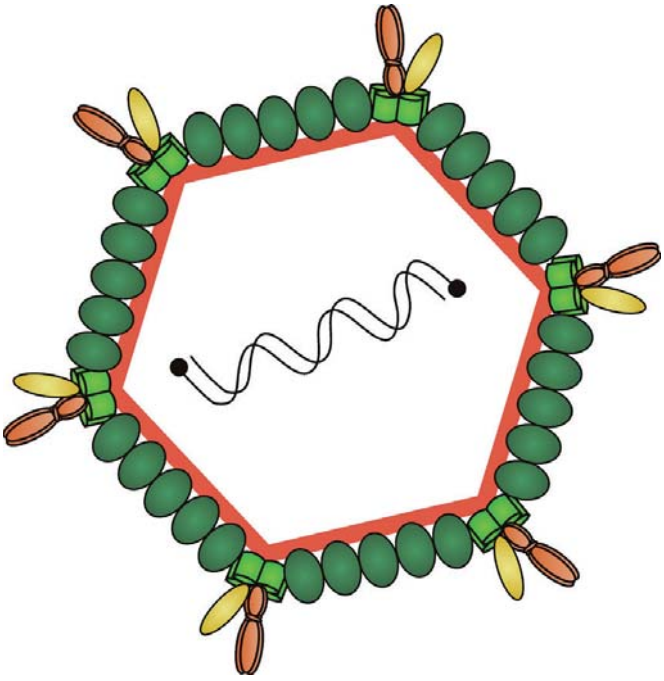
JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE
274

Jenni Karttunen

Interactions of Virus Proteins Within the Host Cell



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 274

Jenni Karttunen

Interactions of Virus Proteins Within the Host Cell

Esitetään Jyväskylän yliopiston matemaattis-luonnontieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Ylistönrinteellä, salissa YAA303
maaliskuun 22. päivänä 2014 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Mathematics and Science of the University of Jyväskylä,
in Ylistönrinne, hall YAA303, on March 22, 2014 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

Interactions of Virus Proteins Within the Host Cell

JYVÄSKYLÄ STUDIES IN BIOLOGICAL AND ENVIRONMENTAL SCIENCE 274

Jenni Karttunen

Interactions of Virus Proteins Within the Host Cell

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2014

Editors
Varpu Marjomäki
Department of Biological and Environmental Science, University of Jyväskylä
Pekka Olsbo, Timo Hautala
Publishing Unit, University Library of Jyväskylä

Jyväskylä Studies in Biological and Environmental Science

URN:ISBN:978-951-39-5634-9
ISBN 978-951-39-5634-9 (PDF)

ISBN 978-951-39-5633-2 (nid.)
ISSN 1456-9701

Jyväskylä University Printing House, Jyväskylä 2014

ABSTRACT

Karttunen, Jenni
Interactions of virus proteins within the host cell
Jyväskylä: University of Jyväskylä, 2014, 55 p.
Jyväskylä Studies in Biological and Environmental Science
ISSN 1456-9701; 274)
ISBN 978-951-39-5633-2 (nid.)
ISBN 978-951-39-5634-9 (PDF)
Yhteenveto: Virusproteiinit ja niiden vuorovaikutukset isäntäsolussa
Diss.

Viruses are ancient parasites that predate all three domains of life: Eukarya, Bacteria and Archaea. Viruses usually specifically infect a determined cell type, largely defined by the receptors they recognise. Canine parvovirus is a small, non-enveloped animal virus that infects cells in dividing cells, especially in puppies. PRD1 is a bacteriophage infecting a wide range of Gram-negative bacteria. It is a well-known model virus of the *Tectiviridae* family and its structure has been the subject of numerous studies. In this thesis, interactions between virus proteins and the host cell were studied using both CPV and PRD1 as model viruses. The study of the complex inner protein localization in bacterial cells is in its infancy, yet particularly important for the understanding of virus-cell interactions. We constructed a vector library for the production of fluorescent fusion proteins and used it to study the localization of several PRD1 viral proteins inside *E. coli*. Results revealed variations in the localization patterns; one dividing characteristic appeared to be the multimericity of proteins. The vector library, together with complementation assays, was further utilized in the study of PRD1 proteins P33 and P17. The obtained results indicated that these proteins interact with the host cell chaperonin machinery. The GroEL/GroES machinery is vital for both the host cell and several phages including PRD1. Finally, the interactions of the CPV capsid with lipid membranes were studied in order to elucidate the virus escape mechanism from the endocytic vesicles. Results suggested that, together with known phospholipase A2 activity, there is an additional, membrane-induced mechanism to promote the escape.

Keywords: Bacteriophage PRD1; canine parvovirus; confocal microscopy; green fluorescent protein; virus-host interactions.

Jenni Karttunen, University of Jyväskylä, Department of Biological and Environmental Science, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

Author's address Jenni Karttunen
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland
jenni.m.karttunen@jyu.fi

Supervisors Professor Jaana Bamford
Department of Biological and Environmental Science
P.O. Box 35
FI-40014 University of Jyväskylä
Finland

Docent Hanna Oksanen
Institute of Biotechnology and Department of Biosciences
P.O. Box 56
FI-00014 University of Helsinki
Finland

Reviewers PhD Johanna Laakkonen
University of Eastern Finland
A. I. Virtanen Institute for Molecular Sciences
Department of Biotechnology and Molecular Medicine
P.O. Box 1627
FI-70211 Kuopio
Finland

Docent Vesa Hytönen
BioMediTech / Protein Dynamics
FI-33014 University of Tampere
Finland

Opponent Professor Markku Kulomaa
BioMediTech / Molecular Biotechnology
FI-33014 University of Tampere
Finland

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following original papers, which will be referred to in the text by their Roman numerals I-III.

I Karttunen J., Mäntynen S., Ihalainen T.O., Lehtivuori H., Tkachenko N.V., Vihinen-Ranta M., Ihalainen J.A., Bamford J.K. & Oksanen H.M. 2014. Subcellular localization of bacteriophage PRD1 proteins in *Escherichia coli*. *Virus Research* Jan 22;179:44-52.

II Karttunen J*, Mäntynen S*, Ihalainen TO, Oksanen HM & Bamford JKH. Membrane-containing bacteriophage PRD1 protein P33 complements the defect in GroES of Escherichia coli. *Submitted manuscript*.

III Pakkanen K., Karttunen J., Virtanen S. & Vuento M. 2008. Sphingomyelin induces structural alteration in canine parvovirus capsid. *Virus Research* Mar;132(1-2):187-91.

* Equal contribution

RESPONSIBILITIES OF JENNI KARTTUNEN IN THE THESIS ARTICLES

Article I — I supervised the Master Thesis projects in which the plasmid library was constructed and I cloned some of the plasmids used in the study. Sari Mäntynen was responsible for the preparation of the expression plasmids. I made the rate zonal centrifugation assays and analysis together with Sari Mäntynen. I performed the complementation analysis. I performed the fluorescence lifetime measurements and analysis under the supervision of Janne Ihalainen, Heli Lehtivuori and Nikolai Tkachenko. I performed the confocal microscopy and data analysis. I made the figures for the article. I wrote the article together with the co-authors.

Article II — I performed the confocal microscope imaging and the FRAP-assay. Teemu Ihalainen performed the FRAP simulations. Sari Mäntynen prepared the plasmids for overexpression of the fluorescent fusion proteins. I made the rate zonal centrifugation assay and analysis. I wrote the article together with Sari Mäntynen. I made most of the figures for the article.

Article III — I performed the tryptophan fluorescence measurements, which were repeated by Salla Virtanen. Kirsi Pakkanen performed all other measurements and wrote the article.

1 INTRODUCTION

Viruses are ancient parasites that predate all three domains of life: Eukarya, Bacteria and Archaea. Viruses can be found in every habitat where life exists and they are a strong driving force behind the evolution of organisms. In particular, viruses that infect bacteria have been isolated from the most diverse and extreme locations, ranging from the desert to the arctic sea ice through oceans and even hot springs. Despite their wide distribution, individual viruses are extremely specific and they infect only a restricted range of host cells, typically defined by the receptor proteins they recognize.

Animal viruses stand out the most since they cause human and animal diseases, frequently leading to dreaded epidemics such influenza, acquired immune deficiency syndrome (AIDS) or severe acute respiratory syndrome (SARS). This is one reason why animal viruses are extensively studied, the ultimate goal being the development of efficient vaccines. This has been successfully put into practice with the eradication of smallpox in the late 70s.

Bacteriophages are viruses infecting bacterial cells. Due to the simpler structure of bacteria and their easiness of manipulation, bacteriophages are often used as study model systems for the more complex, but structurally similar viruses. During the past decades, the rising tide of antibiotic resistance in bacteria has become a growing concern in the developed world. This has led to a renewed interest in bacteriophages as they are natural enemies of bacteria and a potential tool for antibacterial therapy. For both successful development of vaccines against animal viruses and use of bacteriophages against bacteria, a comprehensive understanding of their developmental cycles is an imperative.

In this doctoral thesis, virus protein interactions within the host cell were studied using bacteriophage PRD1 and canine parvovirus (CPV) as model viruses. This was done utilizing traditional genetics combined with fluorescence microscopy and spectroscopy. The localization of several PRD1 proteins was observed in the host cell, revealing interesting localization patterns. The unknown role of PRD1 proteins P33 and P17 was also studied, revealing their potential role in the function of a chaperone complex. Finally, membrane

interactions of CPV capsid were probed to find out more about the viral entry and escape from endosomal vesicles.

2 REVIEW OF THE LITERATURE

2.1 Virus structures

In spite of the diversity of viral structures, there are some general features that are shared by all known viral particles. Generally, viruses contain a genome, which stores the genetic information, and a protein capsid to shelter the genome. In addition, viruses can have an inner or an outer membrane, multiple enzymatic proteins and other complex structures. The viral genome can be as small as 3.5 kb (e. g. rous sarcoma virus) and the smallest capsids are around 20 nm in diameter (e. g. parvoviruses). While so far considered to represent the smallest entities on Earth, unexpectedly large viruses, termed giant viruses (genome >1 Mb and capsid size >400 nm), are continuously being isolated: Mimivirus was found in 1992 (La Scola *et al.* 2003) and become the largest virus indentified until the discovery of Megavirus chilensis in 2011 (Arslan *et al.* 2011). Followed then the Pandoravirus, the discovery of which was reported in the Journal Science in 2013 (Philippe *et al.* 2013).

One practical way of categorizing viruses is according to the nature of their genome. The genome can be single (ss) or double stranded (ds) and be made of DNA or RNA. Another classification criterion is by way of the symmetry of their capsid. Viral capsid structures are usually highly symmetrical. Two common symmetry types are the helical symmetry, which is employed by rod-like or filamentous viruses (e. g. tobacco mosaic virus), and the icosahedral symmetry of spherical viruses (e. g. PRD1 and canine parvovirus). The third group encompasses viruses that display no clear symmetry and possibly, have a very complex structure (e. g. mimivirus). There are also combinations of the previously mentioned structures, for example tailed bacteriophages with an icosahedral or elongated icosahedral head and a helical tail attached to one vertex (phage T4). In addition to these shapes, viruses infecting archaea can possess extremely complex structures not seen in viruses infecting eukaryotes or prokaryotes (Krupovic *et al.* 2012).

The virus capsid is composed of only a few capsid protein types, limiting the virus genome size to the necessary minimum. For example, a helical capsid can form from identical subunits and the simplest icosahedral symmetry is made of 60 pieces of identical pentameric subunits (Fig. 1A), usually coat protein trimers. With larger icosahedral structures, identical pentamers alone cannot form a rigid structure and they occupy only the vertices of the capsid. In this case, the addition of hexamers is necessary (Fig. 1B and C). The precise structure of an icosahedral capsid can be determined by its triangulation number (T-number). This number describes the ratio between pentagons and hexagons in the capsid and it can be calculated with the formula: $T = h^2 + kh + k^2$, where h and k are coordinates of the distance between two adjacent 5-fold vertices (Caspar and Klug, 1962).

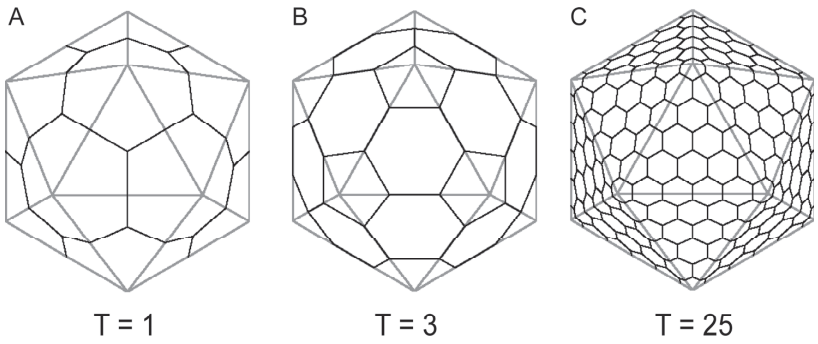


FIGURE 1 Illustration of icosahedral symmetry and T-number. The complexity of virus structure grows together with the T-number. a) A virus with T = 1 symmetry can be composed of only one type of pentameric subunits. b) and c) Larger viruses require hexameric protein structures in addition to pentamers. Icosahedrons were modelled from the Virus particle explorer (Viper) database (Carrillo-Tripp *et al.* 2009).

2.2 Viral entry

The first stage of viral entry is the binding to the receptor protein. This can be mediated through unspecific binding to cells surface helping the virus particle to reach the correct receptor protein. Some viruses require also co-receptors to trigger the penetration (e.g. human immunodeficiency virus). There are three distinct mechanisms by which the virus can initiate cell penetration. One of them involves the fusion of the virus envelope with the cellular membrane. Virus binding to the receptor protein triggers the virus envelope to physically fuse with the host cell plasma membrane. In this process, the capsid is released into the cytosol and continues its progression inside the cell. For example, retroviruses (e.g. human immunodeficiency virus, murine leukemia virus) infect their hosts using this mechanism.

A second common way of entry is through endocytosis: all non-enveloped animal viruses exploit this route. The virus utilizes cell receptors involved in the uptake of useful extracellular proteins and triggers the uptake in order to later escape the endocytic vesicles. There are many different endocytic pathways that viruses can exploit according to their nature; some viruses can even exploit several different routes. Size of the viral particle is one determining factor that influences to the cellular pathway that the virus can exploit. Possible routes are e.g. the clathrin-mediated pathway, the caveolin-mediated pathway, clathrin- and caveolin- independent pathways, phagosytosis or macropinocytosis (Mercer *et al.* 2010). The clathrin-mediated route is explained in more detail in section 2.3.

In general, bacteriophages use the third mechanism of infection, which consists in the injection of the genome inside the cell while the capsid remains outside. This entry is usually a complex process, involving for example multi-step binding and the use of lytic enzymes that will locally digest the cell wall to facilitate access to the internalization machinery. In many cases, the tail is used as a platform for injection. Tailless viruses such as tectiviruses, on the other hand, use their inner membrane as a tube for injecting the genome inside the cell. A newly found delivery mechanism consists of the DNA pilot protein of bacteriophage ϕX174. Here, the protein forms a tube during infection that delivers the DNA into the cytoplasm (Sun *et al.* 2013). One known exception for the bacteriophage entry mechanism is phage ϕ6 (*Cystoviridae* family). Cystoviruses have two separate capsids as well as an outer membrane, and they enter the host cell by using both membrane fusion and lytic enzyme digestion. As a result, only the inner capsid is transported into the cytosol. These viruses appear to be more closely related to animal viruses, such as families *Reoviridae* and *Totiviridae*, than to bacteriophages and this also applies to other aspects of their infectious cycle (Bamford *et al.* 1987, Poranen *et al.* 1999, Cvirkaite-Krupovic *et al.* 2010).

The differences in entry mechanisms between bacteriophages and animal viruses are attributable to the different host cell types they infect. Bacterial cells are their own entities whereas eukaryotic cells have the multicellular organism as a protection. In addition in order to infect the host, animal viruses must overcome epithelial barriers and evade the immune system. Eukaryotic cells lack the outer membrane and cell wall present in many bacterial cells and thus, the inner membrane is easier to reach. On the other hand, eukaryotic cells display a more complex inner structure and the virus might need additional capsid proteins in the subsequent journey to the nucleus or other places of replication. If the virus encompasses lipids, they always play a role in the entry process.

2.3 Canine parvovirus

Canine parvovirus (CPV-2) is a small, non-enveloped animal virus. It belongs to the autonomous *Parvovirus* genus of the *Parvoviridae* family and replicates in dividing cells, especially in puppies (Cotmore and Tattersall 1987). The 5kb long

genome is composed of a linear single stranded DNA molecule encoding four different proteins from two open reading frames (ORFs): non-structural proteins NS-1 and NS-2 are coded by the first ORF, and structural proteins VP-1 and VP-2 by the second. Non-structural proteins have separate reading frames whereas the structural proteins have almost identical sequences, apart from the additional N-terminal extension that distinguishes the VP-1. A third structural protein, VP-3, is formed in DNA-containing capsids as a result of host cell proteolytic cleavage of exposed N-terminal ends of VP-2 (Paradiso *et al.* 1982, Reed *et al.* 1988, Weichert *et al.* 1998).

The non-enveloped capsid of CPV has a diameter of approximately 26 nm and is structurally quite simple. It has icosahedral T=1 symmetry (Fig 1A) and the 60 subunits are mainly composed of the capsid protein VP-2. The minor capsid proteins VP-1 and VP-3 are also present in the capsid structure. The unique N-terminus of VP-1 harbors a nuclear localization signal and a phospholipase A2 (PLA_2) motif, which has been shown to be essential for infection (Tsao *et al.* 1991). The N-terminus is enclosed within the virus capsid, but becomes exposed in the acidic conditions encountered in endosomes. Also, heat or urea treatment can expose the protein terminus (Weichert *et al.* 1998, Vihinen-Ranta *et al.* 2002, Suikkanen *et al.* 2003).

CPV binds to transferrin receptors (TfR1) on the host cell and uses the clathrin-mediated endocytosis route for entry (Fig. 2). Clathrin-mediated endocytosis maintain cell and serum homeostasis and is involved in signal transduction (McMahon and Boucrot 2011). First, the cargo-bound receptors move along to form clathrin-coated pits that will evolve in to a clathrin-coated vesicle. The vesicle is then uncoated by cellular enzymes and it fuses with early endosomes, which are also responsible for sorting the endocytic cargo. The sorted cargo moves forward, either to recycling endosomes or to late endosomes, depending on its final destination. From the late endosomes, the remaining cargo moves to lysosomes where it is degraded. Along the endosomal route, the pH decreases from mildly acidic (pH of 6.5-6.0 in early endosomes) trough acidic (pH = 6.0-5.0 in late endosomes) to finally reaching a pH of 5.0-4.5 in the lysosomes. In addition to this simplified presentation, the molecules taken in can be transported from the vesicles to different parts of the cell, for example, to the nucleus or Golgi apparatus, depending on where in the cell they are going to be utilized.

CPV particles remain bound to the receptor molecules for several hours in the endocytic vesicles after infection and the precise mechanism and place of escape are not thoroughly known. Viral particles do not completely disrupt the endosomal membranes during release into the cytosol and PLA_2 has been suggested to play a role in this process when exposed to acidic conditions (Parker and Parrish 2000, Suikkanen *et al.* 2003). From the cytosol, viral particles travel along microtubules using motor protein dynein to the nucleus, which they enter through its pores (Suikkanen *et al.* 2003). The process is controlled by the nuclear localization signal found near the PLA_2 domain in VP-1. DNA replication, gene expression and particle assembly all take place in the nucleus. (For more detailed review of CPV lifecycle see Vihinen-Ranta *et al.* 2004.)

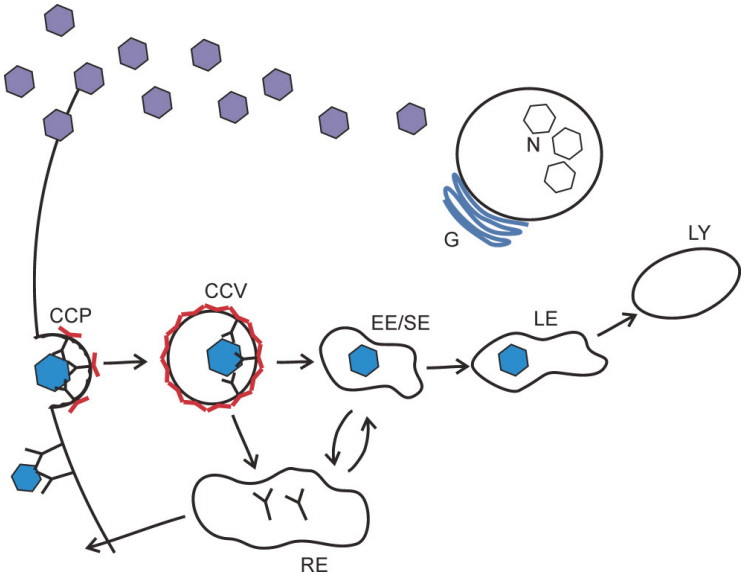


FIGURE 2 CPV clathrin-mediated endocytosis and lifecycle. After binding to transferrin receptors viral particles move through the clathrin-coated pit (CCP), clathrin-coated vesicle (CCV), early (sorting) endosome (EE/SE), late endosome (LE) and lysosome (LY) stages. CPV particles have been observed all the way up to lysosomes. After release into the cytosol, viral particles travel on microtubules to enter the nucleus (N) through nuclear pores. In the nucleus empty particles are formed into which the genome is packaged and newly formed CPV particles escape from the cell. G = Golgi apparatus. The transferrin receptors are recycled back to the cell surface by recycling endosome (RE).

2.4 The *Tectiviridae* family

The *Tectiviridae* family comprises icosahedral (66 nm from facet to facet), tailless bacteriophages infecting either gram-negative or gram-positive bacteria. The characteristic feature that distinguishes this virus family from other membranous viruses is the position of the membrane underneath the protein capsid. A similar inner membrane is only found in few virus groups including for example the *Corticoviridae*, which however has only one representative virus, *pseudoalteromonas* phage PM2. Other characteristic features of tectiviruses are a linear dsDNA genome of about 15 kb and the use of the inner membrane as a tubular structure in DNA delivery (Oksanen and Bamford 2012).

The family includes six Gram-negative bacteria infecting viruses: PR3, PR4, PR5, L17, PR772 and PRD1, of which PRD1 is the study model of the family (Bamford *et al.* 1981). Despite the high degree of similarity (up to 98% of sequence identity), these viruses have been isolated from distant locations (Olsen *et al.* 1974, Stanisich 1974, Wong and Bryan 1978, Coetzee and Bekker 1979, Bamford *et al.* 1981). Gram-positive bacteria infecting tectiviruses are more

diverse in sequence and include *Bacillus anthracis* phage AP50 and *B. thuringiensis* phage Bam35, GIL01 and GIL16 (Nagy *et al.* 1976, Ackermann *et al.* 1978, Ravantti *et al.* 2003, Verheust *et al.* 2003, Verheust *et al.* 2005).

The last members to have proposed to join the family are several *Thermus* species infecting phages isolated from alkaline hot springs (Yu *et al.* 2006). However, at least with phage P23-77, closer studies have revealed that it is fundamentally different from other tectiviruses (structure of the major coat protein, T-number, circular genome) (Jalasvuori *et al.* 2009, Rissanen *et al.* 2012) leading to its classification under a distinct family (Pawlowski *et al.* 2014).

2.4.1 Model virus PRD1

PRD1, the type virus of *Tectiviridae*, was isolated from Michigan sewers in 1974. Its host range includes various Gram-negative bacteria such as *Escherichia coli*, *Salmonella enterica* and *Pseudomonas aeruginosa*. PRD1 recognizes as receptor a component of the bacterial conjugative apparatus that is encoded on multidrug-resistant plasmids of the IncP, IncN or IncW incompatibility groups. The widespread occurrence of those plasmids in enteric bacteria is responsible for the large host spectrum of PRD1. The dsDNA genome is 14.9 kb long, delineated by identical inverted repeats and has terminal proteins covalently attached to the 5′ ends. The genome is organized into five operons, of which two are expressed early during infection and three are late operons. Altogether, these operons encode at least 27 proteins (Olsen *et al.* 1974, Bamford *et al.* 1991, Bamford *et al.* 1995).

A schematic representation of the PRD1 virion is given in Figure 3. The viral capsid has icosahedral symmetry with a T-number of 25 (see also Fig. 1C). The capsid is formed of hexagonally-shaped trimers of the major capsid protein P3 (total amount of 240). They are united by the tape-measure protein P30, which lies underneath the capsid and controls the size of the capsid during assembly (Rydman *et al.* 2001). The vertices are occupied by pentameric protein P31. It forms the base to which the spike complex of receptor binding protein P2 and spike protein P5 are attached (Rydman *et al.* 1999, Huiskonen *et al.* 2007). The transmembrane glue protein P16 is also found on the vertices and it connects proteins P31 and P3 to the inner membrane (Jaatinen *et al.* 2004, Abrescia *et al.* 2004). The capsid includes one special vertex for genome packaging and this vertex is presumed to also serve as a DNA delivery gate, although this has never been directly demonstrated. The special vertex contains the packaging ATPase P9, which is linked to the capsid through small membrane proteins P20 and P22, and capsid protein P6 (Stromsten *et al.* 2003). The structure of several structural proteins of PRD1 have been solved and in addition, the crystal structure of a virion defective for the receptor protein P2 was determined at 4 Å resolution in 2004 (Abrescia *et al.* 2004).

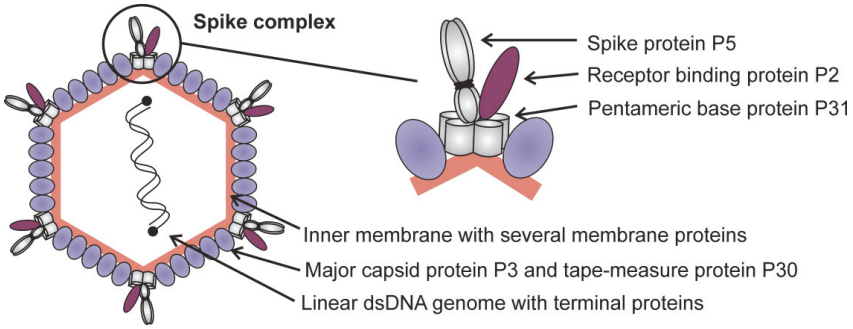


FIGURE 3 Schematic representation of PRD1 virion and its spike complex. The inner membrane encloses the dsDNA genome with its terminal proteins. Major capsid protein P3 trimers are held together by tape-measure protein P30. Pentameric vertices are formed of protein P31, which is connected to the membrane by membrane protein P16 and from which spike protein P5 and receptor binding protein P2 are protruding. The special vertex involved in genome packaging and likely, its delivery, is not illustrated.

The viral inner membrane originates from the cellular cytoplasmic membrane and is composed of lipids (approximately 60%) and phage-specific proteins. The phospholipid composition may vary depending on the host cell. Still, the membranes are not identical and for instance, phosphatidylglycerol is enriched in the virus when compared to the host cytoplasmic membrane (Laurinavicius *et al.* 2004).

Similar to several other bacteriophages, PRD1 infects the host by injecting the DNA into the cell by the tube formed from inner membrane, leaving the capsid outside. The injection is not a simple task: the virus must overcome three layers of cell envelope (outer membrane, peptidoglycan layer and cytoplasmic membrane) and at least proteins P7, P11, P14, P18 and P32 are involved in the process (Rydman and Bamford 2000, Grahn *et al.* 2002a, Grahn *et al.* 2002b). Being a virulent phage, PRD1 engages immediately into the lytic pathway upon infection. The assembly of new viral particles starts with lipid vesicles being covered with non-structural protein P10 and small amounts of capsid protein P3. Vesicles are proposed to form from virus-specific membrane rafts in the same manner as clathrin-coated pits are formed in the clathrin-mediated endocytosis process. Together with tape-measure protein P30, the major capsid protein eventually substitutes P10 proteins resulting in a procapsid including an inner membrane and all the structural proteins, except the packaging ATPase P9 and DNA-bound terminal protein P8. (Butcher *et al.* 2012) The DNA is actively packaged into these procapsids by the packaging ATPase and accessory proteins (Strömsten *et al.* 2003, Strömsten *et al.* 2005). The progeny phages are released after host cell lysis caused by lytic muramidase P15 and holin protein P35.

The lifecycle and structure of PRD1 have been extensively studied and yet, many details remain unsolved. PRD1 proteins P33 and P17 are two examples for which the functions remain unknown. The use of mutants deficient for P17 showed that it is an essential factor that most likely acts at the level of virus

assembly (Caldentey *et al.* 1999). It has also been observed to be a soluble tetramer, which binds to positively charged membranes when studied in its purified form (Holopainen *et al.* 2000). As for P33, there is no antibody or defective virus mutant available and therefore, the current knowledge of its function remains limited. The location of these two genes within the same operon, along with packaging and assembly genes, suggests that P17 and P33 might also have a role in the assembly of the viral particles.

2.5 Chaperone proteins and virus life cycle

Chaperones are housekeeping proteins involved in several essential cellular processes including folding of proteins, disaggregation of protein aggregates, protection of developing polypeptide chains from premature aggregation, polypeptide transport across biological membranes and proteolysis. Chaperone proteins are found in all biological organisms: eukaryotes, prokaryotes and archaea. They are also commonly utilized in both bacterial and animal virus life cycles.

Chaperonin proteins are a subgroup of chaperones, defined by sequence similarity, and responsible for the folding of several cellular proteins. The most thoroughly studied chaperonin proteins are the GroEL/GroES complex in *E. coli*. The structures of all chaperonin complexes are quite conserved. In the GroEL/GroES complex, two heptameric GroEL-rings form a barrel-like structure, which acts as a folding chamber for the substrate proteins. When the substrate binds, the chamber is sealed by the heptameric GroES hat-like structure and the folding process is powered by adenosine triphosphate (ATP) hydrolysis (Fig. 4). GroES interacts with GroEL through a mobile loop structure. The loop consists of three hydrophobic residues surrounded by flexible glycine residues (Landry *et al.* 1993). The chaperonin complex goes through several conformational changes in the process of folding proteins. The crystal structures of these different states have been determined in recent decades, revealing extremely complex protein machinery (Braig *et al.* 1995, Xu *et al.* 1997, Chaudhry *et al.* 2004, Bartolucci *et al.* 2005, Fei *et al.* 2013).

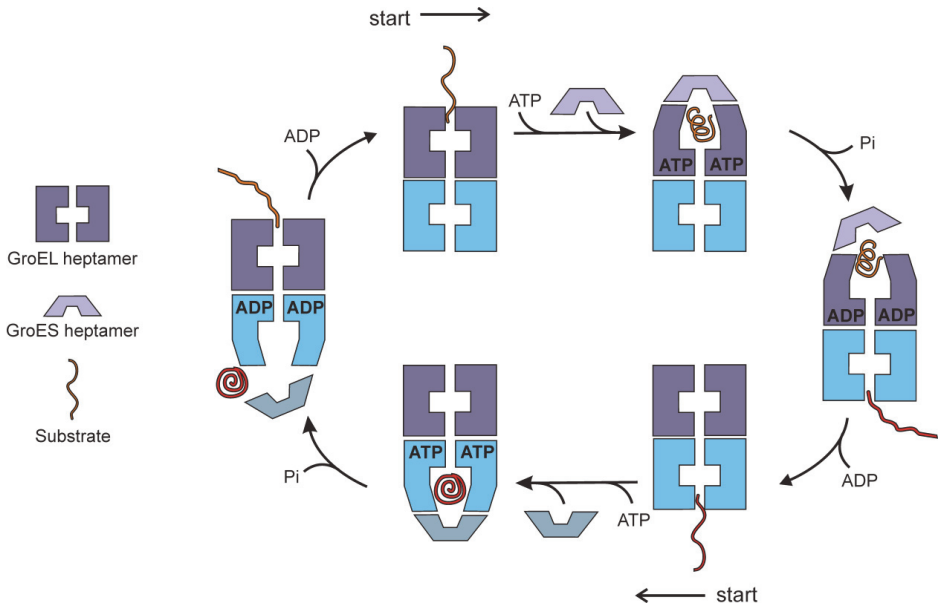


FIGURE 4 Schematic representation of the GroEL-GroES chaperonin complex. The unfolded substrate binds to the hydrophobic residues of one of the GroEL chambers. The substrate induces binding of first, ATP, and second, GroES, to the complex. Due to GroES binding, hydrophobic residues fold away and expose a larger polar chamber, which induces the protein to fold correctly into its native state. The binding of ATP locks the opposite GroEL ring, which remains unbound until the ATP is hydrolysed into ADP. After hydrolysis, another unfolded protein can bind to the opposite GroEL ring, resulting in the release of the folded substrate from the first ring. Figure has been prepared based on recet reviews (Xu *et al.* 1997, Fei *et al.* 2013).

Chaperonins are one of the many cellular proteins that viruses exploit during their lifecycle. In fact, chaperonins were first discovered through bacterial mutants blocking the growth of phage λ (Georgopoulos 2006). Such as in other viruses, chaperonin proteins also play an essential role in the lifecycle of PRD1. The GroEL/ES complex is needed for proper folding and trimerization of the capsid protein P3 and spike protein P5. Also, the assembly of P11 and small membrane protein(s) onto the viral particle is affected in the absence of GroEL or GroES (Hänninen *et al.* 1997).

Several bacteriophages encode co-chaperonin proteins capable of replacing GroES in the chaperonin complex. Examples are protein Gp31 of bacteriophage T4 and protein CocO of phage RB49 (van der Vies *et al.* 1994, Ang *et al.* 2001). Interestingly, these co-chaperonins can also substitute GroES in the folding of *E. coli*'s own essential proteins (Keppel *et al.* 2002). The production of its own co-chaperonin protein can be vital for the virus: for example, the coat protein of T4 is too large for the host cells natural chaperonin chamber and the viral GroES-like protein is proposed to sufficiently expand the chamber for the virus protein to fold (Hunt *et al.* 1997). Lately, genetic comparisons have revealed putative GroEL-like proteins in other viruses, and protein gp146 of bacteriophage EL has

indeed been shown to have a function similar to that of GroEL (Kurochkina *et al.* 2012).

In addition to chaperone-like proteins, some phage-borne proteins, such as gene product Gp39.2 of phage RB69, manipulate the chaperone function in an alternative manner. Similarly functioning homologues have also been found in phages T4, RB43 and RB49. These proteins act as a switch changing the equilibrium of GroEL from closed to open state and thus, inducing the binding of the substrate. The gene product of T4 is found to be essential, to the point that its deletion from the genome inhibits the growth of the phage in certain *E. coli* hosts (Ang and Georgopoulos 2012).

2.6 Protein localization in bacteria

The bacterial cell structure differs in many ways from the structure of eukaryotic cells. Instead of a membrane-enclosed nucleus found in eukaryotes, the genetic material in bacteria is organized into a nucleoid, which is not surrounded by a membrane. There are no mitochondria or Golgi apparatus in bacterial cells and ribosomes are slightly smaller compared to the ones found in animal or plant cells. The structure of peptidoglycan in the cell wall distinguishes Gram-negative from Gram-positive bacteria. Gram-negative cells have relatively thin cell wall surrounded by an outer lipid membrane, compared to Gram-positive bacteria where the cell wall is much thicker and the outer membrane is absent.

The complexity of animal cells has been a subject of study for a long time. Animal viruses exploit this complexity and specialize into hijacking cellular mechanisms for their own development. Unlike eukaryotic cells, bacterial cells were long considered to be reaction vessels containing proteins in an even mixture. Detailed research using high resolution microscopy methods has proven this assumption to be wrong and revealed that bacterial cells are almost as complex as their eukaryotic counterparts. One of the first findings was that, contrary to the common belief, bacterial cells have homologs of all known eukaryotic cytoskeletal elements and even a few unique ones. These cytoskeleton proteins are found to be responsible, for example, for maintaining the bacterial shape (Celler *et al.* 2013). In addition, very large and structurally sophisticated microcompartments (e.g. carboxysome) and protein clusters has been identified from bacterial cells. The revealed complexity has led to the proposition that bacterial cells are suitable model organisms in the study of universal mechanisms behind spatial regulation of cellular processes (Amster-Choder 2011).

One of the most studied bacterial models, in general and with regard to protein localization, is *E. coli*. A summary of its common protein localization patterns is depicted in Figure 5. Polar localization is exploited, for example, by the chemotaxis system (Greenfield *et al.* 2009) and the phosphotransferase system (Lopian *et al.* 2010), which is responsible for cells carbon uptake and metabolism. These two systems communicate extensively, and together they form the equivalent of a metabolic nervous system in bacteria. Therefore, their close

cellular localization does not strike as surprising. One example of midcell localization is FtsZ, tubulin-like GTPase involved in the cytokinesis of bacteria. It forms Z-ring which defines the division site of the cell (Meier and Goley 2014). The helical protein localization is another commonly observed pattern and is used for example by protein translocation machinery (Sec) (Shiomi *et al.* 2006). Helical localization is typical also for cytoskeletal proteins. MreB, the bacterial counterpart of actin, was first considered to display a helical organization but this view has later been challenged and the protein might in fact be organized in discrete patches (Swulius and Jensen 2012). In addition to proteins, different mRNAs have also been shown to have helical, membrane and polar localizations (Nevo-Dinur *et al.* 2011).

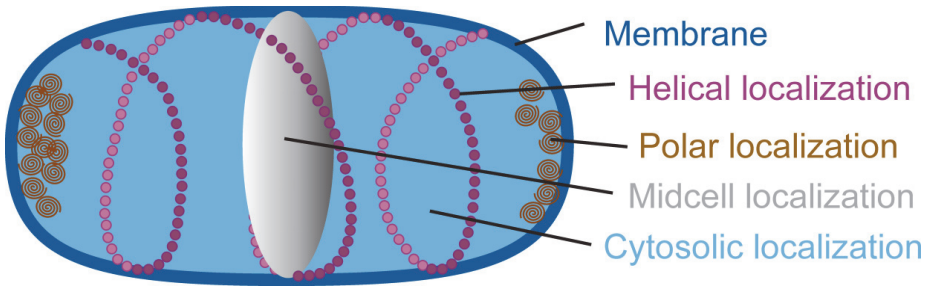


FIGURE 5 The most common protein localization patterns in *E. coli*.

The mechanism behind protein localization within bacterial cells exploits at least three possible routes. (1) *Diffusion and capture*. Proteins diffuse rapidly inside the cell and eventually encounters the destination proteins or surroundings and bind to the target. (2) *Active protein targeting systems*. Proteins are actively delivered to their target place by filamentous structures, for example, the bacterial cytoskeleton. (3) *mRNA targeting*. Some bacterial mRNAs are known to be capable of migrating to positions inside the cell where their protein products are required (Amster-Choder 2011). Currently, the general mechanisms behind protein localization are only vaguely known and new insights are being gained as research progresses. Especially the polar area of bacterial cell is an interesting subject for study. The poles are not identical resulting from the cell division and the old and new pole can hold separate functions relating to the survival of the bacterial cell.

2.6.1 Viral protein localization in bacteria

The complexity of bacterial cells raises the question whether bacteriophages exploit cellular features for their development in the same way as animal viruses. Animal viruses utilize the structure of the host cells very specifically and it is presumed that bacteriophages are as efficient. Still, the localization of viral proteins inside bacterial cells is poorly described.

The cellular poles are linked to viral infection and DNA intake since phages are known to prefer the polar areas for their entry. The poles are also known favorite sites for DNA intake in natural competent cells (Chen *et al.* 2005, Edgar *et al.* 2008). The *E. coli* phage λ (Rothenberg *et al.* 2011) and the *B. subtilis* phage SPP1 (Jakutyte *et al.* 2011), for instance, locate viral DNA injection and replication to the polar regions.

A favorable specific place for virus replication is likely to be the cytoskeleton as the MreB system is needed for efficient viral DNA replication of at least phages PRD1, SPP1 and *B. subtilis* phage ϕ29. An in-depth study has revealed the helix-like localization of several proteins and DNA involved in replication of phage ϕ29 near the cellular membrane (Munoz-Espin *et al.* 2009). Terminal proteins of ϕ29 and PRD1 have been shown to independently associate with the bacterial nucleoid without assistance from other phage-encoded proteins. These terminal proteins also harbor localization signals that lead them to the nucleus when expressed in mammalian cells (Munoz-Espin *et al.* 2010, Redrejo-Rodriguez *et al.* 2012).

The growing interest in viral protein localization in bacteria stems from a better knowledge of bacterial cells and their own protein localization. Those are raising study subjects opening multiple avenues to explore.

2.7 Fluorescence methods in virus research

The isolation of the green fluorescent protein (GFP) from jellyfish in the 60s (Shimomura *et al.* 1962), followed by its sequence determination and gene cloning in 1992 (Prasher *et al.* 1992), the first successful recombinant protein production in 1994 (Chalfie *et al.* 1994, Inouye and Tsuji 1994), and crystal structure determination in 1996 (Ormo *et al.* 1996) opened a new era for biological research. The significance of Shimomura's and Chalfie's work was acknowledged when they received the Nobel Prize in chemistry in 2008 for the discovery and development of the green fluorescent protein. Today, several modified versions of GFP are available with, for example, modified emission or excitation wavelengths, accelerated folding or improved folding efficiency (Merola *et al.* 2013, Campanini *et al.* 2013). Together with other fluorescent dyes and powerful fluorescence detection methods, GFP is nowadays considered an elementary tool in biological research. For example, a search for "fluorescence microscopy" on the NCBI PubMed-database will show that the amount of research papers and review articles has dramatically increased between publication years 2000 and 2012 (research papers: 2503/6769 and review articles: 66 /235).

Fluorescence microscopy has been extensively used for studying animal viruses. For example, the lifecycle of CPV has been studied in detail (Ihalainen *et al.* 2007, Ihalainen *et al.* 2009, Ihalainen *et al.* 2012, Niskanen *et al.* 2013). While the small size of bacterial cells used to be a limiting factor in microscopy, the continuous improvement of instrumentation and resolution are now offering

new perspectives for research. Both fluorescence based methods and electron microscopy are today widely used. In addition to basic imaging methods, several advanced techniques have been developed to gain more information from biological phenomena in both animal and bacterial cells.

Förster resonance energy transfer (FRET) is a useful phenomenon for the study of molecular interactions. By fusing together two proteins with two spectrally overlapping chromophores such as the GFP variants cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP), it is possible to monitor the protein-protein interactions *in vivo* (Pollok and Heim 1999). When the emission spectrum of the donor chromophore overlaps with the absorption spectrum of the acceptor chromophore, the FRET can occur (Fig. 6A). The efficiency of transfer is proportional to the inverse of the sixth power of the distance between the donor and acceptor molecules. The maximum distance between the fluorophores required for FRET is in the range of 1-8 nm (Förster 1965, Clegg 1996, Lakowicz *et al.* 1999).

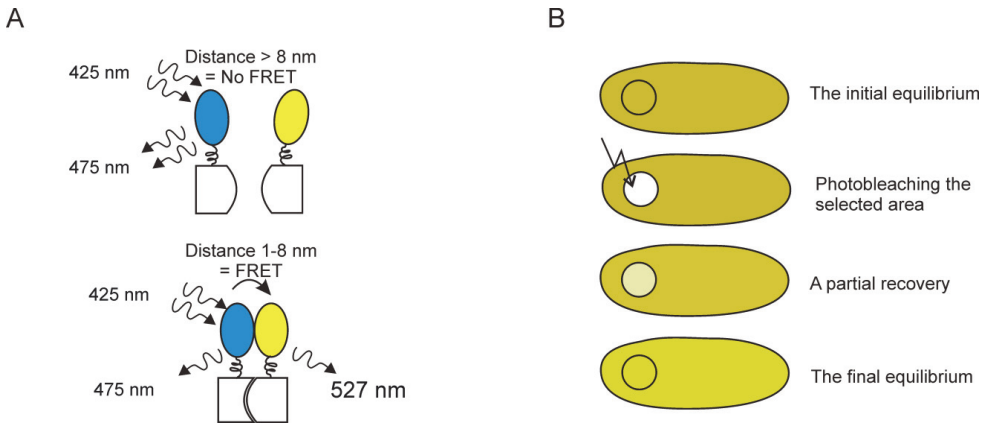


FIGURE 6 Schematic representations of FRET and FRAP. A) FRET. The binding of two proteins can be studied using tagged chromophores with overlapping emission spectrum of an acceptor molecule and excitation spectrum of a donor molecule. One potential chromophore pair is made of the cyan and yellow fluorescent proteins. The energy transfer can be detected either from the decrease of the CFP (donor) fluorescence lifetime or from the increase of the YFP (acceptor) emission. The corresponding maximal wavelengths are presented. B) FRAP. The fluorescence of a selected area within the cell is photobleached and the recovery of the fluorescence in that area is monitored. Because the photobleaching destroys the fluorophores ability to emit light, the recovery of fluorescence can be assigned to movement of proteins from the area surrounding the photobleached spot.

The FRET efficiency can be assessed by measuring either the decrease or increase of the amount of fluorescence of the donor or acceptor chromophores, respectively. However, this might be a challenge when measuring fluorescence in living cells, since the local chromophore concentrations are unknown. Another way to detect FRET is to monitor the fluorescence lifetime of the donor protein. The energy transfer decreases the lifetime of the donor independently of the

excitation light intensity or fluorophore concentration, and is therefore a more reliable way to study protein-protein interactions (Lakowicz and Szmacinski 1993). Fluorescence lifetime decay can be measured from part of an individual cell when linked to a microscope. Alternatively, a liquid sample with more traditional spectroscopic methods can be utilized.

Fluorescence recovery after photobleaching (FRAP) offers another way to exploit fluorescence and to measure the movement of proteins inside the cell (Fig. 6B). During imaging, a small area of interest is selected from the target cell and the fluorescence is photobleached in that specific area. Photobleaching switches the fluorophore to an excited triple state with long emission time. The fluorescence detected later on the bleached area comes from the diffusion of unbleached proteins into the area. The intensity of fluorescence is monitored and from the recovery curve, the diffusion rate of monitored molecules can be determined (Ishikawa-Ankerhold *et al.* 2012). With computational simulations, the theoretical molecular mass of diffusing molecules or complexes can be estimated, revealing information about the complex size, its movement and possible binding to bigger complexes (Schaff *et al.* 1997).

The increased use of fluorescence-based methods has gradually started to raise concern about the reliability of the research and especially the use of fluorescent proteins as a tag can alter the way the proteins function (Margolin 2012). This has been the case with bacterial cytoskeleton protein MreB, which was reported to display a helical localization pattern in several studies (Shih *et al.* 2005, van Teeffelen *et al.* 2011). A follow-up study combining several different methods has revealed the helical localization to be only an artifact, resulting from the fusion with the yellow fluorescent protein (Swulius and Jensen 2012). The same article raised concern about the very straightforward interpretation of the localization patterns. Especially, the difference between helical and patch-like localization in bacterial cells is easy to detect using focal stack analysis but frequently this method is omitted. The reliability of research can be at risk when reasonable unreliable human eye is used as a tool for analysis. It can lead to misrepresentation of the samples or distortion of the results. Advanced analytical methods are constantly being developed to improve quantitation of the microscopical data. One example is BioImageXD, an open source platform developed specifically to the three-dimension visualization and analysis of the biological data (Kankaanpää *et al.* 2012). In spite of the minor downsides fluorescence methods offer a unique addition to the traditional techniques in virus research and they should be further utilized in the future.

3 AIMS OF THE STUDY

In this thesis, the interactions of virus proteins with the host cell were studied using bacteriophage PRD1 and canine parvovirus as model viruses. The detailed aims of this research were as follows:

I. Construct a vector library for the production of fluorescent fusion proteins in bacterial cells.

II. To use the above vector library to express PRD1 and cellular fusion proteins and explore their localization inside *E. coli* cells.

III. To extend the use of fluorescent fusion proteins to the study of protein-protein interactions in bacterial cells using FRET.

IV. To study potential interactions of PRD1 proteins P33 and P17 with the cellular chaperone complex in *E. coli*.

V. To further describe the CPV entry mechanism by exploring the interactions of CPV particles with liposomes resembling cell membrane.

4 MATERIALS AND METHODS

The methods used in this thesis are summarized in the table below (Table 1). Detailed descriptions are found in the publications indicated by Roman numerals.

TABLE 1 Methods used in this thesis.

Method	Publication
Molecular cloning	I and II
Mutagenesis	I and II
Sanger sequencing method	I and II
SDS-PAGE	I and II
Western blotting	I and II
Complementation analysis	I and II
Rate-zonal centrifugation	I and II
Confocal microscopy	I and II
Virus culture and purification	I, II and III
Recombinant protein expression	I and II
Fluorescence lifetime microscopy	I
Time-correlated single photon counting	I
Fluorescence recovery after photobleaching	II
Fluorescence spectroscopy	III
Circular dichroism spectroscopy	III
Kinetics of PLA_2 enzyme activity	III
Small unilamellar vesicles (SOV), sonication method	III
Student's t-test	III
Protein concentration assay (Bradford method)	III

5 RESULTS AND DISCUSSION

5.1 Fluorescent fusion proteins

5.1.1 Vector library for production of bacterial fluorescent fusion proteins

Fluorescent proteins are nowadays routinely used in biological research and microscopy. The advantage of the method is that a gene of interest can be directly fused to the DNA sequence of a fluorescent gene in order to produce a single recombinant protein. This allows tracking down the location or fate of a fusion protein in the cell using fluorescence microscopy, for example. However, a non-negligible disadvantage of studying fluorescent protein fusions is the possible alterations of the protein native fold or function. Also an overexpression of the fusion protein can cause for example aggregation of the proteins.

To construct a library of fluorescent fusion proteins, we designed and built a number of vectors that can be efficiently used to rapidly generate fluorescent fusion proteins. Vectors pSU18 and pET24 were utilized in the process. These two vectors have compatible origins of replication (p15A and ColE1, respectively) and therefore, co-production of two different fluorescent fusion proteins is possible in one same cell. pSU18 is a low-copy number vector widely used in complementation analyses of PRD1 proteins (Bartolome *et al.* 1991, Bamford and Bamford 2000, Rydman *et al.* 2001) and pET24 is a commercial high copy-number expression vector.

The genes coding for fluorescent proteins eGFP, eCFP and eYFP were inserted into pSU18 and pET24, along with a linker sequence of six glycines and cloning sites (Fig. 7). Separate vectors were designed for the production of C- and N-terminal fusion proteins resulting in a final library of twelve different vectors in total (Fig. 7; I, Table 1). The gene coding for a protein of interest can this way be cloned in the chosen orientation, if amplified and delineated by the appropriate restriction sites.

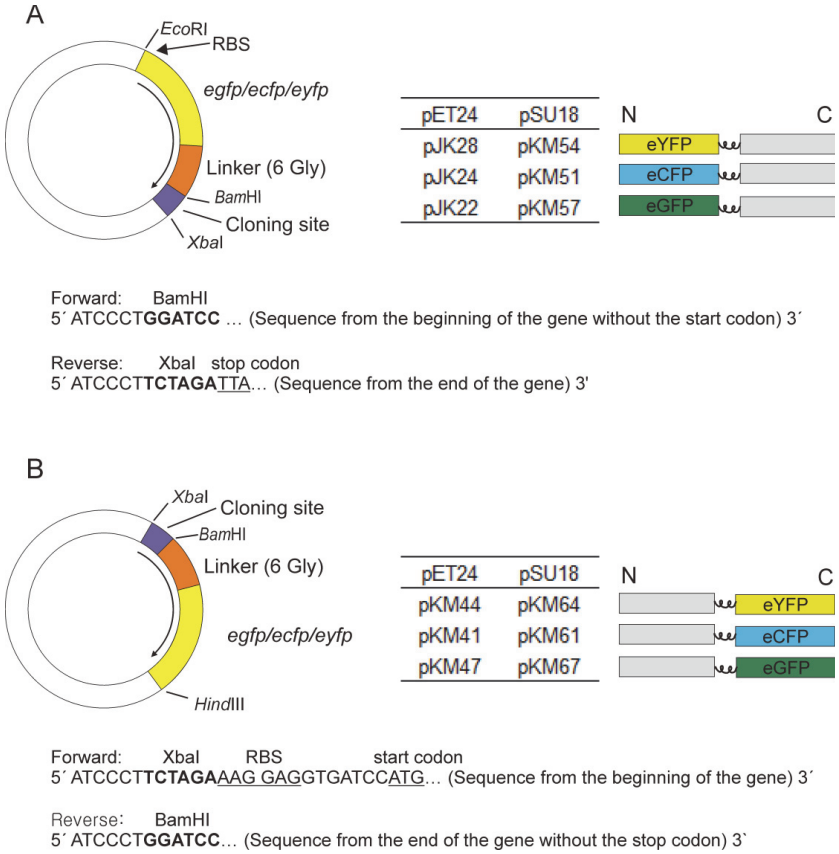


FIGURE 7 Schematic representation of the vector library used in this study. A) Vector for the production of N-terminal fusion proteins. B) Vector for the production of C-terminal fusion proteins. In both examples, a plasmid map with relevant restriction sites are shown on the left. Right, plasmid sources and names are indicated with the different fluorescent genes used in this study. Below each map, the forward and reverse sequences of primers used for gene cloning are given. Modified from I, Fig. 1.

In this work, we utilized the above vector library to generate C- and/or N-terminal fluorescent fusion proteins of several PRD1 proteins (I and II) and the cellular GroEL chaperonin protein (II). The PRD1 collection included multimeric structural proteins (P5, spike protein; P31, penton base protein; P3 major coat protein), a monomeric structural protein (P2, receptor binding protein), a structural integral membrane protein (P16, vertex stabilizing protein) and non-structural proteins (assembly protein P17 and putative assembly protein P33). It is noteworthy that the use of this vector library is not limited to the cloning of viral proteins and that it can serve to generate any fluorescent fusion proteins for as long as the fluorescent proteins and the linker are applicable.

5.1.2 Fluorescent fusion proteins are functional

Protein fusions might affect the nature of the proteins to which they are attached. The folded structure of proteins is not linear and adding a foreign N- or C-terminus can interfere with proper folding and therefore the function of the protein. The added fusion can also prevent the protein function even the protein is folded correctly. The structures of several PRD1 proteins used in this study are solved: proteins P2, P3, P5, P31 and P16 (Xu *et al.* 2000, Benson *et al.* 2002, Abrescia *et al.* 2004, Merckel *et al.* 2005). From these structures, the potential impact of added fluorescent extensions was assessed. With P31 for example, both termini are on the surface of the monomer, but in the pentameric form, the C-termini are located in the middle of the pentamer while the N-termini are pointing outwards (Fig. 8). This suggests that adding the fluorescent protein to the C-terminus is more likely to disrupt the formation of the multimer than when the fusion is present in the N-terminus.

The incorrect folding of proteins can cause aggregation. The produced fusion proteins were confirmed to be mostly soluble when supernatant and pellet of disrupted cells were analyzed by SDS-PAGE and Western blotting with either protein-specific antibodies, when available, or antibodies against the fluorescent protein (data not shown). Soluble supernatants were further analyzed by rate zonal centrifugation and Western blotting to determine the mass of the proteins and the existence of possible multimers (I, Fig. 2; II, Fig. 3). The rate zonal centrifugation separates the elements based on size and shape and the sedimentation of the fusion proteins was compared to selected controls to evaluate their oligomeric state. Fluorescent fusions of protein P2, known to be monomeric in the virion (Grahn *et al.* 1999), were seen to have sedimentation velocity corresponding to the monomeric state of the protein. This was also the case of P33-eYFP, the oligomeric state of which was previously unknown. Sedimentation velocity patterns of multimeric proteins (trimer P5, pentamer P31, trimer P3, tetramer P17 and tetradecamer GroEL) were slightly more diverse. In almost each case, some variation was detected between C- and N-terminal fusions implying that the fluorescent fusion had some effect on the formation of multimers. Still, multimeric forms were detected in all cases. The clearest variation between the two terminal fusions was detected with P31, where eYFP-P31 was multimeric while P31-eYFP displayed a wider mass distribution starting from the monomeric form. Fluorescent fusions of P3 and P17 also had a broad sedimentation distribution, starting from the monomeric state with the majority being in multimeric form.

The distribution of GroEL-eCFP was compared to a sedimentation pattern of *E. coli*'s native GroEL protein (II, Fig. 3B and C). The fluorescent fusion formed a slightly wider distribution but was still maintained as a multimer. It seems that the fluorescence tag did not totally prevent the formation of the barrel-like GroEL complex although the functionality of the complex remained unknown. The cellular natural GroEL is also present when the recombinant GroEL protein

is produced and a combination complex of natural and fusion proteins is possible.

The functionality of P5 and P31 fluorescent fusions was further evaluated by analyzing whether the fused proteins could function in the assembly of the virion. PRD1 virus mutants *sus690* and *sus525* with amber mutations in the genes coding for proteins P5 or P31, respectively were used in complementation analyses. The C- and N-terminal fusions of both proteins were analyzed. The mutant virus titer of a strain with the plasmid producing fusion proteins was compared to the titer of strains with similarly produced native proteins. As a negative control, a strain including pSU18 plasmid without insert was used. According to the complementation analysis (I, Table 2), all fluorescently tagged maintained their ability to assemble into the viral capsid as they complemented the defect of the corresponding proteins. The titers of strains expressing the fusion proteins were similar to those from the suppressor host or positive control. The titers of non-suppressor hosts and negative controls were at least five orders of magnitude lower. It can be concluded that the addition of a fluorescent tag does not interfere with proper folding of PRD1 proteins P5 and P31. The fluorescent fusions of PRD1 seem to correlate with the native protein functions and this was also with other used proteins.

5.2 Localization studies of PRD1 proteins

5.2.1 General PRD1 protein localization inside bacterial cells

The localization of PRD1 proteins P2, P3, P5, P16, P17 and P31 inside the bacterial cell was studied by confocal microscopy. The fluorescent fusion proteins were produced in *E. coli* strain HMS174(DE3), which is also a natural host for PRD1. At least three parallel cultivations were made for each strain and in addition to individual images, the percentage of cells having specifically localized proteins was calculated to gain a broader picture of the fusion protein behaviour (I, Fig. 3). To outline variations between samples, a profile of single representative cells was presented (I, Fig. 3).

Along with cytosolic protein localization, a common detected characteristic was a polar localization spot (i.e. polar locus). These loci were mainly detected with multimeric proteins, for example with P5, P3 and P17 (I, Fig. 3E, F, I, J and K). In line with the results obtained using rate zonal centrifugation, variance between C- and N-terminal fluorescent fusions of P31 were also detected in the localization patterns. The more monomeric P31-eYFP was mostly evenly distributed while the clearly multimeric eYFP-P31 localized to one pole of the cell (I, Fig. 3G and H). Both terminal versions of the monomeric P2 were found evenly distributed in the cytosol like eYFP expressed alone (I, Fig. 3C, D and L). This was also the case with membrane protein P16 (I, Fig. 3A and B).

The amount of loci in individual cells varied between different fusion proteins and the distribution was calculated from samples with clear localization

patterns (I, Fig. 4). From this analysis, the most interesting observations were made for proteins P17 and P3. With P5 and P31, most of the cells (75 – 100 %) had only one polar locus. P17-eYFP had a slightly wider distribution with over 40% of cells having more than one locus. As for the eYFP-P3 fusion, 80 % of cells had more than two loci and the amount of loci rose to the point where eYFP-P3 was present at 4-7 loci per cell (34 %). Also, with P3-eYFP two loci per cell were more frequently detected than a single locus (>60 % of cells). The ongoing cell division in some samples offer one possible explanation especially between one or two polar loci. Here the cell cultures were always prepared the same way and the analyzed images were taken from three parallel cultivations. Additionally, a significant difference in cell size was not detected between samples. It is therefore unlikely that the division of the cells had a notable influence on the results between the fluorescent fusions.

In addition to the localization of single proteins, attention was paid to co-localization of proteins P5 and P31, which are known to interact in the virus particle (Caldentey *et al.* 2000). P5 and P31 seemed to share similar localization patterns. When all loci were analyzed, fluorescent fusions P5-eCFP and eYFP-P31 co-localized in 66 % of the loci (I, 2.4.). The remaining loci without co-localization were mostly in cells where only one of the fluorescent signals was detected and presumably only one of the fusion proteins was produced. At least with these proteins, the localization seemed to concentrate to the same locus.

TABLE 2 Combined results from the rate zonal centrifugation analysis and localization observed by confocal microscopy.

Protein	Oligomeric state in virion	Monomer (kDa)	Fusion protein (kDa)	Observed oligomeric state	Observed localization
P5-eYFP	trimer	34.2	61.1	oligomer	polar locus
eYFP-P5	trimer	34.2	61.1	oligomer	polar locus
P16-eYFP	monomer	12.6	39.5	-	cytosolic
eYFP-P16	monomer	12.6	39.5	-	cytosolic
P2-eYFP	monomer	63.7	90.6	monomer	cytosolic
eYFP-P2	monomer	63.7	90.6	monomer	cytosolic
P31-eYFP	pentamer	13.7	40.6	monomer	cytosolic
eYFP-P31	pentamer	13.7	40.6	oligomer	polar locus
P3-eYFP	trimer	43.1	70	oligomer	polar locus
eYFP-P3	trimer	43.1	70	oligomer	several loci
P17-eYFP	tetramer	9.5	36.4	oligomer	polar loci
eYFP	-	26.9	-	-	cytosolic

Earlier studies have revealed the PRD1 terminal protein P8 interacts with the bacterial nucleoid (Munoz-Espin *et al.* 2010), but the localization of other PRD1 proteins has not been reported. Here, variation was detected between the evenly distibuted monomeric proteins and multimeric proteins which formed more

often polar loci (Table 2). The polar localization might correlate with the oligomeric state of the proteins. Another possible explanation is that the PRD1 particle assembly occurs at a specific area of the cell e. g. in nucleoid-free cell poles, where individual proteins are targeted. It is known that the lipid composition of the PRD1 inner membrane differs from that of the host bacterium. PRD1 assembly is likely to occur at specific locations, possibly controlled by the localization of suitable membranes.

The localization studies by confocal microscopy offer a new perspective for determining the stage of the phage lifecycle and the next step could be to implement similar measurements to study the infection process. A probable downside is that, presently, fluorescence can be detected only from plasmid-produced fusion proteins, which are not expressed in coordination with wild-type viral proteins during assembly. Also, a larger co-localization mapping of the different PRD1 proteins could be carried out to further solve the behavior and interactions of the proteins. At the same time, the restrictions of imaging as a reliable source for information should be taken into account and the possibilities for higher numeric analysis utilized.

5.2.2 Interaction studies of spike proteins P5 and P31 by FRET

One aim behind the construction of the vector library was to expand the use of fluorescent proteins to the interaction studies using the FRET technique. The resolution limit for detection of co-localization with confocal microscope is around 200 nm whereas the distance for detection of FRET is 10 nm. FRET offers therefore more accurate way to study the possible interaction of proteins. The co-localization of proteins P5 and P31 was further analysed using fluorescence lifetime measurements with eCFP as donor and eYFP as acceptor fluorophores. Three different techniques were used to overcome possible technical limitations and to ensure the reliability of the results.

Spike protein P5 and penton protein P31 are both located in the viral capsid (Caldentey *et al.* 2000). The orientation of the P5 spike complex and the structure of P31 have been solved. With P5, the structure of the C-terminal knob is solved, but the central flexible area has restrained solving the structure of the N-terminus (Merckel *et al.* 2005). According to this structural information, an ideal fluorescent protein combination to promote the binding of the proteins is P5-eCFP and eYFP-P31 (Fig. 8). The energy transfer between fluorescently tagged proteins P5 and P31 was tested using samples representing the ideal combination (P5-eCFP + eYFP-P31) and the opposite, where the extensions hinder the interaction (eYFP-P5 + P31-eCFP). As a negative control, P31-eCFP was measured alone. For comparison the measurements were made in two scopes: the locus area and the whole cell.

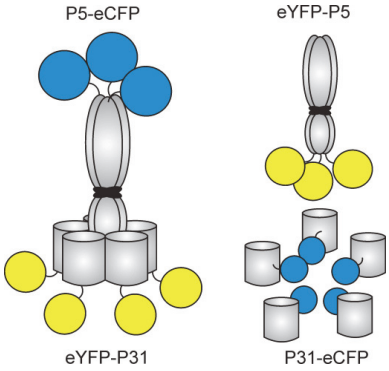


FIGURE 8 Speculative illustration of the effect of fluorescent fusions on the folding of multimeric proteins P5 and P31 and further, on the interaction of the proteins. The proteins are known to interact as a part of the spike vertex complex and the orientations of the proteins are known (Caldentey *et al.* 2000). The location of protein termini influences the location of the fluorescent fusions, which might hinder the proper folding of the monomeric protein or its assembly into a multimer. The fluorescent fusion situated at the N-terminus of P5 might prevent its interaction with P31. If the fusion is at the C-terminus, it does not interfere with the interaction. With P31, the terminus of the fusion might have a critical impact on the formation of multimeric structure.

The first approach was to measure the changes in lifetime of the donor protein eCFP. The FRET between two proteins can be studied by determining the lifetime decrease of the excited state of the donor in the presence of the acceptor. Fluorescence lifetime was measured from individual fixed cell samples using inverse time-resolved fluorescence lifetime microscope (FLIM) and from liquid samples with time-correlated single photon counting (TCSPC). The third approach was to use fluorescence spectra measured by confocal microscopy where the effect of FRET should appear as an additional peak at the eYFP emission wavelength (maximum peak at 525 nm), when eCFP is excited. Samples were prepared in a similar way as those for FLIM measurements and the spectra were compared in the specific polar loci and outside the loci areas.

A minimal energy transfer, i.e. no significant FRET, was observed in the locus area of sample cells where orientations of fluorescent proteins were theoretically ideal (Fig. 8) whereas other samples or areas did not produce a similar signal (I, supplementary material). The results obtained with all three techniques were consistent. Previously, the efficiency of the energy transfer between CFP and YFP has been observed to be significantly better when compared to our results, for example when measured the cleavage of apoptotic facilitator Bid (Onuki *et al.* 2002). It is possible that P5 and P31 do not have direct interaction when tagged with fluorescent fusions and they only localize in same locus. It is also possible that the proteins do not have direct interaction when co-expressed even though they are bound in the virus capsid. The length of the linker has a major influence on the efficiency of the energy transfer (Shimozono and Miyawaki 2008). The fluorescent fusions used here had a linker of six glycines and the observed results could have been different if the length and peptide structure of the linker would have been different. For example,

Lissandron et al. analyzed a FRET based indicator for cyclic adenosine monophosphate (cAMP) and started from somewhat similar levels of energy transfer efficiency as observed by us, but after optimization of the linker they reached improved energy transfer efficiency (Lissandron *et al.* 2005). The orientation of proteins P5 and P31 was theoretically conceivable for detection of FRET, but still, the fluorescent proteins did decidedly not reach each other optimally. A minor energy transfer was systematically observed with all three methods applied, indicating it to be real. In future work for example the optimization of linker might be beneficial next step.

5.3 PRD1 proteins P33 and P17 have influence on the *E. coli* chaperone complex

5.3.1 Genes encoding PRD1 P33 and P17 reside in the assembly operon

The interaction between GroEL and GroES has been defined to occur through a mobile loop of GroES, including three hydrophobic residues surrounding flexible glycine residues (Landry *et al.* 1993). The flexibility has shown to play a crucial role in the function of the loop (Nojima *et al.* 2012). Even though the gene sequences of known bacteriophage-coded co-chaperone proteins do not resemble the sequence of wild-type GroES, a special hydrophobic area has been found from their sequences matching the GroES loop (Hunt *et al.* 1997, Ang *et al.* 2001). In some cases, the protein structure of the co-chaperone protein is highly similar to the GroES structure (Hunt *et al.* 1997). Sequence comparison between GroES, viral co-chaperone proteins Gp31 and CocO, and P33 or P17 does not reveal substantial similarity (II, Fig. 6, for P17 data not shown). Nevertheless, the amino acid sequence of P33 reveals a similar hydrophobic motif with three residues as also detected in the viral co-chaperone proteins. However, surrounding flexible residues observed in GroES were not found in P33 (II, Fig. 6). The structures of proteins P33 and P17 have not been determined and therefore, a structural comparison is not possible.

PRD1 proteins P33 and P17 are encoded in the late operon 2 (OL2) and the precise function of these proteins is unknown. The OL2 includes also genes encoding the packaging efficiency factor P6, non-structural assembly factor P10 and packaging ATPase P9 (II, Fig. 1). With regard to their genome location, the functions of P33 and P17 are expected to be related to the virus assembly process. The location of the genes one after the other, also implies a possible co-operation of P33 and P17. These two proteins are highly conserved in PRD1-like phages indicating a general importance for the phage vitality.

5.3.2 PRD1 proteins P33 and P17 complement the defect of GroES in *E. coli* strain and in PRD1 infection

A mutant *E. coli* strain DW719, carrying a temperature-sensitive *groES619* mutation, was exploited to analyse the functions of P33 and P17. In this mutant, a glycine preceding the hydrophobic loop of GroES is mutated to an aspartic acid (G24D). As a result the mobility of the loop is decreased and the ability of the GroES produced to bind to GroEL is reduced. GroEL/GroES-complex is essential for the bacterial cell and a mutant strain defective for GroES is not viable. Still the heat sensitive mutation can be used. The mutation appears at the non-permissive temperature of 42 °C, whereas in permissive conditions at 37 °C the mutation is blocked and native GroES is produced. In complementation analysis the mutant strain was introduced to a plasmid producing selected protein(s). If the vitality of the strain is improved at a non-permissive temperature the plasmid-produced protein can complement the mutation. The complementation of the weakened GroES was tested with PRD1 proteins P17, P33 and both. A plasmid-produced wt GroES was used as a positive control and the pSU18 plasmid without any insert as a negative control. The wild-type bacterial strain DW720 was used to monitor the wild type levels of colony forming units (cfu) in similar conditions.

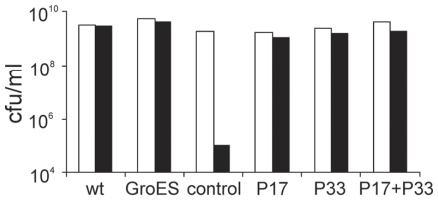


FIGURE 9 The vitality of *E. coli* bacterial strains with temperature-sensitive GroES mutation overexpressing various PRD1 proteins. Bacterial cultures were made at permissive (37 °C, white bars) or non-permissive (42 °C, black bars) temperatures. Modified from II, Fig. 2.

The changes in viability of the mutant strains were analysed in the absence and presence of the above mentioned PRD1 proteins. The levels of colony forming units were similar in the wild-type strain and mutant strain containing the plasmid encoding wt GroES at both temperatures (Fig. 9). The negative control showed six orders of magnitude lower cfu values when cultivated at non-permissive temperature whereas in the permissive temperature the cfu count was comparable with the wt strain and the positive control. These controls showed that the mutation appeared correctly and the mutant strain was functioning as expected. The tested PRD1 proteins seemed to complement the defective GroES in all cases: P33 or P17 expressed alone, or when co-expressed, resulted in cfu levels of the same range as the positive controls at both temperatures.

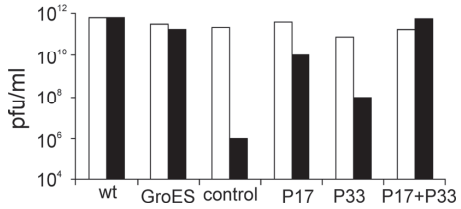


FIGURE 10 PRD1 titers of *E. coli* bacterial strains with temperature-sensitive GroES mutation overexpressing various PRD1 proteins. Bacterial cultures were made at permissive (37 °C, white bars) or non-permissive (40 °C, black bars) temperatures. Modified from II, Fig. 2.

We also studied the function of proteins P33 and P17 during PRD1 infection. The PRD1 infection is dependent on the GroEL/GroES complex, because it is essential for the complete multimerization and protein folding for several virus proteins (Hänninen *et al.* 1997). The above strains were used as host and wt PRD1 titers were determined. In this case 40 °C was used as non-permissive temperature to ensure plaque formation. The controls produced titers as expected: the wt strain and the positive control showed similar titers at both temperatures. The negative control had a similar titer in permissive temperature but it reached a titer approximately six orders of magnitude lower at a non-permissive temperature showing that the mutation blocks the multiplication of the phage (Fig. 10).

Titers of PRD1 on the protein producing strains (P17, P33 and both) showed more variations when compared to the results of bacterial viability studies. At the permissive temperature titers obtained from all PRD1 protein producing strains followed the levels of positive controls, as expected. At the non-permissive temperature only the co-expression of P17 and P33 reached similar levels. The production of P17 alone resulted in a slightly lower titer than the positive control (10^{10} versus 10^{12}). When P33 was produced alone, the titer was even lower (10^{8}), still reaching two orders of magnitude higher than the negative control. Both proteins complemented the defect for GroES at some level and when co-expressed the complementation of GroES was more substantial. The results obtained from complementation analysis suggest that both P17 and P33 have a function relating to the GroEL/GroES complex.

5.3.3 The localization of P33-eYFP is altered when co-expressed either with P17-eCFP or GroEL-eCFP

The fluorescent fusion proteins were utilized in confocal microscopy to study the localization of protein P33 in *E. coli* HMS174(DE3) cells. The yellow fluorescent fusion P33-eYFP was tested with C-terminal cyan fusions of P17 and GroEL. When P33-eYFP was expressed alone, it was evenly distributed in the cytoplasm of the cell (II, Fig. 4). Several cultures were made in parallel and protein localization was always the same. A similar distribution pattern was detected for the monomeric PRD1 protein P2 (see section 5.2.1; I, Fig 3). Surprisingly, when

P33-eYFP was co-expressed with P17-eCFP or GroEL-eCFP, the localization pattern changed to follow the polar localization observed with the production partner (II, Fig. 4). Similar polar localization patterns were also detected with several other multimeric PRD1 proteins (see section 5.2.1; I, Fig. 3). This was not observed when eYFP was expressed alone or co-produced with P17-eCFP or GroEL-eCFP as eYFP always remained evenly distributed in the cytosol (data not shown). This suggests that the change in localization is triggered by the presence of P33. An interaction with GroEL and P17 would be a probable explanation for the changed localization of P33. However, the observed co-localization could also result from another complex cellular process, unidentified at this stage.

5.3.4 The mobility of P33-eYFP within the *E. coli* cell

The behaviour of P33-eYFP in *E. coli* cells was analysed in more detail by FRAP-assay. Using FRAP, the mobility of a protein can be followed and for example, its diffusion coefficient can be determined. Together with computer simulations, estimations of the mass of the detected complex can be made. The principle of the assay and single cell recovery curves are presented in article II, Fig. 5A and B. The fluorescent recovery was detected and averaged from 13 separate cells in all presented results. Based on the simulations, the diffusion coefficient and a mass estimation were obtained for freely moving complexes.

The mobility of proteins eYFP and P33-eYFP was measured separately (II, Fig. 5C). The mobility of eYFP followed previously reported results and based on our simulations, the diffusion coefficient was 5 $\mu m^2 s^{-1}$. For example, Elowitz et al. have earlier measured a diffusion coefficient of 7.7 ± 2.5 $\mu m^2 s^{-1}$ for GFP and Cluzel et al calculated CheY-GFP fusion (40 kDa) to have a diffusion coefficient of 4.6 ± 0.8 $\mu m^2 s^{-1}$ using fluorescence correlation spectroscopy (Elowitz *et al.* 1999, Cluzel *et al.* 2000). The simulated diffusion coefficient is in accordance with the simulated free mobility of eYFP sized protein within bacterial cells. The recovery curve of P33-eYFP differed from that of eYFP beyond the value that an increase in the molecular mass should inflict (26.9 kDa versus 34.4 kDa), indicating alterations in the movement of the protein. The simulations showed two populations with diffusion coefficients of 2 $\mu m^2 s^{-1}$ (91%) and 0.4 $\mu m^2 s^{-1}$ (9%), whereas, based on the theoretical molecular mass, the theoretical diffusion coefficient should be 4.6 $\mu m^2 s^{-1}$. It can be concluded from these observations that protein P33 may have a direct interaction with a cellular factor/complex or there might be another factor hindering its motion. According to the retardation of diffusion of the faster P33-eYFP population and computational simulations, the complex size of P33-eYFP was determined to be approximately 400 kDa. That is within the same range as the single heptameric GroEL-ring (420 kDa). Charbon *et al.* have utilized FRAP-assay for *E.coli* GroEL-complex and gained diffusion coefficient of D_{app} = 0.16 ± 0.15 $\mu m^2\ s^{-1}$ with single population fitting (Charbon *et al.* 2011). This corresponds to the coefficient determined for the slower population of P33-eYFP (0.4 $\mu m^2 s^{-1}$). Interaction of P33 with the cellular endogenous GroEL-GroES complex offers one potential explanation for the detected diffusion.

The diffusion of P33-eYFP decelerated even more when it was co-expressed with either P17-eCFP or GroEL-eCFP (II, Fig. 5D). In spite of the co-localization of these proteins reported in section 5.3.3, only cells with diffuse cytoplasmic localization were selected for the measurement with both combinations. The fluorescence was detected only from the eYFP emission wavelength and the diffusion of P17-eCFP or GroEL-eCFP was not measured. Similar populations and diffusion coefficients (2 $\mu m^2 s^{-1}$ and 0.4 $\mu m^2 s^{-1}$) as with P33-eYFP alone were obtained, but the ratio between the faster and slower migrating populations changed. Co-expression of P33-eYFP and GroEL-eCFP raised the slower population percentage from 9 % to 40 %. Therefore, addition of fluorescently tagged and overproduced GroEL may have an influence on the cellular endogenous GroEL. If P33 binds the chaperonin complex, over expression of GroEL-eCFP offers more binding partners to P33-eYFP and a change in the mobility pattern can be easily explained.

Co-expression of P33-eYFP and P17-eCFP augmented the slower migrating population, with 67 % of P33-eYFP showing hindered motion. The complementation analysis (see section 5.3.2) showed that either P17 or P33 can complement the mutation in the GroES binding loop to some extent. Nevertheless, the mutation was only fully complemented when the proteins were co-expressed. Assuming that the slower migrating population of P33-eYFP is bound to the chaperonin complex, the addition of P17-eCFP raises the binding efficiency further. It suggests that P33 and P17 might have a united function related to the GroEL/GroES complex.

5.3.5 The influence of P33 and P17 on the chaperone complex

The obtained results from complementation analysis and confocal microsopy imply that PRD1 proteins P33 and P17, working either alone or in co-operation, have functions related to the GroEL/GroES complex. The utilization of GroEL/GroES complex by the phage world is not unusual. PRD1 assembly has earlier been shown to be dependent on the chaperonin complex. Especially proteins P3 and P5 remained in monomeric state in the absence of GroEL or GroES and the assembly of P11 and small membrane proteins was affected (Hänninen *et al.* 1997). Several GroES-replacing co-chaperone proteins have been identified from different bacteriophages, recently a GroEL-ortholog has also been characterized (van der Vies *et al.* 1994, Keppel *et al.* 2002, Kurochkina *et al.* 2012). An interesting chaperone related finding is also the gene-product 39.2 of phage RB69 (Ang and Georgopoulos 2012). It does not substitute the function of either chaperone protein but instead, alters the conformation of GroEL into a more favourable form for the phage. In addition to this revealed specific function, there could be other alternative ways, at this stage unidentified, for phage-encoded proteins to alter the chaperone complex for their own benefit.

The heat sensitive mutation *groES619* used in the complementation analysis only reduces the binding efficiency of GroES and GroEL through the loop structure containing hydrophobic residues. Therefore, the complementation of the mutation does not require replacement of the whole function of GroES, but

only to improve the binding of GroES to GroEL. The proteins P33 and P17 might form a GroEL substituting co-chaperone complex in co-operation, but a more likely explanation would be a so far unknown mechanism, possibly similar to Gp39.2 of phage RB69, relating to the complex. P33 and P17 can form a united structure or alternatively, function independently and affect the same GroEL/ES complex.

In addition to the essential viral proteins, there are virus proteins having only an accelerating effect on the viral infectivity, or else, proteins that are only required in special circumstances. For example, there is a structural protein in PRD1 that is crucial for infection of only a specific range of host bacteria. The amber mutation in the gene encoding transmembrane protein P16 narrows down the host range and the mutant viruses can infect only cells with full-length lipopolysaccharides (Jaatinen *et al.* 2004). The host range of PRD1 is extremely broad as the infection is dependent on the presence of a widespread receptor-encoding conjugative plasmid. One can presume that there are proteins functioning only in specific hosts or in determined environments, such as shown for phage T4 (Kai *et al.* 1996). P33 could also exhibit this kind of special function in PRD1 by improving the survival of the phage in novel surroundings.

5.4 Membrane interactions of canine parvovirus

5.4.1 CPV capsid structure alterations in contact with lipid membranes

In order to have successful infection, viruses that enter the host cell through endocytic pathway need to escape from the endosomal vesicles during the infection. The low pH of endosomal vesicles acts usually as a trigger for the escape. Enveloped viruses can penetrate to the lumen by acidification-induced fusion with endosomal membrane and non-enveloped viruses normally undergo conformational changes exposing for example proteolytic activity that can be used in the escape.

CPV is non-enveloped animal virus with icosahedral (T=1) capsid. It enters the host cell through transferring receptors and clathrin-mediated endocytosis route. During the viral entry and escape from endosomal vesicles, CPV viral particles are eventually forced to be in contact with the vesicle membranes. To analyze these contacts in more detail, we applied tryptophan fluorescence measurements and probed conformational changes in the CPV capsid. Vesicles in the endosomal route are relatively small and their structures are unstable because vesicles can rapidly evolve from late endosomes to lysosomes. Therefore, the detailed study of the cell membrane composition required for CPV endosomal escape is difficult using microscopic methods. Novel methods are being developed for the study of endosomal escape, for example a single-cell based imaging assay developed by Suomalainen et al. (2013).

In the CPV capsid, proteins VP-1 and VP-2 encompass approximately 850 tryptophan residues, which are located inside the capsid subunits (Chapman and

Rossmann 1993, Xie and Chapman 1996). In hydrophobic surroundings, these residues have a maximum emission at a wavelength around 330 nm but it shifts to 350 nm when the residues are in a more hydrophilic environment. This feature was exploited by exposing the CPV capsid to acidic conditions in presence of liposomes resembling cell membranes. Previously it is known that the PLA_2 activity located in the N-terminus of VP-1 is exposed in acidic conditions.

The measurements were made in the presence of 1 mM $CaCl_2$, as the binding of Ca^{2+} ions to CPV is known to stabilize the capsid structure (Simpson *et al.* 2000). The difference in tryptophan emission peaks between pH 5.5 and pH 7.4 was determined in the presence of five different membrane vesicles consisting of either sphingomyelin, phosphatidyl serine, phosphatidyl choline, phosphatidyl ethanolamine or phosphatydyl inositol. With sphingomyelin vesicles at pH 5.5, a statistically significant shift of emission peak was observed in the tryptophan fluorescence spectrum (III, Fig. 1A). A similar but statistically insignificant shift was detected at pH 7.4 and also, at both pH values in the phosphatidyl serine liposomes (III, Fig. 1A). Other tested liposomes did not cause noticeable shifts in the emission peak positions (data not shown). The change of the emission peak in the presence of sphingomyelin could be caused by structural changes in the CPV capsid. Those changes bring the tryptophan residues closer to the surface of the capsid or alternatively, weaken the structure in a way that water can enter the capsid. It is notable that these changes require both the acidic pH and the presence of sphingomyelin liposomes. Therefore, the acidic surrounding of late endosomes alone cannot cause the detected changes. In general, model lipid membranes are a good way to mimic cells natural membranes and in addition to the study of viral entry they are utilized in the discovery and testing of novel drugs (Peetla *et al.* 2009).

5.4.2 The alterations in the secondary structure of CPV capsid

The conformational changes in the secondary structures of CPV capsid were further studied with CD spectroscopy in the far UV region of 250-190 nm. The effect of liposomes comprised of sphingomyelin and phospatidyl serine in a molar ratio of 1:1 on CPV particles was measured in the presence of $CaCl_2$ at pH 5.5 and 7.4. The effect of acidic pH in the absence of the membrane vesicles was also measured. Only the characteristic changes in the spectra were observed and the data was not interpreted in terms of detailed secondary structure.

The overview of measured spectra of empty CPV capsids featured a negative valley at 220-200 nm and a positive peak at 190 nm (III, Fig. 1B). In the presence of $CaCl_2$ only, without the membrane vesicles, significant differences in CD spectra were detected between pH values 5.5 and 7.4. The addition of liposomes altered the CD spectra of the capsids in a different way (III, Fig. 1B). In general, in the presence of liposomes, the difference between the two pH values was significantly smaller compared to the situation without membrane vesicles. The CD spectroscopy supports the results gained from tryptophan fluorescence measurements as the CPV capsid alteration was detected in the presence of sphingomyelin-phosphatidyl serine vesicles. On the other hand, the pH change

in absence of the membrane vesicles had an even bigger effect on the capsid structure. The differences caused by the membrane vesicles and acidic pH were not similar, which was the case also with tryptophan fluorescence measurements. This might suggest two separate steps in the conformational changes: one caused by acidic pH and another due to contact with lipid membranes.

Tryptophan fluorescence and CD spectroscopy were utilized to further study the long term effect of acidic environment on CPV. Viral particles were exposed to pH 5.5 for a period of 10 minutes and neutralized. The treated viral particles were compared to native viral particles with both techniques. Tryptophan fluorescence revealed a minor change in the emission peak between the samples but it was proven to be statistically insignificant (III, Fig. 1D). CD-spectra collected from region 240-200 nm revealed a deeper valley of 200-220 nm with native viral particles (III, Fig. 1 E and F). The differences between neutralized native and acid-treated capsids imply long term changes on the capsid supporting the existence of separate conformational changes. Those changes may be needed in the following steps of the infection after the viral particle has penetrated the cytosol.

5.4.3 The escape from endosomal vesicles requires another mechanism in addition to PLA_2 activity

The capsid of CPV is mainly formed of proteins VP-1 and VP-2. The structures of these proteins are almost identical. The only difference is a unique N-terminus of VP-1, where the PLA_2 motif required for the escape of endosomal vesicles and the nuclear localization signal needed in the following steps of the infection are located. The terminus is originally located inside the capsid but it becomes exposed in the presence of acidic pH or urea treatment (Weichert *et al.* 1998, Cotmore *et al.* 1999, Vihinen-Ranta *et al.* 2002, Suikkanen *et al.* 2003).

In general, the escape of non-enveloped viruses from endosomal vesicles can proceed according to two mechanisms: by lysis and total destruction of the vesicles or by the creation of small well-defined pores from which the genome or whole particle is released into the cytosol. The PLA_2 activity of CPV revealed in the acidic conditions of late endosomes would suggest a simple lytic escape. Nevertheless, the escape occurs through small pores because the endosomes of CPV-infected cells remain functional. These pores only allow particles with the maximum mass of 10-20 kDa to pass (Parker and Parrish 2000, Suikkanen *et al.* 2003, Vihinen-Ranta *et al.* 2004). Presumably, there are more complex functions than the PLA_2 activity behind the CPV endosomal escape, which ensure the formation of the pores. The changes observed with tryptophan fluorescence and CD spectroscopy suggest a variety of conformational changes in the CPV capsid. pH changes alter the capsid in a different manner compared to the sphingomyelin and phosphatidyl serine. This can be concluded from CD spectra where the effects of decreased pH differ significantly depending on the presence of membrane vesicles.

To further study the changes in CPV capsid, the phospholipase A2 activity was tested in the presence of sphingomyelin. It was then compared to PLA_2 from

bee venom as a control. The changes in activity induced by sphingomyelin were similar between CPV PLA_2 and the control bee venom PLA_2 (III, Fig. 2). The activity decreased slightly in the presence of liposomes, but the variation was found to be statistically insignificant. The presence of sphingomyelin did not affect the PLA_2 activity of CPV. These findings suggest that the PLA_2 activity is not the only release mechanism in action in the presence of acidic pH and membranes. It is therefore likely that the CPV capsid contains several different mechanisms to assist its escape from the endosomal route.

6 CONCLUDING REMARKS

The main conclusions of this thesis are as follows:

I The constructed vector library offers a useful tool for the expression of fluorescent fusion proteins with a broad range of downstream applications. Its use can be extended to the expression of all kinds of proteins to be investigated through a fluorescent tag.

II PRD1 proteins display various localization patterns inside *E. coli* host cells. Cytosolic localization and polar loci were the most common patterns observed. The localization of viral proteins inside bacterial cells is poorly studied but it can reveal new insights into the viral lifecycle and offer new perspectives for virus research. FRET, once optimized, is a powerful technique to monitor the interactions between proteins.

III The functions of PRD1 proteins P33 and P17 are related to the chaperone complex of the host *E. coli* cell. The proteins can complement for GroES deficiency, both during bacterial growth and PRD1 infection. Also, the mobility of P33 corresponds to the mobility of the GroEL complex, especially when co-expressed with GroEL-eCFP or P17-eCFP.

IV The CPV capsid undergoes structural changes when the acidic environment of endosomes is combined to the presence of sphingomyelin vesicles. The alterations were studied using tryptophan fluorescence and CD-spectroscopy. The presence of membranes did not have an influence on the PLA_2 activity of the capsid, suggesting that there is an additional, still unknown mechanism for the escape from the endosomal vesicles.

Acknowledgements

This work was carried out at the University of Jyväskylä, at the Department of Biological and Environmental Science between years 2007 and 2014. This work was funded by Academy of Finland Centre of Excellence in Virus Research (2006-2011) and Academy of Finland Centre of Excellence in Biological Interactions Research (2012-2017). I received much appreciated travel grants from Finnish Concordia Fund, Oskar Öflund foundation and National Doctoral Programme in Nanoscience (NGS-NANO).

I would like to truly thank everyone that has helped me during my PhD project. First I want to express my greatest gratitude to my supervisors Jaana Bamford and Hanna Oksanen. You have introduced me to the world of bacteriophages and taught me the scientific way to work and think. I want to thank Professor Markku Kulomaa for acting as an opponent. I acknowledge the official reviewers Johanna Laakkonen and Vesa Hytönen, your proficient comments helped me make the thesis better. Thank you Nadine for the valuable comments for the thesis.

I want to thank all the co-authors of the original publications. Kirsi Pakkanen and Matti Vuento, you gave me the very first lessons of scientific work. Teemu Ihalainen, you introduced me to the world of microscopy and kindly helped in any problems I might have encountered. Janne Ihalainen and Heli Lehtivuori, thank you for the introduction to the sometimes very tricky fluorescence lifetime spectroscopy. During my work I have had the privilege of supervising some great Master's thesis students, all thanks to Sari, Kati and Johannes.

I am grateful for the friendship and collaboration of Sari Mäntynen. We have had both scientific and non-scientific adventures together more than I can remember! Petri, Elina and Matti, you have been there from the beginning and we have shared great moments sharing the same office room. I would also extend thank you to all other present and former JB group members. You have provided me with a great working environment. I will miss you, especially the floor ball crowd.

Thank you Outi for all the adventures. My dear friend Eeva, you have walked with me all the over ten years we spent at the University of Jyväskylä. Collectively I want to thank all my other friends, colleagues and important people in my life during these years.

All of my gratitude to my parents and sister, you have always supported me and led me to believe I can achieve anything I want. Finally, my deepest gratitude goes to Oke. Thank you for standing next to me especially in the long nights during the last months. You have always helped and encouraged me, and most of all shared my life, distress and laughs.

YHTEENVETO (RÉSUMÉ IN FINNISH)

Virusproteiinit ja niiden vuorovaikutukset isäntäsolussa

Virukset ovat muinaisia loisia, jotka pystyvät lisääntymään ainoastaan valtaamalla isäntäsolun aineenvaihdunnan omakseen. Tässä väitöskirjatutkimuksessa tutkittiin eri virusproteiinien vuorovaikutuksia isäntäsolun kanssa. Tutkimuksessa käytettiin malleina koiran parvovirusta (CPV) ja tiettyjä Gram-negatiivisia bakteereja infektoivaa bakteriofagia PRD1.

Bakteerisolujen luultua monimuotoisempi rakenne on herättänyt tutkijoiden huomion viime vuosikymmenen aikana. Bakteerisolun monimuotoisuus ja proteiinien sijoittuminen hyvin tarkkoihin paikkoihin solujen sisällä antaa olettaa, että eläinvirusten ohella myös bakteerien virukset käyttävät isäntäsolun monimuotoisuutta hyväkseen luultua enemmän. Tutkimuksessa pystytettiin plasmidikirjasto, jonka avulla voidaan helposti muodostaa fluoresoivia fuusioproteiineja tuottavia plasmideja. Kirjaston ja konfokaalimikroskopian avulla selvitettiin bakteriofagi PRD1:n eri proteiinien sijoittumista bakteerisolun sisällä. Tuloksena useat rakenneproteiinit sijoittuivat bakteerisolun sisällä tarkkoihin pisteisiin. Lokalisaatio vaikuttaisi korreloivan proteiinien multimeerisyyden kanssa. Voidaankin spekuloida bakteerisoluissa olevan oma multimerisoitumiseen vaikuttava rakenne tai mahdollisesti tietty kohta, jossa viruspartikkelien kokoaminen tapahtuu. Tutkimuksessa perehdyttiin vielä tarkemmin kahden samaan kohtaan sijoittuneen proteiinin, P5 ja P31, vuorovaikutukseen ja proteiinien sitoutumista yritettiin todentaa fluoresenssin elinaikamittauksilla. Tuloksena saatiin heikko positiivinen signaali, mutta ei vahvoja todisteita proteiinien vuorovaikutuksesta.

Tutkimuksen toisessa osassa perehdyttiin tarkemmin PRD1 viruksen kahden ei-rakenteellisen proteiinin, P33 ja P17, toistaiseksi tuntemattomiin rooleihin viruksen elinkierrossa. Chaperoniiniproteiinit ovat solulle elintärkeitä proteiineja, jotka ovat vastuussa esimerkiksi muiden proteiinien oikeanlaisesta laskostumisesta ja jotka toimivat kahden alaproteiinin, GroEL ja GroES, muodostaman tynnyrirakenteen kautta. Useat virukset käyttävät tätä rakennetta hyväkseen elinkiertonsa aikana ja joiltakin bakteriofageilta on löydetty proteiineja, jotka joko korvaavat chaperoniproteiineja tai vaikuttavat kyseisen rakenteen toimintaan jotain muuta kautta. Tutkimuksessa saatiin selville, että proteiinit P33 ja P17 voivat sekä yhdessä että erikseen kumota mutaation, joka heikentää GroES ja GroEL alayksiköiden sitoutumista toisiinsa. Lisäksi FRAP -mittausten avulla tutkittiin proteiinin P33 liikkuvuutta bakteerisolun sisällä ja todettiin, että proteiini todennäköisesti on sitoutunut johonkin solun sisäiseen rakenteeseen, joka on chaperoniinikompleksin kanssa samaa kokoluokkaa. Tuloksista voidaan päätellä, että proteiinit P33 ja P17 ovat sidoksissa chaperoniinikompleksin toimintaan.

Tutkimuksen viimeisessä osassa keskityttiin koiran parvoviruksen solun sisään tukeutumiseen. Koiran parvovirus käyttää solun sisään tunkeutumisessa

hyväkseen kohdesolun klatriini-välitteistä endosytoottista reittiä. Tarkkaa mekanismia ja kohtaa viruspartikkeleiden poistumiseen reitiltä ei toistaiseksi tiedetä. Tutkimuksessa selvitettiin viruspartikkeleiden reaktioita eri lipidikalvotyyppeihin ja saatiin selville, että sfingomyosiini kalvorakenteet muuttavat viruksen proteiinikuoren rakennetta. Aiemmin on selvitetty, että alhaisessa pH ympäristössä proteiinikuoresta paljastuu fosfolipaasi A2 entsyymiaktiivisuus, joka on oleellinen endosytoottisista vesikkeleistä postumisen kannalta. Tutkimuksessa havaitut muutokset eivät kuitenkaan vaikuttaneet fosfolipaasi A2:sen toimintaan. Tästä voidaan päätellä, että viruksen proteiinikuoressa on jokin toinenkin piilotettu mekanismi, joka liittyy viruspartikkeleiden endosytoottisista vesikkeleistä vapautumiseen.

REFERENCES

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

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

Amster-Choder O. 2011. The compartmentalized vessel: The bacterial cell as a model for subcellular organization (a tale of two studies). *Cell Logist* 1: 77-81.

Ang D. & Georgopoulos C. 2012. An ORFan no more: the bacteriophage T4 39.2 gene product, NwgI, modulates GroEL chaperone function. *Genetics* 190: 989-1000.

Ang D., Richardson A., Mayer M.P., Keppel F., Krisch H. & Georgopoulos C. 2001. Pseudo-T-even bacteriophage RB49 encodes CocO, a cochaperonin for GroEL, which can substitute for Escherichia coli's GroES and bacteriophage T4's Gp31. *J Biol Chem* 276: 8720-8726.

Arslan D., Legendre M., Seltzer V., Abergel C. & Claverie J.M. 2011. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. *Proc Natl Acad Sci U S A* 108: 17486-17491.

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

Bamford D.H., Romantschuk M. & Somerharju P.J. 1987. Membrane fusion in prokaryotes: bacteriophage phi 6 membrane fuses with the Pseudomonas syringae outer membrane. *EMBO J* 6: 1467-1473.

Bamford D.H., Rouhiainen L., Takkinen K. & Soderlund H. 1981. Comparison of the lipid-containing bacteriophages PRD1, PR3, PR4, PR5 and L17. *J Gen Virol* 57: 365-373.

Bamford J.K., Hanninen A.L., Pakula T.M., Ojala P.M., Kalkkinen N., Frilander M. & Bamford D.H. 1991. Genome organization of membrane-containing bacteriophage PRD1. *Virology* 183: 658-676.

Bamford J.K.H. & Bamford D.H. 2000. A New Mutant Class, Made by Targeted Mutagenesis, of Phage PRD1 Reveals That Protein P5 Connects the Receptor Binding Protein to the Vertex. *J Virol* 74: 7781-7786.

Bartolome B., Jubete Y., Martinez E. & de la Cruz F. 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* 102: 75-78.

Bartolucci C., Lamba D., Grazulis S., Manakova E. & Heumann H. 2005. Crystal structure of wild-type chaperonin GroEL. *J Mol Biol* 354: 940-951.

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

Braig K., Adams P.D. & Brunger A.T. 1995. Conformational variability in the refined structure of the chaperonin GroEL at 2.8 A resolution. *Nat Struct Biol* 2: 1083-1094.

Butcher S.J., Manole V. & Karhu N.J. 2012. Lipid-containing viruses: bacteriophage PRD1 assembly. *Adv Exp Med Biol* 726: 365-377.

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

Caldentey J., Hanninen A.L., Holopainen J.M., Bamford J.K., Kinnunen P.K. & Bamford D.H. 1999. Purification and characterization of the assembly factor P17 of the lipid-containing bacteriophage PRD1. *Eur J Biochem* 260: 549-558.

Campanini B., Pioselli B., Raboni S., Felici P., Giordano I., D'Alfonso L., Collini M., Chirico G. & Bettati S. 2013. Role of histidine 148 in stability and dynamics of a highly fluorescent GFP variant. *Biochim Biophys Acta* 1834: 770-779.

Carrillo-Tripp M., Shepherd C.M., Borelli I.A., Venkataraman S., Lander G., Natarajan P., Johnson J.E., Brooks C.L.,3rd & Reddy V.S. 2009. VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res* 37: D436-42.

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

Celler K., Koning R.I., Koster A.J. & van Wezel G.P. 2013. Multidimensional view of the bacterial cytoskeleton. *J Bacteriol* 195: 1627-1636.

Chalfie M., Tu Y., Euskirchen G., Ward W.W. & Prasher D.C. 1994. Green fluorescent protein as a marker for gene expression. *Science* 263: 802-805.

Chapman M.S. & Rossmann M.G. 1993. Structure, sequence, and function correlations among parvoviruses. *Virology* 194: 491-508.

Charbon G., Wang J., Brustad E., Schultz P.G., Horwich A.L., Jacobs-Wagner C. & Chapman E. 2011. Localization of GroEL determined by in vivo incorporation of a fluorescent amino acid. *Bioorg Med Chem Lett* 21: 6067-6070.

Chaudhry C., Horwich A.L., Brunger A.T. & Adams P.D. 2004. Exploring the structural dynamics of the E.coli chaperonin GroEL using translation-libration-screw crystallographic refinement of intermediate states. *J Mol Biol* 342: 229-245.

Chen I., Christie P.J. & Dubnau D. 2005. The ins and outs of DNA transfer in bacteria. *Science* 310: 1456-1460.

Clegg R.M. 1996. Fluorescence resonance energy transfer vol. 137. In: Wang X. F. and Herman B. (ed.), Fluorescence Imaging Spectroscopy and Microscopy, John Wiley & Sons Inc., New York, pp. 179-251.

Cluzel P., Surette M. & Leibler S. 2000. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science* 287: 1652-1655.

Coetzee W.F. & Bekker P.J. 1979. Pilus-specific, lipid-containing bacteriophages PR4 and PR772: comparison of physical characteristics of genomes. *J Gen Virol* 45: 195-200.

Cotmore S.F. & Tattersall P. 1987. The autonomously replicating parvoviruses of vertebrates. *Adv Virus Res* 33: 91-174.

Cotmore S.F., D'abramo A.M.,Jr, Ticknor C.M. & Tattersall P. 1999. Controlled conformational transitions in the MVM virion expose the VP1 N-terminus and viral genome without particle disassembly. *Virology* 254: 169-181.

Cvirkaite-Krupovic V., Poranen M.M. & Bamford D.H. 2010. Phospholipids act as secondary receptor during the entry of the enveloped, double-stranded RNA bacteriophage phi6. *J Gen Virol* 91: 2116-2120.

Edgar R., Rokney A., Feeney M., Semsey S., Kessel M., Goldberg M.B., Adhya S. & Oppenheim A.B. 2008. Bacteriophage infection is targeted to cellular poles. *Mol Microbiol* 68: 1107-1116.

Elowitz M.B., Surette M.G., Wolf P., Stock J.B. & Leibler S. 1999. Protein Mobility in the Cytoplasm ofEscherichia coli. *Journal of Bacteriology* 181: 197-203.

Fei X., Yang D., LaRonde-LeBlanc N. & Lorimer G.H. 2013. Crystal structure of a GroEL-ADP complex in the relaxed allosteric state at 2.7 Å resolution. *Proceedings of the National Academy of Sciences* 110: E2958-E2966.

Förster T. 1965. Delocalized excitation and excitation transfer. In: Sinanoglu (ed.), Modern Quantum Chemistry, Academic Press Inc., New York, pp. 93-137.

Georgopoulos C. 2006. Toothpicks, serendipity and the emergence of the Escherichia coli DnaK (Hsp70) and GroEL (Hsp60) chaperone machines. *Genetics* 174: 1699-1707.

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

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

Grahn A.M., Caldentey J., Bamford J.K. & Bamford D.H. 1999. Stable packaging of phage PRD1 DNA requires adsorption protein P2, which binds to the IncP plasmid-encoded conjugative transfer complex. *J Bacteriol* 181: 6689-6696.

Greenfield D., McEvoy A.L., Shroff H., Crooks G.E., Wingreen N.S., Betzig E. & Liphardt J. 2009. Self-organization of the Escherichia coli chemotaxis network imaged with super-resolution light microscopy. *PLoS Biol* 7: e1000137.

Holopainen J.M., Saily M., Caldentey J. & Kinnunen P.K. 2000. The assembly factor P17 from bacteriophage PRD1 interacts with positively charged lipid membranes. *Eur J Biochem* 267: 6231-6238.

Huiskonen J.T., Manole V. & Butcher S.J. 2007. Tale of two spikes in bacteriophage PRD1. *Proceedings of the National Academy of Sciences* 104: 6666-6671.

Hunt J.F., van der Vies S.M., Henry L. & Deisenhofer J. 1997. Structural adaptations in the specialized bacteriophage T4 co-chaperonin Gp31 expand the size of the Anfinsen cage. *Cell* 90: 361-371.

Hänninen A.L., Bamford D.H. & Bamford J.K. 1997. Probing phage PRD1-specific proteins with monoclonal and polyclonal antibodies. *Virology* 227: 198-206.

Hänninen A., Bamford D.H. & Bamford J.K.H. 1997. Assembly of Membrane-Containing Bacteriophage PRD1 Is Dependent on GroEL and GroES. *Virology* 227: 207-210.

Ihalainen T.O., Niskanen E.A., Jylhava J., Turpeinen T., Rinne J., Timonen J. & Vihinen-Ranta M. 2007. Dynamics and interactions of parvoviral NS1 protein in the nucleus. *Cell Microbiol* 9: 1946-1959.

Ihalainen T.O., Willman S.F., Niskanen E.A., Paloheimo O., Smolander H., Laurila J.P., Kaikkonen M.U. & Vihinen-Ranta M. 2012. Distribution and dynamics of transcription-associated proteins during parvovirus infection. *J Virol* 86: 13779-13784.

Ihalainen T.O., Niskanen E.A., Jylhava J., Paloheimo O., Dross N., Smolander H., Langowski J., Timonen J. & Vihinen-Ranta M. 2009. Parvovirus induced alterations in nuclear architecture and dynamics. *PLoS One* 4: e5948.

Inouye S. & Tsuji F.I. 1994. Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein. *FEBS Lett* 341: 277-280.

Ishikawa-Ankerhold H.C., Ankerhold R. & Drummen G.P. 2012. Advanced fluorescence microscopy techniques--FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* 17: 4047-4132.

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

Jakutyte L., Baptista C., Sao-Jose C., Daugelavicius R., Carballido-Lopez R. & Tavares P. 2011. Bacteriophage infection in rod-shaped gram-positive bacteria: evidence for a preferential polar route for phage SPP1 entry in Bacillus subtilis. *J Bacteriol* 193: 4893-4903.

Jalasvuori M., Jaatinen S.T., Laurinavicius S., Ahola-Iivarinen E., Kalkkinen N., Bamford D.H. & Bamford J.K. 2009. The closest relatives of icosahedral viruses of thermophilic bacteria are among viruses and plasmids of the halophilic archaea. *J Virol* 83: 9388-9397.

Kai T., Selick H.E. & Yonesaki T. 1996. Destabilization of bacteriophage T4 mRNAs by a mutation of gene 61.5. *Genetics* 144: 7-14.

Kankaanpää P., Paavolainen L., Tiitta S., Karjalainen M., Päivärinne J., Nieminen J., Marjomaki V., Heino J. & White D.J. 2012. BioImageXD: an open, general-purpose and high-throughput image-processing platform. *Nat Methods* 9: 683-689.

Keppel F., Rychner M. & Georgopoulos C. 2002. Bacteriophage-encoded cochaperonins can substitute for Escherichia coli's essential GroES protein. *EMBO Rep* 3: 893-898.

Krupovic M., White M.F., Forterre P. & Prangishvili D. 2012. Chapter 2 - Postcards from the Edge: Structural Genomics of Archaeal Viruses. In: Anonymous Advances in Virus Research, Academic Press, pp. 33-62.

Kurochkina L.P., Semenyuk P.I., Orlov V.N., Robben J., Sykilinda N.N. & Mesyanzhinov V.V. 2012. Expression and functional characterization of the first bacteriophage-encoded chaperonin. *J Virol* 86: 10103-10111.

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

Lakowicz J.R., Gryczynski I., Gryczynski Z. & Dattelbaum J.D. 1999. Anisotropy-based sensing with reference fluorophores. *Anal Biochem* 267: 397-405.

Lakowicz J.R. & Szmacinski H. 1993. Fluorescence lifetime-based sensing of pH, Ca2+, K+ and glucose. *Sensors Actuators B: Chem* 11: 133-143.

Landry S.J., Zeilstra-Ryalls J., Fayet O., Georgopoulos C. & Gierasch L.M. 1993. Characterization of a functionally important mobile domain of GroES. *Nature* 364: 255-258.

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

Lissandron V., Terrin A., Collini M., D'alfonso L., Chirico G., Pantano S. & Zaccolo M. 2005. Improvement of a FRET-based indicator for cAMP by linker design and stabilization of donor-acceptor interaction. *J Mol Biol* 354: 546-555.

Lopian L., Elisha Y., Nussbaum-Shochat A. & Amster-Choder O. 2010. Spatial and temporal organization of the E. coli PTS components. *EMBO J* 29: 3630-3645.

Margolin W. 2012. The price of tags in protein localization studies. *J Bacteriol* 194: 6369-6371.

McMahon H.T. & Boucrot E. 2011. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nat Rev Mol Cell Biol* 12: 517-533.

Meier E.L. & Goley E.D. 2014. Form and function of the bacterial cytokinetic ring. *Curr Opin Cell Biol* 26C: 19-27.

Mercer J., Schelhaas M. & Helenius A. 2010. Virus entry by endocytosis. *Annu Rev Biochem* 79: 803-833.

Merckel M.C., Huiskonen J.T., Bamford D.H., Goldman A. & Tuma R. 2005. The Structure of the Bacteriophage PRD1 Spike Sheds Light on the Evolution of Viral Capsid Architecture. *Mol Cell* 18: 161-170.

Merola F., Fredj A., Betolngar D.B., Ziegler C., Erard M. & Pasquier H. 2013. Newly engineered cyan fluorescent proteins with enhanced performances for live cell FRET imaging. *Biotechnol J* .

Munoz-Espin D., Holguera I., Ballesteros-Plaza D., Carballido-Lopez R. & Salas M. 2010. Viral terminal protein directs early organization of phage DNA replication at the bacterial nucleoid. *Proc Natl Acad Sci U S A* 107: 16548-16553.

Munoz-Espin D., Daniel R., Kawai Y., Carballido-Lopez R., Castilla-Llorente V., Errington J., Meijer W.J. & Salas M. 2009. The actin-like MreB cytoskeleton organizes viral DNA replication in bacteria. *Proc Natl Acad Sci U S A* .

Nagy E., Pragai B. & Ivanovics G. 1976. Characteristics of phage AP50, an RNA phage containing phospholipids. *J Gen Virol* 32: 129-132.

Nevo-Dinur K., Nussbaum-Shochat A., Ben-Yehuda S. & Amster-Choder O. 2011. Translation-independent localization of mRNA in E. coli. *Science* 331: 1081-1084.

Niskanen E.A., Kalliolinna O., Ihalainen T.O., Hakkinen M. & Vihinen-Ranta M. 2013. Mutations in DNA binding and transactivation domains affect the dynamics of parvovirus NS1 protein. *J Virol* 87: 11762-11774.

Nojima T., Ikegami T., Taguchi H. & Yoshida M. 2012. Flexibility of GroES mobile loop is required for efficient chaperonin function. *J Mol Biol* 422: 291-299.

Oksanen H.M. and Bamford D.H. 2012. Family *Tectiviridae*.. In: King A.M.Q., Adams M.J., Carstens E.B. & Lefkowitz E.J. (eds.), Virus taxonomy, Ninth Report of the International Committee on Taxonomy of Viruses., Elsevier, Oxford, pp. 317-318-322.

Olsen R.H., Siak J.S. & Gray R.H. 1974. Characteristics of PRD1, a plasmid-dependent broad host range DNA bacteriophage. *J Virol* 14: 689-699.

Onuki R., Nagasaki A., Kawasaki H., Baba T., Uyeda T.Q. & Taira K. 2002. Confirmation by FRET in individual living cells of the absence of significant amyloid beta -mediated caspase 8 activation. *Proc Natl Acad Sci U S A* 99: 14716-14721.

Ormo M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y. & Remington S.J. 1996. Crystal structure of the Aequorea victoria green fluorescent protein. *Science* 273: 1392-1395.

Paradiso P.R., Rhode S.L.,3rd & Singer I.I. 1982. Canine parvovirus: a biochemical and ultrastructural characterization. *J Gen Virol* 62 (Pt 1): 113-125.

Parker J.S. & Parrish C.R. 2000. Cellular uptake and infection by canine parvovirus involves rapid dynamin-regulated clathrin-mediated endocytosis, followed by slower intracellular trafficking. *J Virol* 74: 1919-1930.

Pawlowski A., Rissanen I., Bamford J.K., Krupovic M. & Jalasvuori M. 2014. Gammasphaerolipovirus, a newly proposed bacteriophage genus, unifies viruses of halophilic archaea and thermophilic bacteria within the novel family Sphaerolipoviridae. *Arch Virol* .

Peetla C., Stine A. & Labhasetwar V. 2009. Biophysical interactions with model lipid membranes: applications in drug discovery and drug delivery. *Mol Pharm* 6: 1264-1276.

Philippe N., Legendre M., Doutre G., Coute Y., Poirot O., Lescot M., Arslan D., Seltzer V., Bertaux L., Bruley C., Garin J., Claverie J.M. & Abergel C. 2013. Pandoraviruses: amoeba viruses with genomes up to 2.5 Mb reaching that of parasitic eukaryotes. *Science* 341: 281-286.

Pollok B.A. & Heim R. 1999. Using GFP in FRET-based applications. *Trends Cell Biol* 9: 57-60.

Poranen M.M., Daugelavicius R., Ojala P.M., Hess M.W. & Bamford D.H. 1999. A novel virus-host cell membrane interaction. Membrane voltage-dependent endocytic-like entry of bacteriophage straight phi6 nucleocapsid. *J Cell Biol* 147: 671-682.

Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G. & Cormier M.J. 1992. Primary structure of the Aequorea victoria green-fluorescent protein. *Gene* 111: 229-233.

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

Redrejo-Rodriguez M., Munoz-Espin D., Holguera I., Mencia M. & Salas M. 2012. Functional eukaryotic nuclear localization signals are widespread in terminal proteins of bacteriophages. *Proc Natl Acad Sci U S A* 109: 18482-18487.

Reed A.P., Jones E.V. & Miller T.J. 1988. Nucleotide sequence and genome organization of canine parvovirus. *J Virol* 62: 266-276.

Rissanen I., Pawlowski A., Harlos K., Grimes J.M., Stuart D.I. & Bamford J.K. 2012. Crystallization and preliminary crystallographic analysis of the major capsid proteins VP16 and VP17 of bacteriophage P23-77. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 68: 580-583.

Rothenberg E., Sepulveda L.A., Skinner S.O., Zeng L., Selvin P.R. & Golding I. 2011. Single-virus tracking reveals a spatial receptor-dependent search mechanism. *Biophys J* 100: 2875-2882.

Rydman P.S. & Bamford D.H. 2000. Bacteriophage PRD1 DNA entry uses a viral membrane-associated transglycosylase activity. *Mol Microbiol* 37: 356-363.

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

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

Schaff J., Fink C.C., Slepchenko B., Carson J.H. & Loew L.M. 1997. A general computational framework for modeling cellular structure and function. *Biophys J* 73: 1135-1146.

Shih Y.L., Kawagishi I. & Rothfield L. 2005. The MreB and Min cytoskeletal-like systems play independent roles in prokaryotic polar differentiation. *Mol Microbiol* 58: 917-928.

Shimomura O., Johnson F.H. & Saiga Y. 1962. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol* 59: 223-239.

Shimozono S. & Miyawaki A. 2008. Engineering FRET constructs using CFP and YFP. *Methods Cell Biol* 85: 381-393.

Shiomi D., Yoshimoto M., Homma M. & Kawagishi I. 2006. Helical distribution of the bacterial chemoreceptor via colocalization with the Sec protein translocation machinery. *Mol Microbiol* 60: 894-906.

Simpson A.A., Chandrasekar V., Hebert B., Sullivan G.M., Rossmann M.G. & Parrish C.R. 2000. Host range and variability of calcium binding by surface loops in the capsids of canine and feline parvoviruses. *J Mol Biol* 300: 597-610.

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

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

Stromsten N.J., Bamford D.H. & Bamford J.K. 2003. The unique vertex of bacterial virus PRD1 is connected to the viral internal membrane. *J Virol* 77: 6314-6321.

Suikkanen S., Aaltonen T., Nevalainen M., Valilehto O., Lindholm L., Vuento M. & Vihinen-Ranta M. 2003. Exploitation of microtubule cytoskeleton and dynein during parvoviral traffic toward the nucleus. *J Virol* 77: 10270-10279.

Suikkanen S., Antila M., Jaatinen A., Vihinen-Ranta M. & Vuento M. 2003. Release of canine parvovirus from endocytic vesicles. *Virology* 316: 267-280.

Sun L., Young L.N., Zhang X., Boudko S.P., Fokine A., Zbornik E., Roznowski A.P., Molineux I.J., Rossmann M.G. & Fane B.A. 2013. Icosahedral bacteriophage PhiX174 forms a tail for DNA transport during infection. *Nature*.

Suomalainen M., Luisoni S., Boucke K., Bianchi S., Engel D.A. & Greber U.F. 2013. A Direct and Versatile Assay Measuring Membrane Penetration of Adenovirus in Single Cells. *Journal of Virology* 87: 12367-12379.

Swulius M.T. & Jensen G.J. 2012. The helical MreB cytoskeleton in Escherichia coli MC1000/pLE7 is an artifact of the N-Terminal yellow fluorescent protein tag. *J Bacteriol* 194: 6382-6386.

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

van der Vies S.M., Gatenby A.A. & Georgopoulos C. 1994. Bacteriophage T4 encodes a co-chaperonin that can substitute for Escherichia coli GroES in protein folding. *Nature* 368: 654-656.

van Teeffelen S., Wang S., Furchtgott L., Huang K.C., Wingreen N.S., Shaevitz J.W. & Gitai Z. 2011. The bacterial actin MreB rotates, and rotation depends on cell-wall assembly. *Proc Natl Acad Sci U S A* 108: 15822-15827.

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

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

Vihinen-Ranta M., Suikkanen S. & Parrish C.R. 2004. Pathways of cell infection by parvoviruses and adeno-associated viruses. *J Virol* 78: 6709-6714.

Vihinen-Ranta M., Wang D., Weichert W.S. & Parrish C.R. 2002. The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection. *J Virol* 76: 1884-1891.

Weichert W.S., Parker J.S., Wahid A.T., Chang S.F., Meier E. & Parrish C.R. 1998. Assaying for structural variation in the parvovirus capsid and its role in infection. *Virology* 250: 106-117.

Wong F.H. & Bryan L.E. 1978. Characteristics of PR5, a lipid-containing plasmid-dependent phage. *Can J Microbiol* 24: 875-882.

Xie Q. & Chapman M.S. 1996. Canine parvovirus capsid structure, analyzed at 2.9 A resolution. *J Mol Biol* 264: 497-520.

Xu L., Butcher S.J., Benson S.D., Bamford D.H. & Burnett R.M. 2000. Crystallization and preliminary X-ray analysis of receptor-binding protein P2 of bacteriophage PRD1. *J Struct Biol* 131: 159-163.

Xu Z., Horwich A.L. & Sigler P.B. 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. *Nature* 388: 741-750.

Yu M.X., Slater M.R. & Ackermann H.W. 2006. Isolation and characterization of Thermus bacteriophages. *Arch Virol* 151: 663-679.

ORIGINAL PAPERS

I

SUBCELLULAR LOCALIZATION OF BACTERIOPHAGE PRD1 PROTEINS IN *ESCHERICHIA COLI*

by

Jenni Karttunen, Sari Mäntynen, Teemu O. Ihalainen, Heli Lehtivuori, Nikolai V. Tkachenko, Maija Vihinen-Ranta, Janne A. Ihalainen, Jaana K. Bamford & Hanna M. Oksanen 2014.

Virus Research 179: 44–52.

Subcellular localization of bacteriophage PRD1 proteins in *Escherichia coli*

Jenni Karttunen[a], Sari Mäntynen[a], Teemu O. Ihalainen[b], Heli Lehtivuori[b], Nikolai V. Tkachenko[c], Maija Vihinen-Ranta[b], Janne A. Ihalainen[b], Jaana K.H. Bamford[a], Hanna M. Oksanen[d],

Virus Research 179 (2014) 44–52

[a] Centre of Excellence in Biological Interactions, Department of Biological and Environmental Science and Nanoscience Center, P.O. Box 35, 40014 University

of Jyväskylä, Finland

[b] Nanoscience Center, Department of Biological and Environmental Science, P.O. Box 35, 40014 University of Jyväskylä, Finland

[c] Department of Chemistry and Bioengineering, Tampere University of Technology, P.O. Box 541, 33101 Tampere, Finland

[d] Institute of Biotechnology and Department of Biosciences, P.O. Box 56, 00014 University of Helsinki, Finland

Keywords: Membrane virus; Confocal microscopy; Protein interactions; Virus assembly; Bacteria

Abstract

Bacteria possess an intricate internal organization resembling that of the eukaryotes. The complexity is especially prominent at the bacterial cell poles, which are also known to be the preferable sites for some bacteriophages to infect. Bacteriophage PRD1 is a well-known model serving as an ideal system to study structures and functions of icosahedral internal membrane-containing viruses. Our aim was to analyze the localization and interactions of individual PRD1 proteins in its native host *Escherichia coli*. This was accomplished by constructing a vector library for production of fluorescent fusion proteins. Analysis of solubility and multimericity of the fusion proteins, as well as their localization in living cells by confocal microscopy, indicated that multimeric PRD1 proteins were prone to localize in the cell poles. Furthermore, PRD1 spike complex proteins P5 and P31, as fusion proteins, were shown to be functional in the virion assembly. In addition, they were shown to co-localize in the specific polar area of the cells, which might have a role in the multimerization and formation of viral protein complexes.

1. Introduction

Bacterial cells have been mainly regarded as amorphous reaction vessels concealing a homogenous solution of proteins. Due to advances in bacterial cell biology, this traditional view has changed dramatically. Similarly to eukaryotes, bacteria deploy macromolecules such as proteins, lipids and nucleic acids into specific subcellular locations. This asymmetric architecture is spatially and temporally dynamic, enabling cells to respond to changing demands during their life cycle (Rudner and Losick, 2010). Accumulated data on bacterial proteins have revealed a variety of localization patterns (Amster-Choder, 2011). Whereas certain proteins oscillate from pole to pole (Gerdes et al., 2010, Leonardy et al., 2010 and Loose et al., 2011), others form clusters on the bacterial cell surface or at specific sub-cellular locations (Amster-Choder, 2011). In addition, it is known that certain bacterial proteins, especially cytoskeletal, assemble into helical structures extending along the cell or construct ring-like structures at the mid-cell position (Vats et al., 2009). However, interpretation of the localization pattern data with fluorescent tagged proteins has been challenging and some artifacts have emerged (Swulius and Jensen, 2012).

Recent studies have elucidated factors governing the asymmetric protein distribution in bacteria, which is presumably most commonly mediated by 'diffusion and capture', when proteins diffuse freely until interacting with other, so-called target proteins (Deich et al., 2004 and Rudner and Losick, 2002). This raises the question about the primary factors directing the target proteins to their specific cellular sites, and emphasizes the need to reveal other mechanisms for protein targeting. For instance, self-assembly is a unique variation of the 'diffusion and capture' positioning, which does not require any pre-existing anchor structures. Cellular factors such as geometric cues and physical constrictions have a role in positioning a number of bacterial proteins into their specific intracellular sites (Rudner and Losick, 2010). Also localization signals can be found in certain bacterial proteins, similarly to their eukaryotic analogs (Russell and Keiler, 2007). Correspondingly, there is evidence of subcellular localization of certain mRNA transcripts correlating with the localization of their protein products (Nevo-Dinur et al., 2011). Yet another mechanism for protein positioning was brought up by the discovery of cytoskeletal proteins in bacteria (Vats et al., 2009). These structures are also suggested to provide a track for other proteins to locate, resembling again the situation in eukaryotes (Nevo-Dinur et al., 2012). However, the underlying principles of targeting for most bacterial proteins remain elusive.

The studies indicate that the asymmetric protein distribution is particularly conspicuous at the chemically and physically unique cell poles (Lai et al., 2004). For instance, unequally distributed lipid composition and negative curvature of the membrane (Ramamurthi, 2010) are believed to contribute to encompassing proteins into these cellular areas (Nevo-Dinur et al., 2012). It has also been shown that a number of bacteriophages infect preferably at these extreme regions. These viruses bind to distinct cellular receptors on the surface of their Gram negative hosts such as *Escherichia coli, Yersinia pseudotuberculosis*, or Vibrio cholera (Edgar et al., 2008 and Rothenberg et al., 2011) or Gram positive ones such as *Bacillus subtilis* (Jakutyte et al., 2011). This implies that the cell poles contain cellular components essential for DNA intake (Edgar et al., 2008). The hypothesis is supported by the fact that the poles are the preferred site of DNA intake in

natural competent cells (Chen et al., 2005 and Edgar et al., 2008). Moreover, in studies of B. subtilis infecting phage SPP1 (Jakutyte et al., 2011) *and E. coli* phage lambda (Rothenberg et al., 2011) it was observed that in addition to being injected, viral DNA is replicated at the poles. However, several proteins of replication machinery of bacteriophage ϕ29 infecting *B. subtilis* have been found localized in helix-like pattern near the membrane. It was also shown that the ϕ29 replication is dependent on cytoskeleton protein MreB, as also with phage PRD1 (Muñoz-Espín et al., 2009). Terminal protein of these phages has been shown to associate with bacterial nucleoid independently of other phage-coded proteins as well as localize in the nucleus of mammalian cells (Muñoz-Espín et al., 2010 and Redrejo-Rodriguez et al., 2012). It seems that bacteriophages have evolved to exploit the internal asymmetry of their host cells in order to make the infection process more efficient.

One of the most extensively studied bacteriophages is Enterobacteria phage PRD1 (family: *Tectiviridae*, genus: Tectivirus), which infects various Gram negative bacteria, such as *E. coli* and *Salmonella typhimurium*, carrying P-, W- or N-type conjugative plasmid, whereas other tectiviruses can also infect Gram positive bacteria such as *Bacillus* (Grahn et al., 2006 and Oksanen and Bamford, 2012). The PRD1 virion is formed by an icosahedral protein capsid surrounding a protein-rich membrane which, in turn, encloses the linear dsDNA genome (Abrescia et al., 2004, Cockburn et al., 2004 and Olsen et al., 1974). Based on X-ray crystallographic analyses, PRD1 belongs to a certain structure based viral lineage with several other viruses such as: adenovirus, Paramecium bursaria chlorella virus 1 (PBCV-1) and Sulfolobus turreted icosahedral virus (STIV) (Abrescia et al., 2012, Benson et al., 1999, Khayat et al., 2005 and Nandhagopal et al., 2002). All these viruses have a major capsid protein with a double β-barrel fold and similar virion architecture. In PRD1, the capsid is mainly composed of the major capsid protein P3, which forms pseudo-hexameric trimers (Abrescia et al., 2004 and Benson et al., 1999). The receptor binding spike complex at the virion vertices contains the pentameric penton protein P31 forming the base structure from which the trimeric spike protein P5 and the monomeric receptor binding protein P2 protrude (Caldentey et al., 2000, Merckel et al., 2005, Rydman et al., 1999 and Xu et al., 2003). The spike structure complex is stabilized by the integral membrane protein P16 linking the vertex to the underlying viral membrane (Jaatinen et al., 2004). In addition to several other PRD1 structural proteins, also a number of non-structural proteins have been identified, such as the tetrameric assembly protein P17 required for virion formation (Caldentey et al., 1999, Holopainen et al., 2000 and Mindich et al., 1982). Despite the intensive structural and functional characterization, the interactions of a number of predicted PRD1 proteins are yet to be revealed.

Our aim was to analyze the localization and interactions of PRD1 proteins in its native host bacterium *E. coli*. The study included viral monomeric and multimeric structural proteins, an integral membrane protein and a soluble assembly protein. Special attention was paid to the receptor binding spike complex proteins P5 and P31 for which the structures at atomic resolution are known (Abrescia et al., 2004, Caldentey et al., 2000 and Rydman et al., 1999). We localized the proteins in living cells by exploiting fluorescent fusion protein technology and confocal microscopy.

2. Results and discussion

2.1. Construction of bacterial expression vector library to produce fluorescent fusion proteins

We created a bacterial vector library for convenient production of fluorescent fusion proteins (Fig. 1; Table 1). The vectors were constructed by cloning genes encoding eGFP and its cyan and yellow variants eCFP and eYFP. We used two bacterial vectors pSU18 and pET24 bearing replicons p15A and ColE1, respectively, enabling simultaneous expression of two proteins. According to confocal microscopy the expression of the fluorescent protein genes in bacterial cells produced functional proteins (for eYFP see Section 2.3, for eGFP and eCFP data not shown). Using these vectors, it is possible to insert any gene of interest into either end of the fluorescent protein gene, thereby creating N-terminal fluorescent fusion protein (the fluorescent protein is linked to the N-terminus of the target protein) or C-terminal fluorescent fusion protein (the fluorescent protein is linked to the C-terminus of the protein). A linker sequence of six glycines was designed to these vectors to separate the fluorescent protein from the protein of interest reducing steric hindrance. In this study, we exploited the vector library to create cyan and/or yellow fluorescent fusion proteins of bacteriophage PRD1 proteins P2 (receptor binding protein), P3 (major capsid protein), P5 (spike protein), P16 (vertex stabilizing integral membrane protein), P17 (non-structural assembly protein), and P31 (penton protein) (Table 1). Genes were cloned into both vector types (Fig. 1) using either pSU18 or pET24 to produce both N-terminal and C-terminal fusion proteins, except gene *XVII*, for which only fusion P17-eYFP was available. Sequencing of the vectors revealed only minor changes in PRD1 genes (Supplementary Table S1).

2.2. Solubility and multimericity of viral fluorescent fusion proteins

In the fusion protein studies, the first concern is whether the fusion affects on the folding and functionality of the native protein. One way to evaluate this is to monitor changes in the protein solubility and find out whether the known multimeric proteins form multimers with fluorescent protein tags. The majority of the fusion proteins (P2, P3, P5, P17 and P31) were expressed as soluble (data not shown). These proteins were directed to sedimentation assay by a rate zonal centrifugation for the size determination.

Monomeric receptor binding protein P2 (Grahn et al., 1999 and Xu et al., 2003) was expressed as a fusion protein in a monomeric form (Fig. 2A). Small fraction of smaller side-product was detected with both P2 fusions. More variation in the molecular mass distribution was detected with proteins, which can be released as multimers from the virion. The individual PRD1 spikes composed of the trimeric protein P5 form an elongated structure (Bamford and Bamford, 2000, Caldentey et al., 2000, Huiskonen et al., 2007 and Merckel et al., 2005), and there is no obvious reason that the fluorescent tag at the C-terminus of the protein would interfere the folding. The N-terminal fusion protein eYFP-P5 (~61 kDa) sedimented as a monomer (Fig. 2B), but also a smaller multimeric side product (~45 kDa) was detected by an antibody against P5 (data not shown). The C-terminal fusion protein P5-eYFP showed two separate peaks in the sedimentation assay indicating that the protein was in two different forms (monomer and multimer) (Fig. 2B).

It also had a smaller P5-specific side-product (~55 kDa) in fractions representing monomeric and multimeric proteins (data not shown). The rate zonal centrifugation indicated that the C-terminal fusion of the penton protein P31 was mainly monomeric whereas the fluorescent protein attached to the N-terminus of P31 formed larger multimers (Fig. 2C). This correlates well with the known P31 X-ray structure as its C-termini are located in the middle of the pentamer (Abrescia et al., 2004) and therefore the added fluorescent tag can hinder the formation of the multimeric complex. However, the N-termini of P31 are pointing outwards from the pentamer (Abrescia et al., 2004). Thus, the N-terminal fluorescence tag should not interfere with the formation of the multimer.

Both fluorescent protein fusions with the major capsid protein P3 were broadly distributed in the multimericity assay starting from the monomeric forms, but also trimeric molecules were detected (P3 is a trimer; Benson et al., 1999) (Fig. 2D). Sedimentation analysis of the assembly protein P17 fusion (P17-eYFP) revealed both monomeric and multimeric forms (Fig. 2E).

The functionality of the produced fluorescent fusion proteins P5 and P31 was tested by complementation assay using PRD1 virus mutants, sus690 and sus525, having amber mutation either in gene *V* (protein P5) or gene *XXXI* (protein P31), respectively. Both N- and C-terminal versions of P5 and P31 fusion proteins complemented the defect in the corresponding gene at the same level as with the plasmid-produced wt protein or when the mutant was grown on suppressor host (Table 2) showing that the folding of the proteins was not compromised. According to the results fluorescent fusions did not seem to interfere the viral proteins significantly and multimerization was altered mostly in cases where the protein structures suggested steric hindrance between protein subunits.

2.3. Multimeric PRD1 proteins localize in the cell poles of *E. coli*

Localization of the fluorescent fusion proteins in *E. coli* cells was studied by confocal microscopy using living cells at the stationary phase of the bacterial growth. As a fusion protein, P16 was expressed all over the cytoplasm (Fig. 3A and B). Localization around the circumference of the cells, which is typical for membrane proteins, was not observed (Li and Young, 2012 and Maier et al., 2008). In the virion, protein P16 locks the vertex complex to the inner membrane stabilizing the vertex structure and is found in the virus membrane (Abrescia et al., 2004 and Jaatinen et al., 2004). During virus assembly, the interaction of P16 with the virus membrane might require other viral proteins, which may explain the localization of the P16 fusion protein. It is known that the formation of the procapsid including also P16 is dependent on the non-structural scaffolding protein P10 and assembly factor(s) P17 (and most probably P33) (Bamford et al., 2002, Mindich et al., 1982 and Rydman et al., 2001).

Although both fusions of the monomeric receptor binding protein P2 were also evenly distributed inside the bacterium, P2 was occasionally found specifically localized in the cell poles (Fig. 3C and D). Clear loci (a locus is used here to describe the specific localization of fluorescence in a cell) were detected mainly with multimeric fusion proteins, especially with the spike protein P5 (Fig. 3E and F). Both fusions were clearly localized in one specific polar locus in the majority of the cells (~64% and ~81%; Fig. 3E and F).

More notable variations between C- and N-terminal fusion proteins were detected with the penton protein P31 and the major capsid protein P3. Like P2 fusion proteins, the fluorescence of P31-eYFP was evenly spread in the cytoplasm in the most of the cells, but about 5% of the cells had specific fluorescence locus (Fig. 3G). Most of the cells producing eYFP-P31 (~90%) had very low intensity level and the fluorescence was spread throughout the cytoplasm (Fig. 3H). However, rest of the cells (~10%) were having high fluorescence intensity and eYFP-P31 was specifically localized in the polar end of the bacteria (Fig. 3H). It can be concluded that the multimeric P31 fusion proteins were found mostly localized in specific polar regions more frequently than the monomeric ones (Fig. 2 and Fig. 3).

The C-terminal fusion of the major capsid protein P3 (P3-eYFP) formed clear polar loci in half of the analyzed cells in all cultivations (Fig. 3I). With the eYFP-P3 fusion protein, the specific localization was detected only in around half of the parallel cultures. The other half of the cultivations had cells with fluorescence equally distributed. The parallel samples used for data collection on eYFP-P3 (Fig. 3J) were taken from the cultures with clear localization. Specific loci were detected not only in the polar region, but eYFP-P3 was also found specifically localized in several other regions inside the cell (~69% of the cells had loci) (Fig. 3J). Fluorescence was also found in specific loci in the fusion of assembly protein P17 (P17-eYFP) on most of the cells (Fig. 3K). P17 is tetrameric (Caldentey et al., 1999), but its function in the virus assembly is rather unknown. As a control, eYFP was produced alone and it was distributed evenly across the bacterial cell, as also previously reported (Fig. 3L) (Deich et al., 2004 and Edgar et al., 2008).

The amount of loci in one cell was calculated from the samples, in which ~10% or more of the cells were with loci (Fig. 4). Protein P31 and most of the protein P5 were localized in one specific locus. The remaining cells producing P5 had two loci. With P17-eYFP the number of loci in one cell was slightly more diverse (from one to five loci), but majority of the cells had one locus (60% of cells with localization). Both P3 fusion proteins were often found in several specific positions (Fig. 4). Especially, eYFP-P3 was localized more often in three or more loci per cell than in one specific locus. In the virion, the major capsid protein P3 has a connection to the viral inner membrane by its N-terminus (Abrescia et al., 2004 and Benson et al., 1999). Both N- and C- termini are located outwards from the trimeric protein capsomer, nevertheless they have roles in locking trimers together to form larger capsid facets (Abrescia et al., 2004). The membrane connection might partly explain the observed differences in the localization of P3 fusion proteins. In addition, the loci of eYFP-P3 were occasionally in shuffling motion (data not shown).

These observations indicated that the viral proteins had a specific intracellular distribution and the multimeric ones seemed to accumulate into polar areas of the host cells similarly with bacterial proteins, for example the chemotaxis protein CheA (Sourjik and Berg, 2000) and chaperon protein GroES (Li and Young, 2012) (Fig. 3, Supplementary Fig. S1). However, there were differences in the number of loci between viral proteins thus the process leading to localization might not be the same for all proteins. The specific polar regions for protein localization might be the assembly sites of protein multimers. It has also been reported that the protein aggregation has led to a similar polar localization (Lindner et al., 2008 and Lloyd-Price et al., 2012). However, the low copy-number plasmid pSU18 used here has been widely utilized for the production of functional PRD1 structural

proteins (Bamford and Bamford, 2000, Bartolome et al., 1991 and Rydman et al., 2001). In addition, the fluorescent fusion proteins of P5 and P31 complemented the defect of virus mutants (Table 2). Also the detected localization varied between proteins and for example the clear difference between monomeric and multimeric proteins indicates that the proteins were produced as soluble (Fig. 2).

2.4. Proteins P5 and P31 co-localize within a specific locus area

The observed protein localization was studied further with proteins P5 and P31. In the virion, P5 and P31 are known to interact as a part of the spike vertex complex (Caldentey et al., 2000). To find out whether the proteins co-localize in the same locus, P5 and P31 were fused with eCFP and eYFP, respectively, and their co-expression was imaged by confocal microscopy. The background was manually removed, yielding images only from the higher intensity loci. The loci were considered to be co-localized if the locus emission had contribution from both eCFP and eYFP labels. When P5-eCFP and eYFP-P31 were co-produced, 66% of the loci were identical (n = 140/211). Rest of the loci contained only either eYFP-P31 (16%, n = 34/211) or P5-eCFP (18%, n = 37/211), which is mostly explained by a production of only one type of fusion protein in a cell.

Based on PRD1 structural data the theoretical maximum distance observed in the virion between the C-terminus of P5 and the N-terminus of P31 is around 30 nm (Abrescia et al., 2004 and Huiskonen et al., 2007). Protein P5 is an elongated and flexible trimer with a collagen like region in the middle of the protein (Bamford et al., 1991, Bamford and Bamford, 2000, Caldentey et al., 2000 and Sokolova et al., 2001). In addition, the glycine linker region in the fusion protein also allows the fluorescent protein tag to move and interact suggesting that protein–protein interaction could be followed by förster resonance energy transfer (FRET).

We tested the energy transfer between proteins P5 and P31 by fluorescence lifetime microscope (FLIM) measurements using three samples: (i) co-expression of P5-eCFP and eYFP-P31, which represents the ideal combination for the interaction studies, (ii) co-expression of eYFP-P5 and P31-eCFP, where the fluorescent proteins hinder the interaction, and (iii) expression of P31-eCFP, used as a control to observe the lifetime of eCFP in the absence of FRET. With samples (i) and (iii) the lifetime of eCFP was measured from locus area, outside the locus area and from entire cell (Supplementary Fig. S1A). For sample (ii) only entire cells were measured. The results showed a minimal energy transfer in the locus area of sample (i) comparing to other measured samples, and so no significant FRET could be observed (Supplementary Fig. S1B and C). Similar results were obtained when fluorescence spectra and decays were measured by confocal microscopy and time-correlated single photon counting (TCSPC) from liquid cell samples, respectively (Supplementary Fig. S1D and E). In studies by others (Onuki et al., 2002) the efficiency of the energy transfer between CFP and YFP has been observed to be significantly more efficient than we observed here, although the orientation of P5 and P31 as fusion proteins was theoretically ideal for detection FRET based on their X-ray structures and orientation in the virion.

3. Conclusions

Numerous bacterial proteins and functions have been localized to the poles of bacteria (Gestwicki et al., 2000, Li and Young, 2012 and Maddock et al., 1993) and in other specific regions (Maier et al., 2008, Nevo-Dinur et al., 2012 and Russell and Keiler, 2008). The current knowledge of the complexity of bacterial cells provides also a new aspect to the study of functions and life cycle of bacterial viruses. Animal viruses exploit the organization of host cells in very efficient way and it is likely that bacteriophages do the same. We designed and created a vector library (Fig. 1) and utilized it in expression of virus specific proteins, but the approach can be used easily for other research frames to produce proteins with fluorescent tags. We observed mainly polar localization of several PRD1 viral proteins (Fig. 3). The clear localization was observed only with multimeric proteins as monomeric proteins seemed to be evenly distributed. We also showed that the host receptor recognition vertex associated proteins, the spike protein P5 and the penton protein P31 co-localize within specific cell areas in *E. coli*. These polar areas might play a role in the multimerization and formation of viral protein complexes.

During the virus life cycle, viral proteins are expressed and function according to precise scheme in a close interaction with each other. When a single viral protein is produced from a plasmid, the protein loses these interactions occurring during the viral replication cycle and this might have an influence on the protein functions. However, in the absence of the natural virus infection context viral proteins tested here showed a clear polarized localization in the host cells. The specific localization pattern of these virus proteins suggests that localization is determined either by the proteins themselves and/or by their interactions with other proteins. This indicates that viral proteins are not randomly distributed in the host cell and their (polar) localization might be explained by interaction with specific bacterial proteins. Based on our results, we suggest that viral proteins are interacting with specific bacterial proteins essential for the viral infection. However, many fundamental aspects regarding the molecular mechanisms of interactions as well as the specific bacterial interactions partners remain to be elucidated.

4. Materials and methods

4.1. Bacteria, plasmids, molecular cloning and protein expression

The strains and plasmids used in this study are listed in Table 1. Cells were grown on Luria-Bertani (LB) medium with chloramphenicol (25 µg/ml) and kanamycin (25 µg/ml) when appropriate. *E. coli* HB101 was used as a host for plasmid propagation and molecular cloning. Genes *egfp*, *eyfp* and *ecfp* were amplified by PCR using pEGFP-N3, pECFP-N3 and pEYFP-N3 as templates with primers including a ribosomal binding site (RBS) and restriction enzyme cleavage sites. The fragments were cloned between *Eco*RI and *Hind*III restriction sites in pSU18 and pET24 vectors. Site-directed mutagenesis was used to insert a linker encoding six glycines and new restriction enzyme cutting sites in one end of fluorescent protein genes resulting in 12 new vectors (Fig. 1; Table 1). These vectors were used to construct plasmids for production of C- and N-terminal fusions with yellow or cyan fluorescent proteins. For fusion construction PRD1 genes *II*, *III*, *V*, *XVI*, *XVII* and *XXXI* were amplified by PCR using the phage genome as a template and specific

primers, containing restriction enzyme cutting sites and when needed RBS. The fragments were cloned between *Xba*I and *Bam*HI restriction sites in pKM41, pJK24, pKM54 or pKM64 (Fig. 1; Table 1). The plasmids were sequenced using the Sanger sequencing method with an automatic sequencer (Applied Biosystems 3130xl Genetic Analyzer) and a BigDye Terminator, version 3.1, Cycle Sequencing Kit (Applied Biosystems). Base calling and sequence refining were performed with Sequencing Analysis, version 5.2.0 (Applied Biosystems). Plasmids were transformed to *E. coli* HMS174(DE3) cells, which were used for protein expression.

The cells were grown at 28 °C. The protein production was induced at A550 = 0.75 by adding isopropyl β-d-1-thiogalactopyranoside (IPTG, final concentration of 1 mM) and the growth was continued for 16–20 h at 18 °C. The strains used for the co-expression of P5-eCFP and eYFP-P31 and for co-expression of eYFP-P5 and P31-eCFP were HMS174(DE3)(pSSM21)(pSSM22) and HMS174(DE3)(pJK10)(pJK12), respectively.

4.2. Solubility and multimericity of the proteins

Cells were grown as described above and concentrated 1:100 by centrifugation (Sorvall SLA3000, 4200 g, 10 min, 5 °C) in 50 mM Tris–HCl, pH 7.2. Cells were disrupted by a French pressure cell. To analyze the solubility of the proteins, the protein ratio between the supernatant and pellet was determined after centrifugation (Sorvall SLA3000 rotor, 10,800 g, 15 min, 5 °C) by SDS-PAGE and Western blotting. SDS-PAGE was performed according to previously reported method (Olkkonen and Bamford, 1989) and for Western blotting the proteins were transferred onto a PVDF membrane (Millipore). Monoclonal 16A201 (anti-P16) serum (Hänninen et al., 1997) polyclonal antisera agaist PRD1 proteins P2, P3, P5 and P31 antisera (Grahn et al., 1999, Hänninen et al., 1997, Rydman et al., 1999 and Rydman et al., 2001) or anti-GFP (Invitrogen) were used as primary antibodies. Proteins were visualized with the Thermo Scientific Supersignal West Pico Chemiluminescent Substrate kit using HRP-conjugated swine anti-rabbit Igs (Dako) as a secondary antibody.

For the protein multimericity assay the supernatant was applied on a top of a 10–40% (w/v) linear sucrose gradient in 50 mM Tris–HCl, pH 7.2 and centrifuged (Beckmann SW41 rotor, 210,000 × g, 42 h, 15 °C). Lysozyme (14.3 kDa), albumin (66 kDa), bovine serum albumin (68 kDa), lactate dehydrogenase (140 kDa), aldolase (158 kDa), catalase (232 kDa), ferritin (440 kDa) and thyroglobulin (669 kDa) were used as molecular mass markers. After centrifugation twelve 1 ml-fractions and the pellet were collected and analyzed by SDS-PAGE and Western blotting (see above).

PRD1 mutants sus690 (amber mutation in gene *V*) and sus525 (amber mutation in gene *XXXI*) were propagated on Salmonella enterica suppressor strain PSA (*sup*E) or DB7156 (supF30) harboring pLM2, respectively (Table 1). The titers of the viruses were determined on their suppressor strain and on the non-suppressing strain *S. enterica serovar Typhimurium* LT2 DS88 (wt host; Table 1). The functions of the fluorescent fusion proteins (eYFP-P5, P5-eYFP, P31-eYFP and eYFP-P31) were tested with the mutant viruses by in vivo complementation assay using plasmids carrying either the genes for the fusion proteins or the corresponding PRD1 wt genes *V* and *XXXI* (Table 1). The PRD1 sensitive strain carrying only the cloning vector was used as a negative control.

4.3. Confocal microscopy and localization

The cells were grown as described and diluted 1:1 in phosphate-buffered saline buffer (PBS). Plates were coated with poly-l-lysine (0.01%, MW 70,000–150,000) and cell suspension was applied to the plates and incubated for 10–20 min. Excess of cell suspension was removed. The samples were covered with LB-soft-agar and imaged immediately.

The imaging was performed with an Olympus FV1000 laser scanning confocal microscope attached to an IX81 inverted microscope frame (Olympus, Japan) with an UPLSAPO 60x water immersion objective having a numerical aperture (NA) 1.20 for live samples or UPLSAPO 60× oil immersion objective (NA = 1.3) for fixed samples. eGFP was excited with 488 nm laserline, eCFP with the 405 nm laserline, and eYFP with 515 nm laserline, and the fluorescence signals were collected with 500–600 nm, 425–525 nm, and 530–630 nm band-pass filters, respectively. In co-localization imaging 458 nm laserline was used for eCFP excitation and the fluorescence was collected with 465–505 nm band-pass filter. Images were captured with an image size of 512 pixels × 512 pixels. For presentation purpose (Fig. 3) the images were cropped further to the size of 18 µm × 18 µm. The number of cells with localization loci and the average amount of loci in one cell were calculated from the images from three separate cultivations.

Acknowledgements

We thank Kati Mökkönen for the help in DNA cloning and Dr. Vladimir Chukharev for technical support. This study was supported by the Academy of Finland Centre of Excellence Program in Virus Research (11296841, 2006–2011 J.K.H.B.), the Centre of Excellence Program in Biological Interactions (252411, 2012–2014 J.K.H.B.), Academy of Finland grants 127665 (H.M.O.), 138063 (J.A.I.) and 251106 (J.K.H.B.) and a grant from the Finnish Cultural Foundation (S.M.).

References

Abrescia, N.G., Bamford, D.H., Grimes, J.M., Stuart, D.I., 2012. Structure unifies the viral universe. *Annu. Rev. Biochem*. 81, 795-822.

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

Amster-Choder, O., 2011. The compartmentalized vessel: The bacterial cell as a model for subcellular organization (a tale of two studies). *Cell.* Logist 1, 77-81.

Bamford, J.K., Bamford, D.H., 1990. Capsomer proteins of bacteriophage PRD1, a bacterial virus with a membrane. *Virology* 177, 445-451.

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

Bamford, J.K., Hänninen, A.L., Pakula, T.M., Ojala, P.M., Kalkkinen, N., Frilander, M., Bamford, D.H., 1991. Genome organization of membrane-containing bacteriophage PRD1. *Virology* 183, 658-676.

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

Bartolome, B., Jubete, Y., Martinez, E., de la Cruz, F., 1991. Construction and properties of a family of pACYC184-derived cloning vectors compatible with pBR322 and its derivatives. *Gene* 102, 75-78.

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

Caldentey, J., Hänninen, A.L., Holopainen, J.M., Bamford, J.K., Kinnunen, P.K., Bamford, D.H., 1999. Purification and characterization of the assembly factor P17 of the lipid-containing bacteriophage PRD1. *Eur. J. Biochem.* 260, 549-558.

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

Chen, I., Christie, P.J., Dubnau, D., 2005. The ins and outs of DNA transfer in bacteria. *Science* 310, 1456-1460.

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

Deich, J., Judd, E.M., McAdams, H.H., Moerner, W.E., 2004. Visualization of the movement of single histidine kinase molecules in live Caulobacter cells. *Proc. Natl. Acad. Sci. U. S. A.* 101, 15921-15926.

Edgar, R., Rokney, A., Feeney, M., Semsey, S., Kessel, M., Goldberg, M.B., Adhya, S., Oppenheim, A.B., 2008. Bacteriophage infection is targeted to cellular poles. *Mol. Microbiol.* 68, 1107-1116.

Gerdes, K., Howard, M., Szardenings, F., 2010. Pushing and pulling in prokaryotic DNA segregation. *Cell* 141, 927-942.

Gestwicki, J.E., Lamanna, A.C., Harshey, R.M., McCarter, L.L., Kiessling, L.L., Adler, J., 2000. Evolutionary conservation of methyl-accepting chemotaxis protein location in Bacteria and Archaea. *J. Bacteriol.* 182, 6499-6502.

Grahn AM, Butcher SJ, Bamford JKH, Bamford DH., 2006. PRD1 – dissecting the genome, structure and entry. In: Calendar R. (Ed.), The Bacteriophages. Oxford University Press, New York, pp. 161-170.

Grahn, A.M., Caldentey, J., Bamford, J.K., Bamford, D.H., 1999. Stable packaging of phage PRD1 DNA requires adsorption protein P2, which binds to the IncP plasmid-encoded conjugative transfer complex. *J. Bacteriol.* 181, 6689-6696.

Hänninen, A.L., Bamford, D.H., Bamford, J.K., 1997. Probing phage PRD1-specific proteins with monoclonal and polyclonal antibodies. *Virology* 227, 198-206.

Hellwig, D., Münch, S., Orthaus, S., Hoischen, C., Hemmerich, P., Diekmann, S., 2008. Live-cell imaging reveals sustained centromere binding of CENP-T via CENP-A and CENP-B. *J. Biophotonics* 1, 245-254.

Holopainen, J.M., Saily, M., Caldentey, J., Kinnunen, P.K., 2000. The assembly factor P17 from bacteriophage PRD1 interacts with positively charged lipid membranes. *Eur. J. Biochem.* 267, 6231-6238.

Huiskonen, J.T., Manole, V., Butcher, S.J., 2007. Tale of two spikes in bacteriophage PRD1. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6666-6671.

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

Jakutyte, L., Baptista, C., Sao-Jose, C., Daugelavicius, R., Carballido-Lopez, R., Tavares, P., 2011. Bacteriophage infection in rod-shaped gram-positive bacteria: evidence for a

preferential polar route for phage SPP1 entry in *Bacillus subtilis*. *J. Bacteriol.* 193, 4893-4903.

Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M., Johnson, J.E., 2005. Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc. Natl. Acad. Sci. U. S. A.* 102, 18944-18949.

Lai, E.M., Nair, U., Phadke, N.D., Maddock, J.R., 2004. Proteomic screening and identification of differentially distributed membrane proteins in *Escherichia coli*. *Mol. Microbiol.* 52, 1029-1044.

Leonardy, S., Miertzschke, M., Bulyha, I., Sperling, E., Wittinghofer, A., Sogaard-Andersen, L., 2010. Regulation of dynamic polarity switching in bacteria by a Ras-like G-protein and its cognate GAP. *EMBO J.* 29, 2276-2289.

Lindner, A.B., Madden, R., Demarez, A., Stewart, E.J., Taddei, F., 2008. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc. Natl. Acad. Sci. U. S. A.* 105, 3076-3081.

Lloyd-Price, J., Häkkinen, A., Kandhavelu, M., Marques, I.J., Chowdhury, S., Lihavainen, E., Yli-Harja, O., Ribeiro, A.S., 2012. Asymmetric disposal of individual protein aggregates in *Escherichia coli*, one aggregate at a time. *J. Bacteriol.* 194, 1747-1752.

Li, G., Young, K.D., 2012. Isolation and identification of new inner membrane-associated proteins that localize to cell poles in *Escherichia coli*. *Mol. Microbiol.* 84, 276-295.

Loose, M., Kruse, K., Schwille, P., 2011. Protein self-organization: lessons from the min system. *Annu. Rev. Biophys.* 40, 315-336.

Maddock, J.R., Alley, M.R., Shapiro, L., 1993. Polarized cells, polar actions. *J. Bacteriol.* 175, 7125-7129.

Maier, K.S., Hubich, S., Liebhart, H., Krauss, S., Kuhn, A., Facey, S.J., 2008. An amphiphilic region in the cytoplasmic domain of KdpD is recognized by the signal recognition particle and targeted to the *Escherichia coli* membrane. *Mol. Microbiol.* 68, 1471-1484.

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

Mindich, L., Cohen, J., Weisburd, M., 1976. Isolation of nonsense suppressor mutants in *Pseudomonas*. *J. Bacteriol.* 126, 177-182.

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

Muñoz-Espín, D., Daniel, R., Kawai, Y., Carballido-Lopez, R., Castilla-Llorente, V., Errington, J., Meijer, W.J., Salas, M., 2009. The actin-like MreB cytoskeleton organizes viral DNA replication in bacteria. *Proc. Natl. Acad. Sci. U. S. A.*

Muñoz-Espín, D., Holguera, I., Ballesteros-Plaza, D., Carballido-Lopez, R., Salas, M., 2010. Viral terminal protein directs early organization of phage DNA replication at the bacterial nucleoid. *Proc. Natl. Acad. Sci. U. S. A.* 107, 16548-16553.

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

Nevo-Dinur, K., Govindarajan, S., Amster-Choder, O., 2012. Subcellular localization of RNA and proteins in prokaryotes. *Trends Genet.* 28, 314-322.

Nevo-Dinur, K., Nussbaum-Shochat, A., Ben-Yehuda, S., Amster-Choder, O., 2011. Translation-independent localization of mRNA in *E. coli. Science* 331, 1081-1084.

Oksanen H.M., Bamford D.H., 2012. Family *Tectiviridae*, in: King A.M.Q., Adams M.J., Carstens E.B., Lefkowitz E.J. (Eds.), Virus taxonomy, Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier, Oxford, pp 317-322.

Olsen, R.H., Siak, J.S., Gray, R.H., 1974. Characteristics of PRD1, a plasmid-dependent broad host range DNA bacteriophage. *J. Virol.* 14, 689-699.

Onuki, R., Nagasaki, A., Kawasaki, H., Baba, T., Uyeda, T.Q., Taira, K., 2002. Confirmation by FRET in individual living cells of the absence of significant amyloid beta -mediated caspase 8 activation. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14716-14721.

Ramamurthi, K.S., 2010. Protein localization by recognition of membrane curvature. *Curr. Opin. Microbiol.* 13, 753-757.

Redrejo-Rodriguez, M., Muñoz-Espín, D., Holguera, I., Mencia, M., Salas, M., 2012. Functional eukaryotic nuclear localization signals are widespread in terminal proteins of bacteriophages. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18482-18487.

Rothenberg, E., Sepúlveda, L.A., Skinner, S.O., Zeng, L., Selvin, P.R., Golding, I., 2011. Single-virus tracking reveals a spatial receptor-dependent search mechanism. *Biophys. J.* 100, 2875-2882.

Rudner, D.Z., Losick, R., 2002. A sporulation membrane protein tethers the pro-sigmaK processing enzyme to its inhibitor and dictates its subcellular localization. *Genes Dev.* 16, 1007-1018.

Rudner, D.Z., Losick, R., 2010. Protein subcellular localization in bacteria. *Cold Spring Harb Perspect. Biol.* 2, a000307.

Russell, J.H., Keiler, K.C., 2007. Peptide signals encode protein localization. *J. Bacteriol.* 189, 7581-7585.

Russell, J.H., Keiler, K.C., 2008. Screen for localized proteins in *Caulobacter crescentus*. *PLoS One* 3, e1756.

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

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

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

Shimozono, S., Hosoi, H., Mizuno, H., Fukano, T., Tahara, T., Miyawaki, A., 2006. Concatenation of cyan and yellow fluorescent proteins for efficient resonance energy transfer. *Biochemistry* 45, 6267-6271.

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

Sourjik, V., Berg, H.C., 2000. Localization of components of the chemotaxis machinery of Escherichia coli using fluorescent protein fusions. *Mol. Microbiol.* 37, 740-751.

Swulius, M.T., Jensen, G.J., 2012. The helical MreB cytoskeleton in *Escherichia coli* MC1000/pLE7 is an artifact of the N-Terminal yellow fluorescent protein tag. *J. Bacteriol.* 194, 6382-6386.

Tkachenko, N.V., 2006. Optical Spectroscopy: Methods and Instrumentations. Elsevier, Amsterdam, pp. 115.

Tramier, M., Gautier, I., Piolot, T., Ravalet, S., Kemnitz, K., Coppey, J., Durieux, C., Mignotte, V., Coppey-Moisan, M., 2002. Picosecond-hetero-FRET microscopy to probe protein-protein interactions in live cells. *Biophys. J.* 83, 3570-3577.

Vats, P., Yu, J., Rothfield, L., 2009. The dynamic nature of the bacterial cytoskeleton. *Cell Mol. Life Sci.* 66, 3353-3362.

Winston, F., Botstein, D., Miller, J.H., 1979. Characterization of amber and ochre suppressors in *Salmonella typhimurium*. *J. Bacteriol.* 137, 433-439.

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

Figures and Tables

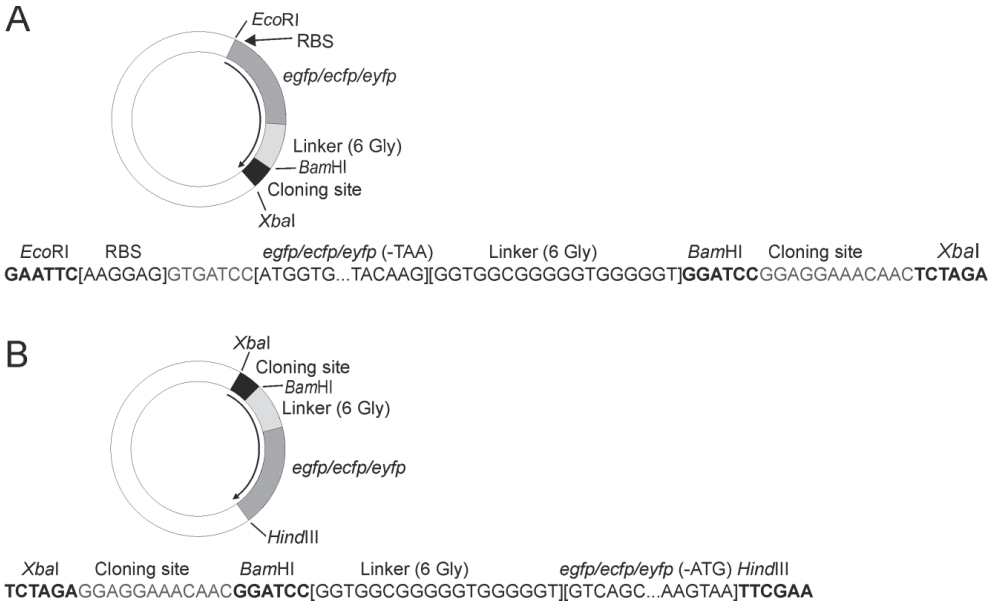


Fig. 1. Vector library. Vectors and sequence of mutated area shown by arrow for production of fluorescent fusion proteins. (A) Vectors for production of N-terminal fusion proteins containing egfp (pKM57 and pJK22), ecfp (pKM51 and pJK24) or eyfp (pKM54 and pJK28). Genes for fluorescent proteins are indicated by dark gray. The cloning site (BamHII-XbaI) and the glysine linker are shown by black and light gray, respectively. (B) Vectors for C-terminal fusion proteins containing egfp (pKM47 and pKM67), ecfp (pKM41 and pKM61) or eyfp (pKM44 and pKM64). Colors are like in (A). See also Table 1.

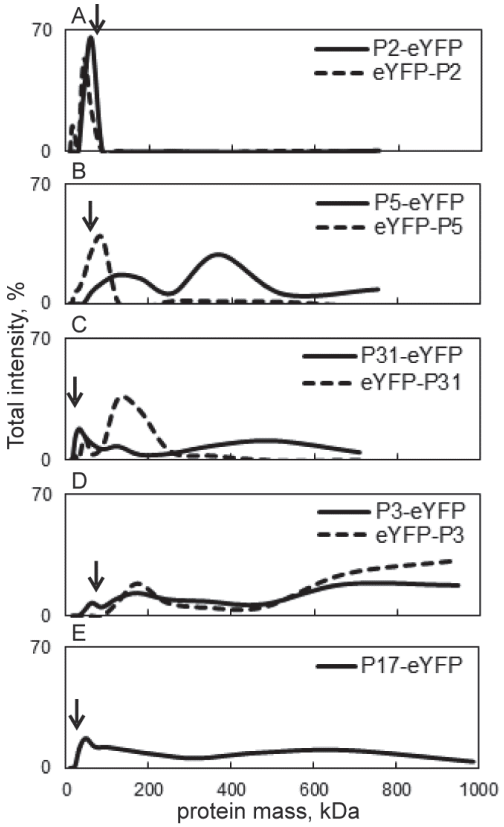


Fig. 2. Sedimentation assay to determine the multimericity of the fluorescent fusion proteins. The masses of the expressed proteins from the soluble fraction of the cell extracts were analyzed by rate zonal centrifugation using standard proteins as a control (see Section 4), SDS-PAGE and Western blotting using specific antibodies against P2, P5, P31 and P3. The P17 fusion protein was identified with antibody against GFP. Calculated monomeric masses are shown by arrows. Only fusion proteins with correct monomeric molecular mass were taken account when creating the image, smaller side products were seen with proteins P5 and P2. (A) Receptor binding protein P2 (monomeric in the virion). (B) Spike protein P5 (trimeric in the virion). (C) Penton protein P31 (pentameric in the virion). (D) Major capsid protein P3 (trimer in the virion). (E) Assembly factor P17 (tetrameric in its native form).

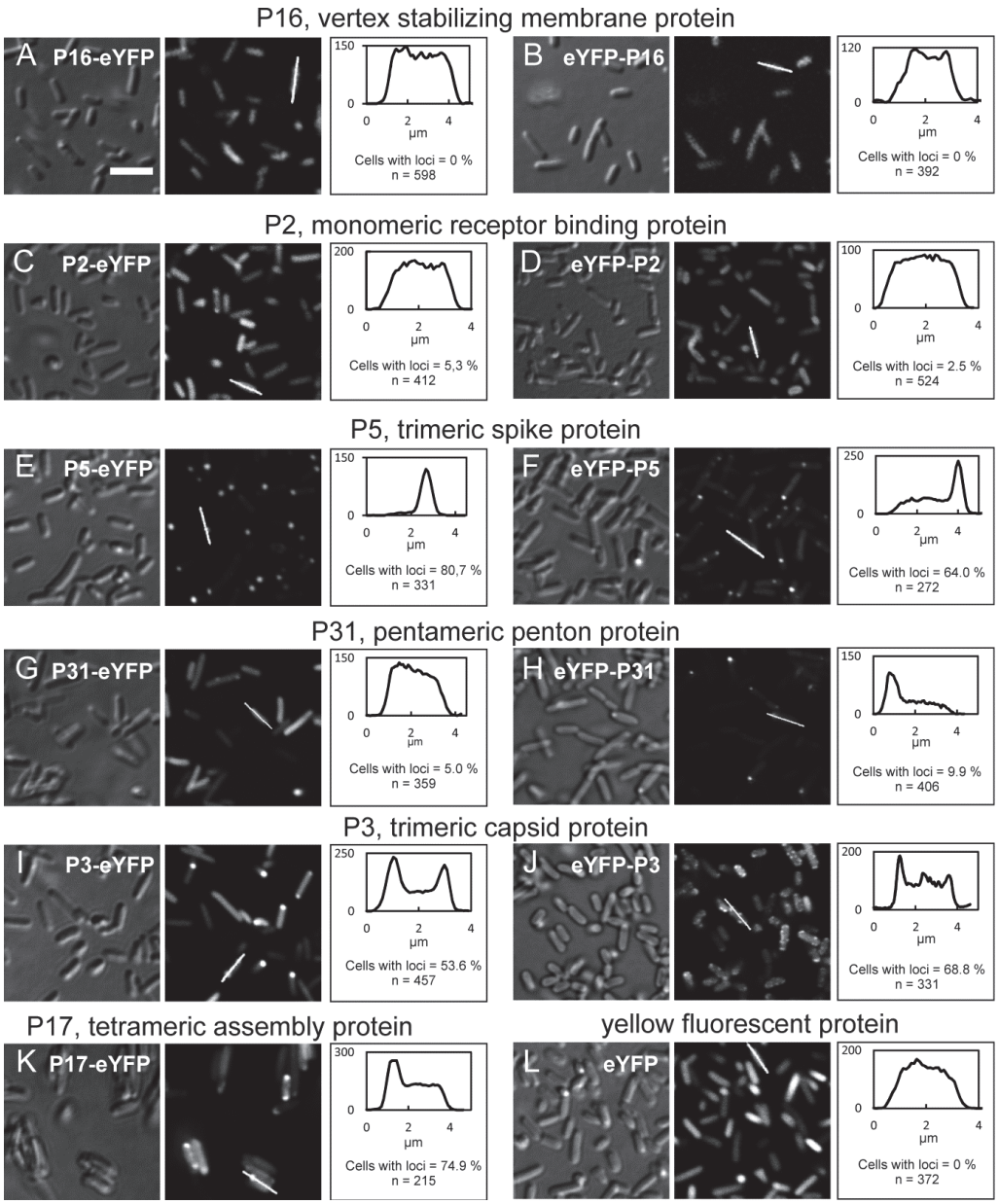


Fig. 3. Localization of PRD1 proteins. Confocal microscope images of different fluorescent fusion proteins of PRD1 overproduced in E. coli HMS174 (DE3) cells. In the left column differential interference contrast (DIC) image, in the middle fluorescence image and on the right intensity (y axes) profile of a single cell marked with a white line (see the fluorescence image) and percentage of cells with loci are shown. The scale bar applicable to all images in (A) is 5 µm.

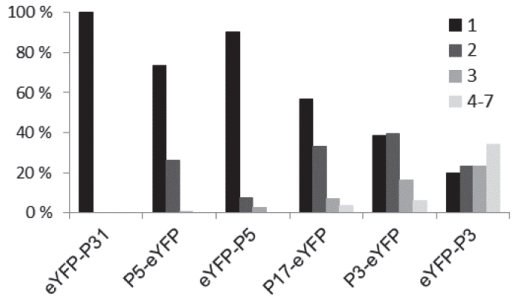


Fig. 4. Amount of loci in single cells. The distribution of amount of loci in one cell calculated from images used in Fig. 3. Only cells with loci were taken account.

Table 1. Phages, bacterial strains and plasmids used in this study

Phages, bacterial strains and plasmids	**Relevant genotype or Description (nt coordinates in PRD1 genome)[a]**	**Relevant phenotype**	**Source or reference**
Phages			
PRD1 wt			Olsen et al., 1974
PRD1 *sus690*	Amber mutation in gene *V*		Bamford and Bamford, 2000
PRD1 *sus525*	Amber mutation in gene *XXXI*		Rydman et al., 1999
Bacterial strains			
Escherichia coli K-12			
HB101	*supE44 hsdS20* ($r_B^-m_B$) *recA13 ara14 proA2 lacY1 galK2 rpsL20 xyl5 mtl1*	Cloning host	Bolivar and Backman, 1979
HMS174(DE3)	*recA1 hsd* Rf^r	Expression host	Campbell et al., 1978
Salmonella enterica serovar Typhimurium LT2			
DS88	*SL5676 ΔH2 H1-i::Tn 10 (TcS)* (pLM2)	Non-suppressor host for PRD1	Bamford and Bamford, 1990
PSA(pLM2)	*supE*	Suppressor host for *sus690*	Mindich et al., 1976
DB7156(pLM2)	*leuA414(Am) hisC527(Am) supF30*	Suppressor host for *sus525*	Winston et al., 1979
Plasmids			
pLM2		Encodes PRD1 receptor	Mindich et al., 1976
pSU18	Low-copy-number cloning vector; p15A replicon, Cm^R		Bartolome et al., 1991
pET24	High-level-expression vector; ColE1 replicon, Km^R		Novagen
pJB500	pSU18 + PRD1 *XXXI* + *V*	P5 and P31	Bamford and Bamford, 2000
pEGFP-N3			Clontech
pECFP-N3			Clontech
pEYFP-N3			Clontech
pJK5	pSU18Δ(*Eco*RI-*Hind*III)Ω(T7 RBS + *egfp* from pEGFP-N3)	eGFP	This study

pJK6	pSU18Δ(*Eco*RI-*Hind*III)Ω(T7 RBS + *ecfp* from pECFP-N3)	eCFP	This study
pJK7	pSU18Δ(EcoRI-HindIII)Ω(T7 RBS + *eyfp* from pEYFP-N3)	eYFP	This study
pSSM1	pET24Δ(*Eco*RI-*Hind*III)Ω(T7 RBS + *egfp* from pEGFP-N3)	eGFP	This study
pSSM2	pET24Δ(*Eco*RI-*Hind*III)Ω(T7 RBS + *ecfp* from pECFP-N3)	eCFP	This study
pSSM3	pET24Δ(*Eco*RI-*Hind*III)Ω(T7 RBS + *eyfp* from pEYFP-N3)	eYFP	This study
pKM57[b]	pJK5Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)	Cloning vector for fusion protein production	This study
pKM51[b]	pJK6Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)		This study
pKM54[b]	pJK7Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)		This study
pKM67[b]	pJK5Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pKM61[b]	pJK6Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pKM64[b]	pJK7Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pJK22[b]	pSSM1Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)		This study
pJK24[b]	pSSM2Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)		This study
pJK28[b]	pSSM3Ω(*Bam*HI-*Xba*I)Ω(6 x Gly)		This study
pKM47[b]	pSSM1Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pKM41[b]	pSSM2Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pKM44[b]	pSSM3Ω(*Xba*I-*Bam*HI)Ω(6 x Gly)Δ(T7 RBS)		This study
pSSM22	pKM41Δ(*XbaI-BamHI*)Ω(PRD1 gene *V* (5287-6309))	P5-eCFP	This study
pJK8	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *V* (5287-6309))	P5-eYFP	This study
pJK10	pKM54Δ(*Bam*HI-*Xba*I)Ω(PRD1 gene *V* (5287-6309))	eYFP-P5	This study
pJK12	pKM41Δ(*Xba*I-*Bam*HI)Ω(PRD1 gene *XXXI* (4907-5287))	P31-eCFP	This study
pSSM20	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *XXXI* (4907-5287))	P31-eYFP	This study
pSSM21	pKM54Δ(*Bam*HI-*Xba*I)Ω(PRD1 gene *XXXI* (4907-5287))	eYFP-P31	This study
pSSM30	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *II* (3128-4903))	P2-eYFP	This study
pSSM32	pKM54Δ(*Bam*HI-*Xba*I)Ω(PRD1 gene *II* (3128-4903))	eYFP-P2	This study
pSSM39	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *III* (8595-9782))	P3-eYFP	This study
pSSM41	pKM54Δ(*Bam*HI-*Xba*I)Ω(PRD1 gene *III* (8595-9782))	eYFP-P3	This study
pSSM43	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *XVII* (6328-6588))	P17-eYFP	This study
pSSM49	pKM64Δ(*Xba*I-*Bam*HI)Ω(T7 RBS + PRD1 gene *XVI* (11836-12189))	P16 - eYFP	This study
pSSM34	pKM54Δ(BamHI-XbaI)Ω(PRD1 gene *XVI* (11836-12189))	eYFP-P16	This study

[a] Gene Bank Acc No AY848689 (Bamford et al., 1991; Saren et al., 2005)

[b] See details in Fig. 1.

Table 2. Complementation titers of PRD1 mutants on strains producing either P5 or P31 fusion proteins

Strain	Description	Titers (pfu / ml)	
		sus690 (gene *V* mutant)	*sus525* (gene *XXXI* mutant)
DS88	non-suppressor host	2×10^5	1×10^5
PSA	suppressor host	3×10^{11}	
DB7156	suppressor host		2×10^{11}
HMS174(pLM2)(pSU18)	negative control	2.6×10^4	3.3×10^4
HMS174(pLM2)(pJB500)	positive control	1×10^{10}	1.6×10^{10}
HMS174(pLM2)(pJK10)	eYFP-P5	1×10^{10} [a]	
HMS174(pLM2)(pJK8)	P5-eYFP	2.3×10^{10}	
HMS174(pLM2)(pSSM20)	P31-eYFP		2.4×10^{10}
HMS174(pLM2)(pSSM21)	eYFP-P31		4.9×10^{10}

[a] A bit weaker plaques than in other plates

II

MEMBRANE-CONTAINING BACTERIOPHAGE PRD1 PROTEIN P33 COMPLEMENTS THE DEFECT IN GROES OF *ESCHERICHIA COLI*

by

Jenni Karttunen*, Sari Mäntynen*, Teemu O. Ihalainen, Hanna M. Oksanen & Jaana K. H. Bamford

* Equal contribution

Submitted manuscript

III

SPHINGOMYELIN INDUCES STRUCTURAL ALTERATION IN CANINE PARVOVIRUS CAPSID

by

Kirsi Pakkanen, Jenni Karttunen, Salla Virtanen & Matti Vuento 2008

Virus Research 132: 187–91.

Sphingomyelin induces structural alteration in canine parvovirus capsid

Kirsi Pakkanen, Jenni Karttunen, Salla Virtanen, Matti Vuento

Virus Research 132 (2008) 187–191

Nanoscience Center, Department of Biological and Environmental Science, University of Jyväskylä, Finland

Keywords: Canine parvovirus; Sphingomyelin; Acidic pH; Tryptophan fluorescence; Circular dichroism

Abstract

One of the essential steps in canine parvovirus (CPV) infection, the release from endosomal vesicles, is dominated by interactions between the virus capsid and the endosomal membranes. In this study, the effect of sphingomyelin and phosphatidyl serine on canine parvovirus capsid and on the phospholipase A2 (PLA2) activity of CPV VP1 unique N-terminus was analyzed. Accordingly, a significant ($P \leq 0.05$) shift of tryptophan fluorescence emission peak was detected at pH 5.5 in the presence of sphingomyelin, whereas at pH 7.4 a similar but minor shift was observed. This effect may relate to the exposure of VP1 N-terminus in acidic pH as well as to interactions between sphingomyelin and CPV. When the phenomenon was further characterized using circular dichroism spectroscopy, differences in CPV capsid CD spectra with and without sphingomyelin and phosphatidyl serine were detected, corresponding to data obtained with tryptophan fluorescence. However, when the enzymatic activity of CPV PLA2 was tested in the presence of sphingomyelin, no significant effect in the function of the enzyme was detected. Thus, the structural changes observed with spectroscopic techniques appear not to manipulate the activity of CPV PLA2, and may therefore implicate alternative interactions between CPV capsid and sphingomyelin.

Canine parvovirus (CPV) is a member of the autonomous Parvovirus genus of Parvoviridae family (Tattersall and Cotmore, 1988) and has a diameter of approximately 26 nm (Agbandje et al., 1993 and Chapman and Rossmann, 1993), that is, of 260 Å. The capsid surface features 15 Å depressions at the twofold axes, 22 Å protrusions (spikes) at the threefold axes and channels surrounded by 15 Å depressions at fivefold axes (Tsao et al., 1991 and Wu and Rossmann, 1993). Like other parvoviruses, CPV has not been shown to contain lipids or carbohydrates (Berns, 1990).

The capsid is composed of 60 subunits, 90% of which are VP2 proteins (584 amino acid residues) and 10% VP1 proteins (727 amino acid residues), and arranged in icosahedral symmetry (T = 1). The third structural protein, VP3, is formed only in DNA containing capsids following cleavage of VP2 amino terminus by host cell proteases (Paradiso et al., 1982, Cotmore and Tattersall, 1987 and Reed et al., 1988). The monomers of the capsid are formed with eight-stranded antiparallel β-barrel, "jelly roll", structure. The loops connecting barrel strands form partly the surface of the capsid (Tsao et al., 1991). One of the loops in CPV capsid is the site of an essential structural difference between CPV and its closest relative feline panleukopenia virus. The loop between residues 359 and 375 in these viruses differs in orientation and flexibility (Agbandje et al., 1995). The glycine residues at the ends of the loop have been shown to contribute to the flexibility. Also the three calcium-binding sites of the capsid affect the conformation of the loop in pH and Ca2+ concentration-dependent ways (Simpson et al., 2000).

The capsid protein VP1 contains the entire amino acid sequence of VP2 protein in addition to the unique N-terminus of 143 amino acid residues in length. While some of the VP2 N-termini are external, their glycine-rich sequence residing along the fivefold pore, all the VP1 N-termini are buried inside the capsid (Paradiso, 1981 and Weichert et al., 1998). The unique part of VP1 can be exposed *in vitro* using limited heating or urea (Weichert et al., 1998, Cotmore et al., 1999 and Vihinen-Ranta et al., 2002). In cells, the unique part of VP1 is known to change location relative to the capsid surface in the acidic environment of the late stages of endocytic route (Suikkanen et al., 2003 and Farr et al., 2006). In addition to this, Farr et al. have shown that acidity-triggered exposure of minute virus of mice VP1 N-terminus is sensitive to proteolytic cleavage of VP2 to form VP3 (Farr et al., 2006). Importantly, the exposure of VP1 N-termini has been shown to be critical for successful CPV infection (Suikkanen et al., 2003). The VP1 unique region of many parvoviruses, including CPV, has been shown to harbor two motifs. Firstly, the unique part of CPV contains basic sequences resembling classical nuclear localization signals (Vihinen-Ranta et al., 1997). Secondly, the unique part contains two motifs conserved in secretory phospholipase A_2 ($sPLA_2$): HDXXY (amino acid residues 46–51) found in catalytic sites and YXGXG (amino acid residues 20–25) known to be associated with Ca^{2+}-binding loops (Zádori et al., 2001). The unique terminus itself, as well as capsids exposed to acidic pH or heat, have been shown to exhibit phospholipase A_2 (PLA_2) activity (Suikkanen et al., 2003, Zádori et al., 2001 and Canaan et al., 2004). Certain properties of parvoviral PLA_2s differ from $sPLA_2$s. Parvoviral PLA2s do not contain cysteine, the sequence connecting the two catalytic site helices is shorter and the YXGPG sequence of Ca^{2+}-binding loop is strictly conserved (Zádori et al., 2001). Parvoviral PLA_2s do not exhibit substantial specificity to substrates in terms of lipid species or the degree of saturation of sn − 2 acyl fatty acid chains, the latter being a typical feature of $sPLA_2$s (Zádori et al.,

2001 and Canaan et al., 2004). Another characteristic of sPLA$_2$s, Ca^{2+} dependence, is characteristic also for parvoviral PLA$_2$s. In addition, the PLA$_2$ region of VP1 unique part has been proposed to be able to bind one Ca^{2+} ion (Canaan et al., 2004). PLA$_2$ enzymes have higher activity towards membranes or aggregated lipids than lipids free in solution (Verger et al., 1973). This interfacial activation has also been shown with parvoviral phospholipases (Canaan et al., 2004).

Since the exposure of VP1 N-terminus is known to be closely associated with penetration of CPV through endosomal membranes (Weichert et al., 1998, Cotmore et al., 1999, Suikkanen et al., 2003 and Farr et al., 2006), it is fair to assume that important structural changes occur in the acidic conditions of the endosomal vesicles. However, the possible involvement of membranes during the exposure of the VP1 N-terminus or as a facilitator of other structural changes is currently unknown. Here, to gain insight into the effects of sphingomyelin and phosphatidyl serine bilayers combined with neutral and acidic pH on CPV capsids, tryptophan fluorescence and circular dichroism spectroscopy were utilized. In addition, the effect of sphingomyelin on CPV PLA$_2$ activity was characterized.

Canine parvovirus (CPV-2d) used in this study, originally derived from infectious plasmid clone of the virus (Parrish, 1991), was cultured and purified as previously described (Suikkanen et al., 2003) with the exception that the final pellet from ultracentrifugation was dissolved in a small volume of non-buffered saline and diluted with potassium phosphate for analysis. Pooled preparations of infectious CPV capsids containing DNA were used in all experiments except in CD spectroscopy, where pooled empty capsids (not containing DNA) were used to avoid contribution of DNA to CD signal. The purified virus stocks were characterized by polyacrylamide gel electrophoresis using 10% gels and coomassie brilliant blue staining. Concentration of virus was determined with Bradford assay using bovine serum albumin as a standard (Bio-Rad, Hercules, CA). Bovine brain phosphatidyl serine and bovine brain sphingomyelin were from Sigma Aldrich (St. Louis, MO). Lipids were dissolved in chloroform and dried with nitrogen-flow. To prepare small unilamellar vesicles (SUVs), lipids were resuspended in 10 mM potassium phosphate buffer (or in the case of sphingomyelin for PLA$_2$ activity studies: assay buffer) of desired pH to a final lipid concentration of 4 mg/ml. The suspension was vortexed for 20 min and sonicated with a tip sonifier (Branson, Danbury, CT) in an ice-water bath (Lee et al., 2001).

CPV capsid intrinsic tryptophan fluorescence was recorded at +25 °C using a Perkin Elmer LS55 fluorescence spectrophotometer and a cuvette with 10 mm path length (Perkin Elmer Industries, Wellesley, MA). Temperature was maintained within ±1 °C with Thermo Haake water bath and circulator (Thermo Electron Corporation,Waltham, MA). Samples were excited at 290 nm and non-corrected emission was measured at 310–430 nm. Final concentration of CPV in a cuvette was 2.5 µg/ml diluted in 10 mM potassium phosphate buffer at appropriate pH. Final concentration of CaCl2 was 1 mM and of lipids (as SUVs) approximately 0.05 mg/ml. For the acidification studies CPV capsids were acidified with 0.1 M HCl (resulting as final pH of 5.5) for 10 min and subsequently neutralized with 0.1 M NaOH and diluted to measurement volume, resulting in final concentration of 2.5 µg/ml CPV. Excitation and emission slits were set at 2.5 nm. Background fluorescence of liposomes, $CaCl_2$ and buffer were subtracted from corresponding sample data. The

spectra were averaged from three scans and emission maxima taken from three sets of averaged results. The emission maxima data were analyzed using the unpaired Student's t-test with a two-tailed P-value, and statistical significance was determined relative to the control ($CaCl_2$ only) samples (**$P < 0.05$).

The CPV capsid contains approximately 850 tryptophan (Trp) residues (14 in each VP2 protein and 15 in each VP1 protein), mostly located inside the capsid subunits in positions most likely shielded from water molecules (Chapman and Rossmann, 1993, Reed et al., 1988 and Xie and Chapman, 1996). Tryptophan residues have emission maximum near 330 nm when located in hydrophobic environment and closer to 350 nm when in the hydrophilic areas of the protein. In some cases this red-shift is due to certain parts of the protein being moved closer to the surface of a protein complex as a result of conformational changes. On other occasions, the red-shift is influenced by access of aqueous buffer inside a protein or protein subunits of a larger multiprotein complex (Lacowicz, 1999 and Vivian and Callis, 2001), indicating loosening of binding interactions between subunits. Fluorescence of tryptophan residues, fully exposed to water, is not affected by changes in pH, as demonstrated with N-acetyl tryptophan amide (Luykx et al., 2004).

Here, environment-induced changes in CPV virions were monitored using the intrinsic fluorescence of Trp residues. CPV capsid is known to bind Ca^{2+} ions (Simpson et al., 2000), which may provide stabilization to the capsid structure. To mimic this effect of intracellular environment, fluorescence measurements with lipids were performed using $CaCl_2$-containing solutions. Red-shifts in emission peak position of Trp fluorescence of CPV were detected when the effect of sphingomyelin and phosphatidyl serine membranes were examined (Fig. 1A). At pH 5.5, sphingomyelin induced a statistically significant red-shift (3.8 ± 0.3 nm;$P = 0.002$) in tryptophan fluorescence. Similar, but statistically insignificant, red-shifts were observed also with sphingomyelin at pH 7.4 (5.9 ± 3.8 nm) and in the presence of phosphatidyl serine in both acidic and neutral conditions (4.2 ± 2.5 and 0.9 ± 1 nm, respectively). Phosphatidyl choline, phosphatidyl ethanolamine, and phosphatidyl inositol induced no apparent shifts in the emission peak positions (not shown). These results suggest that sphingomyelin- and possibly phosphatidyl serine-containing membranes induce structural changes to CPV capsid, leading to possible water penetration into the virion. To understand the effect of acidic conditions on CPV capsid more profoundly, CPV capsids were exposed shortly to pH 5.5 and neutralized after 10 min and analyzed with fluorescence spectroscopy (Fig. 1D). Despite the seeming decrease in emission peak wavelength, no statistically significant change in Trp fluorescence emission peak positions were detected ($P = 0.458$).

Results obtained with tryptophan fluorescence were further analyzed by circular dichroism spectroscopy in far UV region using a Jasco J-720 spectropolarimeter (JASCO Corporation, Tokyo, Japan) and a cylindrical quarz cuvette (Helma GmbH & Co KG, Müllheim, Germany) with a pathlength of 1 mm. Before use in CD measurements empty (not containing DNA) capsids of CPV were dialysed against 10 mM potassium phosphate buffer to avoid the effect of sodium ions on CD signal. 20 μg micrograms of capsids were used in each measurement. The spectra were recorded at 240–190 nm with bandwidth of 2.0 nm, 0.5 s response time. Scan speed was 10 nm/min. The spectra shown are

representative averages of three scans. Standard deviations shown in Fig. 1C are from five experiments. Final concentration of $CaCl_2$ was 1 mM in the cuvette. SUVs (prepared as described above) used in CD measurements were composed of phosphatidyl serine and sphingomyelin (molar ratio 1:1) with total lipid concentration of 4 mg/ml. Final concentration on lipids (as liposomes) in the cuvette was approximately 0.05 mg/ml. Baseline was measured with or without SUVs. The capsids for acidification studies were treated as described with tryptophan fluorescence measurements. 40 µg of capsids were used in each CD measurement. The spectra were drawn from data point values after subtraction of baseline, and fitted to Stineman function in the Kaleida Graph software Version 3.6.4 (Synergy Software, Reading, PA).

CD spectra of empty CPV capsids with Ca^{2+} and sphingomyelin-phosphatidyl serine liposomes or with Ca^{2+} only are presented in Fig. 1B and standard deviations of the 220–200 nm region of the spectra are in Fig. 1C. The overall form of the spectra features a wide negative valley at 220–200 nm and a positive peak at 190 nm. With Ca^{2+} only, the negative valley at 220–200 nm was less profound and the positive peak at 190 nm was lower at pH 5.5 than at pH 7.4. However, the presence of sphingomyelin-phosphatidyl serine liposomes with Ca^{2+} markedly changed the CD spectra of the capsids. At pH 7.4, the negative valley at 220–200 nm became less negative and the positive peak at 190 nm became more negative in the presence of liposomes and Ca^{2+} in comparison with Ca^{2+} only. On the contrary, at pH 5.5, the presence of liposomes shifted the 190 nm peak to a positive and the 220–200 nm valley to a more negative direction (and especially the 220 nm side of the valley). The CD data were not interpreted in terms of secondary structure. However, these data suggest that sphingomyelin-phosphatidyl serine membranes induce structural changes in CPV in both acidic and neutral conditions, supporting the results obtained with tryptophan fluorescence. The spectral change observed in acidic pH in the presence of Ca^{2+} without liposomes is likely to be at least partly associated with VP1 N-terminus exposure, but some other structural modifications need to be present to justify the differences between spectra with liposomes and with Ca^{2+} only. It is possible that these structural changes detected are, at least to some extent, responsible for the differences in water penetration into the capsid as detected with tryptophan fluorescence measurements. Also, the results of acidification/neutralization studies were further characterized with CD spectroscopy. The CD spectra (240–200 nm region) of native and acid-treated capsids are presented in Fig. 1E and standard deviations of the 220–200 nm region in Fig. 1F. The spectrum of acid-treated capsids resembles that of native capsids, but has some distinguishing features. The 220–200 nm valley is less profound in acid-treated capsids and there is a distinct shoulder at 225 nm. The differences in data obtained with acidified capsids with tryptophan fluorescence spectroscopy and CD spectroscopy are interesting and may relate to previously unknown long-term effects of the low pH step in CPV infection.

PLA_2 activities were measured using a commercial kit (Cayman Chemicals, Ann Arbor, MI, USA) with bee venom PLA2 as a control. For each sample, 35 µg CPV or bee venom PLA2 was used. In the case of added lipids, the reaction mixture was supplemented with 100 µg sphingomyelin as SUVs prepared (as above) in assay buffer. In reaction mixture without added sphingomyelin only assay buffer was used. Absorbances were measured using a Perkin Elmer Lambda 850 spectrophotometer. Temperature was maintained at

+25 °C using a water bath and circulator (Thermo Electron Corporation, Waltham, MA). PLA2 activities were measured and calculated according to manufacturer's instructions. Enzyme activity in the presence of sphingomyelin was analyzed using the unpaired Student's t-test with a two-tailed P value, and statistical significance was determined relative to the control (no added lipid) samples (**$P < 0.05$).

To exclude the potential effect of these putative structural changes on PLA_2 function, the enzymatic activity of CPV PLA_2 was tested using thiol-phosphatidyl choline substrate supplemented with sphingomyelin. The results suggested CPV PLA_2 to function in the presence of sphingomyelin similarly as the bee venom PLA_2 used as a control. Since, the activity of CPV PLA_2 was only slightly decreased, from 0.0019 ± 0.0003 μmol (ml min)$^{-1}$ without sphingomyelin to 0.0016 ± 0.00031 μmol (ml min)$^{-1}$ with sphingomyelin (Fig. 2). This decrease was found to be statistically insignificant ($P = 0.2621$). It therefore seems that the addition of sphingomyelin into the reaction mixture has no significant effect on the activity of either CPV PLA_2. This further implies that the sphingomyelin-related conformational changes detected with spectroscopic methods most likely are not involved in facilitation of CPV PLA_2 function and may therefore suggest some other so far unknown mechanisms related to the endosomal escape process of CPV.

During its endocytic entry, CPV encounters the acidic intraendosomal environment. The virus travels to late endosomes and possibly even further to lysosomes. CPV is then released from these compartments to the cytoplasm, to undertake a further voyage to the nucleus (Suikkanen et al., 2003 and Reed et al., 1988). The exact mechanism of the release is unknown, although it has been shown that the CPV-containing endosomes are not broken down and become permeable only to particles smaller than approximately 10–20 kDa, as shown with dextran beads (Suikkanen et al., 2003) and with sarcin (Parker and Parrish, 2000). During the release, at the latest, CPV comes into contact with endosomal membranes. In addition, the virions appear to have affinity for some membrane lipids (Suikkanen et al., 2003). The low pH of endosomes is necessary for CPV release, as inhibitors of V-type ATPase, responsible for acidification of the endosomal lumen, retain CPV in the endosomes (Suikkanen et al., 2003 and Parker and Parrish, 2000). The release of CPV from late endosomal compartments is inhibited by PLA_2 inhibitors, strongly suggesting that the enzymatic activity has a crucial role during viral escape. However, exposure of the N-terminal domain and activation of PLA_2 assisted by endosomal conditions are probably not enough to promote the release, since endocytosed, pre-activated capsids are retained in the endosomes in the presence of V-ATPase inhibitors similarly to intact capsids (Suikkanen et al., 2003). Thus, other conformational rearrangements in the capsid that facilitate the endosomal escape may occur. Such changes could be related to low pH, but also to contact of capsids with endosomal membranes. In this study we have characterized lipid or membrane induced changes as well as acid-induced changes in CPV capsids. The data suggest that sphingomyelin induces changes in CPV capsids, which, at least to some extent, differ from those induced by acidic conditions. However, these changes do not seem to be related to PLA_2 activity of the capsid, which may indicate that membrane–capsid interactions have a role in the membrane penetration process of the virus even outside the scope of lipase action. In addition to this, we have shown that acidic conditions induce long-term changes in CPV

capsids. Even though the permeability to water seems to recover to almost original level after the capsid moves to neutral environment; some changes in the structure still remain. This may have significance in later steps of the viral life cycle.

Acknowledgements

We are especially grateful to Ms. Anna R. Mäkelä, MSc, for her valuable advice during the writing process. We also thank Dr. Tuula O. Jalonen for critically reading the manuscript, Pirjo Kauppinen for technical assistance and Dr. Ilkka Kilpeläinen and Dr. Tero Pihlajamaa for the use of CD spectroscopy facilities. This study was supported by a grant from the Rector of University of Jyväskylä and the National Graduate School in Nanoscience (K.P.) and by Academy of Finland (contract # 102161).

References

M. Agbandje, R. Mckenna, M.G. Rossmann, M.L. Strassheim, C.R. Parrish. Structure determination of feline panleukopenia virus empty particles. *Proteins*, 16 (2) (1993), pp. 155–171.

M. Agbandje, C.R. Parrish, M.G. Rossmann. The structure of parvoviruses. *Semin. Virol.*, 6 (1995), pp. 299–309

K.I. Berns, Parvovirus replication. *Microbiol. Rev.*, 54 (1990), pp. 316–329

S. Canaan, Z. Zádori, F. Ghomashchi, J. Bollinger, M. Sadilek, M.E. Moreau, P. Tijssen, M.H. Gelb. Interfacial enzymology of parvovirus phospholipases A2. *J. Biol. Chem.*, 279 (15) (2004), pp. 14502–14508

M.S. Chapman, M.G. Rossmann. Structure, sequence and function correlations among parvoviruses. *Virology*, 94 (1993), pp. 491–508

S.F. Cotmore, P. Tattersall. The autonomously replicating parvoviruses of vertebrates. *Adv. Virus Res.*, 33 (1987), pp. 91–169

S.F. Cotmore, A.M. D'abramo Jr., C.M. Ticknor, P. Tattersall. Controlled conformational transitions in the MVM virion expose the VP1 N-terminus and viral genome without particle disassembly. *Virology*, 254 (1999), pp. 169–181

G.A. Farr, S.F. Cotmore, P. Tattersall. VP2 cleavage and the leucine ring at the base of the fivefold cylinder control pH-dependent externalization of both the VP1 N-terminus and the genome of minute virus of mice. *J. Virol.*, 80 (1) (2006), pp. 161–171

J.R. Lacowicz. Principles of Fluorescence Spectroscopy. *Plenum Press*, NY, USA (1999)

S.-K. Lee, C. Dabney-Smith, D.L. Hacker, B.D. Bruce. Southern cowpea mosaic virus coat protein: the role of basic amino acids, helix-forming potential, and lipid composition. *Virology*, 291 (2001), pp. 299–310

D.M.A.M. Luykx, M.G. Castelejn, W. Jiskoot, J. Westdijk, P.M.J.M. Jongen. Physicochemical studies on the stability of influenza haemagglutinin in vaccine bulk material. *Eur. J. Pharm. Sci*., 23 (2004), pp. 67–75

P.R. Paradiso. Infectious processes of the parvovirus H-1: correlation of protein content, particle density, and viral infectivity. *J. Virol.*, 39 (3) (1981), pp. 800–807

P.R. Paradiso, S.L. III Rhode, I. Singer. Canine parvovirus: a biochemical and ultrastructural study. *J. Gen. Virol.*, 62 (1982), pp. 113–125

J.S. Parker, C.R. Parrish. Cellular uptake and infection by canine parvovirus involves rapid dynamin-regulated clathrin-mediated endocytosis, followed by slower intracellular trafficking. *J. Virol*., 74 (4) (2000), pp. 1919–1930

C.R. Parrish. Mapping specific functions in the capsid structure of canine parvovirus and feline panleukopenia virus using infectious plasmid clones. *Virology*, 183 (1991), pp. 195–205

A.P. Reed, E.V. Jones, T.J. Miller. Nucleotide sequence and genome organization of canine parvovirus. *J. Virol*., 62 (1) (1988), pp. 266–276

A.A. Simpson, V. Chandrasekar, B. Hebert, G.M. Sullivan, M.G. Rossmann, C.R. Parrish. Host range and variability of calcium binding by surface loops in the capsids of canine and feline parvoviruses. *J. Mol. Biol*., 300 (2000), pp. 597–610

S. Suikkanen, M. Antila, A. Jaatinen, M. Vihinen-Ranta, M. Vuento. Release of canine parvovirus from endocytic vesicles. *Virology*, 316 (2) (2003), pp. 267–280

P. Tattersall, S.F. Cotmore. The nature of parvoviruses. *Parvoviruses and Human DiseaseCRC Press*, Boca Raton, FL (1988)

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

R.M. Verger, C.E. Mieras, G.H. de Haas. Action of phospholipase A at interfaces. *J. Biol. Chem.*, 248 (1973), pp. 4023–4034

M. Vihinen-Ranta, L. Kakkola, A. Kalela, P. Vilja, M. Vuento. Characterization of a nuclear localization signal of canine parvovirus proteins. *Eur. J. Biochem.*, 1 (1997), pp. 389–394

M. Vihinen-Ranta, D. Wang, W.S. Weichert, C.R. Parrish. The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection. *J. Virol.*, 76 (4) (2002), pp. 1884–1891

J.T. Vivian, P.R. Callis. Mechanisms of tryptophan fluorescence shifts in proteins. Biophys. *J.*, 80 (2001), pp. 2093–2109

W.S. Weichert, J.S. Parker, A.T.M. Wahid, S.F. Chang, E. Meier, C.R. Parrish. Assaying for structural variation in the parvovirus capsid and its role in infection. *Virology*, 250 (1998), pp. 106–117

H. Wu, M.G. Rossmann. The canine parvovirus empty capsid structure. *J. Mol. Biol.*, 233 (1993), pp. 231–244

Q. Xie, M.S. Chapman. Canine parvovirus capsid structure, analyzed at 2.9 Å resolution. *J. Mol. Biol.*, 264 (1996), pp. 497–520

Z. Zádori, J. Szelei, M.-C. Lacoste, Y. Li, S. Gariepy, P. Raymond, M. Allaire, I.R. Nabi, P. Tijssen. A viral phospholipase A2 is required for parvovirus infectivity. *Dev. Cell*, 1 (2001), pp. 291–302

Figures

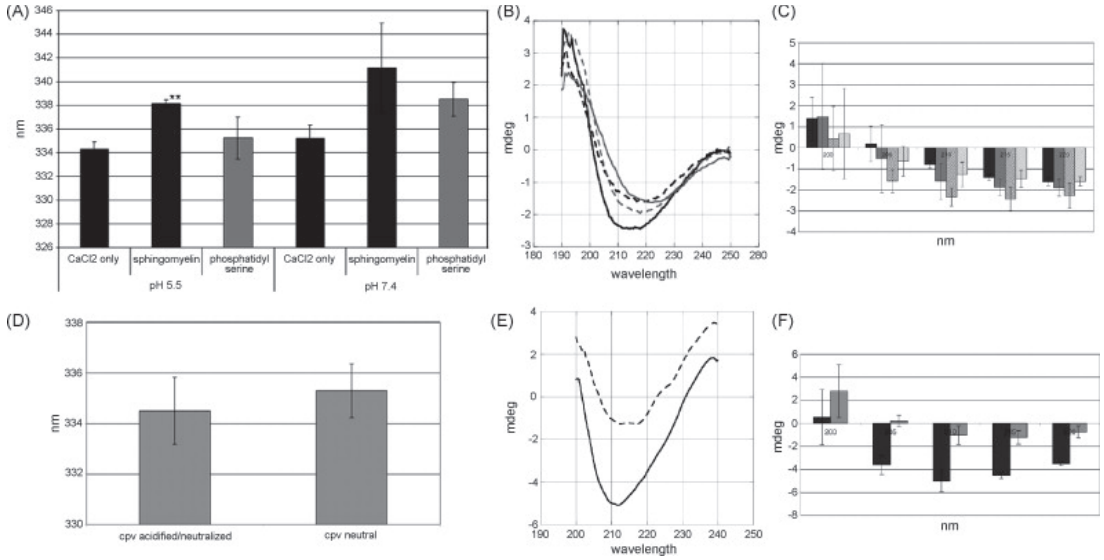


Figure 1 Intrinsic tryptophan fluorescence emission peak positions and CD spectra of CPV capsids with and without sphingomyelin and/or phosphatidyl serine. (A) Emission peak positions of CPV capsid intrinsic tryptophan fluorescence in 10 mM potassium phosphate buffer in +25 °C in the presence of 1mM CaCl2 with and without sphingomyelin and phosphatidyl serine liposomes at pH 5.5 and 7.4. The data were compared using the unpaired Student's t-test with a two-tailed P value, and statistical significance was determined **P < 0.05. (B) CD spectra of empty CPV capsids in the presence of CaCl2 with and without liposomes (sphingomyelin and phosphatidyl serine). CPV with CaCl2 and liposomes at pH 5.5: (--) grey, CPV with CaCl2 at pH 5.5: (—) grey, CPV with CaCl2 and liposomes at pH 7.4: (--) black and CPV with CaCl2 at pH 7.4: (—) black. (C) A blow up of the 220–200 nm region of the CD spectra in B with error bars indicating standard deviations. CPV with CaCl2 at pH 5.5: dark grey, CPV with CaCl2 and liposomes at pH 5.5: intermediate grey, CPV with CaCl2 at pH 7.4: lighter grey and CPV with CaCl2 and liposomes at pH 7.4: lightest grey. (D) Emission peak positions of CPV capsid intrinsic tryptophan fluorescence in neutral conditions with native capsids and capsids acidified at pH 5.5 for 10 min (and subsequently neutralized). (E) CD spectra of acidified and native CPV capsids in neutral conditions. Native CPV: (—) black and acidified CPV: (--) grey. (F) A blow up of 220–200 nm region of the CD spectra in (E) with error bars indicating standard deviations. Native CPV: dark grey and acidified CPV: light grey.

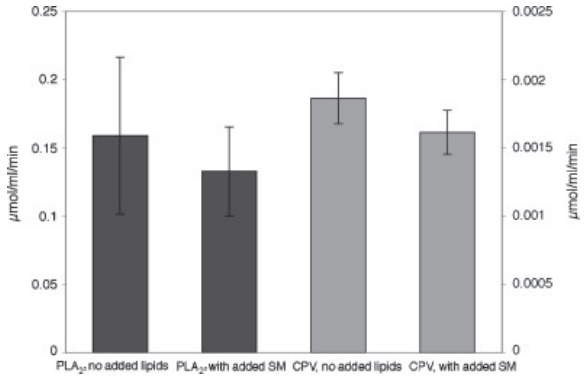


Figure 2 Activity of CPV PLA2 with and without sphingomyelin. Activities of CPV PLA2 and bee venom PLA2 with added and sphingomyelin and without added lipids. The vertical axis on the left corresponds to scaling of the values of PLA2 (dark grey) and the vertical axis on the right-hand side corresponds to scaling of the values of CPV (light grey).