical studies shed light on the molecular mechanisms of hexameric helicases. Several models exist. One of them is based on the crystallographic structures of replicative hexameric helicase of SV40 large tumor antigen (LTag) at different stages of the catalytic cycle (Gai et al., 2004; Li et al., 2003). The SV40 helicase, an AAA^+ protein, is a member of SF3 superfamily. LTag functions as an efficient molecular machine powered by ATP binding and hydrolysis for origin DNA melting and fork unwinding in viral replication. It was demonstrated that motion of β-hairpins within the central channel is coupled to ATP binding via a domain movement, and consequently the hairpins were proposed to bind DNA and effect translocation.

Two models for the power stroke mechanism of T7 gp4 helicase exist: a "subunit rotation" model (Singleton et al., 2000) and a "direct drive" model (Falson et al., 2004; Mavroidis et al., 2004). The "subunit rotation" model was based on the information derived from helicase crystal structure with and without dTTP. The relative rotational angles of the adjacent subunits of the hexamer are different for the empty state and the dTTP bound state. When dTTP binds to the catalytic site, it triggers the alignment of the positively charged groups constituting the so-called "arginine fingers" to the γ-phosphate of dTTP (Crampton et al., 2004). This alignment drives the relative rotation between neighbouring subunits, and the residues at the tip of the subunit translocate the bound DNA. In the "direct drive" model the dTTP binding directly deforms the conserved β-sheets, as in F1-ATPase (Abrahams et al., 1994; Itoh et al., 2004; Kinosita et al., 2004). The deformation propagates to the DNA binding region to move DNA (Liao et al., 2005). Two models are different in a way stress radiates outward from the catalytic site to effect the translocation power stroke.

A.4.4. Nucleotide hydrolysis coordination within hexameric ring

First, we address the question: are the power strokes of subunits simultaneous, random or sequential? Knowladge the order of power strokes will help to clarify how subunits communicate with each other. T7 hexameric helicase most likely utilizes the sequential mechanism (Liao et al., 2005). In this mechanism, power strokes are

carried out in a strict sequential order: (1->2->3->4->5->6). After one subunit finishes the power stroke, the adjacent subunit binds DNA and executes the next power stroke. This scheme requires cooperativity between adjacent subunits to coordinate sequential hydrolysis.

Set of structures of SV40 helicase corresponding to ATP, ADP, and Nt-free states supports a concerted (or synchronised) ATP binding and hydrolysis mechanism (Gai et al., 2004). There is no other evidence that supports this first example of all-or-none Nt binding and hydrolysis mode for other hexameric helicases or any other AAA-family molecular machine.

A.4.5. Mechanism of DNA/RNA unwinding and translocation

The detailed molecular mechanism of NA unwinding by helicases is still not clear. However, there are certain features of unwinding and translocation that are probably common to all helicases. Two popular models for the general mechanism of helicases exist: the 'active rolling' and the 'inchworm' model (Lohman and Bjornson, 1996). In both, ATP hydrolysis is required and the helicase contacts and translocates along the sugar–phosphate backbone of one strand.

Examples of the "active rolling" model could be drawn from Rep helicase. The Rep protein exists as a stable monomer in the absence of DNA. Once the Rep monomer binds to DNA, it changes its conformation and forms a homodimer, becoming functionally active in translocation and unwinding. The dimeric helicase unwinds by interacting directly with both dsDNA and ssDNA. Each subunit alternates binding to dsDNA as the dimer translocates when one subunit releases ssDNA and rebinds to dsDNA. In this model, translocation along ssDNA is coupled to ATP binding, whereas ATP hydrolysis drives the unwinding of multiple DNA base pairs per each catalytic event (Hsieh et al., 1999).

"Inchworm" model is consistent with either the monomeric or oligomeric state of the protein. The enzyme monomer first binds to ssDNA and then translocates along the DNA strand until it encounters the duplex region at the fork. Unwinding and release of one of the ssDNA strands follows this event. If the enzyme is a hexamer, such as the SV40 large T antigen, T7 gp4, then one of the subunits remains associated with the fork through the unwinding cycle (Gai et al., 2004; Singleton et al., 2000; Soultanas et al., 2000).

Bianco & Kowalczykowski established the ‘quantum inchworm’ model for translocation and unwinding of duplex DNA by RecBCD DNA helicase (Bianco and Kowalczykowski, 2000; Singleton et al., 2004). Translocation and unwinding are two separate and consecutive events in the mechanism and are brought about by two different domains (leading ‘L’ and trailing ‘T’) within the RecBCD enzyme. In this model the L domain anchors the enzyme to only one strand of duplex DNA and translocates along it, whereas the T domain is responsible for unwinding. During the translocation and unwinding reaction, the L domain binds up to 23 nt ahead of the T domain, and the T domain uses energy derived from ATP binding and hydrolysis to open the duplex.

Recently, Levin et al. proposed an alternative model in which HCV helicase translocates as a Brownian motor with a simple two-stroke cycle (Levin et al., 2005). The directional movement step is fuelled by single-stranded DNA binding energy while ATP binding allows for a brief period of random movement that prepares the helicase for the next cycle.

B. AIMS OF THE STUDY

When this work was initiated only limited information was available on the mechanism of genome packaging in double-stranded RNA bacteriophages from the *Cystoviridae* family. In particular, it was shown, that in bacteriophage phi6, packaging of single-stranded genomic precursors into preformed empty capsid (procapsid) requires a hexameric NTPase, P4. However, translocation properties and mechanism of P4 were elusive.

The aim of this study was to determine mechanism of genome packaging in dsRNA viruses from *Cystoviridae family*. Specifically:

- express, purify and characterize soluble recombinant P4 enzymes from different cystoviruses;
- identify the role of P4 through the viral life-cycle (transcription, packaging and procapsid assembly);
- assay RNA translocation activity of isolated P4 hexamers and P4 within procapsid;
- crystallize P4 proteins from different cystoviruses and obtain high-resolution X-ray structures of different complexes with nucleotides and divalent cations;
- determine the mechanism of RNA loading;
- define the role of conserved amino acid residues in P4 packaging NTPases.

C. MATERIALS AND METHODS

C.1. BACTERIAL STRAINS AND PLASMIDS

E. coli strains DH5a (Gibco-BRL) and XL1-Blue (Stratagene) were used as the hosts for the plasmid propagation and molecular cloning. Expression strains were based on *E. coli* BL21 or BL21 (DE3) (Washington et al., 1996). Phi12 P4 selenomethionyl derivative was produced in the methionine auxotroph *E. coli* B834 (DE3) as described (Mancini et al., 2004). The plasmids used in this study are listed in Table 2.

Table 2. Plasmids used in this study

Name	Comments [a)]	Reference
pLM2424	phi8 PC production	(Hoogstraten et al., 2000)
pHY1	phi8 P2 production	(Bain et al., 2001)
pT7-7	*E.coli* expression vector; T7 promoter and SD [b)]	(Tabor, 1990)
pDK5	phi8 P1 production	I
pET32b(+)	*E.coli* expression vector; T7 promoter and SD	Novagen
pSJ1b	phi8 P4 production	I
pDK3	phi13 P4 production	II
pLM2200	complete l^+ phi13; cDNA	(Qiao et al., 2000)
pJTJ7	phi6 P4 production	(Juuti et al., 1998)
pLM659	complete s^+ ; T7 promoter	(Frilander et al., 1992)
pLM1771	s^+ with deleted 11-23 nt; T7 promoter	(Qiao et al., 1997)
pLM1772	s^+ with deleted 11-32 nt; T7 promoter	(Qiao et al., 1997)
pLM1773	s^+ with deleted 11-43 nt; T7 promoter	(Qiao et al., 1997)
pEM21	for RNA1 transcription, T7 promoter	II
pGEM3Zf (+)	vector ; SP6 promoter	Promega
pPG27	phi12 P4 production	(Mancini et al., 2004)
pDK27	phi12 P4 with TTS 202-204 LKK substitution	This thesis
pDK49	phi12 P4 with deleted TTS202-204	This thesis
pDK45	phi12 P4 with T203C mutation	This thesis
pDK44	phi12 P4 with N234G mutation	VI
pDK37	phi12 P4 with K241A mutation	VI
pDK46	phi12 P4 with K241C mutation	VI
pDK47	phi12 P4 with R251A mutation	VI
pDK33	phi12 P4 with S252Q mutation	VI
pDK32	phi12 P4 with S252I mutation	VI
pDK56	phi12 P4 with S252A mutation	VI
pDK48	phi12 P4 with N253A mutation	VI
pDK35	phi12 P4 with R272A mutation	VI
pDK30	phi12 P4 with Q278A mutation	VI
pDK36	phi12 P4 with R279A mutation	VI
pDK29	phi12 P4 with Y288A mutation	VI
pDK31	phi12 P4 with S292A mutation	VI
pDK28	phi12 P4 with K310G mutation	This thesis
pDK50	phi12 P4 with deleted C-terminus (I312Stop)	This thesis
pDK21	phi8 P4 with deleted LKK184-186	V
pDK85	phi8 P4 with K185A mutation	V
pDK86	phi8 P4 with K186A mutation	V

a) pDK86 contains kanamicin resistance marker, other plasmids contain ampicillin resistance markers.

b) SD, Shine-Dalgarno ribosome-binding sequence.

C.2. EXPERIMENTAL METHODS

The methods used in this work are listed in Table 3. References to published methods and modifications are described in attached articles. New methods are described in the original publications.

Table 3. List of methods used in this study

Methods	Articles
Standard DNA and RNA techniques (Sambrook and Russell, 2001)	I, II, V, VI
Virus purification	I, IV
Virus assembly	I
Protein expression and purification	I, II, V, VI
Protein crystallization and X-ray diffraction analysis	III, VI
Thin Layer Chromatography	II, V, IV
EnzChek Phosphate assay	II, IV, VI
Gel-shift assay	II, IV, VI, V
Complementary oligonucleotide displacement assay	II, IV, VI, V
Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and Fourier transform ion cyclotron resonance mass spectrometry	I, V
Hydrogen-deuterium exchange	V
Electron microscopy	I, II
Analytical gel-filtration	I, II
Light Scattering	I, II
Circular Dichroism Spectroscopy	I
Raman Spectroscopy	II
Protein structure alignments (Holm and Sander, 1993)	II, III, VI, V
Programs: Tina, SigmaPlot, CorelDraw, RasMol and others	I, II, III, IV, V, VI

D. RESULTS AND DISCUSSION

D.1. ISOLATION AND CHARACTERISATION OF P4 PROTEINS FROM DIFFERENT CYSTOVIRUSES

P4 enzymes from three newly identified cystoviruses (phi8, phi12 and phi13) were isolated and characterised in details (I-VI). Biophysical and biochemical properties of P4 proteins are summarised in Table 4.

The average molecular weight of P4 subunit is 35 kDa. P4 proteins show low sequence identity (~25%), however all four proteins are hexameric, ssRNA-stimulated NTPases (Juuti et al., 1998; Paatero et al., 1995; I, II, IV). Phi12 P4 exhibits strict purine specificity, whereas P4 proteins from the other bacteriophages hydrolyse both purine and pyrimidine nucleotide triphosphates (II, III). P4 hydrolases from phi8 and phi13 exhibit high affinity for RNA and possess RNA-specific helicase activity (II). P4 proteins from phi6 and phi12 bind RNA with lower affinity and require association with the procapsid to perform helicase activity (II, IV). P4 proteins translocate ssRNA in the 5' to 3' direction (II). These properties and the overall hexameric morphology group P4 proteins with hexameric helicases (Delagoutte and von Hippel, 2002).

D.2. STRUCTURE OF P4 PROTEINS

To understand the mechanism of coupling of ATP hydrolysis and nucleic acid translocation in hexameric helicases, in general, and virus packaging motors, in particular, we solved the crystal structures of P4 proteins from phi12, phi13 and phi6 bacteriophages (Fig. 6; III and unpublished).

Table 4. Morphology and biochemical activities of P4 proteins *in vitro*

P4	Phi6	Phi8	Phi12	Phi13
Mol Weight (kDa)	35.0	34.1	35.1	37.6
Oligomerization state	Hexamer (+ATP/ADP)	Hexamer	Hexamer	Hexamer
ATPase activity without RNA	yes	none	yes	yes
ssRNA stimulation of ATPase	weak	strong	weak	weak
ssRNA binding	none	strong	none	strong
ssRNA translocation	none	strong	weak	strong
COD activity	none	strong	none	weak
Structure determination	2.8Å, Mol Rep	ND	1.9Å, SAD	1.7Å, SAD

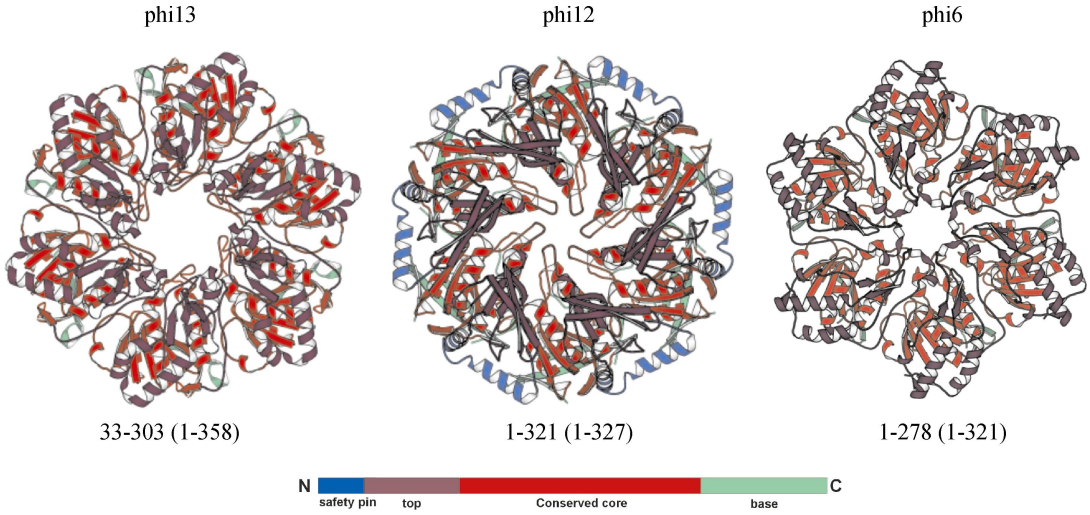


Figure 6. Hexameric packaging NTPases P4 contain a structurally conserved RecA-like catalytic core. The P4 hexamers are shown in terms of secondary structural elements in top views. The secondary structural elements are coloured according to the bar where different colors distinguish subdomains or segments of the P4 monomers: N-terminal safety pin motif (blue), all β domain (dark purple), conserved RecA-like ATP binding domain (red), and antiparallel β strands and C-terminal α-helical base (green).

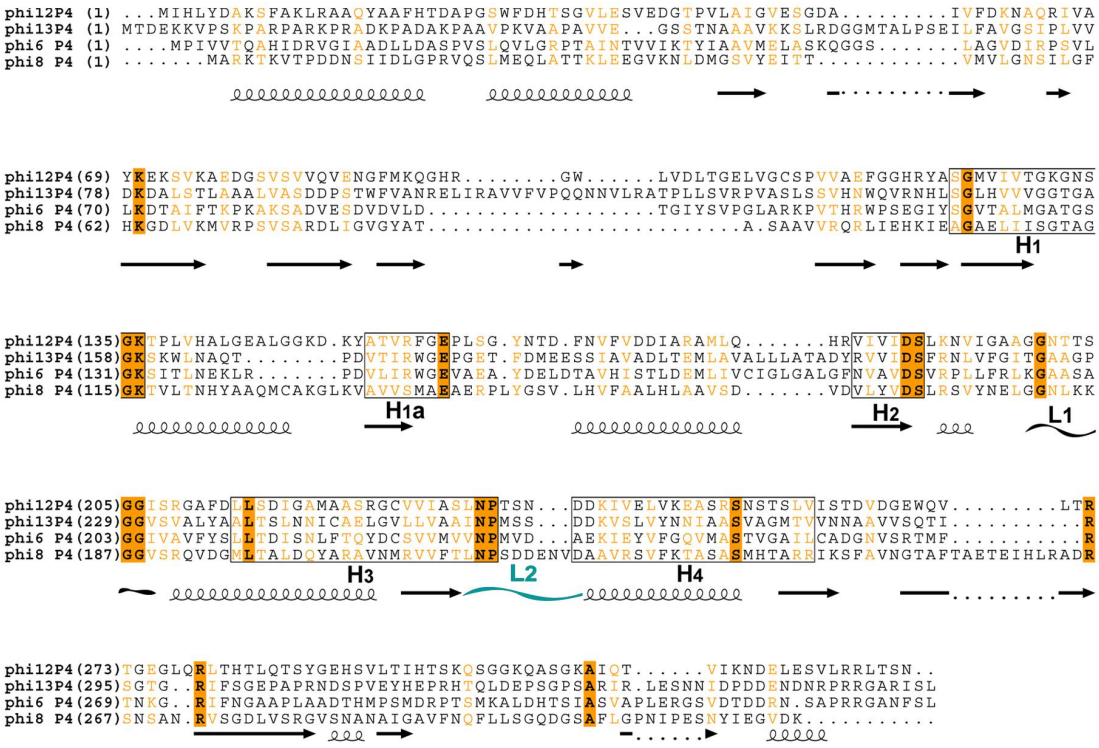


Figure 7. Sequence alignment of cystoviral P4 proteins. P4 subunits contain five motifs (orange filled boxes) that are conserved among F4 helicase family, according to (Hall and Matson, 1999). The positions of loops L1 (disordered) and L2 (ordered) are indicated by black twirls. Secondary structure elements of phi12 P4 are shown below the aligned sequences.

Table 5. RMS deviations, Z scores and structural alignment statistics for P4 and structurally related proteins as obtained from DALI search (Holm and Sander, 1993; Mancini and Tuma, unpublished).

Helicase/ PDB	Z score	RMSD (Å)	N of aligned residues	Sequence Identity %
T7 helicase/ 1CR1	12.2	3.7	175	15
RecA/ 2REB	10.2	4.4	169	12
RepA/ 1G8Y	9.3	3.7	168	17
F1-ATPase (α)/1SKYB	8.5	3.8	172	14
F1-ATPase (β)/1SKYE	8.1	4.2	176	12
TrwB/ 1G8Y	7.3	3.9	155	18
NSF/ 1D2N	7.1	2.9	119	14

Subunits of P4 hexamers contain the conserved RecA-like catalytic core, which is structurally equivalent to the ATPase domain in RecA, T7 gp4, RepA and other RecA-like NTPases (Table 5 and Fig. 6; Wang, 2004; Ye et al., 2004). The RecA-like core of P4 proteins contains the conserved motifs (H1, H1a, H2, H3, H4) of F4 helicase family which play role in transducing the chemical energy of NTP hydrolysis to mechanical work (Fig. 7; Hall and Matson, 1999, IV and VI).

D.3. NUCLEOTIDE BINDING AND SPECIFICITY

We have determined high-resolution structures of P4 phi12 hexamers in distinct nucleotide binding states, including apo, substrate analog bound and product bound (IV). In substrate analog and product bound states six identical nucleotides occupy six nucleotide-binding pockets. Nucleotide-binding pockets of P4 lie at the interface of adjacent subunits (Fig. 8A and B; IV) essentially on the outer surface, some 25 Å from the inner channel. Biochemical and structural analysis revealed that residues in the binding pocket have roles: (1) in catalysis, (2) in stabilization of substrate binding, and (3) as γ-phosphate sensors (VI).

Phi12 P4 is the only purine specific hydrolase among P4 proteins (Fig. 8C). The nucleotide base is sandwiched between Tyr288 from the catalytic subunit (i.e. the one providing residues of the Walker A and B motifs) and Gln278 from the neighboring subunit. Although these residues do not belong to the conserved helicase motifs, similar stacking interactions were found in RepA helicase (Arg85 and Tyr242) (Ziegelin, 2003). In addition, in phi12 P4 a hydrogen bond is observed between Ser292 and N7 of the adenine ring. These three residues are determinants of base specificity. Comparison of P4 structures with bound UTP and AMPcPP revealed that purine specificity is achieved by a combination of base-specific stacking and hydrogen bonding to the N7 site of the base (Fig. 8D).

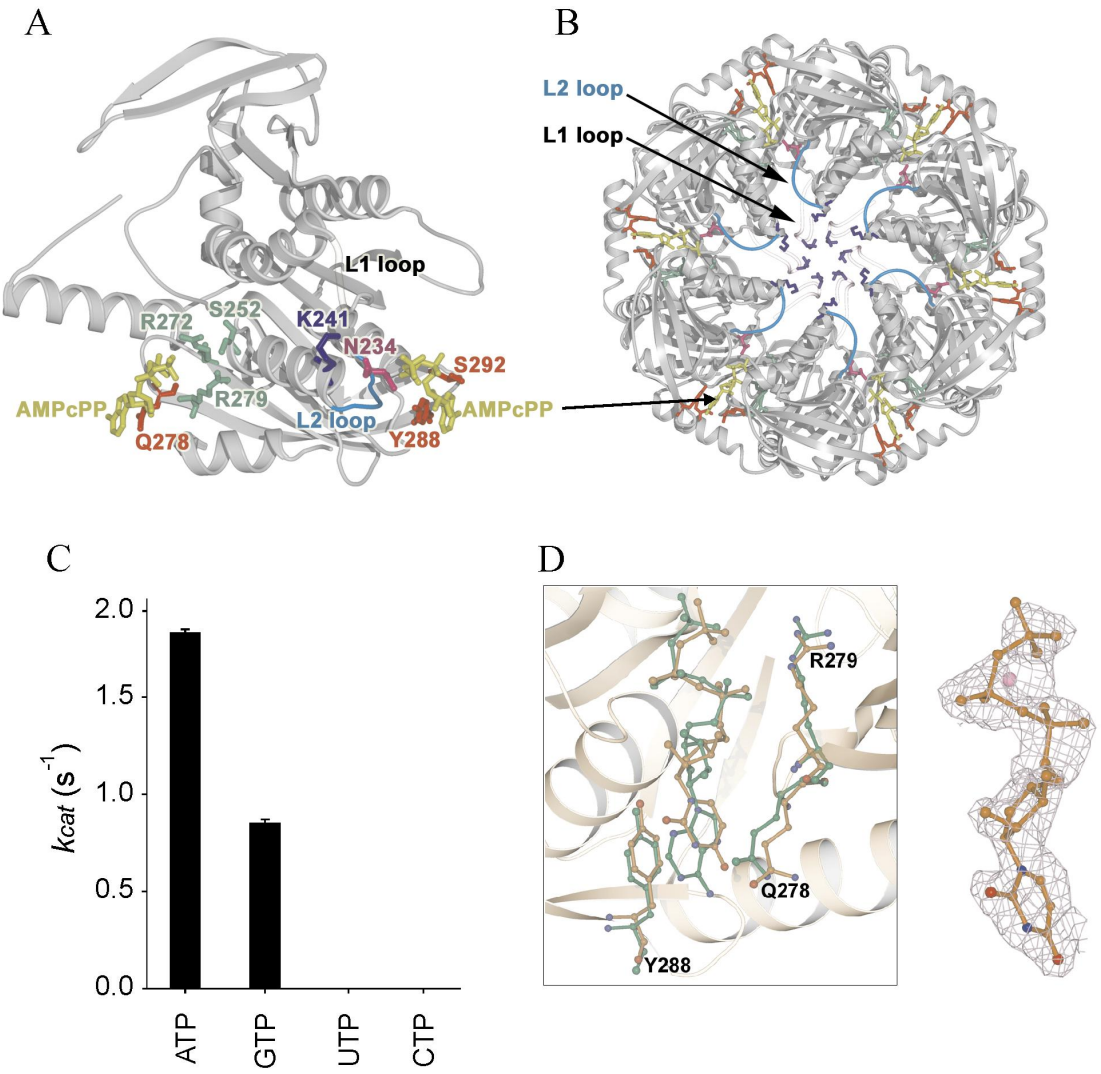


Figure 8. Nucleotide binding and specificity of P4 phi12. (A) P4 monomer and (B) P4 hexamer viewed from the bottom to expose the nucleotide binding cleft between the subunits. Essential residues are shown as ball-and-stick models and colored according to their functions (red: base specific binding, green: insertion of arginine fingers, blue: RNA binding, magenta: coupling of movement to hydrolysis). The bound nucleotide triphosphates are shown in yellow. (C) Hydrolysis of different nucleotides under optimal conditions. P4 NTPs turnover (k_{cat}) was measured using steady-state kinetics of P_i release. (D) Comparison the P4 nucleotide binding cleft with bound UTP (orange) and bound AMPcPP (green). Electron density map around the modeled UTP is also shown.

The correct coordination of the base results in the precise alignment of the nucleotide inside the cleft that is in turn essential for catalysis, whereas UTP can bind but is not hydrolysed, due to misalignment.

D.4. MECHANISM OF COUPLING OF ATP HYDROLYSIS TO RNA TRANSLOCATION

Insight into the P4 mechanism was obtained based on P4 phi12 structures in key catalytic states (IV). The nucleotide binding sites, which are located at the

perimeter of the hexamer, are connected to the central channel via helix α6 (motif H4). Two loops (L1 and L2) protrude into the central channel. These loops are equivalent of the RecA L1 and L2 loops. It was shown by site-directed mutagenesis that both loops in P4 proteins are essential for RNA binding and translocation (V, VI). Thus, RNA was proposed to pass through the central channel (II, VI, V). ATP hydrolysis is associated with switching of the P-loop (conserved motif H1) from a relaxed to a strained conformation. This is accompanied by a swivelling of helix α6 from an "up" position to a "down" position in the product complex (where the translocation channel runs vertically with the particle interior downwards) (Fig. 9A and B).

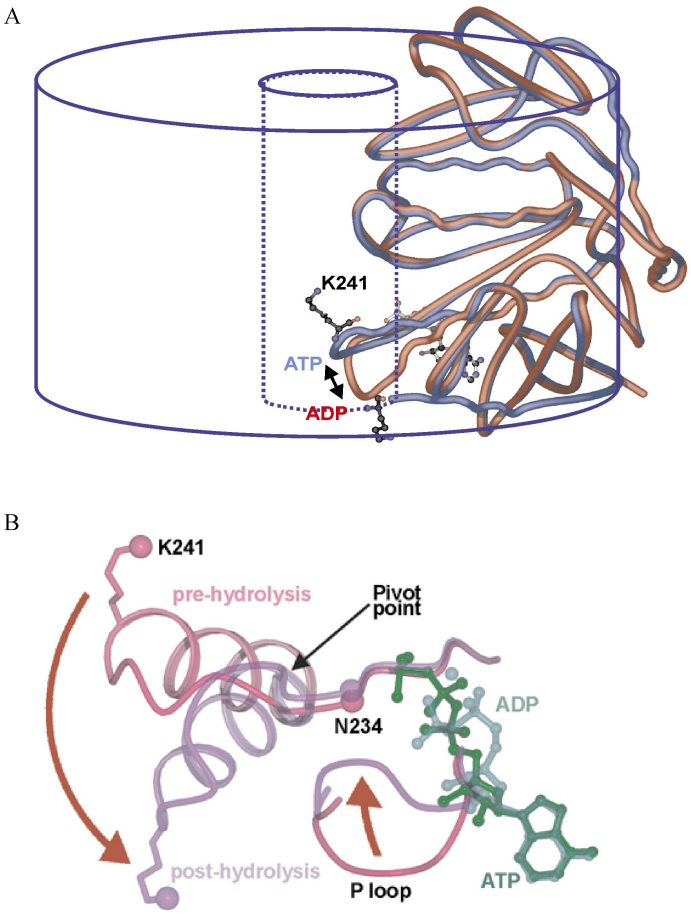


Figure 9. Conformational changes due to nucleotide hydrolysis. The structures of AMPcPP-Mg^{2+} and ADP-Mg^{2+} bound P4 are compared. (A) Ribbon diagrams represent AMPcPP-Mg^{2+} bound P4 monomer (blue) superimposed to the ADP-Mg^{2+} conformation (red) of the same subunit within hexameric P4. The AMPcPP molecule and Lys241 are shown as ball-and-stick models. Black arrow indicates the conformational changes which the α6 helix-L2 loop undergo following hydrolysis of bound ATP. (B) Details of the nucleotide binding domain and corresponding P loop and L2 loop in the AMPcPP-Mg^{2+} conformation superimposed on the ADP-Mg^{2+} conformations. The red arrows indicate the conformational changes which the α6 helix-L2 loop and P loop undergo following hydrolysis of bound ATP.

This movement effects a ~6 Å translocation of the RNA bound to Lys241 and constitutes the mechanical reaction coordinate. Several residues that are conserved among hexameric helicases control motion of L2 loop and helix α6 (H3 and H4 motifs; VI). The proposed mechanism was further substantiated for P4 protein from a related virus phi8 by solution hydrogen-deuterium exchange (this method visualized changes in dynamics associated with ATP hydrolysis and RNA translocation; V). We observed that the concerted changes affect a cooperative unit that encompasses part of the catalytic RecA-like core of the subunit.

D.5. COORDINATION OF THE CATALYSIS BETWEEN SUBUNITS

Enzymatic studies support a novel stochastic-sequential cooperativity model in which P4 without RNA utilizes stochastic mode of catalysis, but with RNA - sequential mode (Lisal and Tuma, 2005). The proposed mechanism requires binding of three nucleotides in a row for hydrolysis to proceed (Lisal and Tuma, 2005). We demonstrated the structural basis for this requirement. Mutations affecting interactions with the nucleotide base, such as Ser292, cause opening of the binding cleft that induces a cooperative distortion of the hexameric ring. Conversely, closing of the cleft upon nucleotide binding may be relayed between the neighboring subunits, facilitating formation of a hydrolysis competent state. However, this mechanism does not account for RNA-induced cooperativity.

Stimulation of ATPase activity by ssRNA arises from an increase in sequential hydrolysis (Lisal and Tuma, 2005). The shortest ssRNA oligonucleotide to stimulate hydrolysis was 5 nucleotides long (Lisal and Tuma, 2005). Assuming a typical ssRNA backbone configuration, this oligonucleotide would span two or three neighboring subunits. Thus, RNA binding to Lys241 residues on neighboring subunits may promote formation of the transition state at subunit i+1 by coordinating the motion of its α6 helix with that of the preceding subunit i (Fig. 10B). In turn, the rate of hydrolysis would increase at high ATP concentrations, when subunit i+1 is likely to contain ATP. Conversely, at low ATP concentrations, the sequential chain of hydrolysis will be broken and the enzyme would have to reinitiate, decreasing the overall rate of hydrolysis as observed (Lisal et al., 2004). Thus, RNA effectively acts as a stress loop and relays energy released from ATP hydrolysis between neighboring subunits.

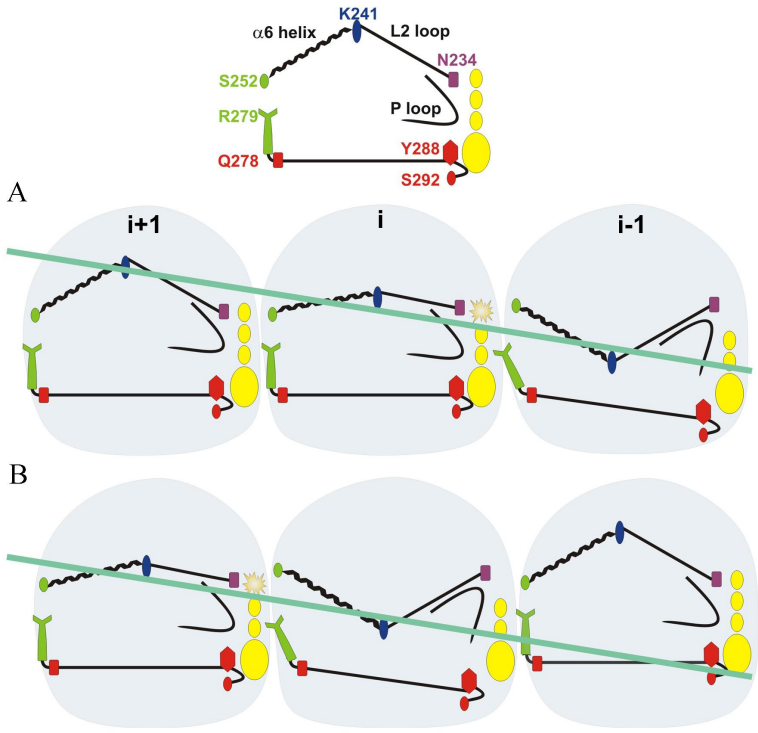


Figure 10. Schematic description of the sequential coordination of hydrolysis. Top panel shows a schematic representation of the conserved motifs in the context of one P4 subunit together with a bound ATP (yellow). (A) After the hydrolysis on subunit i-1 the down movement of helix α6 triggers hydrolysis at subunit i. The L2 loop movement also drags down the bound RNA (cyan). (B) The transition state is reached at subunit i+1 while subunit i-1 exchanges the ADP for ATP, which is sensed by Asn234, concomitantly stabilizing the subunit i L2 loop and helix α6 in the initial up configuration. The subunit i-1 arginine finger R279 disengages from the neighboring active site.

RNA binding to P4 is weak and detachment from Lys241 is likely to result in stochastic fluctuations of the α6 helix to the "up" position, which is then stabilized by ATP binding. Hydrolysis at subunit i would also bring the incoming RNA into vicinity of the L2 loop of the subunit i+1 leading to attachment, at which point the cycle repeats.

Arginine fingers are another structurally conserved feature of hexameric helicases which mediate cooperativity (Crampton et al., 2004). We have demonstrated that the insertion of the P4 arginine finger (Arg279) is dependent on the conformation of Ser252 (VI). Taking in consideration the role of the arginine finger, the sequential ATP hydrolysis and mechano-chemical coupling mechanism can be presented as following. In the first round, stochastic motion of L2 loop at i-1 subunit inserts Arg279 into the active site on subunit i and stabilizes the transition state. In the second round, hydrolysis at subunit i stabilizes the L2 loop in the down position and the hydrolysis cycle is repeated at subunit i+1. Subunit i-1 can return to the initial state upon ATP binding which stabilizes the L2 loop in the up position.

D.6. RNA LOADING MECHANISM

RNA passes through the central channel of the hexamer (II, VI). However, the mechanism of RNA loading remained elusive. To address this issue conformational dynamics of the viral packaging motor P4 phi8 was visualised by hydrogen-deuterium exchange and high-resolution mass spectrometry (V). Deuterium labelling revealed a transition state associated with RNA loading, which proceeds via opening of the hexameric ring. The loading mechanism is similar to that of other hexameric helicases but does not require ATP binding or hydrolysis (Ahnert et al., 2000; Richardson, 2003; Skordalakes and Berger, 2003).

D.7. REGULATION OF P4 ACTIVITY WITHIN THE VIRAL CORE

P4 hexamer is associated with the polymerase complex via its C-terminal face (Paatero et al., 1998). The P4 ATPase activity is required only during packaging (Gottlieb et al., 1992). P4 acts as a passive pore during transcription and its ATPase activity is down regulated (IV). Detailed mapping of sites and nature of P4:PC interactions provides structural basis for the regulation of RNA translocation activity within the PC (Fig. 11).

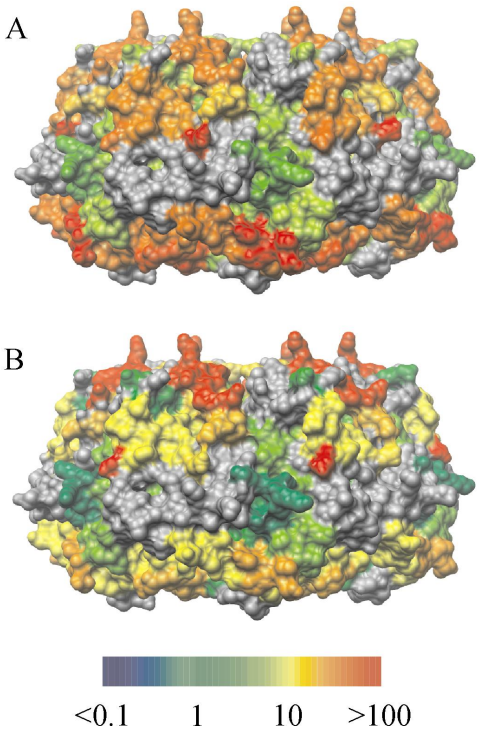


Figure 11. Surface representations of ϕ12 P4 hexamer colored according to the weighted average HDX rates computed from rate distributions. (A) The HDX rate for isolated hexamer. (B) HDX rates for P4 in the procapsid. The color bar in bottom is the color scale for rates (h^{-1}); segments for which no peptic peptides were available are shown in gray. The C-terminal base of P4 molecule is at the bottom of each panel whereas the N-terminal apical domain points up.

The C-terminus appears to be dynamically coupled to the RecA-like catalytic core. C-terminal region (295-310 aa) acts as a brake in the isolated hexamer. In the procapsid this helix may be either withdrawn from the catalytic core to activate the ATPase or "jammed" onto the core to shut off the activity, respectively. One can envision that a similar mechanism may apply to hexameric helicases within the replication and repair machinery (von Hippel and Delagoutte, 2001; von Hippel and Delagoutte, 2003).

D.8. IMPLICATIONS FOR THE VIRAL LIFE CYCLE

Genome packaging

P4 hexamers were visualized by Cryo-EM to reside at the five-fold vertices of the viral procapsid (de Haas et al., 1999). The 5'-end of the viral genome precursor ssRNA (containing the recognition *pac* sequence) is specifically recognised by the major capsid protein P1 and consequently no free 5' end is available for a direct threading through P4 (Pirttimaa and Bamford, 2000). However, the specific binding would bring RNA to the vicinity of P4 and trigger ring opening followed by RNA loading. After loading, the presence of RNA in the central channel would activate the P4 motor resulting in a processive RNA translocation into the virion. When all three segments are packaged, procapsid expands and at least in the case of phi12 virus turns off the ATPase activity of P4 during transcription via interaction of the C-terminal helix with the core (Fig. 12).

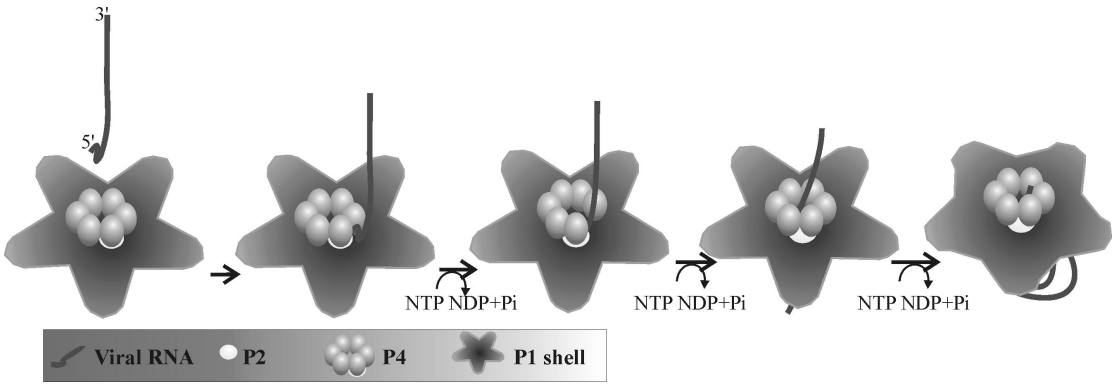


Figure 12. Model of RNA (red) loading into the packaging motor (green) in the context of the viral shell (one five-fold vertex). From left to right: the packaging signal at 5' end of RNA is specifically recognized by capsid protein P1. RNA then binds to a putative primary binding site on P4 surface and ring opening commences. P4 processively translocates captured RNA into the capsid.

Semi-conservative transcription

Switching between the packaging and the transcription mode requires the translocation direction of the P4 motor to reverse. However, our results indicate that while P4 actively translocates RNA during packaging it acts as a passive conduit for RNA export. The directionality switching is accomplished via the regulation of P4 NTPase activity within the polymerase core (III).

Assembly of viral procapsid

Apart from RNA transport P4 hexamer also plays an important role in PC assembly (Poranen and Tuma, 2004). The formation of phi6 PC-like structures *in vitro* is absolutely dependent on P4 (in addition to P1), suggesting that the P4 hexamer is critical for the nucleation of phi6 PC assembly (Poranen et al., 2001). However, P4 phi8 is not essential for phi8 PC nucleation (I). It significantly stabilises assembly intermediates and prevents the formation of aberrant structures, suggesting that phi8 P4 also plays role during PC assembly.

Finally, P4 together with P2 may play role in RNA recombination in the *Cystoviridae* by utilizing putative RNA annealing activity (opposite to helicase activity), creating a pool of new viruses (Mindich, 1999).

E. CONCLUDING REMARKS

Packaging the genome into an empty capsid is central to the assembly of many complex viruses. Hexameric helicase P4 of the *Cystoviridae* family of dsRNA bacteriophages is one of the simplest packaging motors found in nature. The current work focuses on biochemical, biophysical, and structural characterization of P4 proteins from four cystoviruses. It provides detailed explanation of hexameric packaging motor mechanism in atomic details. We found that P4 proteins bear mechanistic and structural similarities to a variety of the pervasive RecA/F1-ATPase like motors involved in diverse biological functions, suggesting that these motors employ common mechanistic principles. The study also highlights the role of P4 in assembly, transcription and replication of dsRNA bacteriophages.

F. ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Roman Tuma for his support allowing me to do my thesis work in the lab of Prof. Dennis Bamford at the Institute of Biotechnology, University of Helsinki in Helsinki, Finland. I would like to thank Prof. Dennis Bamford for giving me the great opportunity to work in his team. The countless helpful discussions and his support were inspiring throughout this work and very much appreciated. Dr. Eugene Makeyev has been a great supervisor for the first half of this work. I would like to thank him for his support and help, without which this work would not have been possible.

I would like to thank the members of the Bamford lab that provided not only a professional scientific but also a personal atmosphere that made working in the lab a pleasure. I'd like to especially thank Sarah Butcher, Jaana Bamford, Vladimir Simonov, Markus Pirttimaa, Jiri Lisal, Pasi Laurinmäki, Riitta Tarkiainen, Sampo Vehma, Risto Tetri, Elina Roine, Petri Papponen, and The Lithuanian Clan (Rimantas Daugelavicius, Ausra Gaidelyte, Gabija Ziedaite). My deepest gratitude is to my collaborators, Drs. Erika Mancini, Jonathan Grimes and Prof. David Stuart from the Oxford University, UK, Drs. TuKiet Lam, Mark Emmett and Prof. Alan Marshall from the Florida State University, USA, Drs. Gulija and Marat Yusupov from IGBMC, France, and Michael Merckel from the Institute of Biotechnology, Finland.

I would like to thank Prof. Mart Saarma, Director of the Institute of Biotechnology for his interest in my work. I am also grateful to Prof. Hannu Saarilahti, the former Head of the Division of Genetics, and Dr. Nina Saris for their help in my academic progress. My thanks are due to the Graduate School in Informational and Structural Biology team: Mrs Kaija Söderlund and Prof. Mark Johnson; Helsinki Graduate School in Biotechnology and Molecular Biology team: Prof. Heikki Rauvala, Drs Erkki Raulo and Anita Tienhaara. I wish to thank Eija Hagfors and Lena Rautavara from CIMO for their kind support.

I thank the reviewers Dr. Alexander Plyusnin and Dr Mikko Frilander, for their critical reading of the manuscript and constructive comments.

I deeply thank my friends Andrey Golubtsov, Natalya and Eugene Kulesski, Tatiana and Sergey Shiriaev, Maxim Bespalov, Maxim Smirnyagin, Konstantin Vagin, Vladimir Tomilov, Alexander Huzin, Michael Fomin, Lidia and Alexander Borkunov, Viltare and my other friends for the great time we had together.

My most sincere gratitude belongs to my parents, Lyubov' and Eugene Kainov, who have loved and supported me throughout my entire life. My sister Tatiana Kainova also gets a big share of my compliments.

This thesis is devoted to my wife Asta for her love and friendship. All the best in me is because of her, my parents and friends. I thank Asta's parents Jurate Zukliene and Alfonsas Zuklys, sister Justina Zuklyte and "brother" Vaidis Skripkauskas for their beautiful gift.

Financial support from the Graduate School in Informational and Structural Biology and CIMO was greatly appreciated.

Denis Kainov

Helsinki, August 2005

G. REFERENCES

Abbate, E.A., Berger, J.M. and Botchan, M.R. (2004) The X-ray structure of the papillomavirus helicase in complex with its molecular matchmaker E2. *Genes Dev*, **18**, 1981-1996.

Ago, H., Adachi, T., Yoshida, A., Yamamoto, M., Habuka, N., Yatsunami, K. and Miyano, M. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Structure Fold Des*, **7**, 1417-1426.

Ahnert, P., Picha, K.M. and Patel, S.S. (2000) A ring-opening mechanism for DNA binding in the central channel of the T7 helicase-primase protein. *EMBO J*, **19**, 3418-3427.

Allison, T.J., Wood, T.C., Briercheck, D.M., Rastinejad, F., Richardson, J.P. and Rule, G.S. (1998) Crystal structure of the RNA-binding domain from transcription termination factor rho. *Nat Struct Biol*, **5**, 352-356.

Bain, D.L., Berton, N., Ortega, M., Baran, J., Yang, Q. and Catalano, C.E. (2001) Biophysical characterization of the DNA binding domain of gpNu1, a viral DNA packaging protein. *J Biol Chem*, **276**, 20175-20181.

Balakrishnan, M., Fay, P.J. and Bambara, R.A. (2001) The kissing hairpin sequence promotes recombination within the HIV-I 5' leader region. *J Biol Chem*, **276**, 36482-36492.

Balakrishnan, M., Roques, B.P., Fay, P.J. and Bambara, R.A. (2003) Template dimerization promotes an acceptor invasion-induced transfer mechanism during human immunodeficiency virus type 1 minus-strand synthesis. *J Virol*, **77**, 4710-4721.

Bamford, D.H. (2000) Virus structures: Those magnificent molecular machines. *Curr Biol*, **10**, R558-561.

Bamford, D.H. and Palva, E.T. (1980) Structure of the lipid-containing bacteriophage phi 6. Disruption by Triton X-100 treatment. *Biochim Biophys Acta*, **601**, 245-259.

Bamford, D.H., Palva, E.T. and Lounatmaa, K. (1976) Ultrastructure and life cycle of the lipid-containing bacteriophage phi 6. *J Gen Virol*, **32**, 249-259.

Barcena, M., Martin, C.S., Weise, F., Ayora, S., Alonso, J.C. and Carazo, J.M. (1998) Polymorphic quaternary organization of the Bacillus subtilis bacteriophage SPP1 replicative helicase (G40 P). *J Mol Biol*, **283**, 809-819.

Beckett, D., Wu, H.N. and Uhlenbeck, O.C. (1988) Roles of operator and non-operator RNA sequences in bacteriophage R17 capsid assembly. *J Mol Biol*, **204**, 939-947.

Bennett, R.J. and Keck, J.L. (2004) Structure and function of RecQ DNA helicases. *Crit Rev Biochem Mol Biol*, **39**, 79-97.

Berkhout, B. and van Wamel, J.L. (1996) Role of the DIS hairpin in replication of human immunodeficiency virus type 1. *J Virol*, **70**, 6723-6732.

Berkowitz, R., Fisher, J. and Goff, S.P. (1996) RNA packaging. *Curr Top Microbiol Immunol*, **214**, 177-218.

Bernstein, D.A., Zittel, M.C. and Keck, J.L. (2003) High-resolution structure of the E.coli RecQ helicase catalytic core. *EMBO J*, **22**, 4910-4921.

Bianco, P.R. and Kowalczykowski, S.C. (2000) Translocation step size and mechanism of the RecBC DNA helicase. *Nature*, **405**, 368-372.

Bienstock, R.J., Skorvaga, M., Mandavilli, B.S. and Van Houten, B. (2003) Structural and functional characterization of the human DNA repair helicase XPD by comparative molecular modeling and site-directed mutagenesis of the bacterial repair protein UvrB. *J Biol Chem*, **278**, 5309-5316.

Bogden, C.E., Fass, D., Bergman, N., Nichols, M.D. and Berger, J.M. (1999) The structural basis for terminator recognition by the Rho transcription termination factor. *Mol Cell*, **3**, 487-493.

Bressanelli, S., Tomei, L., Rey, F.A. and De Francesco, R. (2002) Structural analysis of the hepatitis C virus RNA polymerase in complex with ribonucleotides. *J Virol*, **76**, 3482-3492.

Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R.L., Mathieu, M., De Francesco, R. and Rey, F.A. (1999) Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proc Natl Acad Sci U S A*, **96**, 13034-13039.

Buck, K.W. (1999) Replication of tobacco mosaic virus RNA. *Philos Trans R Soc Lond B Biol Sci*, **354**, 613-627.

Butcher, S.J., Dokland, T., Ojala, P.M., Bamford, D.H. and Fuller, S.D. (1997) Intermediates in the assembly pathway of the double-stranded RNA virus phi6. *EMBO J*, **16**, 4477-4487.

Butcher, S.J., Grimes, J.M., Makeyev, E.V., Bamford, D.H. and Stuart, D.I. (2001) A mechanism for initiating RNA-dependent RNA polymerization. *Nature*, **410**, 235-240.

Caruthers, J.M., Johnson, E.R. and McKay, D.B. (2000) Crystal structure of yeast initiation factor 4A, a DEAD-box RNA helicase. *Proc Natl Acad Sci U S A*, **97**, 13080-13085.

Caruthers, J.M. and McKay, D.B. (2002) Helicase structure and mechanism. *Curr Opin Struct Biol*, **12**, 123-133.

Catalano, C.E. (2000) The terminase enzyme from bacteriophage lambda: a DNA-packaging machine. *Cell Mol Life Sci*, **57**, 128-148.

Catalano, C.E., Cue, D. and Feiss, M. (1995) Virus DNA packaging: the strategy used by phage lambda. *Mol Microbiol*, **16**, 1075-1086.

Clever, J.L. and Parslow, T.G. (1997) Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. *J Virol*, **71**, 3407-3414.

Coulibaly, F., Chevalier, C., Gutsche, I., Pous, J., Navaza, J., Bressanelli, S., Delmas, B. and Rey, F.A. (2005) The birnavirus crystal structure reveals structural relationships among icosahedral viruses. *Cell*, **120**, 761-772.

Crampton, D.J., Guo, S., Johnson, D.E. and Richardson, C.C. (2004) The arginine finger of bacteriophage T7 gene 4 helicase: role in energy coupling. *Proc Natl Acad Sci U S A*, **101**, 4373-4378.

D'Souza, V., Dey, A., Habib, D. and Summers, M.F. (2004) NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus. *J Mol Biol*, **337**, 427-442.

D'Souza, V., Melamed, J., Habib, D., Pullen, K., Wallace, K. and Summers, M.F. (2001) Identification of a high affinity nucleocapsid protein binding element within the Moloney murine leukemia virus Psi-RNA packaging signal: implications for genome recognition. *J Mol Biol*, **314**, 217-232.

D'Souza, V. and Summers, M.F. (2004) Structural basis for packaging the dimeric genome of Moloney murine leukaemia virus. *Nature*, **431**, 586-590.

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 phi6. *J Mol Biol*, **294**, 357-372.

Delagoutte, E. and von Hippel, P.H. (2002) Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Q Rev Biophys*, **35**, 431-478.

Egelman, E.H. (1998) Bacterial helicases. *J Struct Biol*, **124**, 123-128.

Emori, Y., Iba, H. and Okada, Y. (1980) Assignment of viral proteins to the three double-stranded RNA segments of bacteriophage phi 6 genome: translation of phi 6 messenger RNAs transcribed in vitro. *Mol Gen Genet*, **180**, 385-389.

Emori, Y., Iba, H. and Okada, Y. (1982) Morphogenetic pathway of bacteriophage phi 6. A flow analysis of subviral and viral particles in infected cells. *J Mol Biol*, **154**, 287-310.

Enemark, E.J., Chen, G., Vaughn, D.E., Stenlund, A. and Joshua-Tor, L. (2000) Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Mol Cell*, **6**, 149-158.

Enemark, E.J., Stenlund, A. and Joshua-Tor, L. (2002) Crystal structures of two intermediates in the assembly of the papillomavirus replication initiation complex. *EMBO J*, **21**, 1487-1496.

Ewen, M.E. and Revel, H.R. (1988) In vitro replication and transcription of the segmented double-stranded RNA bacteriophage phi 6. *Virology*, **165**, 489-498.

Ewen, M.E. and Revel, H.R. (1990) RNA-protein complexes responsible for replication and transcription of the double-stranded RNA bacteriophage phi 6. *Virology*, **178**, 509-519.

Falson, P., Goffeau, A., Boutry, M. and Jault, J.M. (2004) Structural insight into the cooperativity between catalytic and noncatalytic sites of F1-ATPase. *Biochim Biophys Acta*, **1658**, 133-140.

Fass, D., Bogden, C.E. and Berger, J.M. (1999) Crystal structure of the N-terminal domain of the DnaB hexameric helicase. *Structure Fold Des*, **7**, 691-698.

Feng, Y.X., Fu, W., Winter, A.J., Levin, J.G. and Rein, A. (1995) Multiple regions of Harvey sarcoma virus RNA can dimerize in vitro. *J Virol*, **69**, 2486-2490.

Fletcher, R.J., Bishop, B.E., Leon, R.P., Sclafani, R.A., Ogata, C.M. and Chen, X.S. (2003) The structure and function of MCM from archaeal M. Thermoautotrophicum. *Nat Struct Biol*, **10**, 160-167.

Frilander, M. and Bamford, D.H. (1995) In vitro packaging of the single-stranded RNA genomic precursors of the segmented double-stranded RNA bacteriophage phi 6: the three segments modulate each other's packaging efficiency. *J Mol Biol*, **246**, 418-428.

Frilander, M., Gottlieb, P., Strassman, J., Bamford, D.H. and Mindich, L. (1992) Dependence of minus-strand synthesis on complete genomic packaging in the double-stranded RNA bacteriophage phi 6. *J Virol*, **66**, 5013-5017.

Fritsch, C., Stussi, C., Witz, J. and Hirth, L. (1973) Specificity of TMV RNA encapsidation: in vitro coating of heterologous RNA by TMV protein. *Virology*, **56**, 33-45.

Fu, W. and Rein, A. (1993) Maturation of dimeric viral RNA of Moloney murine leukemia virus. *J Virol*, **67**, 5443-5449.

Fujimura, T. and Esteban, R. (2000) Recognition of RNA encapsidation signal by the yeast L-A double-stranded RNA virus. *J Biol Chem*, **275**, 37118-37126.

Fujimura, T., Esteban, R., Esteban, L.M. and Wickner, R.B. (1990) Portable encapsidation signal of the L-A double-stranded RNA virus of S. cerevisiae. *Cell*, **62**, 819-828.

Fujimura, T., Ribas, J.C., Makhov, A.M. and Wickner, R.B. (1992) Pol of gag-pol fusion protein required for encapsidation of viral RNA of yeast L-A virus. *Nature*, **359**, 746-749.

Gai, D., Zhao, R., Li, D., Finkielstein, C.V. and Chen, X.S. (2004) Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell*, **119**, 47-60.

Goetzinger, K.R. and Rao, V.B. (2003) Defining the ATPase center of bacteriophage T4 DNA packaging machine: requirement for a catalytic glutamate residue in the large terminase protein gp17. *J Mol Biol*, **331**, 139-154.

Gorbalenya, A.E. and Koonin, E.V. (1993) Helicases - Amino-Acid-Sequence Comparisons and Structure-Function-Relationships. *Cur Opin Struct Biol*, **3**, 419-429.

Gorbalenya, A.E., Koonin, E.V., Donchenko, A.P. and Blinov, V.M. (1989) Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Res*, **17**, 4713-4730.

Gottlieb, P., Potgieter, C., Wei, H. and Toporovsky, I. (2002) Characterization of phi12, a bacteriophage related to phi6: nucleotide sequence of the large double-stranded RNA. *Virology*, **295**, 266-271.

Gottlieb, P., Qiao, X., Strassman, J., Frilander, M. and Mindich, L. (1994) Identification of the packaging regions within the genomic RNA segments of bacteriophage phi 6. *Virology*, **200**, 42-47.

Gottlieb, P., Strassman, J., Frucht, A., Qiao, X.Y. and Mindich, L. (1991) In vitro packaging of the bacteriophage phi 6 ssRNA genomic precursors. *Virology*, **181**, 589-594.

Gottlieb, P., Strassman, J. and Mindich, L. (1992) Protein P4 of the bacteriophage phi 6 procapsid has a nucleoside triphosphate-binding site with associated nucleoside triphosphate phosphohydrolase activity. *J Virol*, **66**, 6220-6222.

Gottlieb, P., Strassman, J., Qiao, X.Y., Frucht, A. and Mindich, L. (1990) In vitro replication, packaging, and transcription of the segmented double-stranded RNA genome of bacteriophage phi 6: studies with procapsids assembled from plasmid-encoded proteins. *J Bacteriol*, **172**, 5774-5782.

Grahn, E., Moss, T., Helgstrand, C., Fridborg, K., Sundaram, M., Tars, K., Lago, H., Stonehouse, N.J., Davis, D.R., Stockley, P.G. and Liljas, L. (2001) Structural basis of pyrimidine specificity in the MS2 RNA hairpin-coat-protein complex. *RNA*, **7**, 1616-1627.

Grahn, E., Stonehouse, N.J., Murray, J.B., van den Worm, S., Valegard, K., Fridborg, K., Stockley, P.G. and Liljas, L. (1999) Crystallographic studies of RNA hairpins in complexes with recombinant MS2 capsids: implications for binding requirements. *RNA*, **5**, 131-138.

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, S., Jardine, P.J. and Anderson, D. (2002) Bacteriophage phi 29 DNA packaging. *Adv Virus Res*, **58**, 255-294.

Guo, P. (2002) Structure and function of phi29 hexameric RNA that drives the viral DNA packaging motor: review. *Prog Nucleic Acid Res Mol Biol*, **72**, 415-472.

Guo, P. (2005) Bacterial virus phi29 DNA-packaging motor and its potential applications in gene therapy and nanotechnology. *Methods Mol Biol*, **300**, 285-324.

Haddrick, M., Lear, A.L., Cann, A.J. and Heaphy, S. (1996) Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. *J Mol Biol*, **259**, 58-68.

Hall, M.C. and Matson, S.W. (1999) Helicase motifs: the engine that powers DNA unwinding. *Mol Microbiol*, **34**, 867-877.

Helgstrand, C., Grahn, E., Moss, T., Stonehouse, N.J., Tars, K., Stockley, P.G. and Liljas, L. (2002) Investigating the structural basis of purine specificity in the structures of MS2 coat protein RNA translational operator hairpins. *Nucleic Acids Res*, **30**, 2678-2685.

Hendrix, R.W. (1978) Symmetry mismatch and DNA packaging in large bacteriophages. *Proc Natl Acad Sci U S A*, **75**, 4779-4783.

Higgins, N.P., Kato, K. and Strauss, B. (1976) A model for replication repair in mammalian cells. *J Mol Biol*, **101**, 417-425.

Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol*, **233**, 123-138.

Hoogstraten, D., Qiao, X., Sun, Y., Hu, A., Onodera, S. and Mindich, L. (2000) Characterization of phi8, a bacteriophage containing three double-stranded RNA genomic segments and distantly related to Phi6. *Virology*, **272**, 218-224.

Horn, W.T., Convery, M.A., Stonehouse, N.J., Adams, C.J., Liljas, L., Phillips, S.E. and Stockley, P.G. (2004) The crystal structure of a high affinity RNA stem-loop complexed with the bacteriophage MS2 capsid: further challenges in the modeling of ligand-RNA interactions. *RNA*, **10**, 1776-1782.

Hsieh, J., Moore, K.J. and Lohman, T.M. (1999) A two-site kinetic mechanism for ATP binding and hydrolysis by E. coli Rep helicase dimer bound to a single-stranded oligodeoxynucleotide. *J Mol Biol*, **288**, 255-274.

Hu, W.S. and Temin, H.M. (1990) Retroviral recombination and reverse transcription. *Science*, **250**, 1227-1233.

Ibarra, B., Valpuesta, J.M. and Carrascosa, J.L. (2001) Purification and functional characterization of p16, the ATPase of the bacteriophage Phi29 packaging machinery. *Nucleic Acids Res*, **29**, 4264-4273.

Ikonen, T., Kainov, D., Timmins, P., Serimaa, R. and Tuma, R. (2003) Locating the minor components of double-stranded RNA bacteriophage phi 6 by neutron scattering *J Appl Cryst*, **36**, 525-529

Ilyina, T.V., Gorbalenya, A.E. and Koonin, E.V. (1992) Organization and evolution of bacterial and bacteriophage primase-helicase systems. *J Mol Evol*, **34**, 351-357.

Itoh, H., Takahashi, A., Adachi, K., Noji, H., Yasuda, R., Yoshida, M. and Kinosita, K. (2004) Mechanically driven ATP synthesis by F1-ATPase. *Nature*, **427**, 465-468.

Jayaram, H., Estes, M.K. and Prasad, B.V. (2004) Emerging themes in rotavirus cell entry, genome organization, transcription and replication. *Virus Res*, **101**, 67-81.

Jayaram, H., Taraporewala, Z., Patton, J.T. and Prasad, B.V. (2002) Rotavirus protein involved in genome replication and packaging exhibits a HIT-like fold. *Nature*, **417**, 311-315.

Jonard, G., Guilley, H. and Hirth, L. (1975) Specific encapsidation of TMV RNA fragments by 25S TMV protein: isolation and some properties of the nucleoprotein complexes formed. *Virology*, **64**, 1-9.

Jung, A., Maier, R., Vartanian, J.P., Bocharov, G., Jung, V., Fischer, U., Meese, E., Wain-Hobson, S. and Meyerhans, A. (2002) Multiply infected spleen cells in HIV patients. *Nature*, **418**, 144.

Juuti, J.T. and Bamford, D.H. (1995) RNA binding, packaging and polymerase activities of the different incomplete polymerase complex particles of dsRNA bacteriophage phi 6. *J Mol Biol*, **249**, 545-554.

Juuti, J.T. and Bamford, D.H. (1997) Protein P7 of phage phi6 RNA polymerase complex, acquiring of RNA packaging activity by in vitro assembly of the purified protein onto deficient particles. *J Mol Biol*, **266**, 891-900.

Juuti, J.T., Bamford, D.H., Tuma, R. and Thomas, G.J., Jr. (1998) Structure and NTPase activity of the RNA-translocating protein (P4) of bacteriophage phi 6. *J Mol Biol*, **279**, 347-359.

Kainov, D.E., Simonov, V., Bamford, D.H., Tuma, R., Gottlieb, P., Wei, H., Walsh, M.A., Belrhali, H. and Merckel, M.C. (2004) Crystallization and preliminary X-ray diffraction analysis of bacteriophage varphi12 packaging factor P7. *Acta Crystallogr D Biol Crystallogr*, **60**, 2368-2370.

Kakitani, H., Iba, H. and Okada, Y. (1980) Penetration and partial uncoating of bacteriophage phi 6 particle. *Virology*, **101**, 475-483.

Kanamaru, S., Kondabagil, K., Rossmann, M.G. and Rao, V.B. (2004) The functional domains of bacteriophage t4 terminase. *J Biol Chem*, **279**, 40795-40801.

Karow, J.K., Newman, R.H., Freemont, P.S. and Hickson, I.D. (1999) Oligomeric ring structure of the Bloom's syndrome helicase. *Curr Biol*, **9**, 597-600.

Kim, C.H. and Tinoco, I., Jr. (2000) A retroviral RNA kissing complex containing only two G.C base pairs. *Proc Natl Acad Sci U S A*, **97**, 9396-9401.

Kinosita, K., Jr., Adachi, K. and Itoh, H. (2004) Rotation of F1-ATPase: how an ATP-driven molecular machine may work. *Annu Rev Biophys Biomol Struct*, **33**, 245-268.

Korolev, S., Hsieh, J., Gauss, G.H., Lohman, T.M. and Waksman, G. (1997) Major domain swiveling revealed by the crystal structures of complexes of E. coli Rep helicase bound to single-stranded DNA and ADP. *Cell*, **90**, 635-647.

Ktistakis, N.T., Kao, C.Y. and Lang, D. (1988) In vitro assembly of the outer shell of bacteriophage phi 6 nucleocapsid. *Virology*, **166**, 91-102.

Ktistakis, N.T. and Lang, D. (1987) The dodecahedral framework of the bacteriophage phi 6 nucleocapsid is composed of protein P1. *J Virol*, **61**, 2621-2623.

Laughrea, M., Jette, L., Mak, J., Kleiman, L., Liang, C. and Wainberg, M.A. (1997) Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. *J Virol*, **71**, 3397-3406.

Laurinavicius, S., Kakela, R., Bamford, D.H. and Somerharju, P. (2004) The origin of phospholipids of the enveloped bacteriophage phi6. *Virology*, **326**, 182-190.

Lesburg, C.A., Cable, M.B., Ferrari, E., Hong, Z., Mannarino, A.F. and Weber, P.C. (1999) Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. *Nat Struct Biol*, **6**, 937-943.

Levin, M.K., Gurjar, M. and Patel, S.S. (2005) A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat Struct Mol Biol*, ***5***, 429-435

Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakonyi, G. and Chen, X.S. (2003) Structure of the replicative helicase of the oncoprotein SV40 large tumour antigen. *Nature*, **423**, 512-518.

Liao, J.C., Jeong, Y.J., Kim, D.E., Patel, S.S. and Oster, G. (2005) Mechanochemistry of T7 DNA helicase. *J Mol Biol*, **350**, 452-475.

Lima, C.D., Klein, M.G. and Hendrickson, W.A. (1997) Structure-based analysis of catalysis and substrate definition in the HIT protein family. *Science*, **278**, 286-290.

Linderoth, N.A., Ziermann, R., Haggard-Ljungquist, E., Christie, G.E. and Calendar, R. (1991) Nucleotide sequence of the DNA packaging and capsid synthesis genes of bacteriophage P2. *Nucleic Acids Res*, **19**, 7207-7214.

Lisal, J., Kainov, D.E., Bamford, D.H., Thomas, G.J., Jr. and Tuma, R. (2004) Enzymatic mechanism of RNA translocation in double-stranded RNA bacteriophages. *J Biol Chem*, **279**, 1343-1350.

Lisal, J. and Tuma, R. (2005) Cooperative mechanism of RNA packaging motor. *J Biol Chem*, **280**, 23157-23164

Lohman, T.M. and Bjornson, K.P. (1996) Mechanisms of helicase-catalyzed DNA unwinding. *Annu Rev Biochem*, **65**, 169-214.

Makeyev, E.V. and Bamford, D.H. (2000a) The polymerase subunit of a dsRNA virus plays a central role in the regulation of viral RNA metabolism. *EMBO J*, **19**, 6275-6284.

Makeyev, E.V. and Bamford, D.H. (2000b) Replicase activity of purified recombinant protein P2 of double-stranded RNA bacteriophage phi6. *EMBO J*, **19**, 124-133.

Makeyev, E.V. and Grimes, J.M. (2004) RNA-dependent RNA polymerases of dsRNA bacteriophages. *Virus Res*, **101**, 45-55.

Mancini, E.J., Kainov, D.E., Wei, H., Gottlieb, P., Tuma, R., Bamford, D.H., Stuart, D.I. and Grimes, J.M. (2004) Production, crystallization and preliminary X-ray crystallographic studies of the bacteriophage phi 12 packaging motor. *Acta Crystallogr D Biol Crystallogr*, **60**, 588-590.

Mann, R., Mulligan, R.C. and Baltimore, D. (1983) Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. *Cell*, **33**, 153-159.

Matson, S.W., Bean, D.W. and George, J.W. (1994) DNA helicases: enzymes with essential roles in all aspects of DNA metabolism. *Bioessays*, **16**, 13-22.

Mavroidis, C., Dubey, A. and Yarmush, M.L. (2004) Molecular machines. *Annu Rev Biomed Eng*, **6**, 363-395.

McGlynn, P. and Lloyd, R.G. (2000) Modulation of RNA polymerase by (p)ppGpp reveals a RecG-dependent mechanism for replication fork progression. *Cell*, **101**, 35-45.

Meijer, W.J., Horcajadas, J.A. and Salas, M. (2001) Phi29 family of phages. *Microbiol Mol Biol Rev*, **65**, 261-287.

Mertens, P.P. and Diprose, J. (2004) The bluetongue virus core: a nano-scale transcription machine. *Virus Res*, **101**, 29-43.

Mindich, L. (1999) Precise packaging of the three genomic segments of the double-stranded-RNA bacteriophage phi6. *Microbiol Mol Biol Rev*, **63**, 149-160.

Mindich, L. (2004) Packaging, replication and recombination of the segmented genome of bacteriophage Phi6 and its relatives. *Virus Res*, **101**, 83-92.

Mindich, L., Qiao, X., Qiao, J., Onodera, S., Romantschuk, M. and Hoogstraten, D. (1999) Isolation of additional bacteriophages with genomes of segmented double-stranded RNA. *J Bacteriol*, **181**, 4505-4508.

Mitchell, A.H. and West, S.C. (1994) Hexameric rings of Escherichia coli RuvB protein. Cooperative assembly, processivity and ATPase activity. *J Mol Biol*, **243**, 208-215.

Moore, S.D. and Prevelige, P.E., Jr. (2002) DNA packaging: a new class of molecular motors. *Curr Biol*, **12**, R96-98.

Mougel, M. and Barklis, E. (1997) A role for two hairpin structures as a core RNA encapsidation signal in murine leukemia virus virions. *J Virol*, **71**, 8061-8065.

Moutouh, L., Corbeil, J. and Richman, D.D. (1996) Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure. *Proc Natl Acad Sci U S A*, **93**, 6106-6111.

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

Neuwald, A.F., Aravind, L., Spouge, J.L. and Koonin, E.V. (1999) AAA+: A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res*, **9**, 27-43.

Ni, C.Z., Syed, R., Kodandapani, R., Wickersham, J., Peabody, D.S. and Ely, K.R. (1995) Crystal structure of the MS2 coat protein dimer: implications for RNA binding and virus assembly. *Structure*, **3**, 255-263.

Niedenzu, T., Roleke, D., Bains, G., Scherzinger, E. and Saenger, W. (2001) Crystal structure of the hexameric replicative helicase RepA of plasmid RSF1010. *J Mol Biol*, **306**, 479-487.

Norcum, M.T., Warrington, J.A., Spiering, M.M., Ishmael, F.T., Trakselis, M.A. and Benkovic, S.J. (2005) Architecture of the bacteriophage T4 primosome: electron microscopy studies of helicase (gp41) and primase (gp61). *Proc Natl Acad Sci U S A*, **102**, 3623-3626.

Oertle, S. and Spahr, P.F. (1990) Role of the gag polyprotein precursor in packaging and maturation of Rous sarcoma virus genomic RNA. *J Virol*, **64**, 5757-5763.

Ojala, P.M., Juuti, J.T. and Bamford, D.H. (1993) Protein P4 of double-stranded RNA bacteriophage phi 6 is accessible on the nucleocapsid surface: epitope mapping and orientation of the protein. *J Virol*, **67**, 2879-2886.

Olkkonen, V.M. and Bamford, D.H. (1987) The nucleocapsid of the lipid-containing double-stranded RNA bacteriophage phi 6 contains a protein skeleton consisting of a single polypeptide species. *J Virol*, **61**, 2362-2367.

Olkkonen, V.M., Gottlieb, P., Strassman, J., Qiao, X.Y., Bamford, D.H. and Mindich, L. (1990) In vitro assembly of infectious nucleocapsids of bacteriophage phi 6: formation of a recombinant double-stranded RNA virus. *Proc Natl Acad Sci U S A*, **87**, 9173-9177.

Oroudjev, E.M., Kang, P.C. and Kohlstaedt, L.A. (1999) An additional dimer linkage structure in Moloney murine leukemia virus RNA. *J Mol Biol*, **291**, 603-613.

Paatero, A.O., Mindich, L. and Bamford, D.H. (1998) Mutational analysis of the role of nucleoside triphosphatase P4 in the assembly of the RNA polymerase complex of bacteriophage phi6. *J Virol*, **72**, 10058-10065.

Paatero, A.O., Syvaoja, J.E. and Bamford, D.H. (1995) Double-stranded RNA bacteriophage phi 6 protein P4 is an unspecific nucleoside triphosphatase activated by calcium ions. *J Virol*, **69**, 6729-6734.

Paillart, J.C., Berthoux, L., Ottmann, M., Darlix, J.L., Marquet, R., Ehresmann, B. and Ehresmann, C. (1996) A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. *J Virol*, **70**, 8348-8354.

Paillart, J.C., Shehu-Xhilaga, M., Marquet, R. and Mak, J. (2004) Dimerization of retroviral RNA genomes: an inseparable pair. *Nat Rev Microbiol*, **2**, 461-472.

Pape, T., Meka, H., Chen, S., Vicentini, G., van Heel, M. and Onesti, S. (2003) Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep*, **4**, 1079-1083.

Patel, S.S. and Picha, K.M. (2000) Structure and function of hexameric helicases. *Annu Rev Biochem*, **69**, 651-697.

Pickett, G.G. and Peabody, D.S. (1993) Encapsidation of heterologous RNAs by bacteriophage MS2 coat protein. *Nucleic Acids Res*, **21**, 4621-4626.

Pirttimaa, M.J. and Bamford, D.H. (2000) RNA secondary structures of the bacteriophage phi6 packaging regions. *RNA*, **6**, 880-889.

Pirttimaa, M.J., Paatero, A.O., Frilander, M.J. and Bamford, D.H. (2002) Nonspecific nucleoside triphosphatase P4 of double-stranded RNA bacteriophage phi6 is required for single-stranded RNA packaging and transcription. *J Virol*, **76**, 10122-10127.

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.

Prats, A.C., Roy, C., Wang, P.A., Erard, M., Housset, V., Gabus, C., Paoletti, C. and Darlix, J.L. (1990) cis elements and trans-acting factors involved in dimer formation of murine leukemia virus RNA. *J Virol*, **64**, 774-783.

Przech, A.J., Yu, D. and Weller, S.K. (2003) Point mutations in exon I of the herpes simplex virus putative terminase subunit, UL15, indicate that the most conserved residues are essential for cleavage and packaging. *J Virol*, **77**, 9613-9621.

Putnam, C.D., Clancy, S.B., Tsuruta, H., Gonzalez, S., Wetmur, J.G. and Tainer, J.A. (2001) Structure and mechanism of the RuvB Holliday junction branch migration motor. *J Mol Biol*, **311**, 297-310.

Qiao, J., Qiao, X., Sun, Y. and Mindich, L. (2003a) Isolation and analysis of mutants of double-stranded-RNA bacteriophage phi6 with altered packaging specificity. *J Bacteriol*, **185**, 4572-4577.

Qiao, X., Casini, G., Qiao, J. and Mindich, L. (1995) In vitro packaging of individual genomic segments of bacteriophage phi 6 RNA: serial dependence relationships. *J Virol*, **69**, 2926-2931.

Qiao, X., Qiao, J. and Mindich, L. (1997) Stoichiometric packaging of the three genomic segments of double-stranded RNA bacteriophage phi6. *Proc Natl Acad Sci U S A*, **94**, 4074-4079.

Qiao, X., Qiao, J. and Mindich, L. (2003b) Analysis of specific binding involved in genomic packaging of the double-stranded-RNA bacteriophage phi6. *J Bacteriol*, **185**, 6409-6414.

Qiao, X., Qiao, J., Onodera, S. and Mindich, L. (2000) Characterization of phi 13, a bacteriophage related to phi 6 and containing three dsRNA genomic segments. *Virology*, **275**, 218-224.

Reinisch, K.M., Nibert, M.L. and Harrison, S.C. (2000) Structure of the reovirus core at 3.6 A resolution. *Nature*, **404**, 960-967.

Ribas, J.C. and Wickner, R.B. (1998) The Gag domain of the Gag-Pol fusion protein directs incorporation into the L-A double-stranded RNA viral particles in Saccharomyces cerevisiae. *J Biol Chem*, **273**, 9306-9311.

Richardson, J.P. (2003) Loading Rho to terminate transcription. *Cell*, **114**, 157-159.

Salgado, P.S., Makeyev, E.V., Butcher, S.J., Bamford, D.H., Stuart, D.I. and Grimes, J.M. (2004) The structural basis for RNA specificity and Ca2+ inhibition of an RNA-dependent RNA polymerase. *Structure (Camb)*, **12**, 307-316.

Sambrook, J. and Russell, D. (2001) Molecular cloning: a laboratory manual. 3rd Ed, 1017-1019.

Sands, J.A. and Lowlicht, R.A. (1976) Temporal origin of viral phospholipids of the enveloped bacteriophage phi 6. *Can J Microbiol*, **22**, 154-158.

Sands, J.A., Lowlicht, R.A., Cadden, S.C. and Haneman, J. (1975) Assembly of the enveloped bacteriophage phi 6 in environments which perturb the host cell membranes. *Can J Microbiol*, **21**, 1287-1290.

Sato, M., Gotow, T., You, Z., Komamura-Kohno, Y., Uchiyama, Y., Yabuta, N., Nojima, H. and Ishimi, Y. (2000) Electron microscopic observation and single-stranded DNA binding activity of the Mcm4,6,7 complex. *J Mol Biol*, **300**, 421-431.

Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C. and Ellenberger, T. (1999) Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell*, **99**, 167-177.

Semancik, J.S., Vidaver, A.K. and Van Etten, J.L. (1973) Characterization of segmented double-helical RNA from bacteriophage phi6. *J Mol Biol*, **78**, 617-625.

Serwer, P. (2005) *Encyclopedia of Molecular Cell Biology and Molecular Medicine (R.A. Mayers, ed.)*, **9**, 351-361.

Simpson, A.A., Tao, Y., Leiman, P.G., Badasso, M.O., He, Y., Jardine, P.J., Olson, N.H., Morais, M.C., Grimes, S., Anderson, D.L., Baker, T.S. and Rossmann, M.G. (2000) Structure of the bacteriophage phi29 DNA packaging motor. *Nature*, **408**, 745-750.

Sinclair, J.F., Tzagoloff, A., Levine, D. and Mindich, L. (1975) Proteins of bacteriophage phi6. *J Virol*, **16**, 685-695.

Singleton, M.R., Dillingham, M.S., Gaudier, M., Kowalczykowski, S.C. and Wigley, D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*, **432**, 187-193.

Singleton, M.R., Sawaya, M.R., Ellenberger, T. and Wigley, D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589-600.

Singleton, M.R., Scaife, S., Raven, N.D. and Wigley, D.B. (2001a) Crystallization and preliminary X-ray analysis of RecG, a replication-fork reversal helicase from Thermotoga maritima complexed with a three-way DNA junction. *Acta Crystallogr D Biol Crystallogr*, **57**, 1695-1696.

Singleton, M.R., Scaife, S. and Wigley, D.B. (2001b) Structural analysis of DNA replication fork reversal by RecG. *Cell*, **107**, 79-89.

Skordalakes, E. and Berger, J.M. (2003) Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell*, **114**, 135-146.

Smith, D.E., Tans, S.J., Smith, S.B., Grimes, S., Anderson, D.L. and Bustamante, C. (2001) The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature*, **413**, 748-752.

Sommer, S.S. and Wickner, R.B. (1982) Yeast L dsRNA consists of at least three distinct RNAs; evidence that the non-Mendelian genes [HOK], [NEX] and [EXL] are on one of these dsRNAs. *Cell*, **31**, 429-441.

Soultanas, P., Dillingham, M.S., Wiley, P., Webb, M.R. and Wigley, D.B. (2000) Uncoupling DNA translocation and helicase activity in PcrA: direct evidence for an active mechanism. *EMBO J*, **19**, 3799-3810.

Steely, H.T., Jr. and Lang, D. (1984) Electron microscopy of bacteriophage phi 6 nucleocapsid: two-dimensional image analysis. *J Virol*, **51**, 479-483.

Stewart, L., Schatz, G. and Vogt, V.M. (1990) Properties of avian retrovirus particles defective in viral protease. *J Virol*, **64**, 5076-5092.

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

Stockley, P.G., Stonehouse, N.J. and Valegard, K. (1994) Molecular mechanism of RNA phage morphogenesis. *Int J Biochem*, **26**, 1249-1260.

Subramanya, H.S., Bird, L.E., Brannigan, J.A. and Wigley, D.B. (1996) Crystal structure of a DExx box DNA helicase. *Nature*, **384**, 379-383.

Sun, Y., Qiao, X., Qiao, J., Onodera, S. and Mindich, L. (2003) Unique properties of the inner core of bacteriophage phi8, a virus with a segmented dsRNA genome. *Virology*, **308**, 354-361.

Tabor, S. (1990) *Protocols in molecular biology*, 8322-8328.

Taraporewala, Z., Chen, D. and Patton, J.T. (1999) Multimers formed by the rotavirus nonstructural protein NSP2 bind to RNA and have nucleoside triphosphatase activity. *J Virol*, **73**, 9934-9943.

Taraporewala, Z.F. and Patton, J.T. (2004) Nonstructural proteins involved in genome packaging and replication of rotaviruses and other members of the Reoviridae. *Virus Res*, **101**, 57-66.

Temin, H.M. (1991) Sex and recombination in retroviruses. *Trends Genet*, **7**, 71-74.

Theis, K., Chen, P.J., Skorvaga, M., Van Houten, B. and Kisker, C. (1999) Crystal structure of UvrB, a DNA helicase adapted for nucleotide excision repair. *EMBO J*, **18**, 6899-6907.

Toth, E.A., Li, Y., Sawaya, M.R., Cheng, Y. and Ellenberger, T. (2003) The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol Cell*, **12**, 1113-1123.

Tuma, R., Bamford, J.K., Bamford, D.H. and Thomas, G.J., Jr. (1999) Assembly dynamics of the nucleocapsid shell subunit (P8) of bacteriophage phi6. *Biochemistry*, **38**, 15025-15033.

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

Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S. and Wigley, D.B. (1999) Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell*, **97**, 75-84.

White, C.A., Stow, N.D., Patel, A.H., Hughes, M. and Preston, V.G. (2003) Herpes simplex virus type 1 portal protein UL6 interacts with the putative terminase subunits UL15 and UL28. *J Virol*, **77**, 6351-6358.

von Hippel, P.H. and Delagoutte, E. (2001) A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, **104**, 177-190.

von Hippel, P.H. and Delagoutte, E. (2003) Macromolecular complexes that unwind nucleic acids. *Bioessays*, **25**, 1168-1177.

Abrahams, J.P., Leslie, A.G., Lutter, R., Walker, J.E. (1994) Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. *Nature.* **370,** 621-628

Walker, J.E., Saraste, M., Runswick, M.J. and Gay, N.J. (1982) Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J*, **1**, 945-951.

Wang, J. (2004) Nucleotide-dependent domain motions within rings of the RecA/AAA(+) superfamily. *J Struct Biol*, **148**, 259-267.

Washington, M.T., Rosenberg, A.H., Griffin, K., Studier, F.W. and Patel, S.S. (1996) Biochemical analysis of mutant T7 primase/helicase proteins defective in DNA binding, nucleotide hydrolysis, and the coupling of hydrolysis with DNA unwinding. *J Biol Chem*, **271**, 26825-26834.

Weigelt, J., Brown, S.E., Miles, C.S., Dixon, N.E. and Otting, G. (1999) NMR structure of the N-terminal domain of E. coli DnaB helicase: implications for structure rearrangements in the helicase hexamer. *Structure Fold Des*, **7**, 681-690.

Yamada, K., Kunishima, N., Mayanagi, K., Ohnishi, T., Nishino, T., Iwasaki, H., Shinagawa, H. and Morikawa, K. (2001) Crystal structure of the Holliday junction migration motor protein RuvB from Thermus thermophilus HB8. *Proc Natl Acad Sci U S A*, **98**, 1442-1447.

Yang, H., Gottlieb, P., Wei, H., Bamford, D.H. and Makeyev, E.V. (2003) Temperature requirements for initiation of RNA-dependent RNA polymerization. *Virology*, **314**, 706-715.

Ye, J., Osborne, A.R., Groll, M. and Rapoport, T.A. (2004) RecA-like motor ATPases--lessons from structures. *Biochim Biophys Acta*, **1659**, 1-18.

Yusa, K., Kavlick, M.F., Kosalaraksa, P. and Mitsuya, H. (1997) HIV-1 acquires resistance to two classes of antiviral drugs through homologous recombination. *Antiviral Res*, **36**, 179-189.

Ziegelin, G., T. Niedenzu, Lurz, R., Saenger, W., Lanka, E. (2003). Hexameric RSF1010 helicase RepA: the structural and functional importance of single amino acid residues. *Nucleic Acids Res*, **31,** 5917-5929