universität wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Investigation of the function of gp44 as well as the construciton of ϕCh1::PCNA and its effect on the development of virus ϕCh1"


verfasst von / submitted by

Michael Tschurtschenthaler, BSc


angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2017 / Vienna 2017

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | A 066 830 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Masterstudium Molekulare Mikrobiologie, Mikrobielle Ökologie und Immunbiologie |
| Betreut von / Supervisor: | Ao. Univ.-Prof. Dipl.-Biol. Dr. Angela Witte |

# CURRICULUM VITAE

## EDUCATION

2015 – 2017 **University of Vienna, Austria**
- Master thesis: Laboratory of Prof. Dr. Angela Witte; Department of Microbiology, Immunbiology and Genetics

2013 – 2017 **University of Vienna, Austria**
- Master's degree program Molecular Microbiology, Microbial Ecology and Immunbiology
- Specialization: Molecular Microbiology

2008 – 2013 **University of Vienna, Austria**
- Bachelor's degree program Biology
- Specialization: Microbiology and Genetics

2004 – 2007 **Aufbaulehrgang HAK1 Klagenfurt, Austria**

2001 – 2004 **Handelsschule Lienz, Austria**

1997 – 2001 **Hauptschule Nord Lienz, Austria**

1993 – 1997 **Volksschule Nord Lienz, Austria**

## WORKING EXPERIENCE

2017 – Present **Fresh Soup and Salad, Austria**
- Shop assistant

2016 – Present **University of Vienna, Austria**
- Examination supervisor

SS 2016 **University of Vienna, Austria**
- Tutor of students for practical training in „ Practical Course in Molecular Microbiology, Microbial Ecology and Immunbiology"

WS 2015 **FH Campus Wien (University of Applied Sciences), Austria**
- Tutor of students for practical training in course of the studies „Molekularbiologie II"

**University of Vienna, Austria**
- Tutor of students for practical training in „Übung IIIA"

SS 2015 **FH Campus Wien (University of Applied Sciences), Austria**
- Tutor of students for practical training in course of the studies „Molekularbiologie III“

**University of Vienna, Austria**
- Tutor of students for practical training in „Übung IIIA“

2014 – 2015 **Fresh Soup and Salad, Austria**
- Shop assistant

2012 **Mise en Place, Austria**
- Waiter

2007 – 2008 **Liebherr, Austria**
- Shift worker

## SUMMER INTERSHIP

2011 **Weinphilo, Austria**
- Waiter

**Stadtgemeine Lienz, Austria**
- Mailman

2010 **Weinphilo, Austria**
- Waiter

2008 **Lienzer Schnitzelwelt, Austria**
- Waiter

2007 **Stadtgemeine Lienz, Austria**
- Cleaner

2006 **Liebherr, Austria**
- Shift worker

2005 **Stadtgärtnerrei Lienz, Austria**
- Gardener

## INTERESTS

**Astronomy, Climbing, Hiking, Skiing**

# Acknowledgment

First of all, i want to thank my supervisor Dr. Angela Witte for her great support during my study. She always had an open ear for my problems, questions and requests. I am also very thankfull, that she gave me the oppotrunity to work as a tutor in her practical courses. I have learned alot during this year, not only have i improved my laboratory skills but i also learned alot about myselfe. So thank you for an amazing time!

I also want to thank Agnes Kogler and Kathrin Schönefelder for introducing me to the working practice in the lab. It helped me alot and made the start much easier for me. A big thank you goes to Martin Kaufmann and Yan Gillen for their support and always making me laugh.

Many thanks goes to my family and especially to my mother Erna who supported me during my studies. As a single mother it was not always easy but nevertheless, she made it possible for me to come so far. I also want to thank my aunt Maria who was always there for me and supported me.

Last but definately not least i want to thank my one and only Christian. He was always there for me and listened to all my problems with patience. Thank you for your endless support. Love you!

# TABLE OF CONTENTS

# 1. INTRODUCTION

## 1.1 The three domains of life

Our planet is full of life, whether we can see it or not. No matter how hostile the environment may seem at the first moment, it is there. Microbiology has come a long way, starting from the discovery of the bacteria in 1684 by Antoni van Leeuwenhoek, the finding of penicillin by Alexander Flemming in 1929, spotting the structure of the DNA by James Watson/Francis Crick and Rosalind Franklin in 1953 or the discovery of Archaea as the third domain of life by Carl Woese in 1977 (1).

In the past, life was considered either as plants or as animals. Upon of the discovery of the bacteria, which were first considered as plants, a new Kingdom in 1866 by Ernst Haeckel was created, the Protista. Protists comprised all unicellular organisms. As the light microscope was displaced by the electron microscope, scientists revealed, that there was more to distinguish than between unicellular and multicellular organisms. The nucleus was discovered and the terms prokaryotes and eukaryotes where introduced, together with a 5 Kingdom model. This system was proposed by Rombert Whittaker in 1969 which comprised the kingdom of Monera (all unicellular organisms without nucleus) belonging to the Prokaryotes and the kingdoms of fungi, protista (all unicellular organisms with nucleus), plantae and animalia belonging to the eukaryotes. But all these models didn’t show any phylogenetic relationships and thats where Carl Woese comes into play. He used the 16S rRNA as a molecular marker to create a universal phylogenetic tree and thereby discovered the Archaea. With this finding, the today used model of the three domains was born.

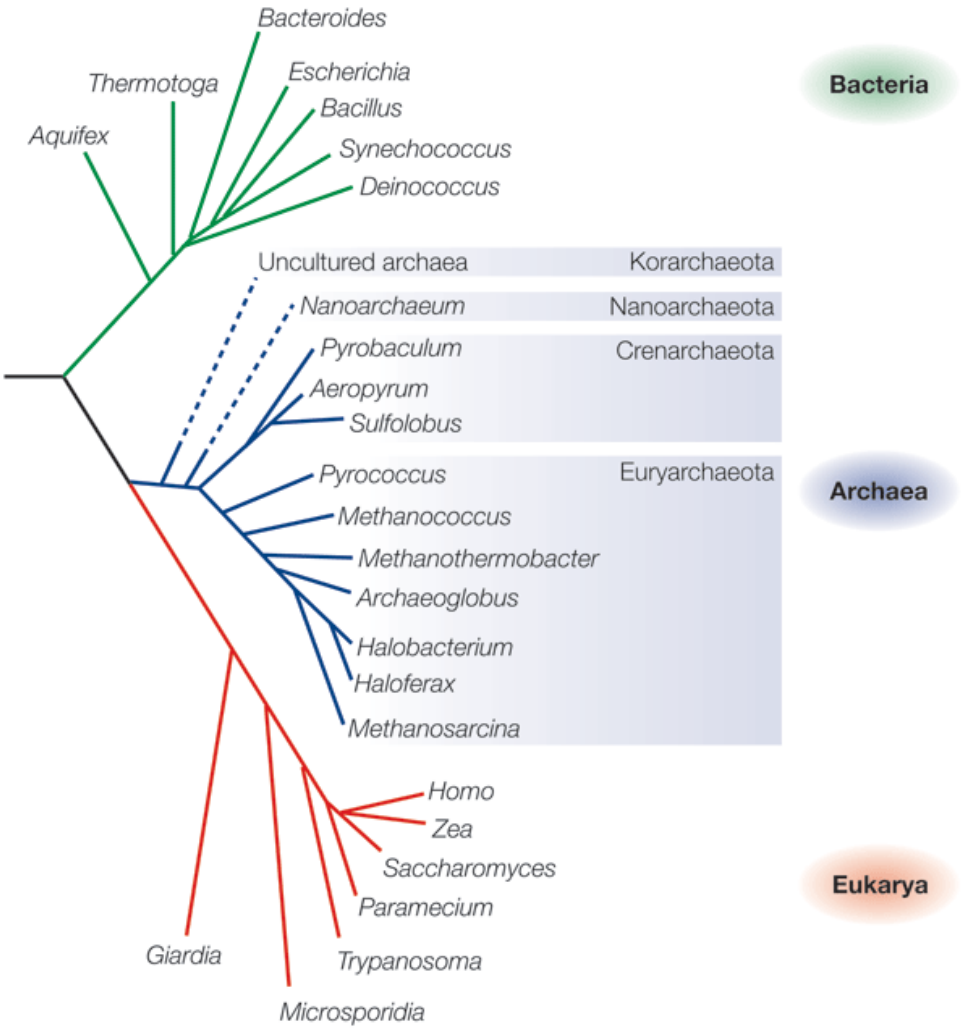


**Figure 1: Tree of life.** The tree of life divides organisms into three domains, bacteria, archaea and eukarya. This classification is based on 16/18S rRNA sequence analysis (71).

## 1.2 Archaea

Carl Woese and George Fox first discovered Archaea in 1977 based on 16S rRNA (3). They belong to the prokaryotes, thus lacking a nucleus. The morphologies of most *Archaea* are similar to those of B*acteria* but also show unique appearance like polygonal structures in halophilic or very irregular cocci in hyperthermophiles (3). Most of the *Archaea* are extremophile and are able to live at temperatures above 95°C (hyperthermophiles), at very low or high pH (acidophiles/alkalophiles), at high salt concentrations (halophiles) or at environments with high pressure (barophiles) (4). *Archaea* are practically abundant in all environments and are

predominated in marine plankton including deep oceans. Nevertheless, their roles in the biogeochemical cycles that they are involved in are still poorly noted. *Archaea* are so far the only organisms that are able to produce methane from $H_2$ and $CO_2$ (41). Methane is the second largest greenhouse gas contributing to climate changes and global warming. Its global warming potential is 25 times higher than of carbon dioxide (42). Up to now, five phyla are known today, the Crenarchaeota, Euryarchaeota, Korachaeota, Thaumarchaeota and the Nanoarchaeota (4).

### 1.2.1 Classification

#### 1.2.1.1Crenarchaeota

The phylum of the Crenarchaeota comprises 26 genera with the only class Thermoprotei or Crenarchaeota. This class can further be divided into 5 orders including *Acidolobales*, *Desulfurococcales*, *Fervidicoccales*, *Sulfolobales* and *Thermoproteales* (4). It was thought that they are hyperthermophilic without any exception (4), but 16S rRNA analysis revealed, that Crenarchaeota are also present in temperature-moderate habitats. With an abundance of 20-40 % of the planktonic communities, they are widely distributed in the picoplankton of the ocean (43). Hyperthermophilic members have been isolated from terrestrial or marine volcanic environments with optimal growth temperatures from 60 °C to 85 °C (44). The phylum Crenarchaeota includes one of the most extreme members of the hyperthermophile, *Pyrolobus fumarii*. It can grow up to 113 °C and was discovered in hydrothermal vents (45). Crenarchaeota are diverse when it comes to metabolism. Species can vary from chemolithoautotroph to chemoorganotroph, from aerobe to anaerobe or facultative anaerobe (44).

#### 1.2.1.2 Euryarchaeota

The phylum Euryrachaeota is the biggest one, with 76 genera (4). They are divided into 8 classes, *Methanobacteria*, *Methanococci*, *Methanomicrobia*, *Methanopyri*, *Halobacteria*, *Thermoplasmata*, *Thermococci* and *Archaeoglobi* (41). Euryarcheaota comprises the biggest phenotypic diversity of cultivable species including hyperthermophiles, thermoacidophiles, methanogens and halophiles (3).

The class ***Halobacteria*** contains 30 genera and all are extremely halophilic microorganisms that grow under high salt concentrations (up to 30-36 % NaCl) and aerobic (few facultative anaerobes) conditions. The range of required NaCl concentrations vary from the lowest minimum of 5 % in *Haloadaptus* and 6 % in *Halobaculum* and the highest minimum of 20 % NaCl in *Halosimplex*. *Halobacteria* are mesophilic and can grow at temperatures of 20-45 °C. With no exception, all species live chemoorganotroph using organic acids, sugars and proteolysis products such as peptone, yeast extracts, casamino acids etc. as an energy source (4).

**Methanogens** are obligate anaerobe and use $CO_2$ to accept an electron excess during respiration. Members of this class are *Methanococci* or *Methanotherma*, *Methanobacteria*, Methanopyri and *Methanomicrobia*. About half of the Methanogens are obligate chemolithoautotroph such as *Methanopyrales* (44) and the other half (especially mesophilic genera) utilize methyl-substrates as an energy source (4).

***Thermoplasmata*** includes only one order called *Thermoplasmatales*. Representatives of this class are all aerobes and extreme acidophiles. In order to receive energy for growth some are lithoautotrophes that oxidize iron compounds (sulfates and sulfides) and others are organotrophs that use some sugars (4). A pH below 2 is required for optimal growth. *Picrophilus* is able to grow at a pH of 0.06, which makes it the most acidophilic organism (44).

The class of ***Thermococci***, also called *Protoarchaea*, includes a single order *Thermococcales* with one family, the Thermococcaceae. They are all organotroph and use carbohydrates, proteolysis products and other organic substrates to generate energy. Some are even capable of fermentation. *Thermococci* requires for optimal growth a neutral pH, marine salinity and no oxygen since it is obligate anaerobe (4).

#### 1.2.1.3 Korarchaeota

This Phylum includes the biggest fraction of unclassified deep-branching *Archaea*. Members of the Korarchaeota are hyperthermophiles and detected in marine and terrestrial environments. The first discovered species was Candidatus *Korarchaeum cryptofilum* (isolated from Yellowstone National Park), which was described as an ultrathin filament (0.18µm in diameter) (49, 50). The enrichment culture was cultivated strict anaerobe at 85 °C with a pH of 6.5 in an organic medium. However, detailed growth requirements were not obtained because Ca. *K. cryptofilum* was only successful propagated in a complex enrichment culture. Nevertheless, predicted protein-coding genes let conclude that the organism lives heterotrophic and obligate anaerobic using peptides as the principal carbon and energy source. This matches with the common metabolic tactic of anaerobic peptide usage in hyperthermophilic Euryrachaeota and Crenarchaeota. Regarding exogenous electron acceptors, Ca. *K. cryptofilum* differentiates from other hyperthermophiles lacking the ability to use nitrate, oxygen, sulfur or iron, respectively (49).

#### 1.2.1.4 Thaumarchaeota

16S rRNA analysis led to the discovery of the linage *Thaumarchaeota* (46). They represent an ancient and very diverse phylum. The ecosystems that are inhabited, ranging from fresh water and marine to soils and hot environments. Some species are capeable to aerobically oxidize ammonia. It makes them to the first *Archaea* ever described to perform nitrification. This was up to now only be reported for a few proteobacterial linages (47). Scientists now think that these ammonium

oxidizing *Archaea* (AOA) play a significant role in the global nitrogen cycle (46). *Nitrosopumilus maritimus* SMC1 was the first AOA to be discovered and belongs to the so called group I.1a *Archaea*. This linage is present in coastal waters and the open ocean with an abundance of 20% to 30% marine microbes. Two thermophilic AOA have been described, Ca. *Nitrosophaera gargensis* and Ca. *Nitrosocaldus yellowstonii* and also the first soil-inhabiting AOA Nitrosophaera viennensis. Latter one is mesophilic and important for upcoming studies demonstrating the worldwide distribution of *Archaea* in soils (48).

#### 1.2.1.5 Nanoarchaeota

The phylum Nanoarchaeota was published by Huber et al. in 2002 with the successful cultivated organism, Nanoarchaeum equitans (44). These hyperthermophiles only grow as an ectosymbiont on the surface of its host Ignicoccus hospitalis (hyperthermophilic Crenarchaeota) (53). N. equitans possesses the smallest genome of cellular organisms, containing only 490.885 bp (52). The small size of the circular genome is a consequence of gene reduction, which indicates to a parasitic relationship to the host (51, 52). The archaeon is not capable to synthesize essential metabolic precursors on its own, such as amino acids, co-factors, lipids and nucleotides. Therefore, a cell-cell contact, where the outer membranes are attached to each other, is mandatory to supply biosynthetic molecules to the parasite (53, 51). The nanoarchaeon *N. equitans* is about 400 nm in diameter and grows at temperatures of 75 °C to 90 °C under strict anaerobic conditions (44).

### 1.2.2 Unique characteristics of Archaea

#### 1.2.2.1 Cell membrane

*Archaea* are able to grow under extreme conditions such as high temperatures. The cytoplasmic membrane plays an important role to survive in such environments. It is not only a barrier between the environment and the cytoplasm, it is also crucial for the generation of energy via proton motive force (PMF). In analogy, a sodium motive force (SMF) via sodium ion pumps can be generated. This energy is used to synthesize ATP, for flagella rotation, the transport of solutes across the membrane or to maintain the intracellular turgor and pH. Normally the permeability of ions is low, but it increases with the temperature. If protons or sodium ions influx too fast into the cell, no sufficient SMF or PMF can be maintained. Cell Membranes have to maintain a liquid crystalline state in order to function. With an increase of temperature, the membrane gets more fluid and is no longer stable (54).

The cell membrane of *Archaea* differs from those of the other two domains. It consists of branched isoprenoid acyl chains, which are ether-linked to the glycerol backbone. Usually a bilayer with 20 carbon branched isoprenyl lipids is formed, but some thermoacidophiles also form a monolayer that consists of 40 carbon branched isoprenyl lipids (6). In *Bacteria* and *Eukaryotes*, the lipids of the cell membrane are composed of linear hydrophobic fatty acid chains that are bound to the glycerol backbone in an ester-linked manner. To conquer high temperatures, the ether linkages in *Archaea* is the key for survivability. The membrane showes a higher stability and rigidity, a greater resistant to high temperature and oxidation and an increased salt tolerance (7).

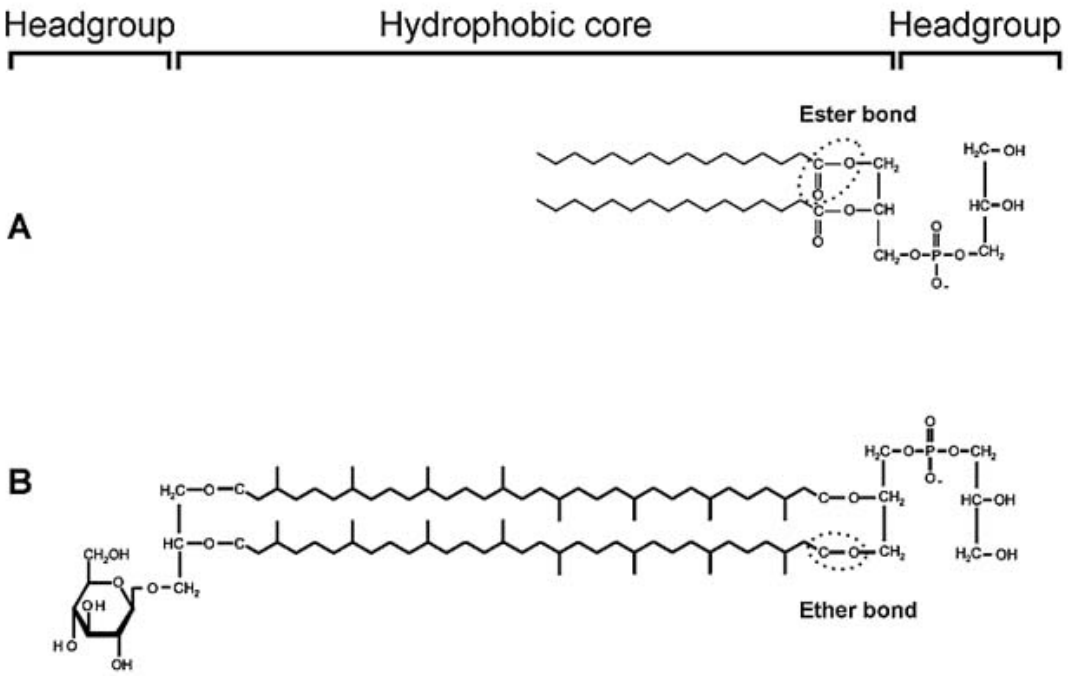


**Figure 2: Lipids of the cell membrane in Archaea and Bacteria. A**: Lipids of Bacteria with linear hydrophobic acid chains bound to glycerol via ester-linkage. **B**: Archaeal lipids with branched isoprenoid acyl chains. Lipids are not ester linked to the glycerol backbone as in Bacteria, but ether linked (54).

### 1.2.2.2 Cell wall

Another unique feature is that the cell wall of *Archaea* does not contain any peptidoglycan as in *Bacteria*, but a very similar polysaccharide called pseudomurein. The backbone is composed of N-acetylglucosamine and N-acetyltalosaminuronic acid which are β-1,3 linked (14). Archaeal cell walls have many variations and even some do not contain any (44). The most common cell wall type is the S-layer (surface layer) (14). It is usually anchored to the cytoplasmatic membrane and these crystals vary between 11 and 30 nm. Depending on how many subunits were arranged in the S-layer, various symmetries such as hexagonal, tetragonal or trimeric structures can occur (5).

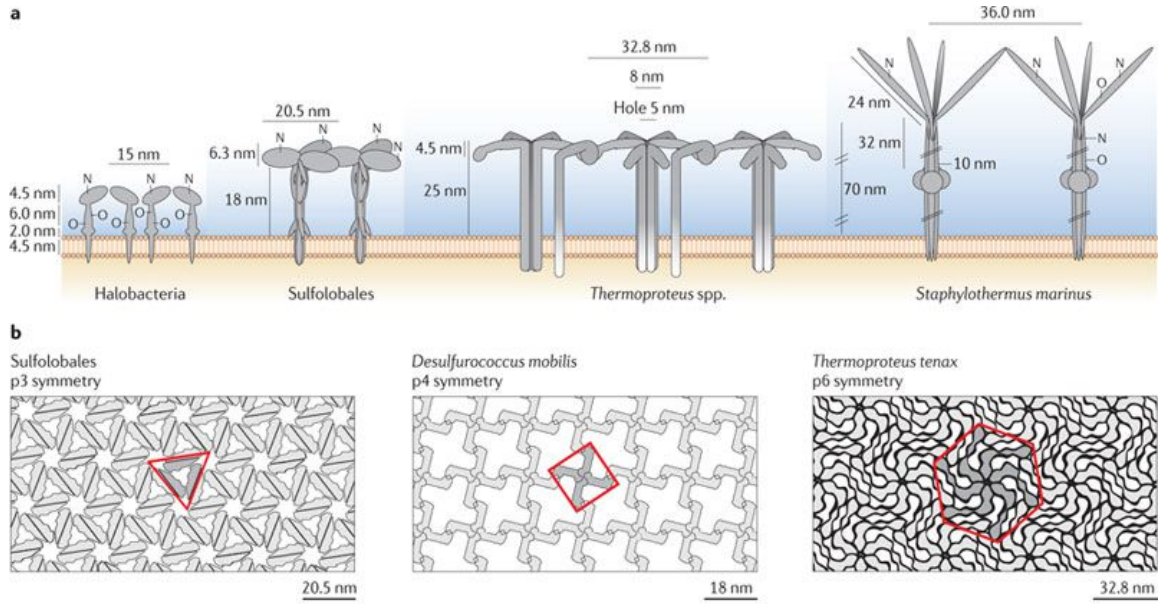


**Figure 3: Different archaeal S-layer types. a**: Side view of S-layer proteins in different *Archaea* (*Halobacteria, Sulfolobales, Thermoproteus spp., Staphylothermus marinus*). **b**: Top view of different S-layers. The surface layer is shown in grey, pores in white and the red outline shows a single unit (72).

### 1.2.3 *Natrialba magadii*

*N. magadii* was first isolated from the alkaline soda lake Magadi in Kenya by Tindall et al in 1984. This haloalkalophilic organism belongs to the family of the *Halobacteriaceae* within the phylum of Euryarchaeota. The rod-shaped cells are about 5 -7 µm in size and appear in an orange to red color because of production of carotinoid pigments in their membrane. For an optimal growth *N. magadii* requires a temperature of 37 to 42 °C and a pH of 8.5 to 10.5. Another important growth condition is the salt concentration, which is about 4 M of NaCl. Concentrations lower than 2 M sodium chloride will result in cell lysis due to the osmotic pressure. Due to the high pH a low magnesium concentration, less than 10 mM is required. As energy source no carbohydrates are needed but amino acids and peptides since *N. magadii* grows proteolytically. This strict aerobic chemoorganotroph archaeon has a rather long generation time of about 9 hours and can have up to 50 copies of its chromosomal DNA. These conditions make it more difficult regarding gene manipulation and creating homozygous mutants.

### 1.2.3.1 Laboratory strains

For *N. magadii*, two strains are available called L11 and L13. *N. magadii* L11 is the wild type strain carrying the provirus ϕCh1 integrated in its genome. This virus was first discovered in 1997 upon spontaneous lysis of *N. magadii*, which is up to now the only known host. Cells start to lyse after they entered the stationary phase and leading to the release of new virus particles. The *N. magadii* L13 strain was cured from the virus and is used as an indicator strain.

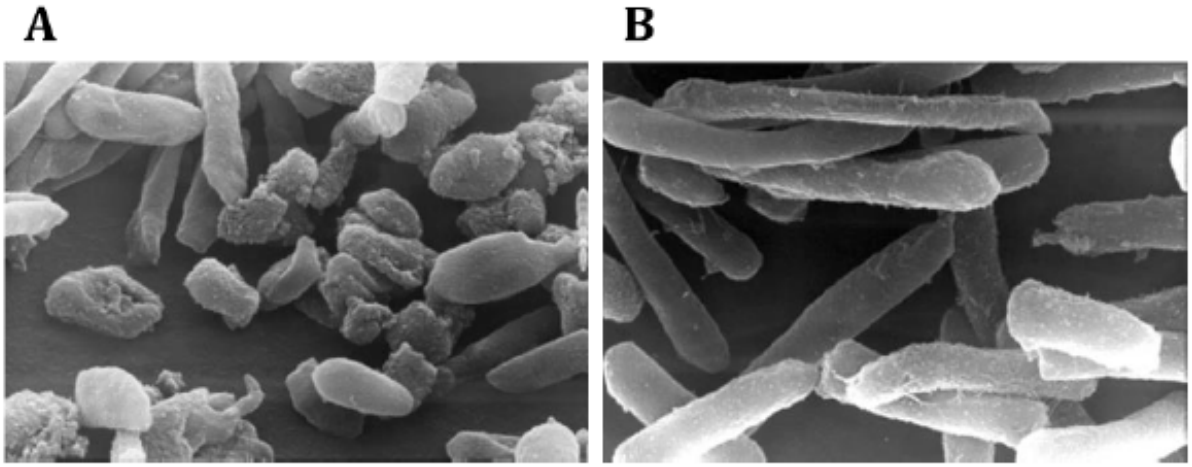


**Figure 4: *N. magadii* L11/L13 morphology. A**: Wildtype L11 carrying the prophage ϕCh1 under the electron micrograph. B: Cured strain L13 (73).

## 1.3 Viruses

### 1.3.1 In general

What are viruses? Are they alive? Where do they come from? As mentioned above, life as we know can be classified as bacterial, archaeal or eukaryal (13). However, where do viruses come into play? Most evolutionary theories were developed without regarding the virosphere (9). There are many questions, but not all of them can be answered yet. It is estimated that about $1.0 \times 10^{30}$ microbes are living on earth and per prokaryotic cell, there comes an average of 10 viruses. Which means that on our planet, $1.0 \times 10^{31}$ viruses are present and thus makes them to the real dominant life form (9). For us humans, viruses play a very important role when it comes to diseases. They

are not only important when it comes to medicine, they can also impact our economy and agriculture (11). Nevertheless, we are not the only one that can be infected. In fact, every life form of the three domains, can be infected by at least one type of virus (10). Most of them are actually phages. These are viruses, that can infect Bacteria (9)

#### 1.3.1.1 What are viruses and how do they look like?

First of all, the difference between the two terms, virus and virions. A virus outside of the host cell is called virion. A virus that is inside of the host, is called virus. (9). Now virions can vary in size and shape. For example, the Poliovirus is about 28 nm in diameter and one of the smallest viruses. The largest Virus is the Smallpox virus with a diameter of about 200 nm. The genomic information can be either consist of DNA or RNA (can be single stranded or double stranded) and some even use both in their replication cycle (14). The genome size varies from a few kilobase pairs (kbp) as in Poliovirus (7.500 bp) up to 1.18 megabase pairs (Mbp) as in Mimivirus (12). To protect their genome, the DNA/RNA is covered by proteins that are called capsid. These proteins are responsible for the virus symmetry. Two kinds of symmetry are found in viruses. Rod-shaped viruses, where the capsid proteins are arranged in a helical configuration and spherical viruses that have an icosahedral symmetry (14). The morphology of many bacteriophages for example, is composed of an icosahedral head and a tail, with tail fibers attached to it (9). Most viruses that infect animal cells have an additional membrane surrounding the nucleocapsid (nucleic acid wrapped with capsid). They are called enveloped viruses and one example is the influenza virus that causes the flu in humans. However, there are also some enveloped bacterial and plat viruses known (14). Therefore, viruses can be very diverse when it comes to their appearance, genetic information and replication strategies. However, all viruses have one thing in common. They can only replicate within a host cell. The reason is, that they are lacking of very important macromolecules called ribosomes. Without ribosomes, the completely translational machinery, where proteins are synthesized, is shut down and a cell is no longer viable (12). Therefore, viruses

abuse a host cell, to force the replication of their own genome. The infection process starts with the attachment to a susceptible host cell and the injection of the foreign DNA. Once the viral DNA entered the cell, the host starts to replicate the viral genome and synthesize proteins (e.g. capsid protein). Then the capsids are assembled and virus DNA is packaged into new virions. The last step is releasing the progeny virions via lysis, leading to the death of the host cell (14). The burst size (number of released virions per cell) can vary considerably from 40 up to 60.000 released virions. The *E. coli* infecting bacteriophage Lambda has a burst size of about 150 and the human immunodeficiency virus (HIV) which infects the memory T cells of H. sapiens has a burst size from 1.000 to 3.000. The Simianes immundefizienz virus (SIV) infecting T cells of R. macaque (rhesus monkey), releases 40.000 to 60.000 new viruses (15). SIV was found to be the origin of HIV (16).

#### 1.3.1.1 Are they alive?

The question, wheatear viruses are living organisms or not are a very controversial one. Biologists established a few criteria that an organism has to comply with, regarding as a living organism:

- Living things are able to grow and evolve
- Living things have its own metabolism
- Living things are able to reproduce
- Living things are able to respond to stimuli of the environment

Viruses perform parts of these criteria. They do reproduce in some way and they indeed evolve, as we know from the annual fluent wave. The greatest difference to living organisms is, that they do not posse own metabolic processes and thus cannot generate ATP (adenosine triphosphate), which is the universal energy currency in a cell. Furthermore, they are not able to synthesize proteins on their own, because they are missing the essential machinery for translation, the ribosomes. To replicate and translate an mRNA (messenger RNA) to proteins, they have to use ribosomes from a

living cell (12). Due to these criteria, viruses are no living organisms. Maybe we have to rethink things. the requirement to grow for example. A great number of plant and animal cells are not growing and several tissue cells of a tree are just resting. When it comes to reproduction, few living cells for example in seeds can persist over years without reproducing via cell division or growing. In addition, movement within a cell is not always easy to identify (17).

In 2008, researchers discovered a giant virus that falls ill by the infection of another virus. The virologist Jean-Michel Claverie from the CNRS UPR laboratories in Marseilles says, “There’s no doubt this is a living organism, the fact that it can get sick makes it more alive”. The so-called mamavirus is closely associated with a small virus named Sputnik.

#### 1.3.1.3 Where do they come from?

The origin of viruses is still an open question, but there have been three main hypotheses created, to get closer to the answer.

##### 1.3.1.3.1 The regressive (reduction) hypothesis

The origin of viruses is caused by a regressive or reductive process, as stated by the regressive hypothesis (12). This was shown in bacteria, which live as parasites in other cells, such as *Rickettsia* or *Chlamydia* species. Studies also suggest that mitochondria of *Rickettsia prowazekii* and other eukaryotic cells, share a mutual ancestor, a former free-living organism (18). This leads to the conclusion, that viruses evolved from organisms, that were free-living, complex lifeforms. The adoption to such a parasitic way to replicate was followed by the loss of genetic material over time. Viruses of the nucleocytoplasmic large DNA viruses (NCLDVs) support this hypothesis. They include one of the largest viruses, the Mimivirus, but they are not only much bigger than an average virus, their genomes are huge. In comparison, the influenza virus has a diameter of about 80 nm; a Mimivirus can have a diameter of 750 nm. When it comes to the genome, the DNA of a

Poliovirus is about 7.500 nt (nucleotides) in length, whereas the length of the Mimivirus DNA is 1.2 million nt. Consequently, this virus is more complex and is more independent than other viruses, when it comes to replication of their genome, within a host cell. These factors may conclude that NCLDVs are progenies of complex ancestors. It was also found, that the Mimivirus encodes for putative genes that can be connected with translation. This may a hint, that there was once a complete translation system and as the parasite got more dependent on the host, genes got lost (12)

#### 1.3.1.3.2 The progressive (escape) hypothesis

This hypothesis suggests that viruses derived from mobile genetic elements. These so-called retrotransposons are parts of most eukaryotic genomes and humans with an abundance of 42% of the total genome (19). Retrotransposons can jump within the genome and use RNA as an intermediate just like retroviruses. The RNA is transcribed into DNA with the enzyme reverse transcriptase and inserted into the host DNA via an integrase. On the other hand, in the case of human retrotransposons the DNA is integrated into a different location on the genome. It is assumed, that such elements escaped one cell, entered another one and became infectious, leading to the evolution of viruses (12).

#### 1.3.1.3.3 The virus first hypothesis

This hypothesis is characterized by the fact, that not cells but viruses came first. Accordingly, in a pre-cellular world, viruses evolved from self-replicating units, which became more and more complex over the time. It is established, that the first replicating molecules were made of RNA. This leads to the assumption, that single-stranded RNA viruses were formed before the first cells appeared on earth and where able to infect them (12).

No one can answer the question where viruses came from. These hypotheses are just three suggestions for their existence. None of them may be true and viruses arose in a different way.

Maybe all of them are true, and viruses were invented in several different ways. Nevertheless, to find the real origin of viruses more research has to be done.

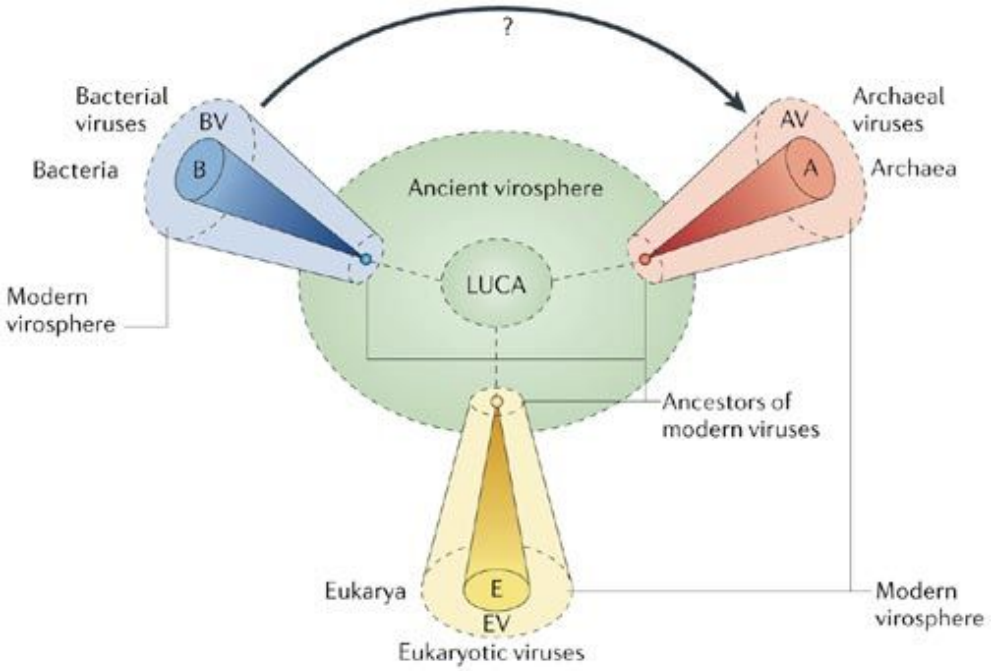


**Figure 5: Virus as ancestor of cellular life.** The virus fist hypothesis suggests that viruses existed before cellular life, where self-replicating units were able to form membranes and cell walls. Viruses than co-evolved with the hosts (12).

### 1.3.2 Archaeal viruses

Viruses of Bacteria and Eukaryotes have been studied for over 10 decades, but the study of archaeal viruses started only 30 years ago. Therefore, the up to now known number of described archaeoviruses represents under 1% of the total known viruses of the two other domains (68). Many virologists consider archaeal viruses as bacteriophages, although most of them have nothing in common with viruses infecting bacteria (69). Archeal viruses indeed share morphotypes with bacterial viruses like icosahedral, pleomorphic and head-tailed virions. However, they do also have very exceptional morphotypes such as fusiform, bottle-shaped and droplet-shaped virions, mainly associated with hyperthermophilic *Archaea*. That is not the only exception, 90% of putative genes in the genomes of hyperthermophilic archael viruses have no homologs or

identifiable functions to other viruses (68). Of the about 40 known archaeal viruses, all of them have double-stranded DNA with one exception, the Aeropyrum coil-shaped virus (ACV) (68, 70). It is the only known hyperthermophilic virus with a single-stranded DNA. Moreover, no virus with an RNA genome from a hyperthermic environment has been reported (68). Mainly, archeal viruses can be classiefied into fusiform viruses, bottle- and droplet-shaped viruses, linear viruses, spherical viruses and head-tail viruses (68).

#### 1.3.2.1 Fusiform viruses

Both kingdoms, Crenarchaeota and Euryarchaeota can serve as a host and members of extreme halophiles, hyperthermophiles and anaerobic methane-producers are infected. These viruses predominate in hypersaline waters and hot, acidic springs. Their fusiform virions are exclusive to *Archaea* but very common (55).

**Fuselloviridae** members infect the hyperthermophilic Crenarchaeon Sulfolobus. Sulfolobus spindle-shaped virus 1 (SSV1) is the best-studied type of this family infecting the the hyperthermophilic host Sulfolobus (Crenarchaeota) (55). The lemon-shaped SSV1 integrates its genome into a tRNA gene of the host chromosome with a size about 15.5 kilobase pairs (kbp). Upon integration, UV irradiation induced virus production and released to the environment without causing lysis of the host cell (56).

**Bicaudaviridae** includes only one member called Acidianus two-tailed virus (ATV) (55). It infects acidophilic, hyperthermophilic *Archaea* like *Acidianus convivator* and other Sulfolobales that live in acidic hot springs with a pH of 1.5 to 3.0 and a temperature between 80 C and 90 °C (57). ATV is released from the host via lysis as fusiform tail-less particles. These virions undergo an extracellular morphological development, where long tails start to grow independent of any

energy source or host cells. The molecular mechanism is still unclear but it seems that these anchor-like structures increase the chances to be adsorbed to new host cells (55).

#### 1.3.2.2 Bottle- and droplet-shaped viruses

Viruses of this family have unique morphological features with the two genera, *Acidianus* bottle-shaped virus (ABV) and *Sulfolobus neozealandicus* droplet-shaped virus (SNDV (55).

The **Ampullaviridae** member ABV has a unique morphology reassembling a bottle that is about 230nm long and 4-75 nm wide. No helical or icosahedral symmetry was found in the cone-shaped virus (55). The broad end is covered with about 20 thin filaments, each 20 nm in length, arranged in a circle (58). Functions of these filaments are unclear, while adsorption to a host cell seems to occur at the narrow end of the virions (55). The hyperthermophilic *Acidianus convivator* (Crenarchaeota) can be infected by ABV and grows optimally at 75 °C. Upon infection of *A. convivator*, neither decrease of the $OD_{600}$ nor formation of cell debris were detectable, conclude a non-lytic release of the progenies (58).

**Guttaviridae** includes the only member of the Guttarvirus, Sulfolobus neozealandicus droplet-shaped virus (SNDV) that infects the hyperthermophilic genus Sulfolobus (Crenarchaeota) (55). The particle has a droplet-shaped form, covered densely with thin tail fibers at pointed ends and the surface is helically ribbed (59, 55). Virus size varies from 110 nm to 185 nm in length and 70 nm to 95 nm in width. SNDV is non-lytic, since no decrease of cell density or cell debris were found. SNDV persists not as a stable prophage in the host, since the virus is not integrated into the host chromosome. Furthermore, the virus gets lost after prolonged incubation of an infected culture (59).

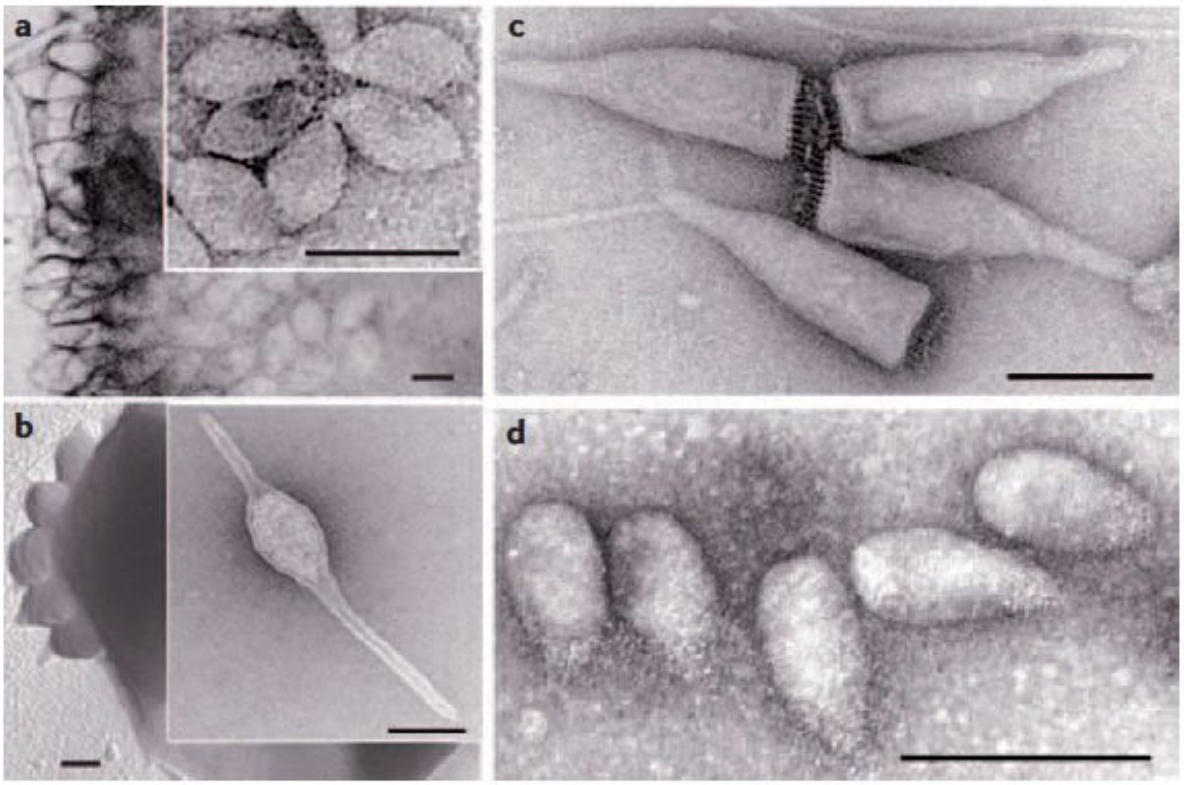


**Figure 6: Exceptional morphologies of archaeal viruses. a:** *Sulfolobus spindle-shaped virus 1* (SSV1). b: *Acidianus two tailed virus* (ATV). c: *Acidianus bottle-shaped virus* (ABV). d: *Sulfolobus neozealandicus droplet-shaped virus* (SNDV). Scale bars represent 100 nm (55).

#### 1.3.2.3 Linear viruses

These linear shaped viruses predominate in environments above 80 C where the Crenarchaea genera *Acidianus*, *Thermoproteus* and *Sulfolobus* can be found. For linear viruses, it has previously not been observed to contain dsDNA but members of the Rudiviridae and Lipothrixviridae do. These viruses linger stable in the host cell, although they do not integrate into the host chromosome, due to the lack of an integrase gene. Furthermore, stress factors like mytomycin C treatment or UV irradiation do not induce virus replication (55).

**Rudiviridae** comprises the three species, Sulfolobus islandicus rod-shaped virus 1 (SIRV1), Sulfolobus islandicus rod-shaped virus 2 (SIRV2) and Acidianus rod-shaped virus 1 (ARV1) (55). These viruses are non-enveloped, with a rod-shaped, helical body containing linear dsDNA (60, 61). Three tail fibers are attached to each end of the filament (60, 61). The lengths of these fibers

are 28 nm in SIRV1/SIRV2 and about 10 nm in ARV1. SIRV1 and SIRV2 have a total size of 830 nm and 900 nm whereas ARV1 is smaller with a size of about 610 nm. They all contain linear dsDNA, which is not integrated into the host chromosome. Furthermore, no induction of virus production can be forced by neither mitomycin C nor UV irradiation. New virions are released without killing the host cell (60, 61).

**Lipothrixviridae** show great diversity regarding replication mechanisms and properties of their linear genomes dividing them in four different genera. The *α-lipothrixvirus* with the specie *Thermoproteus tenax* virus 1(TTV1) has a size of 400 nm x 38 nm with a lipid bilayer envelope with a helical core. The *β-lipothrixvirus* includes the species *Sulfolobus islandicus* filamentous (SIFV), TTV2 and TTV3 (55). SIFV is a flexible rod with a length of 24 x 1,980 nm. On both ends putative fibers are attached (62). The *γ-lipothrixvirus* comprises the *Acidiansu* filamentous virus 1 (AFV1) is 900 x 24 nm long and possesses claw-like structures on both ends for host contact. The *δ-lipotrixvirus* AFV2, which is 1,100 x 24 nm in length and a collar structure at the virion termini (55).

#### 1.3.2.4 Spherical viruses

Two major types of these viruses are found. The spherical viruses *Pyrobaculum* spherical virus (PSV) and *Thermoproteus tenax* spherical virus 1 (TTSV1) and the two icosahedral viruses *Sulfolobus* turreted icosahedral virus (STIV) and *Haloarcula hispanica* virus (SH1) (55).

**Globuloviridae** include the two spherical viruses PSV and TTSV1 (55). The first described virus of this genus is PSV. It is an enveloped virus, that is about 100 nm in diameter. The nucleoprotein core of this dsDNA virion has a helical structure (62). The known hosts *Pyrobaculum* and *Thermoproteus* (both *Crenarchaeota*) were not killed after release of progenies (55). PSV is likely

to exist in the host in a carrier stage, thus reducing the direct contact with the extreme environment (62).

The **Unclassified** species STIV and SH1 are non-tailed icosahedral viruses. SH1 has a linear genome in contrast of STIV that contains circular DNA. Moreover, STIV persists stable in its host with a non-lytic release of virions, whereas SH1 is lytic (55).

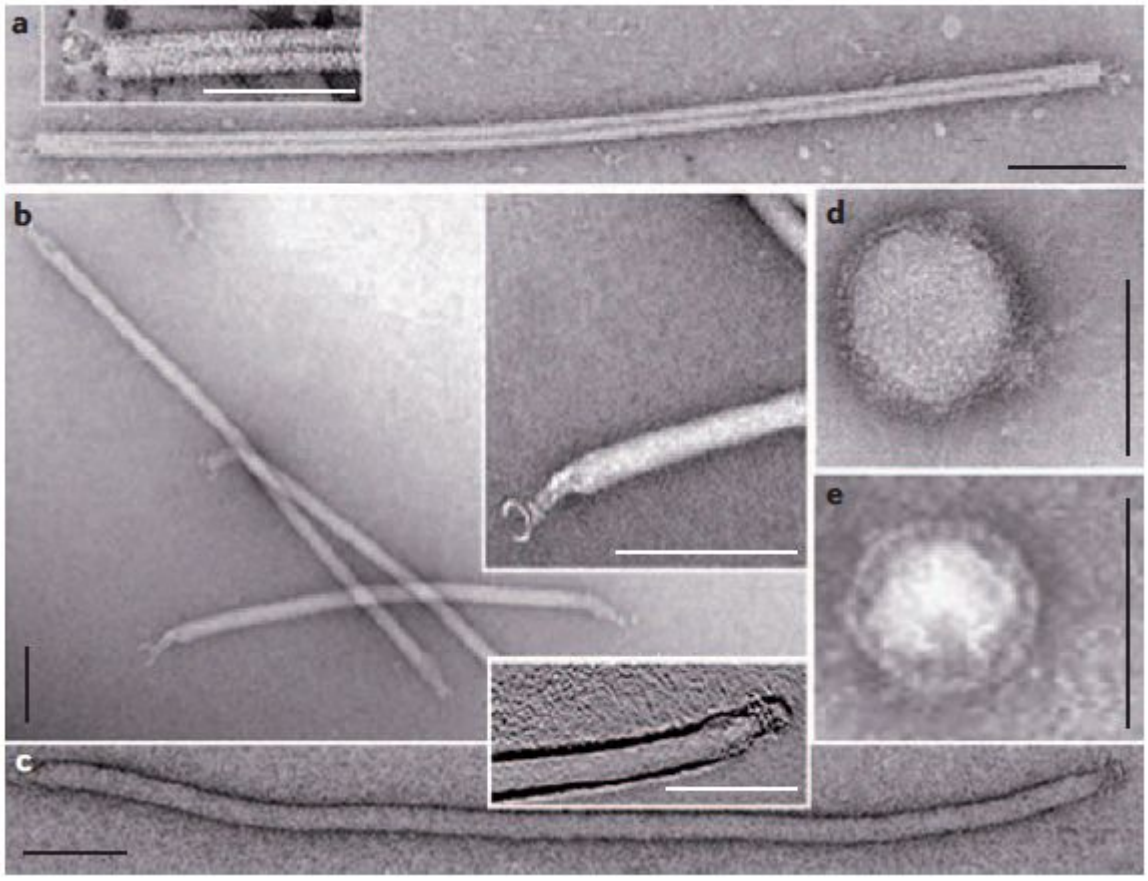


**Figure 7: Linear and spherical viruses of Archaea. a:** *Sulfolobus islandicus rod-shaped virus 1* (SIRV1). **b**: *Acidianus filamentous virus 1* (AFV1). **c**: *Acidianusfilamentous virus 2* (AFV2). **d**: *Pyrobaculum spherical virus* (PSV). **e**: *Haloarcula hispanica virus* (SH1). Scale bars represent 100 nm (55).

#### 1.3.2.5 Head-tail viruses

The morphotype of icosahedral heads and helical tails is very common to viruses of bacteria. They are all associated with methanogens or extreme halophilic Euryarchaeota. Head-tail viruses are divided into the families Siphoviridae, Myoviridae and some unassigned species (55).

The **Myoviridae** include the two best studied viruses ϕCh1 and *ϕH* (55). The latter one was discovered in 1974 as one of the first archaeal viruses (44). ϕCh1 will be discussed later in detail.

### 1.3.3 ϕCh1

#### 1.3.3.1 General features

ϕCh1 is a virus that belongs to the family of the Myroviridae and was first isolated from Witte *et al*. in 1997 upon spontaneous lysis of its host. The temperate virus infects the haloalkaliphilic Archaea *Natrialba Magadii* and it is up to now the only known host of ϕCh1. Furthermore, ϕCh1 is the only characterized virus infecting a haloalkaliphilic *Archaea* (63). The morphology shows a head-tail structure, which is usual for a member of the Myroviridae (55). In total, the virus has a size of approximately 200 nm, whereas the icosahedral head is about 70 nm long and the contractible tails is about 130 nm in length. The tail, where structures were found, so called tail fibre protiens may seem to be responsible for the adsorption and thus involved in infecting the host. The tail is about 20 nm wide. As *Natrialba magadii* lives in an environment with a high salt concentration, ϕCh1 is adapted to these conditions and requires a salt concentration above 2 M. Below 2 M sodium chloride the virus structure becomes instable and is not able to infect its host anymore. (8)

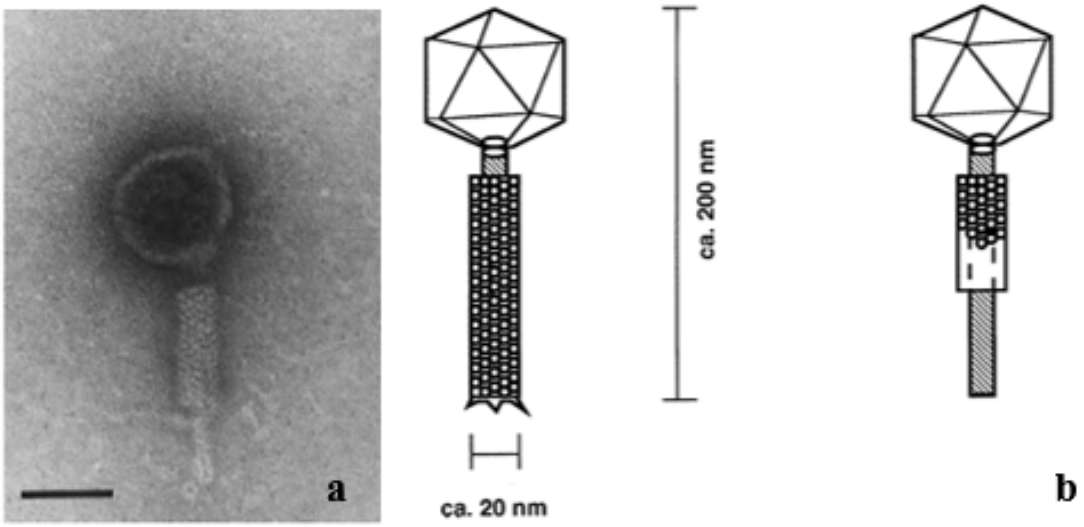


**Figure 8: Morphology of ΦCh1.** a: ϕCh1 under the electron micrograph. **b:** Schematic drawing of ϕCh1 (44).

### 1.3.3.2 Life cycle

Since ϕCh1 is a temperate virus it can switch between a lytic and a lysogenic state. After infecting *Natrialba magadii*, ϕCh1 DNA is integrated into the host chromosome and persists as a provirus. As cells reach the late logarithmic growth phase the virus becomes lytic and virus particles start to assemble. This happens after approximately 4 days of growth, cells were killed and progeny viruses were released, ready for another round of infection. In the lysogenic state, the genome of the provirus is replicated but a repressor protein suppresses the translation of viral proteins that is responsible for cell lysis. This state also prevents superinfection, which means, that only one phage can infect one host at a time (64).

### 1.3.3.3 Genome organization

The genome of ϕCh1 is composed of a double stranded linear DNA, with a total length of 58.498 bp and a total G+C content of 61.9% (8). It contains 98 putative open reading frames and 21 of them, have already been assigned. All of these ORFs start with an ATG, this is not true for the ORFs 3, 41, 79 and 83, their start codon is GTG. As the DNA is terminally redundant and circularly permuted, it seems that the virus DNA is packaged via a headful mechanism. The genome of ϕCh1 forms structural units and therefore can be divided into three parts. The left part containing the ORFs 1 to 34 encoding proteins for structural components and proteins involved in virus assembling. The middle part (ORFs 35 to 55) is responsible for regulation, replication and stabilization of the genome. ORFs 56-98 of the right part encode for proteins with yet unknown functions, but it seems that this part contains enzymes that are involved in DNA modifications (63).

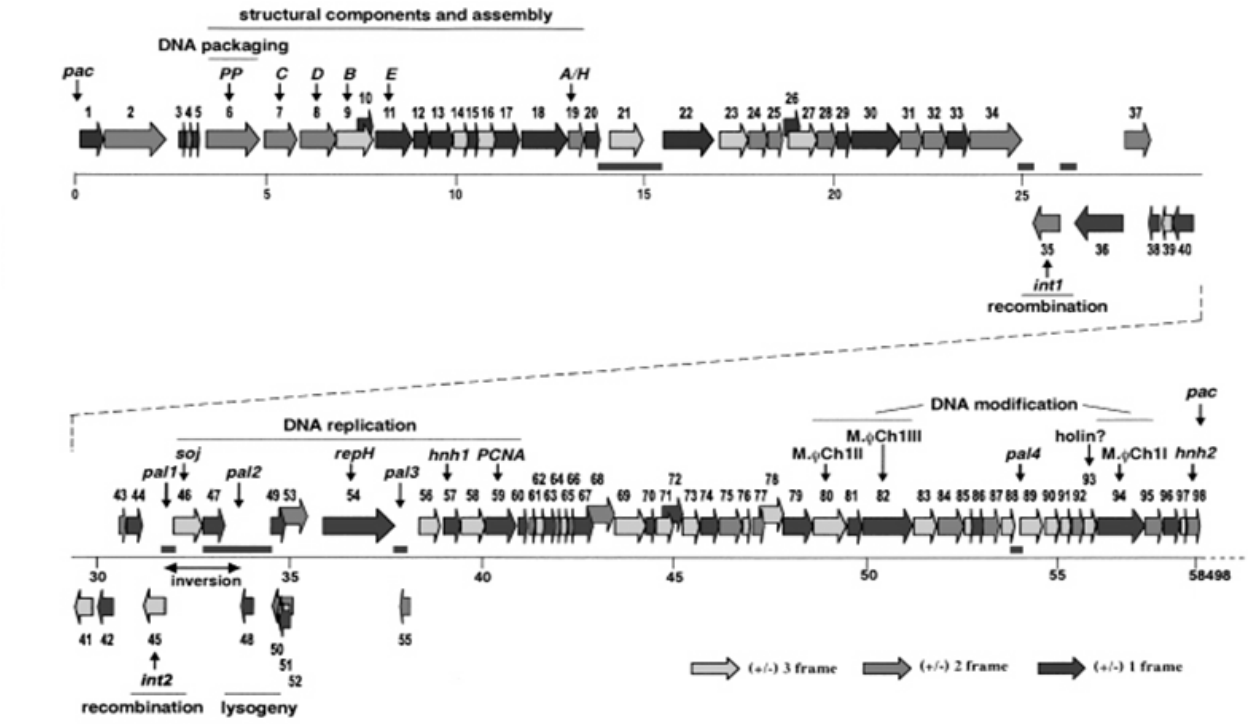


**Figure 9: Linear genome organization of ΦCh1.** This scheme shows the sequenced genome of ϕCh1 with a total length of 58.495 base pairs. It contains 98 putative open reading frames (ORF) indicated by arrows. The genome is organized in three parts: the left part contains ORFs that are responsible for structural components, the middle part is responsible for DNA replication, gene regulation and stabilization. The right part codes for proteins with unknown or DNA modification functions. Known or putative functions are indicated above the arrows. Replication of the genome starts between the ORFs 53 and 54 (40).

### 1.3.3.4 The invertible region

**The integrase Int1**

ORF35 encodes the integrase Int1, which is located between ORF34 and ORF36. This enzyme belongs to the lambda integrase family and therefore, seems to be a site-specific recombinase. Int1 shows great identity to integrases of thermophilic *Archaea* and halophilic *Bacteria*. There are two several groups of site-specific recombinases besides the the lambda integrase and the resolvase/integrase family. The lambda integrase also called tyrosine recombinase uses a tyrosine for DNA cleavage and plays a role in many processes like integration of viral or plasmid DNA into the host genome, DNA excision or inversion and transposition. The C-terminal end of the lambda recombinase consists of three conserved boxes (A, B, C) with four highly conserved amino acids R-H-R-Y (Arg-His-Arg-Tyr), where tyrosin is catalytically active (44, 28, 65).

**The tail fibre protein**

Two ORFs (34 and 36) that encode for the putative tail fibre proteins gp34 and gp36 are located on the genome of ϕCh1. Both ORF34 and ORF36 contain repeat clusters that consist of direct repeats with a length of 30bp. The clusters of ORF34 (IR-L) have 15 repeats and ORF36 (IR-R) has 10 repeats; both are inverse orientated. These two open reading frames are separated from each other by ORF35, which encodes for the lambda integrase Int1. It represents a so-called phase variation system that helps organisms to adapt to changes in the environment or to extant the host range of the virus. In this case, two different types of tail fibre proteins are produced (65, 66). The fibres are trimers that consist of an C-terminal domain which recognizes the host cell receptor, a thin shaft, and an N-terminal domain, that is attached to the viral capsid (67).

Upon inversion, the 3' regions of ORF34 and ORF36 are exchanged via sequence-specific recombination, leading to two different variants of the tail-fibre protein. They were named $ORF34_1$ and $ORF36_1$ with int1 gene in the (-) orientation which represents the unexchanged form and the exchanged version $ORF34_{52}$/$ORF36_{52}$ with the int1 gene in the (+) orientation. A

promoter sequence is only encoded for ORF34, thus it is the only ORF that is transcribed shown by transcription analysis. Only an ORF34-specific fragment was detectable, which indicates, that only ORF34 is transcribed (67). Binding studies with $gp34_1$ and $gp34_{52}$ showed that only the exchanged form $gp34_{52}$ is able to bind to the cell surface of *N. magadii*. Further investigations revealed that the C-terminal parts of $gp34_1$ and $gp34_{52}$ showed similarities to galactose-binding domains but only the C-terminus of $gp34_{52}$ comprises conserved amino acids reassemble the binding pocket for α-D-galactose. Mutation within these conserved regions inhibited the binding to *N. magadii* $gp34_{52}$. Furthermore, incubation with 25 mM α-D-galactose and $gp34_{52}$ prevented binding of the protein to the cell surface of *N. magadii*. Infectivity of ϕCh1 was decreased if 50 mM α-D-galactose was added to the medium. These findings led to the conclusion that the attachment of the virus to *N. magadii* is arranged via a galactose-residue on the cell surface (67). Possible structures might be glycosylated flagellum proteins FlaB1 and FlaB2 or glycoproteins in the S-layer (64).

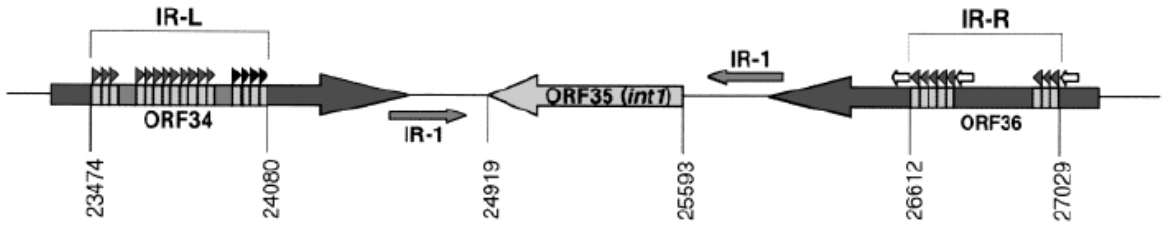


**Figure 10: The invertable region.** This schematic drawing shows the site-specific recombinase Int1, encoded by ORF35. It is flanked by the two ORFs 34 and 36 wich encodes for the putative tail fibre proteins. The 30bp long direct repeats are indicated with the small arrows within the IR-L (ORF34) and IR-R (ORF36). Int1 facilitates the inversion of these repeats, leading to the formation of $ORF34_{52}$ and $ORF36_{52}$ (28).

#### 1.3.3.5 Toxin-antitoxin sytem

The toxin-antitoxin (TA) system was first discovered in 1982 and consists of only two gene products. A stable toxin that can poison the cell, leading to programmed cell death (PCD) and a labile antitoxin that makes the toxin innoxious (20). The transcription of these two genes is autoregulated by the gene products itself. A stabilized concentration of toxin and antitoxin is achieved by negative-feedback loop (21). TA-systems are not only involved in PCD but also in the regulation of biofilm formation, protection against phages, formation of persister cells (antibiotic resistance) and stress resistance (20). Such systems can be encoded either on a plasmid or on the chromosome. The number of TA systems can vary from species to species. *E. coli* for instance, owns 37 of them (23). TA systems encoded on a plasmid are responsible for the so-called post-segregational killing effect. That ensures the propagation of plasmids to the daughter cells. Upon cell division, only those cells that inherit the plasmid will survive because the labile antitoxin, that no longer can be translated, is faster degraded than the toxin (20).

Three different types of TA systems are known. The type I is characterized by a small RNA (sRNA) acting as the antitoxin. The sRNA molecule is able to bind to the mRNA of the toxin, thus preventing the translation of the toxic protein. Once sRNA and mRNA are paired, RNase III is able to degrade the duplex. In the type II system, toxin and antitoxin are both proteins. Inactivation is accomplished by binding of the antitoxin to the toxin. The type III TA system consists of RNA as an antitoxin that directly binds to the toxic protein, resulting in neutralization (20). These three types can further be subdivided into nine families. Some toxins like ParE and CcdB inhibit the DNA gyrase function; some indirectly or directly cleave ribosome-associated mRNAs like Doc, RelE and HigB; the protein kinase HipA targets EF-Tu; MazF and VapC got RNase activity that can cleave mRNA (29).

One of the most studied systems is the mazEF and belongs to the type II TA system. It is a chromosomal located, stress-induced suicide module in *E. coli* (24). It consists of mazF, which represents the toxin, and the upstream-located mazE as the unstable antitoxin. This mediated cell death system can be triggered by various stress factors like antibiotics or sever amino acid starvation (24, 25). Upon starvation, ppGpp is produced by RelA, which leads to the degradation of the labile antitoxin by ClpPA (ATP-dependent serine protease). The Toxin MazF, which is an endoribonuclease, targets mRNAs and cuts at a sequence specific site (ACA) (24). ACA is closely located upstream of the AUG start codon. The nucleotides that are cut off the mRNA include the Shine Dalgarno (SD) sequence (25). The 30S subunit of the 16S rRNA possesses the anti-Shine Dalgarno (aSD) sequence and recognizes the SD on the mRNA. The pairing of the SD-aSD contributes to stabilization of the mRNA-ribosome-imitator tRNA complex, which leads than to the initiation of translation (26). The truncated mRNAs, that are called leaderless mRNAs, can no longer be recognized by Ribosomes and protein synthesis is blocked (about 90%) (24, 25). The other 10% of the proteins were required for cell death (25).

#### 1.3.3.5.1 ORF43/44

The ORF43 and ORF44 with the gene products gp43/44 are encoded by the ϕCh1 genome. The two ORFs represent an operon, with a putative function as a toxin-antitoxin system. They are co-transcribed and co-translated, since their stop and start codons overlap. A promoter sequence upstream of the ORF43 was identified. It has been shown that ORF43/44 has an enhancing effect on the promoter of ORF49. For this promoter studies, a plasmid was constructed, containing the promoter from the intergenic region between ORF48 and ORF49, the gene *bgaH* which serves as a reporter, the ORF48 and ORF43/44. This construct was transformed into *Hfx. volcanii*. Analysis showed, that a mutated start codon in ORF48 has the strongest enhancing effect due to the lack of the repressor Rep and binding of gp43/44 to ORF48 within the coding sequence. This leads to a higher activity of the *bgaH* gene under the ORF49 promoter. With constructs that did not

comprise the repeat sequence or ORF48 no enhancing effect could be detected. Reporter gene activity was completely removed if only ORF44 was cloned into the plasmid (27, 28).

Another experiment with ORF44 indicates, that gp44 degrades the mRNA of ORF48. A plasmid was constructed with the repressor encoded by ORF48, ORF44 and the promoter sequence from the intergenic region. Via RT-PCR, the mRNA level was observed and it was shown, that no mRNA was detectable, leading to the conclusion, that it is degraded by gp44. The sequence of ORF44 was analyzed and showed homology to a VapC PIN domain nuclease (27, 28). VapBC (virulence associated proteins) displays together with MazEF a toxin-antitoxin system that belongs to the type II TA systems and is the most abundant module in *Bacteria* and A*rchaea* (29). PIN-domains are present in *Bacteria*, *Archaea* and *Eukaryotes* and consist of small proteins of about 130 amino acids (AA) whereas three are strictly conserved. These three acidic AA form the active site and are able to cleave single stranded RNA in a sequence specific manner (28, 30).

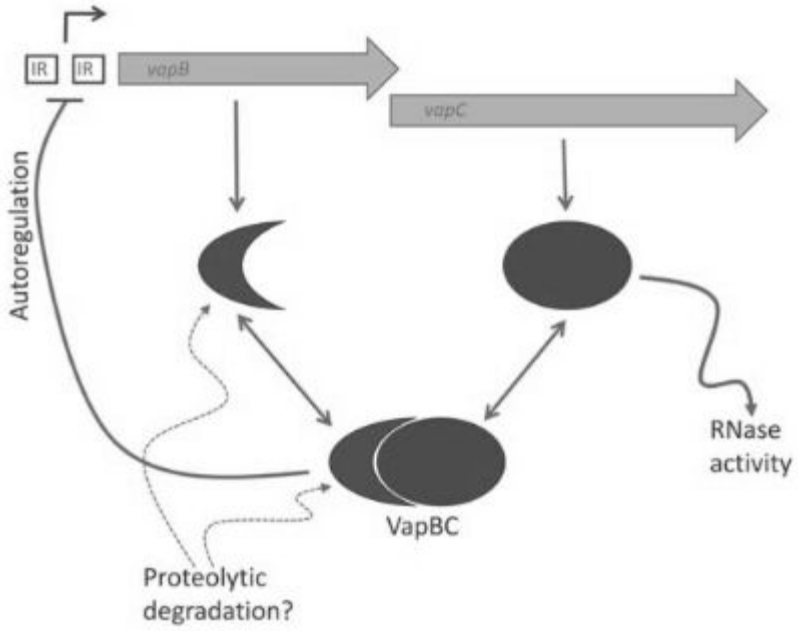


**Figure 11: VapBC toxin-antitoxin system**. VapBC belongs to the type II TA systems and consists of two overlapping genes. The stable toxin (VapC) builds a complex with the instable antitoxin (VapB). Once VapB gets degraded the toxin becomes active. The expression of the toxin and the antitoxin is autoregulated (28).

#### 1.3.3.6 PCNA

The Proliferating cell nuclear antigen (PCNA) is an essential part of the chromosomal DNA replisome in eukaryotes and belongs to the family of DNA sliding clamp (31). Two different groups first discovered PCNA independently. In patients that suffered from lupus erythematosus, Miyachi *et al.* discovered 1978 an auto-antigen that was located in the nucleus of dividing cells. The protein was named proliferating cell nuclear antigen. 1980 Bravo and Celis identified a protein that was expressed in the S-phase of the cell cycle via two dimensional gel electrophoresis. They called it cyclin, it was the same protein, and the name PCNA has prevailed (34). This protein is evolutionarily well conserved and has not only been found in humans, animals, plants and yeasts but also in archaea (32).

##### 1.3.3.6.1 Structure

As mentioned above, PCNA belongs to the family of the DNA sliding clamp together with the DNA polymerase (pol) II β subunit of *E. coli* and the gene 45 protein of the phage T4. The name sliding clamp derived from the ability of the protein, to bind to DNA and slide along the molecule (32), as schown with a crosslinking experiment. As PCNA assembled around a circular plasmid, crosslinking the protein to the DNA was possible but after the plasmid was linearized, crosslinking was no longer feasible. Leading to the conclusion, that PCNA slid off of the plasmid (34). PCNA can either be a homotrimer as in eukaryotes or a heterotrimer as in the crenarchaea *Sulfolobus solfataricus* (35). The ring shaped structure composes of 12 α helices in the inner cavity and a continuous layer of β sheets on the outside. The inner pore is positively charged which ensures a stable interaction between the DNA and the PCNA, whereas the outside is charged negatively that avoids nonspecific interaction with the clamp loader RFC (34). Each monomer consists of an inter-domain connecting loop (IDCL) that connects the N- and C-terminal domains of each monomer. IDCL also have another significant function, performing as a docking site for several proteins such as DNA ligase 1, p21, DNA pol δ, flap endonuclease (Fen1) and

DNA-(cytosine-5) methyl transferase. It was found that the N-terminus of PCNA α-helices interact with cyclinD and the C-terminus interacts with replication factor C, DNA pol ε, cycline-dependent kinase 2 and the protein 45 that is involved in growth arrest and DNA damage (32). Overall the protein is about 80 Å (1 Å = 0,1 nm) wide, 30 Å thick and the pore is approximately 34 Å in diameter.(34).

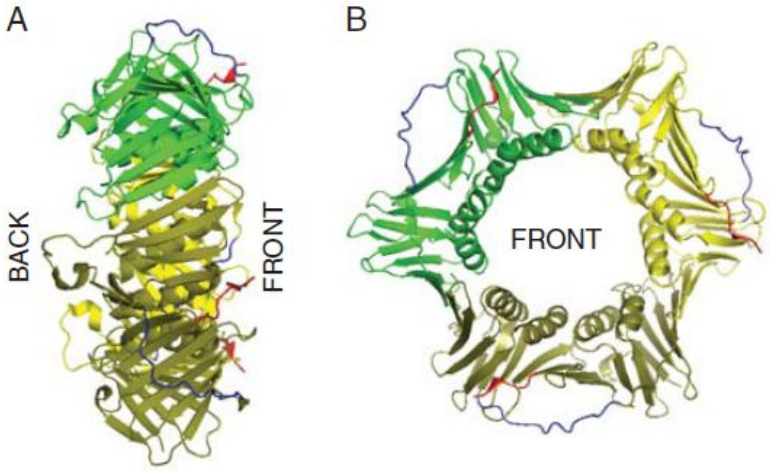


**Figure 12: PCNA structure of Arabidopsis thaliana.** The three subuntis are indicated in different colours (32).

#### 1.3.3.6.2 Functions

##### 1.3.3.6.2.1 DNA replication

Replication in eukaryotes starts with unwinding the double stranded DNA through the protein MCM (helicase). In order to elongate the lagging and leading strand, the enzyme Pol α (primase) synthesizes small RNA primers that can be elongated by the polymerases Pol δ (leading strand) and Pol ε (lagging strand). RPA proteins (replication protein A) bind to the lagging strand and prevent the single stranded DNA from rewinding and nucleases. Now the replication factor C (RFC), also called clamp loader, recognizes these primers and loads the PCNA onto the DNA (33, 36). Now PCNA performs its function and tethers the polymerase to the template and elongation starts. Upon the high speed of the polymerase, the sliding clamp prevents dissociation of the enzyme from the DNA (36).

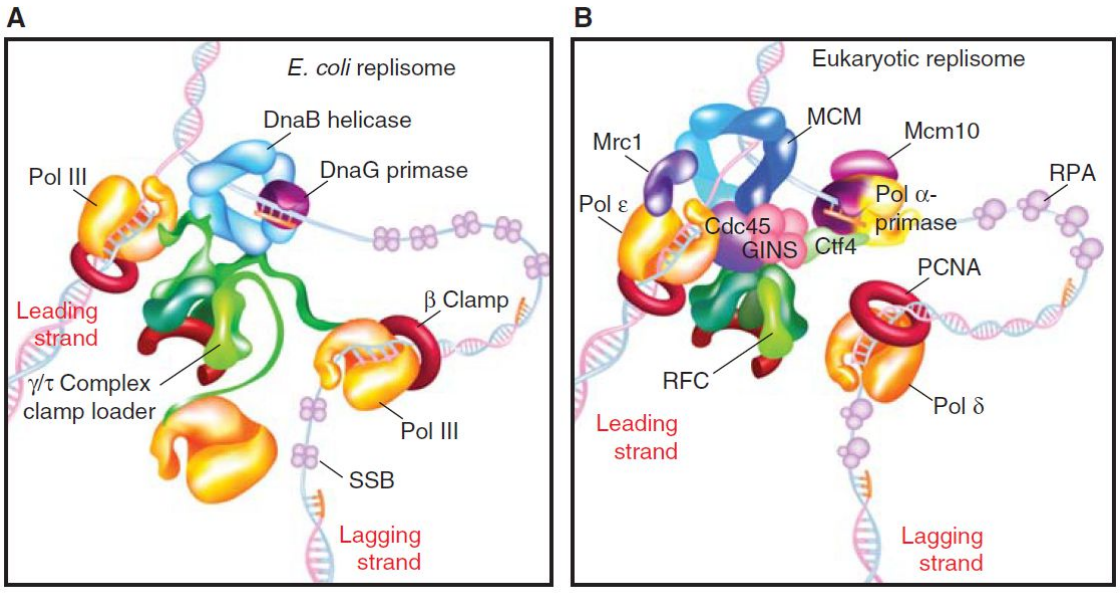


**Figure 13: Schematic drawing of bacterial and eukarytic replisome. A:** In *E. coli* the helicase DnaB unwinds the double stranded DNA. The primase DnaG is now able to synthesize small primers on the lagging strand, which are recognized by the β Clamp followed by binding of the polymerase PolII. **B**: In *Eukaryotes,* the protein complex MCM performes the helicase activity. Pol α-primase synthesizes the small primers on the lagging strand. They are recognized by PCNA which loads the polymerase Pol δ to the DNA and elongation starts (33).

#### 1.3.3.6.2.2 DNA repair

Immunohistochemical studies suggested, that PCNA is involved in DNA repair. The nucleus was incubated with anti-PCNA antibodies and subsequently radiated with UV-light to produce damages in the DNA. Although cells were not in the S-phase, the protein could be detected and led to the assumption that PCNA is involved in the DNA repair process. A key component in the cellular DNA damage response is the protein p53. It is also called "guardian of the genome" and as the concentration of this protein increases, transcription of several genes are affected as well as the cell cycle arrests. Cells lacking p53 are no longer able to arrest after DNA damage. It was shown, that p53 acts as a transcriptional activator of the PCNA promoter (34).

**Nucleotide excision repair (NER)**

UV light exposed to cells can cause severe damages to the DNA resulting in mutations, strand breaks, cell death or transformation to a malignant phenotype (37). Regarding NER, the formation of cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts (6-PPs) are the major targets of this repair mechanism (38). The transcription factor complex TFIIH contains the helicases XP-B and XP-D that unwinds the DNA at the site of the lesion. The damage binding protein XP-A is crucial in verifying the DNA damage and assembles of the repair mechanism. Open DNA is stabilized by the replication protein A and further responsible in positioning the endonucleases XP-G and ERCC1-XPF. Damaged area is removed (24-32 nucleotides) and PCNA mediates re-synthesis of the new DNA fragment by recruiting DNA pol δ/ε, DNA lig1 and Fen 1 (38, 32). PCNA showed to be directly located at the site of lesion. Immunofluorescence studies revealed that localization of PCNA to the damaged site is influenced by a functional XP-A protein that was tested with XP-A mutant cells after UV irradiation. This might conclude that PCNA is also important in other steps of NER except of only DNA re-synthesis (32).

Mismatch repair (MMR)

Base mismatches during DNA synthesis can occur due to the error rate of DNA polymerases, insertion/deletion loops that can form during replication or while recombination events (32). The major sensor complex Mutsα (MSH2-6 complex) recognizes small mismatches whereas MutSβ (MSH2-MSH3 complex) recognizes larger base-base mismatches. The faulty strand is removed by the exonuclease EXO1. It was found that PCNA does not only play a role in DNA repair syntheses but also in formation of MMR machinery. PCNA directly interacts with EXO1, MLH1, MSH6 and MSH3 und thus allow the proper function of this pathway (39).

**Base excision repair (BER)**

This machinery is activated to repair chemical altered nucleotides of bases in DNA (oxidized, reduced, deaminated or alkylated bases) or misincorporated uracils (39). Specific DNA glycosylases recognize these damages and removes the base from the nucleotide (39). These results in formation of apurinic/apyrimidinic (AP) sites also called abasic site. These sites have neither a purin nor a pyrimidine base (32). An AP endonuclease than cleaves the phosphodiester bond and BER can continue in two different ways. In the short-patch repair, pol β is involved that fills up the gap and XRCC1-DNA ligase 3 complex seals the nick (39). In the long-patch repair, PCNA is required to recruit the polymerases pol δ/ε (32). As in the mismatch repair mechanism, PCNA seems to be involved also in the initiation of BER since it also interacts with the glycosylases or the AP endonucleases (39).

#### 1.3.3.6.2.3 Other functions

**Table 1** Proteins known to interact with PCNA, and the roles of the interaction

| *PCNA interacts with* | *Function of protein* | *Role of interaction with PCNA* |
|---|---|---|
| Replication Factor C | Loads and Unloads the sliding clamp | Loading/unloading of PCNA |
| DNA polymerase δ | The catalytic unit of DNA polymerase δ holoenzyme (DNA replication and repair) | Required for processive DNA synthesis |
| DNA polymerase ε | DNA polymerase (lagging strand polymerase [?], DNA repair, check point control) | Required for processive DNA synthesis |
| Fen-1 | Flap endonuclease (Okazaki fragment maturation) | Stimulates the activity of Fen-1 |
| Gadd45 | Unknown; induced upon DNA damage | Unknown |
| D-type cyclins | G1 cyclins | Inhibits DNA replication |
| p21/Cip1 CDK inhibitor | Cyclin-dependent kinase inhibitor | Inhibits processive DNA synthesis |
| RNA polymerase | RNA polymerase | Required for viral late gene transcription (baculovirus) |

**Figure 14: Functions of PCNA.** This protein interacts with many different partners and plays important roles in several mechanisms such as DNA replication and DNA repair (34).

### 1.3.3.6.3 ORF 59

The ORF 59 encodes for the ϕCh1 version of PCNA. It is likely that it is involved in replication of the virus by binding an *N. magadii*-encoded DNA polymerase. PCNA may also bind specialized factors encoded by ϕCh1 that are needed for replication.

# 2 MATERIAL AND METHODS

## 2.1 Material

### 2.1.1 Strains

#### 2.1.1.1 Escherichia coli

| Strain | Genotype | Source |
|---|---|---|
| **Xl1-Blue** | *endA1 gyrA96* ($Nal^R$) *thi-1 recA1 relA1 lac glnV44* F'[::Tn10 *proAB*$^+$ *lacI*$^q$ Δ(lacZ)M15] *hsdR17* ($rK^-mK^+$) | Startagene |
| **Rosetta** | $F^-$. *ompT hsdSB* ($rB^-$ $mB^-$) *gal dcm* λ(DE3) pLysSRARE ($Cam^R$) | Novagene |

#### 2.1.1.2 *Natrialba magadii*

| Strain | Genotype | Source |
|---|---|---|
| **L11** | Wilde type, carry provirus ϕCh1 | Witte *et al*., 1997 |
| **L13** | Cured strain, not infected by ϕCh1 | Witte *et al*., 1997 |

### 2.1.2 Plasmids

| No.* | Construct | Features | Source |
|---|---|---|---|
| **1240** | pBluescritp II KS(+) | mcs, bla, ColE1 ori, lacZa | Startagene |
| **673** | pBAD24 | Bla, araC, rrnB, mcs, $P_{BAD}$ promoter, pBR322 ori | Guzman et al., 1995 |
| **809** | pRO-5 | bla, ColE1 ori, gyrB ($Nov^j$), ϕCh1 derived | Mayrhofer-Iro *et al*., 2015 |

| | | | |
|---|---|---|---|
| **799** | pUC19 | Bla, pMB1ori, lacZa, mcs | Yanisch-Perron et al., 1984 |
| **1440** | pKSII-PCNA-1-2 | pBluescript II KS(+) with upstream flank of PCNA gene | This study |
| **1441** | pKSII-PCNA 1-4 | pBluescript II KS(+) with up- and downstream flank of PCNA gene | This study |
| **1442** | pKSII-PCNA 1-5 | pBluescript II KS(+) with up- and downstream flank of PCNA gene | This study |
| **1449** | pKSII-PCNA 1-5-NovR "forward" | pBluescript II KS(+) with up- and downstream flank of PCNA gene with novobiocin-resistance cassete in forward direction | This study |
| **1450** | pKSII-PCNA 1-5-NovR "reverse" | pBluescript II KS(+) with up- and downstream flank of PCNA gene with novobiocin-resistance cassete in reverse direction | This study |
| **1453** | pUC19-p43-ORF44 | pUC19 plasmid with ORF44 under the promoter of ORF43 | This study |
| **1454** | pKSII-PCNA 1-4-NovR "reverse" | pBluescript II KS(+) with up- and downstream flank of PCNA gene with novobiocin-resistance cassete in reverse direction | This study |
| **1458** | pBAD24-tnaN-ORF44 | pBAD24 with ORF44 under the tnaN promoter | This study |
| **1462** | pKSII-dORF44 1-4 | pBluescript II KS(+) with up- and downstream flank of ORF 44 gene | This study |
| **1489** | pKSII-dORF44-1-4-NovR "reverse" | pBluescript II KS(+) with up- and downstream flank of ORF 44 gene with novobiocin-resistance cassete in reverse direction | This study |
| **1492** | pRO5-pTNA-N-ORF44 | pRO5 with the ORF 44 under the | This study |

| | | | |
|---|---|---|---|
| | | ptnaN promoter | |
| **1495** | pNB102-34up-P1 | pNB102 with downstream fragment of PCNA with the upstream region of ORF34 | This study |
| **1497** | pRSET-PCNA | pRSET-A plasmid with the PCNA gene | This study |
| **1065** | pBAD24-tnaN | pBAD24 with the tryptophanase promoter of *N. magadii* (ptnaN) | Alte, 2011 |
| **783** | pNB102 | bla, CoIE1 ori, hmg ($Mev^R$), pNB101 ori | Zhou et al., 2004 |
| **1347** | pRSET-A | bla, pUC ori, T7 promoter, N-terminal 6x His-tag | Invitrogen |
| **1471** | pKSII-dORF44-1-4-NovR "forward" | pBluescript II KS(+) with up- and downstream flank of ORF 44 gene with novobiocin-resistance cassete in forward direction | This study |

### 2.1.3 Primer

| No.* | Primer | Sequence | Restriction site |
|---|---|---|---|
| 887 | ΔPCNA-1 | GATCTCTAGAGAAGTGCTACCGCTGCA | *Xba*I |
| 888 | ΔPCNA-2 | GAATCCCGGGAGCTCAGCGTGCTCGA | *Sma*I |
| 889 | ΔPCNA-3 | GTATAAGCTTCCAGAGCAACTGAGTGATCA | *Hind*III |
| 890 | ΔPCNA-4 | GATCGGTACCGCGTAGATCGGTCGC | *Kpn*I |
| 891 | ΔPCNA-5 | GATCGGTACCGTCCCGCATGCACTCAG | *Kpn*I |
| 347 | Nov-6 | GGGATCGCAGAGGAGC | |
| 548 | Nov-9 | GATGTCGGTCATCGCGG | |
| 830 | Nov-13 | GACGCCGAATGGGTAGAC | |
| 829 | Nov-12 | GCCGGTGAGTACTTAACGC | |
| 951 | PCNA-3en | GTCGACGATCTCCGCTAGCT | |
| 331 | 56-5 | CAGCAGGATCCATGAGAGAGAACAATCC | *Bam*HI |
| 294 | 43-5 | CAGCAGTCTAGACGTTGTGCCAGCCGT | *Xba*I |
| 311 | 44-Hind | CAGCAAGCTTGATTTAGGACTCGAGGACC | *Hind*III |
| 914 | 44-3-Bam | GTTAGGATCCACAGAACGGACGTACGAC | *Bam*HI |
| 915 | 44-Nde | GATCCATATGACGCTGTTCGTCGACA | *Nde*I |
| 811 | pBAD24-3 | GATCGAATTCGGTACCCTTCTCTCATCCGC | *Eco*RI/ *Kpn*I |

| | | | |
|---|---|---|---|
| **686** | Tna-N-1 | GATTAAGCTTCTGAGGAATCGACCGGTT | *Hind*III |
| **953** | NovR-1s | GATCCCCGGGTGTCGACTGGAACGAGG | *Sma*I |
| **954** | NovR-2p | GTTACTGCAGGCCGTAATATCCAGCTGA | *Pst*I |
| **952** | NovR-1p | GATCCTGCAGTGTCGACTGGAACGAGG | *Pst*I |
| **955** | NovR-2s | GTTACCCGGGCCGTAATATCCAGCTGA | *Sma*I |
| **227** | p30-3 | CAGCAGAAGCTTTCAGTTGCTCTGGATCCG | *Hind*III |
| **229** | PCNA-2 | ACACAGGGTACCGTGTCAATGGTGATGGTGAT GGTTGCTCTGGATCCGC | *Kpn*I |
| **226** | p30-5 | CAGCAGAGATCTATGTTCAAAGCAATCGCTAC | *Bgl*II |
| **101** | 23-Hind | CAGGAAGCTTCTTCGATGGTGACTGTGG | |
| **102** | 23-Kpn | CAGCAGGTACCATGGCGCTCCTCGC | |
| **942** | Δ44-1 | GCAGTCTAGAGTGTCTCATGCATAGGGGT | *Xba*I |
| **810** | pBAD24-5 | GAATTCGGTACCGCACGGCGTCACAC | *Eco*RI/ *Kpn*I |
| **850** | Tna-N1-1 | GATTAAGCTTGCGAGTCTTCAGGCCCTA | |

*: numbers refer to the collection numbers

### 2.11.4 Media

#### 2.1.1 LB-Medium for *E. coli*

| NaCL | 5 g |
|---|---|
| **Yeast extract** | 5 g |
| **Peptone** | 10 g |

add $H_2O$ to a final volume of 1 L and adjust pH to 7.0, subsequently autoclave the medium. For agar plates add 15g/L agar, for softagar 7.5g/L.

#### 2.1.4.2 NVM+ medium for *N. magadii*

| NaCl | 235 g |
|---|---|
| **KCl** | 2.35 g |
| **Tri-Na citrate** | 0.8 g |
| **Yeast extract** | 11.7 g |
| **Casaminoacids** | 8.8 g |

add $H_2O$ to a final volume of 935ml, adjust pH to 9.0 with 4-5 NaOH platelets per liter and autoclave. For agar plates add 8g/L agar or for soft agar 4g/l

after autoclaving, complement the medium with:

| | | |
|---|---|---|
| **0.57 M** | $Na_2CO_3$ | **63 ml** |
| **1.0 M** | $MgSM_4$ | 1 ml |
| **20.0 mM** | $FeSO_4$ | 1 ml |

### 2.1.5 Antibiotics

#### 2.1.5.1 *E. coli*

| Antibiotic | Stock concentration | Final concentration | preparation and storage |
|---|---|---|---|
| **Ampicillin** | 20 mg/ml | 100 µg/ml | dissolved in $ddH_2O$, filtered sterile and stored at 4℃ |
| **Tetracyclin** | 10 mg/ml | 10 µg/ml | dissolved in 1/2 volume $ddH_2O$ and 1/2 volume 96% EtOH, stored protected from the light at -20℃ |
| **Kanamycin** | 25 mg/ml | 50 µg/ml | |

#### 2.1.5.2 *N. magadii*

| Antibiotic | Stock concentration | Final concentration | preparation and storage |
|---|---|---|---|
| **Bacitracin** | 7 mg/ml | 70 µg/ml | dissolved in $ddH_2O$, filtered sterile and stored at 4℃ |
| **Mevinolin** | 10 mg/ml | 7.5 µg/ml | extracted from pulverized tablets and dissolved in 96 % EtOH, stored at -20℃ |

| | | | |
|---|---|---|---|
| **Novobiocin** | 3 mg/ml | 3 µg/ml | dissolved in $ddH_2O$ and filtered sterile, stored protected from light at -20℃ |

### 2.1.6 Enzymes

#### 2.1.6.1 Restriction Enzymes

| Enzyme | Restriction site | Company |
|---|---|---|
| **AhaIII/DraI** | 5′-TTT↓AAA-3′ | Thermo Scientific |
| **BamHI** | 5′-G↓GATCC-3′ | Thermo Scientific |
| **EcoRV** | 5′-GAT↓ATC-3′ | Thermo Scientific |
| **HindIII** | 5′-A↓AGCTT-3′ | Thermo Scientific |
| **KpnI** | 5′-GGTAC↓C-3′ | Thermo Scientific |
| **NdeI** | 5′-CA↓TATG-3′ | Thermo Scientific |
| **NheI** | 5′-G↓CTAGC-3′ | Thermo Scientific |
| **PstI** | 5′-CTGCA↓G-3′ | Thermo Scientific |
| **SmaI** | 5'-CCC ↓G-3' | Thermo Scientific |
| **XbaI** | 5′-T↓CTAGA-3′ | Thermo Scientific |

### 2.1.6.2 DNA polymerases

| Enzyme | Company |
|---|---|
| **GoTaq® Green Master Mix** | Promega |
| **GoTaq® DNA Polymerase** | Promega |
| ***Pfu* DNA Polymerase** | Promega |
| **T4 DNA Polymerase** | Thermo Scientific |

### 2.1.6.3 DNA-modifying Enzymes

| Enzyme | Company |
|---|---|
| **Klenow Fragment** | Thermo Scientific |
| **T4 DNA Ligase** | Promega |

### 2.1.7 Antibodies

| Primary Antibody | Dilution | Target | Source |
|---|---|---|---|
| **α-E (from rabbit)** | 1: 5000 | Protein E of ϕCh1 | Lab. stock |
| **α-gp34 (from rabbit)** | 1:2500 | Tail fibre protein of ϕCh1 | Lab. stock |
| **α-MTase (from rabbit)** | 1: 5000 | | Lab. stock |
| **α-ΔPCNA (from rabbit)** | 1:500 | PCNA of ϕCh1 | Lab. stock |

| Secondary Antibody | Dilution | Target | Source |
|---|---|---|---|
| **ECL Anti-mouse IgG, horseradish peroxidase linked (from donkey)** | 1:5000 | mouse IgG | GE Healthcare Life Science |

### 2.1.8 Marker

#### 2.1.8.1 DNA Ladders

| DNA Ladder | Size of DNA fragments in base pairs | Company | Product Number |
|---|---|---|---|
| **λ*Bst*EII** | 8453, 7242, 6369, 5687, 4822, 4324, 3675, 2323, 1929, 1371, 1264, 702 | Thermo Scientific | λ DNA: SD0011 |
| **GeneRuler 1 kb DNA Ladder** | 10000, 8000, 6000, 5000, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 750, 500, 250 | Thermo Scientific | SM0311 |

#### 2.1.8.2 Protein Ladders

| Protein Ladder | Size of Proteins in kDa | Company | Product Number |
|---|---|---|---|
| **Page Ruler Ptrestained Protein Ladder** | 170, 130, 100, 70, 55, 40, 35, 25, 15, 10 | Thermo Scientific | 26616 |
| **Unstained Protein Molecular Weight Marker** | 116, 66.2, 45, 35, 25, 18.4, 14.4 | Thermo Scientific | 26610 |

### 2.1.9 Kits

| Kit | Company | Product Number |
|---|---|---|
| **QIAquick Gel Extraction Kit** | Qiagen | 28704 |
| **GeneJET Plasmid Miniprep Kit** | Thermo Scientific | K0502 |
| **GeneJET PCR Purification Kit** | Thermo Scientific | K0701 |
| **Super Signal West Pico Chemiluminescent Substrate** | Thermo Scientific | 34080 |

### 2.1.10 Other Additives

| Additive | Stock concentration | Final concentration | |
|---|---|---|---|
| **IPTG** | 1 M | 0.5 - 1 mM | dissolved in $ddH_2O$ and stored at -20℃ |
| **Tryptophan** | 0.6 M | 2 mM | dissolved in 1 M NaOH and stored at -20℃ |
| **X-Gal** | 100 mg/ml | 40 µg/ml | dissolved in DMF and stored at -20℃ |

### 2.1.11 Buffers and Solutions

#### 2.1.11.1 Competent cells

##### 2.1.11.1.1 *E. coli*

| MOPS I | |
|---|---|
| $CaCl_2$ | 10 mM |
| RbCl | 10 mM |
| MOPS | 100 mM |
| adjust the pH to 7 with KOH | |

| MOPS II | |
|---|---|
| $CaCl_2$ | 70 mM |
| RbCl | 10 mM |
| MOPS | 100 mM |
| adjust the pH to 6.5 with KOH | |

| MOPS IIa | |
|---|---|
| $CaCl_2$ | 70 mM |
| RbCl | 10 mM |
| Glycerol | 15 % |
| MOPS | 100 mM |
| adjust the pH to 6.5 with KOH | |

2.1.11.1.2 *N. magadii*

| Buffered high salt spheroblasting solution with glycerol | |
|---|---|
| **NaCl** | 2 M |
| **KCl** | 27 mM |
| **Tris/HCl (pH9)** | 50 mM |
| **Glycerol** | 15 % |
| **after autoclaving add 15 % sterile filtered sucrose** | |

| Buffered high salt spheroblasting solution without glycerol | |
|---|---|
| **NaCl** | 2 M |
| **KCl** | 27 mM |
| **Tris/HCl (pH8)** | 50 mM |
| **after autoclaving add 15% sterile filtered sucrose** | |

| Unbuffered high salt spheroblasting solution | |
|---|---|
| **NaCl** | 2 M |
| **KCl** | 27 mM |
| **after autoclaving add 15% sterile filtered sucrose** | |

### 2.1.11.2 Virus isolation

| High-salt alkaline solution | |
|---|---|
| Tris/HCl, pH 9.5 | 50 mM |
| NaCl | 4 M |

| 1.5 CsCl solution | |
|---|---|
| Tris/HCl, pH 9.5 | 50 mM |
| NaCl | 23.4 g |
| CsCl | 135 g |
| fill up to 200 ml with $ddH_2O$ | |

| 1.3 CsCl solution | |
|---|---|
| Tris/HCl, pH 9.5 | 50 mM |
| NaCl | 23.4 g |
| CsCl | 90 g |
| fill up to 200 ml with $ddH_2O$ | |

| 1.1 CsCl solution | |
|---|---|
| Tris/HCl, pH 9.5 | 50 mM |
| NaCl | 23.4 g |
| CsCl | 20 g |
| fill up to 200 ml with $ddH_2O$ | |

### 2.1.11.3 Protein extracts

| 2x Laemmli buffer | |
|---|---|
| **SDS** | 2 % |
| **ß-mercaptoethanol** | 5 % |
| **Glycerol** | 10 % |
| **Bromphenol blue** | 0.01 % |
| **Tris-HCl pH 6.8** | 60 mM |

| 5 mM Sodium phosphate buffer pH 6.8 | |
|---|---|
| **$NaH_2PO_4$** | 0.2 M |
| **$Na_2HPO_4$** | 0.2 M |

### 2.1.11.4 SDS-PAGE

| 4x Separating gel buffer | |
|---|---|
| **SDS** | 0.4 % |
| **Tris-HCl, pH8.8** | 1.5 M |

| 4x Separating gel buffer | |
|---|---|
| **SDS** | 0.4 % |
| **Tris-HCl, pH8.8** | 1.5 M |

| 12 % Separation gel | |
|---|---|
| **$ddH_2O$** | 1.75 ml |
| **4x Separation gel buffer** | 1.25 ml |
| **30 % PAA** | 2 ml |
| **10 % APS** | 60 µl |
| **TEMED** | 10 µl |

| 4 % Stacking gel | |
|---|---|
| **$ddH_2O$** | 1.23 ml |
| **4x Separation gel buffer** | 500 µl |
| **30 % PAA** | 267 µl |
| **10 % APS** | 20 µl |
| **TEMED** | 5 µl |

| 10x SDS Running buffer | |
|---|---|
| **Glycine** | 1.92 M |
| **Tris-base** | 0.25 M |

| 30 % Acrylamide | |
|---|---|
| N,N'-methylenebisacrylamide | 1 % |
| Acrylamide | 29 % |

### 2.1.11.5 Coomassie staining

| Coomassie staining solution | |
|---|---|
| Acetic acid | 10 % |
| Methanol | 25 % |
| Coomassie Brilliant Blue R-250 | 0.15 % |

| Destaining solution | |
|---|---|
| Acetic acid | 10 % |
| Methanol | 25 % |

### 2.1.11.6 Western blot

| Transfer buffer | |
|---|---|
| **Glycine** | 39 mM |
| **Tris-base** | 48 mM |
| **Methanol** | 20 % |
| **SDS** | 0.037 % |

### 2.1.11.7 Ponceau S staining and blocking

| Ponceau S staining solution | |
|---|---|
| **TCA** | 3 % |
| **Ponceau S** | 0.5 % |

| 10x TBS | |
|---|---|
| **NaCl** | 1.37 M |
| **KCl** | 27 mM |
| **Tris-HCl pH8** | 0.25 M |

| Blocking solution | |
|---|---|
| **1x TBS** | x ml |
| **Milk powder** | 5 % |

### 2.1.11.8 Protein purification

| 10x PBS | |
|---|---|
| **$NaH_2PO_4$** | 14.7 mM |
| **$Na_2HPO_4$** | 81.0 mM |
| **KCl** | 27.0 mM |
| **NaCl** | 1.37 mM |
| **pH 7.4** | |

| Buffer B pH 8- Lysis buffer | |
|---|---|
| **Urea** | 8 M |
| **Tris base** | 10 mM |
| **$NaH_2PO_4$** | 100 mM |
| **pH 8.0, adjusted with NaOH** | |

| Buffer D pH 5.9 - Elution buffer | |
|---|---|
| **Urea** | 8 M |
| **Tris base** | 10 mM |
| **$NaH_2PO_4$** | 100 mM |
| **pH 5.9, adjusted with HCl** | |

| Buffer C pH 6.3 - Wash buffer | |
|---|---|
| **Urea** | 8 M |
| **Tris base** | 10 mM |
| **$NaH_2PO_4$** | 100 mM |
| **pH 6.3, adjusted with HCl** | |

| Buffer E pH 4.5 - Elution buffer | |
|---|---|
| **Urea** | 8 M |
| **Tris base** | 10 mM |
| **$NaH_2PO_4$** | 100 mM |
| **pH 4.5, adjusted with HCl** | |

The pH of buffers B-E was adjusted before use, due to the dissociation of urea that causes pH changes.

### 2.1.11.9 Southern blot

| 20 x SSC | |
|---|---|
| NaCl | 3.0 M |
| Sodium citrate | 0.3 M |
| pH 7.2 | |

| 50 x Denhardt´s solution | |
|---|---|
| BSA | 0.1 g/l |
| Polyvinylpyrolidone | 0.1 g/l |
| Ficoll 400 | 0.1 g/l |

| Hybridization buffer | |
|---|---|
| $ddH_2O$ | 55 ml |
| 20 x SSC | 25 ml |
| Denhardt´s solution | 10 ml |
| 10 % BSA | 5 ml |
| 1 M $Na_2HPO_4$ | 5 ml |
| 20 % SDS | 500 µl |
| 0.5 M EDTA | 200 µl |

| Blocking solution | |
|---|---|
| NaCl | 7.3 g |
| $Na_2HPO$ | 2.41 g |
| $NaH_2PO_4$ | 0.96 g |
| SDS | 49,89 g |
| pH 7.2 | |

| 10 x Wash solution II | |
|---|---|
| Tris | 12.1 g |
| NaCl | 5.85 g |
| $MgCl_2$ | 2.03 g |
| pH 9.5 | |

| 20 x SSC | |
|---|---|
| NaCl | 3.0 M |
| Sodium citrate | 0.3 M |
| pH 7.2 | |

## 2.2. Methods

### 2.2.1 Transformation of *E. coli*

#### 2.2.1.1 Preparation of competent cells

For the generation of competent *E. coli* cells, 100 ml LB-medium (supplemented with the required antibiotics, see 2.1.1) was inoculated with a fresh overnight culture of *E. coli* strain. Starting with an OD600 of 0.1 the culture was incubated with agitation to an OD600 of 0.6 at 37 °C. Afterwards, they were harvested by centrifugation at 10.000 rpm at 4 C for 10 minutes. The supernatant was discarded, the pelleted cells were resuspended in 40 ml MOPS I solution (see 2.1.11.1.1) and incubated for 10 minutes on ice. The cells were collected (10.000 rpm and 4 C for 10 minutes), the pellet was dissolved in 40 ml MOPS II (see 2.1.11.1.1) and incubated on ice for 30 minutes. The third and last centrifugation step was performed as previous at 10.000 rpm at 4 C for 10 minutes. Supernatant was discarded again and 2 ml MOPS IIa (see 2.1.11.1.1) was added. In order for a later use of the generated competent cells, they were aliquoted in Eppendorf tubes (100µl each) and stored at -80 °C.

#### 2.2.1.2 Transformation of competent cells

For the transformation of competent cells, the prepared aliquot (100 µl) stored at -80 °C was thawed for 10 minutes on ice. 15 µl of the desired ligation batch was added and incubated again on ice for 30 minutes. As a next step, the cells were incubated for 2 minutes at 42 °C and placed back on ice for a few minutes. In order to let the cells regenerate at 37 °C for 30 minutes, 300 µl of LB-medium without any antibiotics were added. After the cells were recovered, they where plated out on 3 LB agar plates (100 µl each) with the appropriate antibiotics and incubated over night at 37 °C. Once, colonies have formed, they were picked and each single one was incubated in test tubes containing 5 ml LB-medium and the suitable antibiotics over night (37 °C).

#### 2.2.1.3 Screening

In order to analyse the clones, 300 µl of the previous incubated culture, derived from a single colony were centrifuged at 13.000 rpm for 3 minutes. The supernatant was discarded and the pellet was resuspended in 30 µl 5X DNA loading buffer. 14 µl Phenol/Chloroform (1:1) were added and vortexed for approximately 20 seconds. The sample was centrifuged for 5 minutes at 13.000 rpm. 12 µl of the aqueous phase were loaded immediately onto a 0.8 % agarose gel. For better interpretation of the gel, also the corresponding plasmid without the insert was applied. Based on the separation behavior, smaller fragments migrate faster through the gel than larger ones. Thus, the plasmid with the insert derived from a putative positive clone, should run higher on the gel than the empty control plasmid. To further verify the putative positive clone, the gene of interest, that has been cloned, was than detected via analytical PCR.

### 2.2.2 Transformation of *Natrialba magadii*

#### 2.2.2.1 Preparation of competent cells

For the generation of competent *N. magadii* cells, three baffled Erlenmeyer flasks with 60 ml NVM rich medium (see 2.1.4.2) were inoculated with a dense grown culture of *N. magadii*. The

amount of fresh culture that was added was increased with each flask, 2 ml, 4 ml and 6 ml. The reason for this is, to ensure that one of the flasks has the right $OD_{600}$ for the next steps. However, before the cells were grown over night at 37 °C, 70 µg/ml of bacitracin was added. Once the cells reached an $OD_{600}$ of 0.5 to 0.6 they were harvested by centrifugation at 6.000 rpm for 15 miutes at room temperature. The supernatant was discarded and the pellet was resuspended carefully in 30 ml high salt-buffered spheroblasting solution with glycerol. After that, 30 µl of proteinase K was added for the digestion of the S-layer. The mixture was transferred to a 100 ml flask and was incubated for 1 to 2 days at 42 °C. After complete digestion of the S-layer, the cells were no longer rod shaped but spheroblasts. This was monitored by microscope analysis. These competent cells were used immediately for the subsequent transformation. If they were not used right after they were made competent, cells were stored at -80 °C, but only for a short storage time of about 7 days; otherwise they were no longer useful for proper transformation.

#### 2.2.2.2 Transformation of competent cells

For the transformation 1.5 ml of the competent cells were centrifuged at 10.000 rpm for 3 minutes at room temperature. Supernatant was discarded and the pellet was resuspended in 150µl high salt buffered spheroblasting solution without glycerol (2.1.11.1.2). 15 µl of 0.5M EDTA were added and incubated for 10 minutes at room temperature. Afterwards, a maximum of 10 µl DNA, with a DNA concentration of 0.5 to 10 µg was added and incubated for 5 minutes at room temperature. Since the DNA is dissolved in $ddH_2O$, 10 µl should not exceed to prevent lysis of the cells, due to the changing NaCl concentrations. Now 150µl of 60% PEG 600 (600µl of PEG 600 was mixed with 400 µl of unbuffered high salt spheroblasting solution without glycerol) was added and incubated for 30 minutes at room temperature. 1 ml of NVM+ rich medium was carefully added and centrifuged at 10.000 rpm for 5 minutes, this step was repeated 2 times to wash the cells and get rid of the PEG 600. After the supernatant was discarded the pellet was very carefully resuspended with 1 ml of NVM+ rich medium and regenerated at 37 °C. During the regeneration,

where the S-layer is build up and cells become rod shaped again, the medium was changed every day. This was monitored via microscope.

Once the cells were regenerated, they where plated on NVM+ agar plates (100 µl per plate) and incubated at 42 °C. To prevent evaporation of the water and drying out of the plates, they were sealed in plastic bags. Colonies could be detected after 2 to 3 weeks.

#### 2.2.2.3 Screening

After colonies had formed on the plates, single colonies were picked and inoculated in 700 µl NVM+ medium (in Eppendorf tubes) and incubated at 42 °C. For the better growth of *N. magadii*, tubes were ventilated every day for about 10 seconds. After the cultures were grown they were tested by PCR (see 2.2.3.2.2).

### 2.2.3 DNA Methods

#### 2.2.3.1 Agarose gel electrophoresis

A 0.8 % agarose gel was poured for the separation of DNA fragments by length. The required amount of agarose was dissolved in 1x TAE buffer and melted. After all, the solution had to be cooled down before pouring into the gel electrophoresis apparatus. After separation, the gel was stained with ethidium bromide, in order to visualize the DNA with an UV transilluminator. The ethidium bromide bath had a concentration of 10µg/ml

#### 2.2.3.2 Polymerase chain reaction

##### 2.2.3.2.1 Preparative PCR

The preparative polymerase chain reaction (PCR) was performed using the *Pfu*-polymerase. This enzyme derived from the thermophilic archaea *Pyrococcus furiosus*; it has proofreading activity and is able to incorporate 500 nucleotides per minutes throughout the elongation step. These properties ensure a sequence exact amplification.

| **Batch for *Pfu* DNA polymerase** | |
|---|---|
| **ddH2O** | 67 µl |
| **10x Pfu buffer** | 10 µl |
| **dNTP's (0.2M)** | 10 µl |
| **forward primer (0.05µg/µl)** | 5 µl |
| **reverse primer (0.05 µg/µl)** | 5 µl |
| **Pfu Polymerase (2U/µl)** | 1 µl |
| **DNA Template** | 1 µl |
| | 100 µl |

| PCR-Programm | | | |
|---|---|---|---|
| **Step** | Temperature (°C) | Duration (minutes) | Number of cycles |
| **Initial Denaturation** | 95 | 5 | 1 |
| **Denaturation** | 95 | 1 | 35 |
| **Annealing** | X | 1 | |
| **Extension** | 72 | Y | |
| **Final Extension** | 72 | 5 | 1 |

X... annealing temperature is dependent of the melting temperature of the primers. 4 °C were substracted from the calculated annealing temperature.

Y... elongation time is dependent on the length of the PCR product and the used polymerase

#### 2.2.3.2.2 Analytical PCR

For analytical PCR the GoTaq DNA polymerase was used. It is able to incorporate 1000 nucleotides per minute but has a relatively high error rate of approximately 1 x $10^{-5}$ per base. This is due to the enzyme's lack of proofreading activity.

| Batch for GoTaq DNA polymerase | |
|---|---|
| **dd$H_2O$** | 6.5 µl |
| **2X GoTaq Green Master Mix** | 10 µl |
| **forward primer (0.1 µg/µl)** | 10 µl |
| **reverse primer (0.1µg/µl)** | 5 µl |
| **DNA Template** | 1 µl |
| | 25 µl |

GoTaq Green Master Mix allready contains nucleotides and reaction buffer.

| PCR-Programm | | | |
|---|---|---|---|
| **Step** | Temperature (C) | Duration (minutes) | Number of cycles |
| **Initial Denaturation** | 95 | 5 | 1 |
| **Denaturation** | 95 | 1 | 20 |
| **Annealing** | X | 1 | |
| **Extension** | 72 | Y | |
| **Final Extension** | 72 | 2 x Y | 1 |

X... annealing temperature is dependent of the melting temperature of the primers. 4 °C were substracted from the calculated annealing temperature.

Y... elongation time is dependent on the length of the PCR product and the used polymerase

#### 2.2.3.3 Template preparation of E. *coli*

To produce a template from *E. coli*, 5 µl of a culture was mixed with 100 µl $ddH_2O$ and vortexed about 10 seconds. Then the cells were heated up to 95 °C for 10 minutes. DNA is set free and ready to use as a template.

#### 2.2.3.4 Template preparation of *N. magadii*

For the preparation of an *N. magadii* template, 30 µl of a culture was centrifuged 5 minutes at 13.000 rpm. The supernatant was discarded and the pellet was resuspended in 100 µl $ddH_2O$ which causes cells to burst and thus setting free the DNA.

#### 2.2.3.5 Purification of PCR

After the PCR was finished and controlled on a agarose gel for proper working, the PCR had to be purified to get rid of the buffer, dNTPs and polymerase to ensure optimal conditions for upcoming restrictions. Therefore, the GeneJet PCR Purification kit from Thermo Scientific was used. All steps were done according to the manufactures protocol, but instead to elute the DNA with elution buffer, $ddH_2O$ was used.

#### 2.2.3.6 Gel elution

Sometimes it may happen, that after the PCR product is applied on a gel, more than the desired DNA fragment appears and unspecific bands occur. Therefore, the fragment of interest was cut out of the gel and was purified with the QIAquick Gel Extraction Kit from QIAGEN. All steps were done according to the manufactures protocol, but instead to elute the DNA with elution buffer, $ddH_2O$ was used.

#### 2.2.3.4 Plasmid isolation from *E. coli*

To isolate the plasmid DNA from *E. coli*, the GeneJet Plasmid Miniprep Kit was used. All steps were performed according to the manufactures protocol; elution was carried out with $ddH_2O$ instead of elution buffer.

#### 2.2.3.5 Isolation of chromosomal DNA from *N. magadii*

500 ml of a dense culture of *N. magadii* was centrifuged for 20 minutes at 8.000 rpm at room temperature. The supernatant was discarded and the pellet was resuspended in 5 ml of 4 M NaCl/ 50 mM Tris-HCl (pH 9.5). 2.5 ml of 14mM deoxycholate were added and mixed gently. The mixture was diluted with 7.5 ml of $ddH_2O$ and extracted with 12.5 ml of phenole/chlorophorm (1:1) After centrifugation at 12.000 rpm for 30 minutes at 4 °C, the supernatant was transferred to a 100 ml Erlenmeyer flask and overlayed with 0.6 volume of isopropanol. The DNA was precipitated and dissolved in 5 ml sterile $ddH_2O$. EDTA with a final concentration of 10 mM was added. 7 g of CsCl per 5 ml dissolved DNA were added as well as 15 µl Ethidiumbromid (10 mg/ml). The mixture was centrifuged for about 16 hours with 60.000 rpm at room temperature. The DNA was isolated and ethidiumbromid was removed with water saturated $ddH_2O$. CsCl was removed by dialysing the sample against sterile $ddH_2O$ over night.

#### 2.2.3.9 DNA modifications

##### 2.2.3.9.1 DNA restriction

Two different types of restriction enzymes were used. DNA that was used for cloning steps was restricted with conventional enzymes as recommaned by the manufacturer´s and incubated either over night or for 3 hours at 37 °C. For analytical analysis, fFast digest restriction enzymes were used as recommended and were incubated for 1 hour at 37 °C.

#### 2.2.3.9.2 Fill-in 5' overhangs

5' overhangs occuring after restriction were filled in with the Klenow fragment from Thermo Scientific. Enzyme was used according to the manufacturer's protocol.

#### 2.2.3.9.3 Ligation

To connect a plasmid with a DNA fragment that were restricted with the same enzymes, were ligated with the T4 DNA ligase from Promega. This enzyme was used for both blunt end and sticky end ligation. According to the manufacturer's protocol, an insert to vector ratio of 3:1 was respected. For sticky end ligation, reactions were carried out at room temperature or over night at 4 °C. Blunt end ligation was performed over night at 16 °C.

| Batch for T4-DNA Ligase | |
|---|---|
| **Plasmid DNA** | 1 µl |
| **Insert DNA** | 11.5 µl |
| **Reaction buffer** | 1.5 µl |
| **T4-DNA Ligase** | 1 µl |
| | 15 µl |

## 2.2.4 Southern Blot

### 2.2.4.1 Synthesis of the probe

For the southern blot, virus DNA was isolated as described in (virus DNA isolation). A probe for the 5´-end (1214 bp) and the 3´-end (649 bp) was created with biotinylated dUTP via PCR.

| Batch for probe preparation | | | |
|---|---|---|---|
| | | 5´-end | 3´-end |
| **Primer 1** | 10 µl | ΔPCNA-1 | ΔPCNA-3 |
| **Primer 2** | 10 µl | ΔPCNA-2 | ΔPCNA-5 |
| **GoTaq Green Mastermix** | 50 µl | | |
| **Template** | 1 µl | | |
| **dd$H_2O$** | 26.5 µl | | |
| **Biotinylated dUTP** | 2.5 µl | | |
| | 100 µl | | |

PCR program see...

#### 2.2.4.1.1 Blotting the DNA to the membrane

As a first step, the DNA was denatured in 0.4 N NaOH/ 0.6 M NaCl for 30 minutes and neutralized 30 minutes in 1.5 M NaCl/ 0.5 Tris-HCl pH 7.5. Membrane was equilibrated in 10 x SSC and the capillary blot was build up were the DNA was transferred over night. As a next step, the membrane was incubated for one minute in 0.4 N NaOH and 0.2 M Tris-HCl pH 7.5 each. Now the blotted DNA was fixed by UV-crosslinking.

#### 2.2.4.1.2 Prehybridization and Hybridization

Before hybridization, the membrane was blocked for 3 hours in 12 ml of hybridization buffer and 120 µl salmon sperm DNA (10 mg/ml) at 65 °C. Now the prepared probe was denatured for 5 minutes at 95 °C and added to the hybridization reaction. This step took place over night at 65 °C. On the next day the membrane was washed twice with 2 x SSC/0.1% SDS at room temperature for 5 minutes. After that, two wash steps with 0.1 x SSC/ 0.1% SDS at 65 °C for 15 minutes followed.

#### 2.2.4.1.3 Developing the blot

The blot was blocked with blocking solution for 5 minutes and afterwards incubated 5 minutes with streptavidin solution (7 ml blocking solution + 7 µl streptavidin). Then the membrane was washed three times for 5 minutes with wash solution I (1:10 delution of the blocking solution). The next step was to incubate the blot 5 minutes with biotinylated alkaline phosphatase solution (7 ml blocking solution + 7 µll alkaline phosphatase) and wash it for 5 minutes with the blocking solution. Another three washing steps were done for 5 minutes with wash solution II. The membrane was shrink-wraped and incubated for 5 min in CDP-Star™ reagent (3 ml dilution buffer + 6 µl CDP Star reagent). Finally, the liquid was removed from the membrane followed by the exposure to an x-ray film.

### 2.2.5 Creating homozygote *N. magadii* mutants

Since *N. magadii* can have up to 50 copies of its genome in one cell, it was important to replace all the wild type copies with the mutated ones. Two strategies to achieve this were performed.

#### 2.2.5.1 Passaging

The transformation of suicid plasmids into *N. magadii* and successfully recombination leads to a heterogzygote due to the high copy number of chromosomes. The strains had to be passaged to get in homozygous cultures. Therefore, 25 ml of NVM medium were inoculated with 200 µl of a

dense preculture. Selection was performed with 25 µl novobiocin. The culture was grown on 37 °C untill an OD600 of 1. This culture was used to inoculate frech medium again. This process was repeated until no wild type copies coud be detected anymore. A PCR analysis for homozygote mutants was performed every fifth passage.

#### 2.2.5.2 Infection with ϕCh1

Homozygous cultures of lysogenic strains were performed as follows: the respective culture was inoculated in NVM medium and incubated at 37 °C until it was lysed completely. The cultre was centrifuged for 15 minutes at 13.000 rpm on room temperature and afterwards filtered sterile. A *N. magadii* L13 culture was grown to an $OD_{600}$ 0.5. 500 µl of this culture where than infected with 100 µl, 200 µl, 300 µl, 400 µl and 500 µl of phage lysate and incubated for 1 hour at 37 °C. 200 µl of each mixture were plated out on NVM (with novobiocin) plates and incubated at 42 °C until colonies were visible. These colonies were than tested by PCR for homozygous mutants.

### 2.2.6 Protein Methods

#### 2.2.6.1 Preparation of crude protein extracts

To analyze proteins of *E. coli* and *N. magadii* via SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) and Western blot, crude protein extracts had to be generated. Therefore, 1.5 ml of a culture was centrifuged at 13.000 rpm for 3 minutes. The supernatant was discarded and the pellet was resuspended in x µl 5 mM sodium phosphate buffer and x µl of 2x Laemmli buffer. The amount of the buffers was calculated as following: $OD_{600}$ of the culture multiplied by 75. For *E. coli* protein extracts, samples were boiled at 95 °C for 10 minutes and ready to be loaded onto a protein gel. For *N. magadii* protein extracts, the samples were incubated at 37 °C for about 24 hours until they were not viscous anymore and probes were able to load onto the protein gel. Samples not used directly were stored at - 20 °C.

### 2.2.6.2 SDS-PAGE

Sodiumdodecylsulfate polyacrylamide gel electrophoresis is a method to separate proteins according to their size. SDS is a detergent and has the ability to denature proteins and cover up the intrinsic charge of them. Separation of the now denatured and constantly negative charged proteins is possible. The preparation of the gel was performed with the Bio-Rad Mini-PROTEAN apparatus. First the separation gel (see 2.1.11.4) with 12 % polyacrylamid was mixed on ice and poured between two glas plates. It was covered with about 1 ml of isopropanol. Once the separation gel was polymerized, the isopropanol was removed. The stacking gel (see 2.1.11.4) with 4 % polyacrylamid was mixed on ice and poured directly above the separating gel, followed by the immediate insertion of the combs. The polymerized gel was set into the electrophoresis chamber, covered with 1x SDS Running buffer (see 2.1.11.4) and the comb was removed carefully. Finally, samples were loaded and separated (40 V until the samples reached the stacking gel, 60 V until the samples reached the separation gel and 100 V for the separation gel. For *N. magadii* proteins the voltage was turned up to only 30 V for a better separation of samples with high salt concentrations. After the electrophoresis was finnished, the separation gel was stained with Coomassie for 5 to 10 minutes. The gel was than destained (see 2.1.11.5) until the background was completely clear and the protein bands were visible.

### 2.2.6.3 Western Blot

This technique allows the detection of specific proteins using antibodies. First, proteins were separated by size via SDS-PAGE as described above, but instead of staining the gel with Coomassie it was transferred to a nitrocellulose membrane (Protran BA85). Therefore, 6 Whatman filter papers and the nitrocellulose membrane were cut into the size of the gel and soaked with transfer buffer. The blot was build up as following: at the bottom 3 layers of Whatman papers, followed the nitrocellulose membrane, then the PAA-Gel and on top again 3 layers Whatman paper. The transfer was performed into a blotting apparatus for 20 minutes with 20 V. The negatively charged proteins move to the positive electrode right into the nitrocellulose

membrane. After the blot was finished, the membrane was stained with Ponceau S for a couple of minutes and destained with dd$H_2O$. The next step was the antibody reaction, but first, the membrane had to be blocked to avoid unspecific binding of the antibodies. Therefore, the membrane was incubated in blocking solution (see 2.2.6.3) over night at 4 °C. The next day, the blocking solution was removed and the blot was washed 3 times with 1x TBS for 10 minutes. TBS was removed and the membrane was incubated with the primary antibody for 1 hour at room temperature. Now the blot was washed 3 times with 1x TBS for 10 minutes to remove antibodies that had not bound to the proteins. Then the secondary antibody was added and also incubated for 1 hour. Secondary antibodies cannot be reused and thus was discarded. Again the blot was washed 3 times for 10 minutes to get rid of the excess of secondary antibodies. The final step of the western blot was the detection via chemiluminescence (SuperSignal West Pico Chemiluminescent Substrat). Secondary antibodies were linked to an enzyme, called horse redish peroxidase (HRP). This enzyme is capable to catalyze the oxidation of luminole in the presence of H2O2 to 3-aminoptalate and light. This light that is produced in this reaction is caught on an X-ray film (Amersham Hyperfilm ECL) in the darkroom

### 2.2.7 Construction of an expression strain

For the over-expression of a gene the *E. coli* strain Rosetta was used. This strain carries the T7 polymerase of bacteriophage T7 that was cloned under the control of the lac promoter which is inducible by IPTG. The gene that encodes for the protein of interest was cloned into the expression vector pRSET-A, that has the strong T7 promoter region. This constructed plasmid was transformed into Rosetta. Upon inducing with IPTG the T7 polymerase gets synthesized, thus leading to the expression of the protein under the control of the T7 promoter region on the plasmid. Before producing large quantities, the clone was analysed regarding the expressing the gene of interest as well as the time point of the highest expression rate. Therefore, the culture was inoculated to an OD*600* of 0.1 and was grown to an OD*600* of 0.3 at 37 °C. Then induction was induced by adding 1 mM of IPTG. The culture was incubated for 2 hours at 37 C. Meanwhile

samples were taken at several time points (one before induction, after 15 min, 30 min, 45 min, 60 min, 90 min, 120 min) to monitor the level of expression. These samples were prepared like in 2.2.6.1 and separated via SDS-PAGE. Proteins were detected via Wester-blot.

### 2.2.8 Over-expression and purification

Once a clone was producing the protein of interest, 1 L LB-medium was inoculated with this strain and induced as mentioned above. Since the western blot showed that after 60 minutes, expression was not getting stronger, cells were harvested by centrifugation with 6.000 rpm for 15 minutes at 4 °C. The pellet was frozen over night at -20 C. The pellet was thawed on ice for 10 minutes, resuspended in buffer B (see 2.1.11.8) and stirred for 4 hours at room temperature. Lysis of the cells was checked under the microscope. The samples were sonicated when unlysed cells were remaining. After that, the lysate was centrifuged at 10.000 x g for 30 minutes at room temperature. The supernatant was incubated with Ni-NTA over night while stirring. This step causes the binding of the Ni-NTA beads to the His-Tag of the proteins und subsequently allows purification by affinity chromatography.

After stirring over night, the mixture was loaded onto a chromatography column. The flow through was collected, while the proteins remain bound to the matrix. Proteins were washed twice with 4 ml of buffer C (see 2.1.11.8) and eluted with 4 x 500 µl of buffer D and E see (2.1.11.8). Each step was collected in an Eppendorf tube for analysis in which fraction the protein is via SDS-PAGE. Samples were mixed with 2x Laemmli buffer, heated for 10 minutes at 95 C to unfold proteins and applied on the gel.

Fractions containing the proteins they were pooled and dialyzed against 1x PBS (see 2.1.11.8). First, 1 hour at room temperature and then over night after replacing the buffer with fresh one.

### 2.2.9 Growth curve experiment

For growth curve experiments, a fresh and dense culture of either the strain *N. magadii* L11 or *N. magadii* L13 was inoculated with 60 ml NVM to an OD600 of 0.1. Now over a period of 8-10 days, the optical density was measured every day, as well as taking samples (2.2.6.1).

### 2.2.10 Cloning strategies

#### 2.2.10.1 PCNA deletion mutant.

For the construction of a PCNA deletion mutant, first the upstream region of ORF59 that encodes for the β-sliding clamp protein, was amplified using the ϕCh1 DNA as template. With the primers ΔPCNA-1 and ΔPCNA-2 a 1240bp long fragment was generated. These primer pair created restriction sites *Xba*I (5' end) and *Sma*I (3' end). This fragment was ligated with the pKSII+ vector, upon restriction with the same enzymes. Positive clones were named pKSII-PCNA1-2. The next step was the amplification of the downstream region of ORF59 with the primers PCNA-3 (5' end) and PCNA-4 (3' end), creating a fragment with 780 bp in length and restriction sites for *Hind*III and *Kpn*I. Since the ORFs are very close to each other on the genome of ϕCh1, a second downstream fragment was amplified using the primers PCNA-3 (5' end) and PCNA-5 (3' end), avoiding to amplify parts of the neighboring ORF60. This fragment was 629 bp long, with the same restriction sites. The both fragments were restricted with *Hind*III and *Kpn*I as well as the pKSII-PCNA1-2 and ligated. Positive clones were named pKSII-PCNA-1-4 and pKSII-PCNA-1-5. The novobiocin resistance cassette (2453 bp) in forward and reverse position was isolated from the vector pUC19-NovR "forward" and pUC19-NovR "reverse" with the restriction enzymes *Sma*I and *Pst*I. Novobiocin cassettes were cloned into the vectors. Positive clones were named pKSII-PCNA-1-4-novR "reverse", pKSII-PCNA-1-5-novR "forward" and pKSII-PCNA1-5-novR "reverse". The construct pKSII-PCNA1-4-novR "forward" could not be obtained during this thesis. These three plasmids were transformed into *N. magadii* L11.

### 2.2.10.2 Over-expression/complementation of ORF59

For the over expression, pQE-34upP1 (with His-Tag) was isolated from *E. coli*. The vector was restricted with XbaI and KpnI and the insert was eluted from an agarose gel. The plasmid pNB102 was digested with the same enzymes and ligated with purified insert. Positive clones were named pNB102-34up-P1 with His-Tag. For over expression the plasmid was transformed in *N. magadii* L11. As a control, the plasmid was also transformed in the strain L13. For the complementation, the construct was transformed into L11-ΔPCNA.

### 2.2.10.3 ORF44 deletion mutant

The upstream region of the ORF44 was amplified using the ϕCh1 DNA as template to construct a deletion mutant. With the primers Δ44-1 and Δ44-2 a 480 bp long fragment was generated. These primers inserted the restriction sites *Bam*HI (5' end) and *Xba*I (3' end). The fragment was ligated into the pKSII+ vector, which was restricted with the same enzymes. Positive clones were named pKSII-ΔORF44 1-2. The downstream region was amplified with the primers Δ44-3 and Δ44-4 with a length of 659 bp. The fragment and the construct pKSII-ΔORF44 1-2 were both restricted with *Hind*III and *Kpn*I and ligated. Positive clones were named pKSII-ΔORF44 1-4. The novobiocin resistance cassette (2453 bp) in forward and reverse orientation was isolated from the vector pUC19-NovR "forward" and pUC19-NovR "reverse" with the restriction enzymes *Sma*I and *Pst*I. Novobiocin cassettes were cloned into the vectors. Positive clones were named pKSII-ΔORF44-1-4-novR "forward" and pKSII-ΔORF44-1-4-novR "reverse.

These plasmids were trasformed into *N. magadii* L11.

#### 2.2.10.4 ORF44 under the pTNA promoter.

The ORF44 was amplified with the primers 44-3-Bam and 44-Nde using the ϕCh1 as template. The fragment was 416 bp in length and restricted with the restriction enzymes *Bam*HI and *Nde*I. The plasmid with the tnaN promoter was isolated, pBAD24-tnaN. Vector was digested with same enzymes and ligated with the ORF44 fragment. Positive clones were named pBAD24-tnaN-ORF44. Now ORF44 with the pTNA promoter was isolated via PCR with the primers pBAD24-5 and Tna-N1-1, resulting in a 920 bp long fragment. The vector pRo-5 and the insert were both restricted with *Kpn*I/*Hind*III and ligated. Positive clones were named pRo5-pTNA-N-ORF44 and transformed together with a second construct pNB102-Bgb34$_{52}$ into *N. magadii* L13. As a control, the empty vector pRo5 together with pNB102-Bgb34$_{52}$ was transformed in *N. magadii* L13.

### 2.2.11 ϕCh1 methods

#### 2.2.11.1 Isolation of virus particles

The first step for the isolation of ϕCh1 virus particels was to inoculate 4.5 L NVM medium with 40 ml of *N. magadii* L11 culture. They were incubated at 37 °C until all the cells were lysed. This was monitored by measuring the optical density every day. The typical growth curve of *N. magadii* L11 shows its peak after 3 days where lyses starts and new viruses are released. After about 8 days, where no lysis was detected anymore, cells were centrifuged at 8.000 rpm for 20 minutes at room temperature. Cells and cell debris are now pelleted and separated from the virus particles, which are still in the supernatant. To precipitate the particles, 10% PEG-6000 was added to the supernatant and stirred over night. Now virus particles are coated with PEG-6000 and able to be pelleted by centrifugation at 8.000 rpm for 20 minutes at room temperature. The pellet was resuspended in high-salt alkaline solution (see 2.1.11.2). To purify the virus, this mixture was applied on a discontinuous CsCl-gradient. A centrifuge tube was filled with 3.5 ml of 1.5 CsCl solution, then 4 ml of 1.3 CsCl solution was added carefully to avoid mixing the gradient. On top 5 ml of the virus suspension was pipetted. The 1.1 CsCl solution (see 2.1.11.2) was used for

balancing the tubes against each other. The centrifugation step was performed at 30.000 rpm for 20 hours at room temperature. Now 2 light blue bands showed up, the upper band were archaeal flagellum und the lower band the desired virus particles. This band was carefully removed with the pipette and applied on a continuous gradient for another round of purification. For this gradient, only the 1.3 CsCl solution (see 2.1.11.2) was used and centrifuged again at 30.000 rpm for 20 hours at room temperature. The light blue band was collected. The next step was, to get rid of the CsCl in the sample via dialysis against high-salt alkaline solution. First, the virus was dialyzed for 1 hour, then the buffer was changed and the second round took place over night at room temperature. The isolated virus was stored at room temperature.

### 2.2.11.2 Virus DNA isolation

For the isolation of virus DNA, 100 µl of purified viruses were mixed with 400µl dd$H_2O$ and 400 µl phenole/chlorophorm (1:1). After short vortexing, the mixture was centrifuged 2 minutes at 13.000 rpm at room temperature. The supernatant was collected and again mixed with 400µl of phenole/chlorophorm (1:1) and centrifuged as before. This step was repeated until no white interphase was visible. The DNA was precipitated with 800µl isopropanol and centrifuged for 30 minutes at 13.000 rpm and 4 C. Pellet was dried and resuspended in 10µl ddH2O. As a control, 3 µl of DNA was applied on an agarose gel.

### 2.2.11.3 Virus titer assay

To determine the infectivity of the ϕCh1 virus, a virus titer or also called plaque assay was performed. For this experiment, a *N. magadii* L13 culture grown to the stationary phase and isolated viruses were needed. The pure virus was diluted with NVM medium ($10^{-2}$, $10^{-4}$, $10^{-6}$, $10^{-8}$, $10^{-10}$) and added to 5 ml soft agar as well as 300 µl of the uninfected *N. magadii* L13 culture. This mixture was poured onto NVM agar plates and incubated at 37 C in sealed plastic bags. Once plaques have formed, they were counted and the plaque forming units (PFU) per ml virus were calculated.

# 3 RESULTS AND DISCUSSION

## 3.1 Construciton of ϕCh1::PCNA and its effect on the development of virus ϕCh1

### 3.1.1. Aim

The aim of this project was, to generate a mutant strain with a deleted PCNA gene of ϕCh1 and to analyse effects of this mutation regarding the development of the virus. Alignment analysis of the PCNA of *N. magadii* and ϕCh1 showed a sequence identity of 65% over 323 AA. Furthermore, the virus encoded PCNA is 21 AA larger than the one of *N. magadii*, thus makes it possible to distinguish between these two proteins. As described above, PCNA is involved in the DNA replication process and many other steps. It encircles the single stranded DNA followed by binding of the polymerase to the protein. This interaction prevents the dissociations of the polymerase from the DNA during the replication process (36). The reason why the virus is not using the PCNA of the host and instead encoding its own is not known yet. A possible explanation could be that the virus-encoded form of PCNA has a greater affinity to its target. Therefore, the question arose what are the differences of the two strains: the wild type and the mutated one.

### 3.1.2 Experimental setup

For the deletion of the PCNA wild type gene, a plasmid was constructed that contained a novobiocin resistance cassette flanked by the up- and downstream region of the PCNA gene. Via these regions, a homologous recombination event occurred leading to the replacement of the PCNA gene with the novobiocin cassette. The plasmid that was used is called pBluescript II KS+. This suicide plasmid is able to replicate in *E. coli* but not in *N. magadii*, since it lacks an archaeal origin of replication. So after possible successful homologous recombination in *N. magadii*, pBluescript II KS+ cannot propagate and gets lost over time. The finished construct was transformed into *N. magadii* and positive clones were subsequently screened for homologous recombination and further made homozygous.

Three different plasmids were created. PCNA 1-4 contains a 5' end of the PCNA gene (1,194 bp), the NovR cassette (2,453 bp) in reverse orientation and the 3' end of PCNA (780 bp). PCNA 1-5 has the same 5' ends as the one before but a shorter 3' end of the PCNA gene (629 bp). PCNA 1-5 contains the NovR cassette in forward and reverse orientation. The 3' end in the PCNA 1-5 construction was shortened to avoid cloning parts of the neighboring ORF since they are very close to each other. Since the NovR cassette was cloned without a terminator sequence, both orientations were used to avoid influencing the neighboring genes upon transcription. In this master thesis, no positive clone with the construct PCNA 1-4 NovR in forward orientation could be obtained.

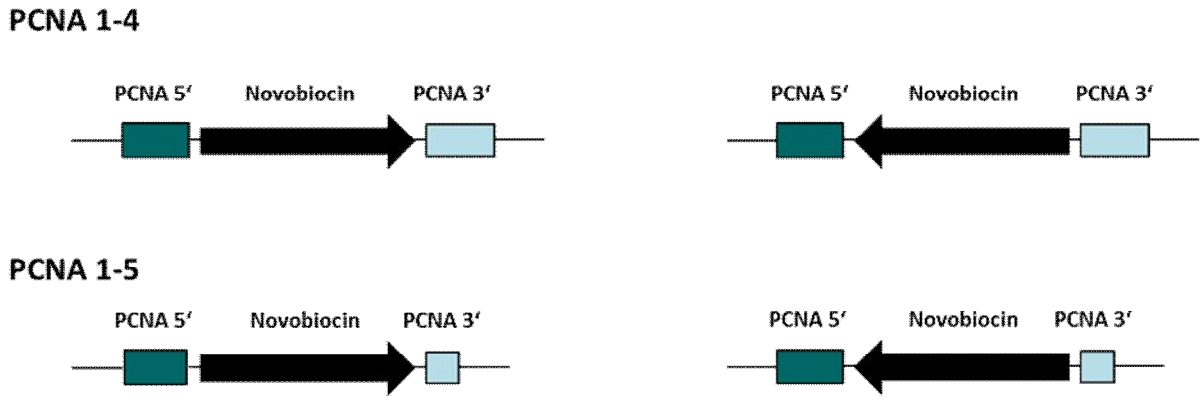


**Figure 15: schematically presentation of the 4 clones which should be cloned during this thesis.** The novobiocin cassette is indicated by a black arrow, the upstream and downstream sequences of ORF59 with a green or light green box and marked with PCNA 5´and PCNA 3´, respectively. Top clone PCNA 1-4, bottom: clone PCNA 1-5. Both variants differ in their 3´-prime recombination sites. The drawing is not to shale.

### 3.1.3 Screening for positive clones

After the transformation of the plasmids into the strain *N. magadii* L11, single colonies were picked and inoculated in NVM medium. The uptake of the desired plasmid was screened via PCR with primers within the 5' region of PCNA and within the NovR cassette (Primer ΔPCNA1 + Nov-13).

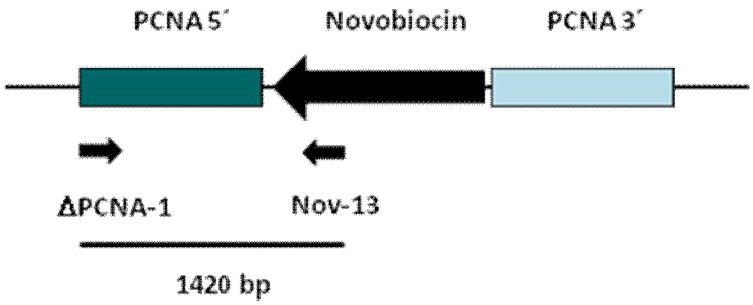


**Figure 16: schematically drawing of the expected region inserted into the chromosome of *N. magadii* L11.** The up- and downstream sequences are given in colors; the novobiocin resistance cassette is indicated as an arrow. Primes used for the verification are indicated as arrows and marked with ΔPCNA1 and Nov-13. The desired fragment length is given beneath the primer names. The drawing is not to scale.

The analyses of 50 different transformants showed that at least with one colony a positive signal using the PCR analysis could be obtained (Fig. 17, lane 5). A signal is visible with the same length compared to the positive control used in Fig 17, lane 12. Here plasmid "pKSII-PCNA 1-5-NovR" reverse used for the transformation of *N. magadii* L11 was used.

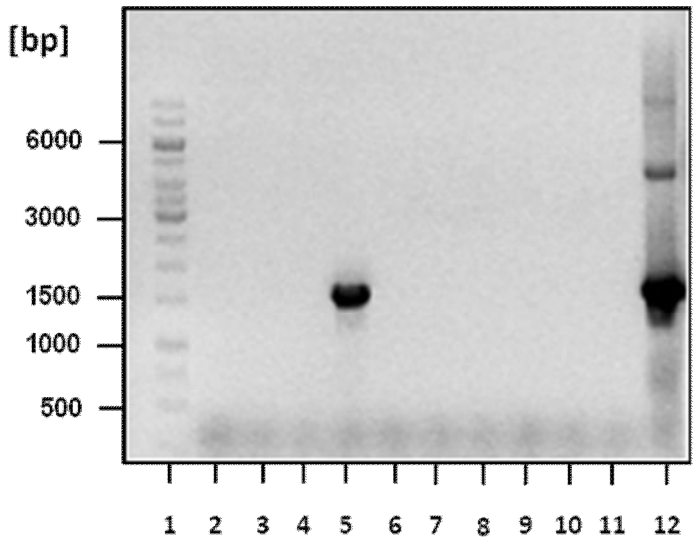


**Figure 17: Positive clone screening of a PCNA deletion mutant**. Lane 1: 1-kb-ladder, lane 2-10: different clones of the transformation in *N. magadii* L11, Lane 11: plasmid pKSII-PCNA 1-5-NovR "reverse" as positive control. Clone number 44 in lane 5 showed a positive signal.

### 3.1.4 Homozygation of *N. magadii* L11-ΔPCNA and double cross over

The genome of ϕCh1 is integrated in the chromosomal DNA of the lysogenic strain *N. magadii* L11 (8). Halophilic *Archaea* are known to be polyploid. It was determined that for example *Halobacterium salinarum* contains 25 copies of the chromosome during logarithmic growth. It is assumed that haloalkaliphilic *Archaea* are polyploidy, too. As shown by Derntl *et al*. (2015), passaging of mutant strains to generate a homozygous strain is necessary. In case of deletion mutants located on the chromosome, 15 to 20 passages are needed to generate homozygous strains. For ϕCh1 mutants the infection of *N. magadii* L13 with virus particles should result in homozygous strains. In general, only one particle should infect a cell of *N. magadii* L13 since super-infection does not occur (8). In order to generate a homozygous mutant, the positive clone was transferred into NMV medium and incubated until all cells were lysed. Virus particles of the supernatant were collected containing both wild type and the mutated gene. The cured strain of *N. magadii* L13 was incubated with the virus lysate. Infected cells were plated on rich medium agar plates with novobiocin and only cells which are infected by the mutated virus will be able to grow. Once colonies had formed they were tested for a double crossover via PCR. To analyze if recombination occurred at the 5' end, the primers 56-5 and Nov 13 were used. For the 3' end the

primers PCNA3-en and Nov-12 were used (Fig. 18). Here, besides the internal novobiocin primers, primer 56-5 and PCNA-3en were used. Both primers, specific for the ϕCh1 sequence, are located outside the fragments used for the recombination.

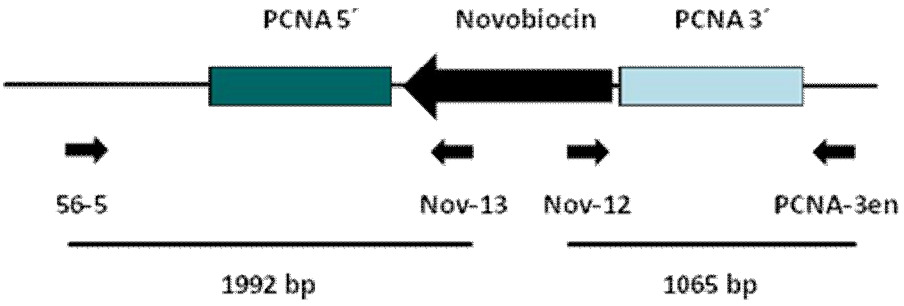


**Figure 18: schematically drawing of the expected region inserted into the chromosome of *N. magadii* L11.** The up- and downstream sequences are given in colors; the novobiocin resistance cassette is indicated as an arrow. Primes used for the verification are indicated as arrows and marked with ΔPCNA1 and Nov-13. The desired fragment lengths are given beneath the primer names. The drawing is not to scale.

The PCR analyses of a possible double crossover were performed with 10 different colonies inoculated in rich medium after the lysogenization of *N. magadii* L13 (Fig. 19).

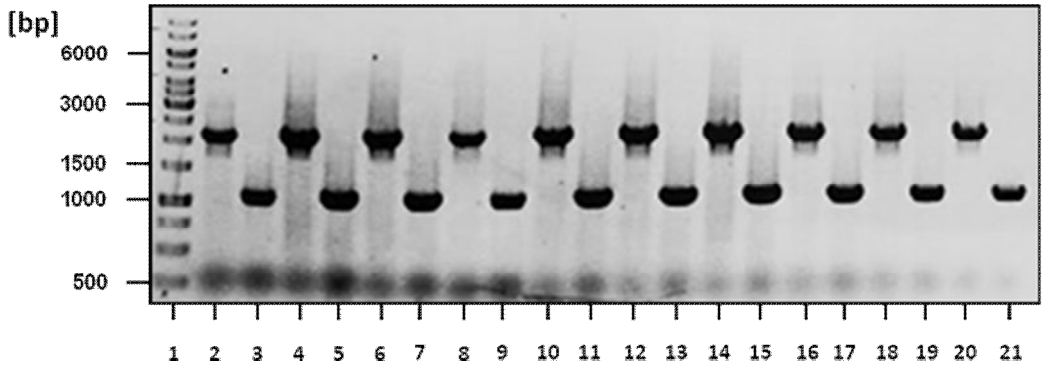


**Figure 19: Screening for double-crossover**. Lane 1: 1-kb-ladder, lane 2-21: different clones of *N. magadii* L11-ΔPCNA. Lanes 2, 4, 6, 8, 10, 12, 14,16, 18 and 20 represent clones 1 – 10 using primers 56-5 and Nov-13 to verify the 5´-crossover, lanes 3, 5, 7, 9, 11, 13, 15, 17, 19 and 21 represent the same clones using primers Nov-12 and PCNA-3en to confirm the 3´-crossover. All clones showed the expected fragments with 1992 bp for the 5´-crossover and 1065 bp for the 3´-crossover. The fragment sizes of the used marker are indicated on the left.

As shown in Fig. 19, the PCR reactions of all cultures resulted in a signal using primers 56-5 and Nov-13. As primer 56-5 is specific for the ϕCh1 genome located upstream of the cloned fragment used for recombination, the crossover of this fragment into the genome of ϕCh1 could be confirmed. On the other hand, the crossover of the 3´-fragment could also be verified, as the calculated fragment length of 1065 bp could by amplified with primers Nov-12 and PCNA-3en. Again, primer PCNA-3en is specific for ϕCh1 and outside the cloned fragment. Therefore, all 10 candidates seem to be positive for a double crossover. In the following chapter, the absence of ORF59, encoding the PCNA protein, was tested by PCR and Southern blot hybridization.

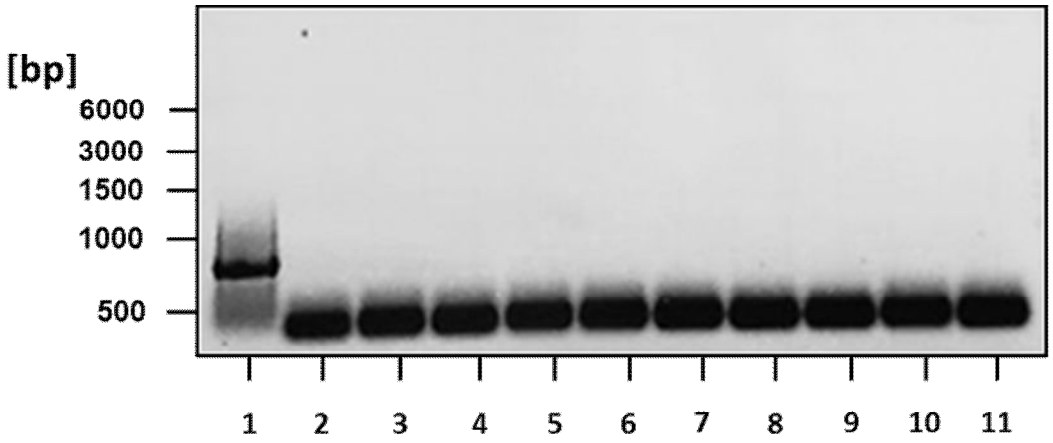


**Figure 20: ORF59 in wild type and mutant strains of *N. magadii* L11.** *N. magadii* L11 and the 10 different mutant strains of *N. magadii* L11-ΔPCNA were used as templates in a PCR reaction with primers p30-5 and p30-3 resulting in a 764 bp fragment of ORF59. Lane 1: *N. magadii* L11, lanes 2 – 11: 10 different mutant strains of *N. magadii* L11-ΔPCNA. The fragment sizes of the used marker are indicated on the left.

In order to amplify ORF59, encoding the PCNA protein, a PCR reaction with primers p30-5 and p30-3 was performed. Thereby, the complete coding region for PCNA can be amplified. As shown for *N. magadii* L11, a fragment of 764 bp could be amplified (Fig. 20, lane 1). However, no signal was detected for all 10 different mutant strains, generated within this word. Therefore, all used mutants were deleted for ORF59.

### 3.1.4 Confirmation with Southern Blot

In order to verify the results of the PCR analyses, mutant strains of ϕCh1-ΔPCNA were analyzed by Southern hybridization. Therefore, the corresponding lysogenic strains of *N. magadii* L11 and *N. magadii* L11-ΔPCNA were inoculated in rich medium and incubated at 37°C until lysis occurred. Virus particles were subsequently isolated and purified using different CsCl step gradients (see Materials and Methods for details). DNA was isolated and restricted with *Bam*HI. The samples were separated on a 0.8% agarose gel (Fig. 21, lanes 1 and 2).

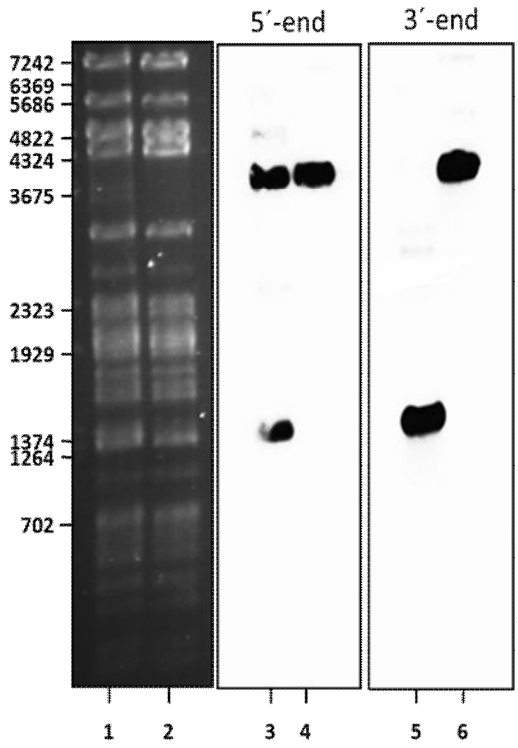


**Figure 21: Southern blot of ΦCh1 wt and ΦCh1-ΔPCNA.** DNA was isolated from virus particles and digested with *Bam*HI. The samples were separated on a 0.8% agarose gel, transferred to a nylon membrane and used for hybridization using a 5´and a 3´probe. Lanes 1, 3, 5: ϕCh1 wt, lanes 2, 4 and 6: ϕCh1-ΔPCNA. Lanes 1 and 2: agarose gel stained with ethidium bromide, lanes 3 and 4: hybridization with the 5´probe, lanes 5 and 6: hybridization with the 3´probe. The sizes of the used DNA marker are indicated at the left.

As a probe the upstream region of ORF59, encoding the PCNA protein, was used. The fragment was amplified using primers ΔPCNA-1 / ΔPCNA-2 and ϕCh1 DNA as a template. The reaction was supplemented with biotin-dUTP, which can be recognized using streptavidin coupled with

horse radish peroxidase. For the 3´-end the upstream sequence of ORF59 was amplified with primers ΔPCNA-3 / ΔPCNA-5, resulting in a 649 bp fragment. The localization of the different fragments is indicated in Fig. 22. The hybridization was performed with restricted DNA isolated from ϕCh1 wt and ϕCh1-ΔPCNA (Fig. 21).

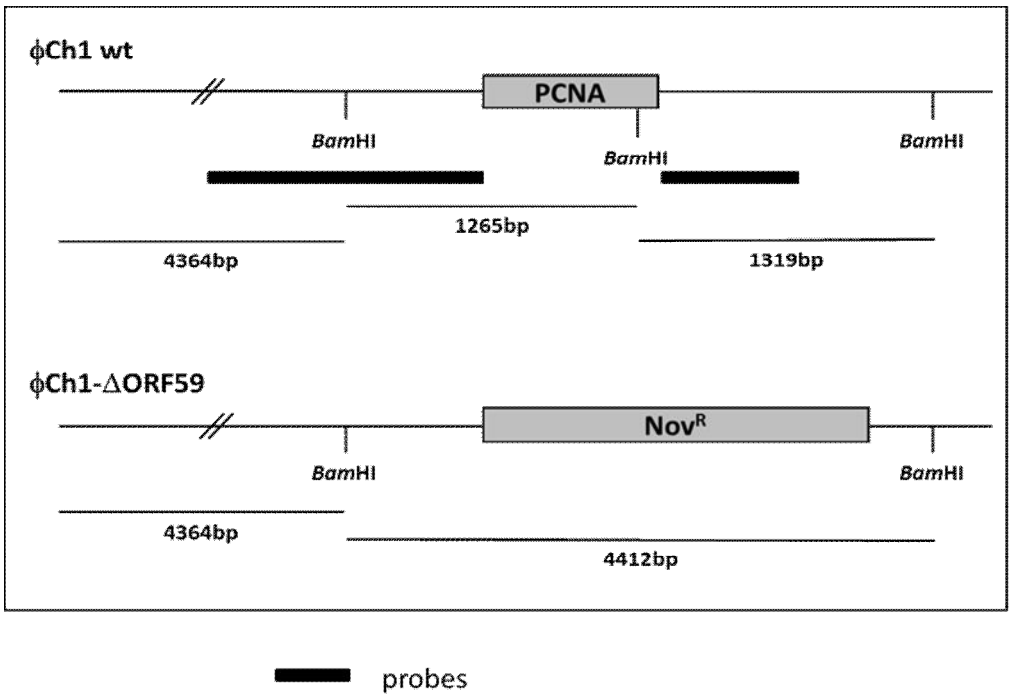


**Figure 22: Schematically drawing of ORF59 (PCNA) region of ϕCh1.** The gene of ORF59 (PCNA) as well as the novobiocin resistance cassette ($Nov^R$) is given as grey boxes. *Bam*HI restriction sites are indicated. The different fragments are marked at the bottom of both areas including the length of the fragments. Probes are indicated as black boxes only for the wild type. Drawing is not to scale.

As can be seen in Fig. 21, the expected signals for the wild type DNA of ϕCh1 (4364 bp, 1265 bp) using the 5´-probe are visible (lane 3) as well as a signal for the DNA of ϕCh1-ΔPCNA (4364 and 4412 bp) (lane 4). Due to the small difference between the two signals, only one is detectable. The hybridization signal for the 3´-probe changes from 1319 bp (ϕCh1 wt) to 4412 bp (ϕCh1-ΔPCNA) due to the loss of one *Bam*HI restriction site based on the exchange of ORF59 with the novobiocin resistance cassette (Fig. 21, lane 5 and 6). Taken these results together, the mutant of ϕCh1-ΔPCNA was verified to be correct concerning the deletion of ORF59 and the recombination sites.

Therefore, this clone was termed ϕCh1-ΔPCNA$_{1_5}$ and used for further studies to elucidate the function of PCNA, encoded by ORF59.

### 3.1.5 Construction of ϕCh1-ΔPCNA$_{1_4}$.

ORF59 is embedded in a region of ϕCh1 characterized by overlapping genes and ORFs encoded by the (+) and (-) strand. Therefore, the recombination sites chosen for clone ϕCh1-ΔPCNA$_{1\text{-}5}$ could interfere with other or yet unknown genes. In order to overcome this problem, a second construct for the deletion of ORF59 was created. Here the 3´end of the recombination fragment was changed and a shorter fragment was used (for details see Material and Methods). This plasmid, was transformed in *N. magadii* L11 and colonies were selected on novobiocin. Single colonies were incubated in rich medium and pools of 10 different cultures were analyzed by PCR for the presence of the deletion cassette (data not shown). As for *N. magadii* L11-ΔPCNA$_{1\text{-}5}$ a positive signal could be observed and the single colonies of such a pool were analyzed, again using PCR (Fig. 23).

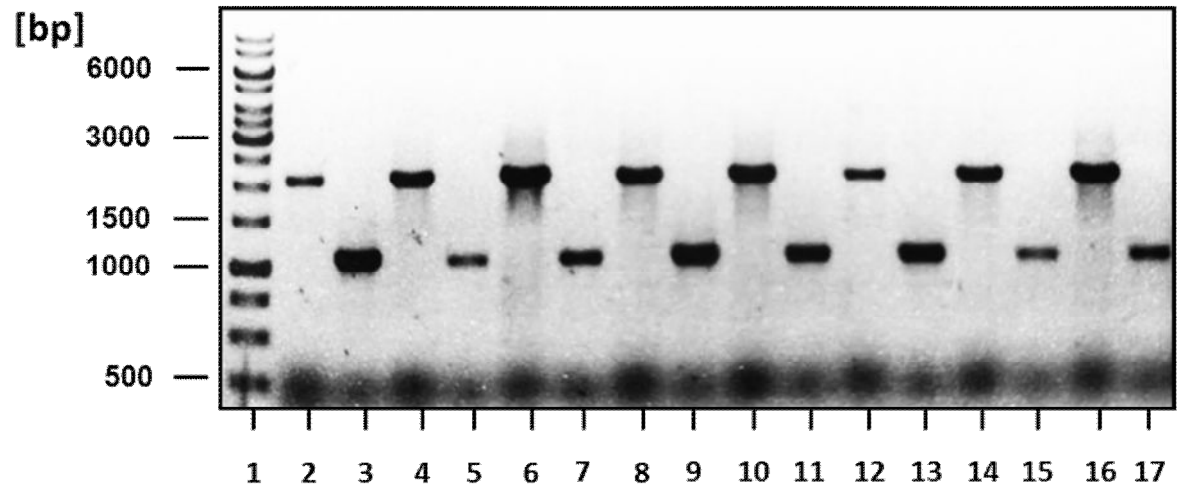


**Figure 23: Screening for double-crossover**. Lane 1: 1-kb-ladder, lane 2-17: different clones of *N. magadii* L11-ΔPCNA. Lanes 2, 4, 6, 8, 10, 12, 14, and 16 represent clones 1 – 10 using primers 56-5 and Nov-13 to verify the 5´-crossover, lanes 3, 5, 7, 9, 11, 13, 15, and 17 represent the same clones using primers Nov-12 and PCNA-3en to confirm the 3´-crossover. All clones showed the expected fragments with 1992 bp for the 5´-crossover and 1195 bp for the 3´-crossover. The fragment sizes of the used marker are indicated on the left.

One positive culture was grown until lysis occurred and the supernatant was used to infect *N. magadii* L13 in order to homogenize the deleted copy of ϕCh1 (for details see Material and Methods). The infected cells were spread on agar plates containing novobiocin for selection. Single colonies of these plates were inoculated in rich medium and analyzed by PCR as described for mutant strain *N. magadii* L11-ΔPCNA$_{1\text{-}5}$.

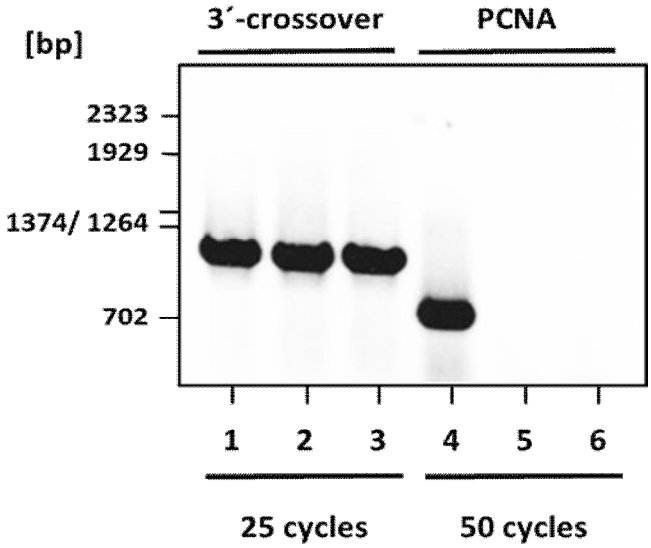


**Figure 24: Analysis of homozygote cultures of *N. magadii* L11 concerning the loss of ORF59.** 3 different mutant strains of *N. magadii* L11-ΔPCNA were used as templates in a PCR reaction with primersΔPCNA-3 and PCNA-3en to identify cells being infected with ϕCh1-ΔPCNA (lanes 1 to 3) and with primers p30-5 and p30-3 to ensure the loss of ORF59 (lanes 4 to 6). Lanes 1, 4: culture no. 1, lanes 2 and 5: culture no. 2, lanes 3 and 6: culture no. 3. The PCR using primers ΔPCNA-3 and Nov-12 resulted in a 1065 bp fragment and with primers p30-5 and p30-3 in a 764 bp fragment of ORF59. To verify the 3´-crossover, 25 cycles were used, for the deletion of ORF59 50 cycles, as indicated at the bottom. The fragment sizes of the used marker are indicated on the left.

All 3 cultures tested showed a signal using PCR for the 3´-crossover (Fig. 24, lanes 1 to 3). Therefore, it can be expected that all of these strains contain the mutated ϕCh1 virus. In addition, the deletion of ORF59 was analyzed using primers p30-5 and p30-3, specific for ORF59. However, as shown in Fig. 24, lanes 4 to 6, only two cultures did not show a signal for ORF59 (no. 2 and 3), whereas culture no. 1 contain a provirus with ORF59 and a copy containing the mutated one, detected with the primers specific for the 3´-crossover. Cells containing a ϕCh1

provirus cannot be infected by a second ϕCh1 particle (8). The occurrence of both the wild type and the mutated strain of ϕCh1 would imply an infection of both forms simultaneously. Therefore, strain no. 1 was discarded and strains no. 2 and 3 were used for further analyses.

In order to verify the correct recombination of the novobiocin resistance cassette into the ϕCh1 genome, virus DNA was isolated from purified particles and restricted with BamHI (Fig. 25, lanes 1 and 2). The restricted DNA of ϕCh1 wild type DNA as well as DNA from ϕCh1-ΔPCNA$_{1\text{-}4}$ was transferred to a nylon membrane. As probes for the 5´end and the 3´end the cloned up- and downstream sequences were used (Fig. 25, lanes 3 – 6).

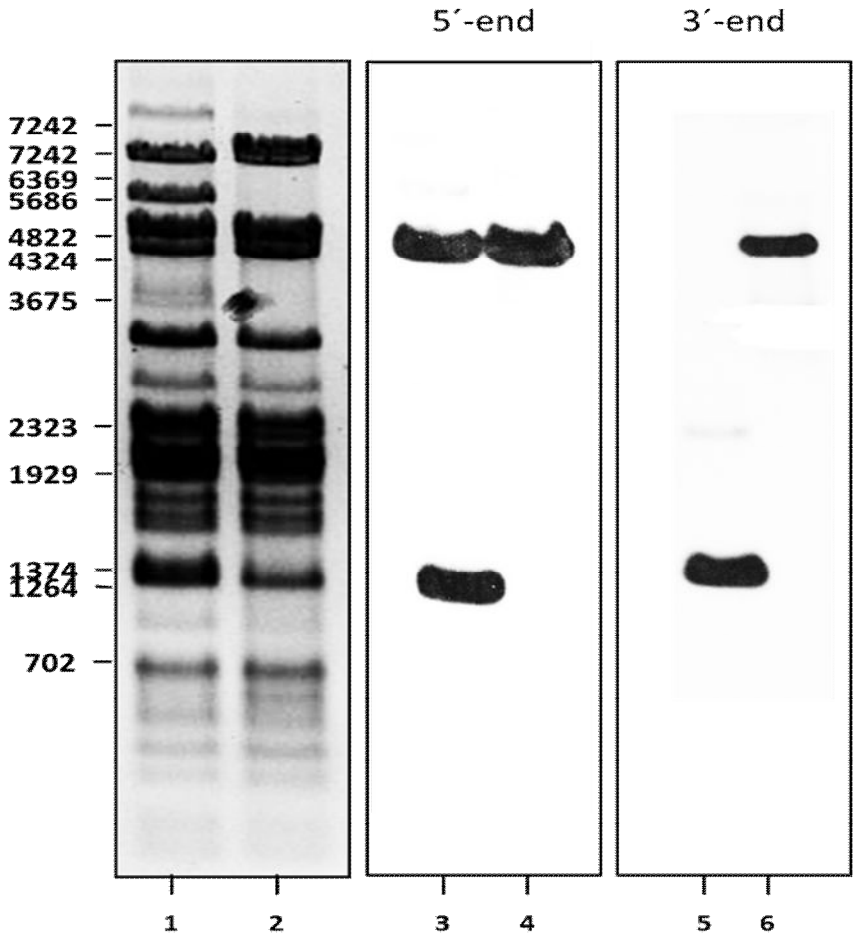


**Figure 25: Southern blot of ΦCh1 wt and ΦCh1-ΔPCNA.** DNA was isolated from virus particles and digested with *Bam*HI. The samples were separated on a 0.8% agarose gel, transferred to a nylon membrane and used for hybridization using a 5´and a 3´probe. Lanes 1, 3, 5: ϕCh1 wt, lanes 2, 4 and 6: ϕCh1-ΔPCNA. Lanes 1 and 2: agarose gel stained with ethidium bromide, lanes 3 and 4: hybridization with the 5´probe, lanes 5 and 6: hybridization with the 3´probe. The sizes of the used DNA marker are indicated on the left.

As shown in Fig. 25, the anticipated signals for the wild type DNA of ϕCh1 (4364 bp, 1265 bp) using the 5´-probe are visible (lane 3) as well as a signal for the DNA of ϕCh1-ΔPCNA (4364 and 4412 bp) (lane 4). Due to the small variation between the two signals, only one is veritably. The hybridization signal for the 3´-probe changes from 1312 bp (ϕCh1 wt) to 4412 bp (ϕCh1-ΔPCNA) due to the loss of one *Bam*HI restriction site based on the exchange of ORF59 with the novobiocin resistance cassette (Fig. 25, lane 5 and 6). Taken these results together, the second mutant of ϕCh1-ΔPCNA was verified to be correct concerning the deletion of ORF59.

Both mutant strains were used to analyze the phenotype of this deletion by determine the lysis time point as well as the expression of ORF59.

### 3.1.6 Phenotypically characterization of the ϕCh1-ΔPCNA mutants.

Growth and lysis characteristics were first analyzed in order to elucidate an influence of the mutation on the development of ϕCh1. Therefore, all three strains, *N. magadii* L11, *N. magadii* L11-ΔPCNA$_{1_4}$ and *N. magadii* L11-ΔPCNA$_{1_5}$ were inoculated in rich medium with an optical density ($OD_{600}$) of 0.1 and incubated at 37°C with agitation. Every day the $OD_{600}$ was measured.

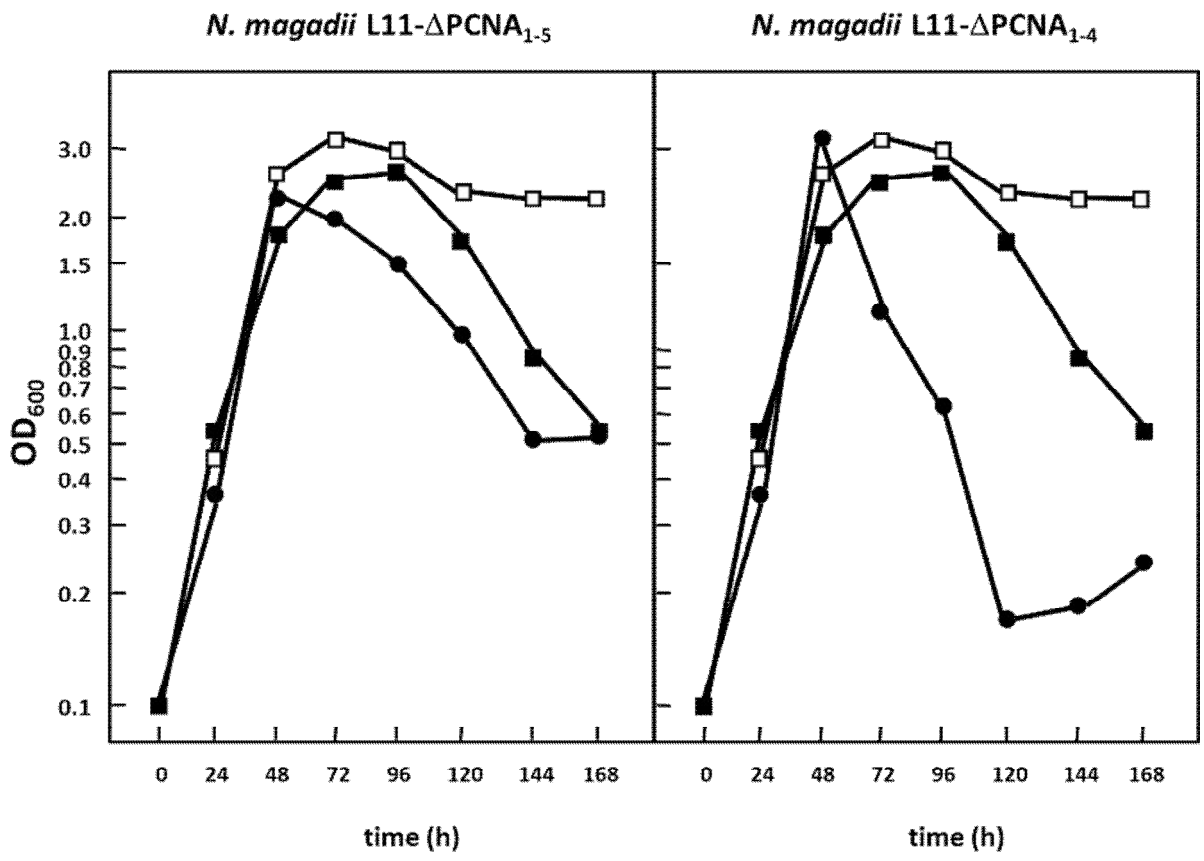


**Figure 26: Growth and lysis of the lysogenic strains *N. magadii* L11, *N. magadii* L11-ΔPCNA$_{1_4}$ and *N. magadii* L11-ΔPCNA$_{1_5}$ as well as the cured strain *N. magadii* L13.** *N. magadii* L11 (■), *N. magadii* L11-ΔPCNA$_{1_4}$ (●), *N. magadii* L11-ΔPCNA$_{1_5}$ (●) and *N. magadii* L13 (□) were grown in rich medium at 37°C with agitation. The optical density was measured at 600 nm (OD600). Left panel: *N. magadii* L11-ΔPCNA$_{1_5}$, right panel: *N. magadii* L11-ΔPCNA$_{1_4}$.

The wild type strain *N. magadii* L11 is characterized by an onset of lysis between day 4 and day 5 after inoculation, meaning the culture has to enter the stationary phase to start lysis (Fig. 26). This has been reported before (8). In contrast to the wild type strain, the mutant strain *N. magadii* L11-ΔPCNA started lysis between day 2 and 3, meaning 48h earlier then the wild type strain (Fig. 26). As discussed before, it was expected that the loss of ORF59 led to a delayed onset of lysis. This was estimated concerning the function of ORF59 encoding a putative PCNA molecule. Therefore, the second gene encoding a PCNA molecule which is located on the genome of ϕCh1 should enhance the production of progeny particles beside the PCNA copy located on the chromosome of *N. magadii*.

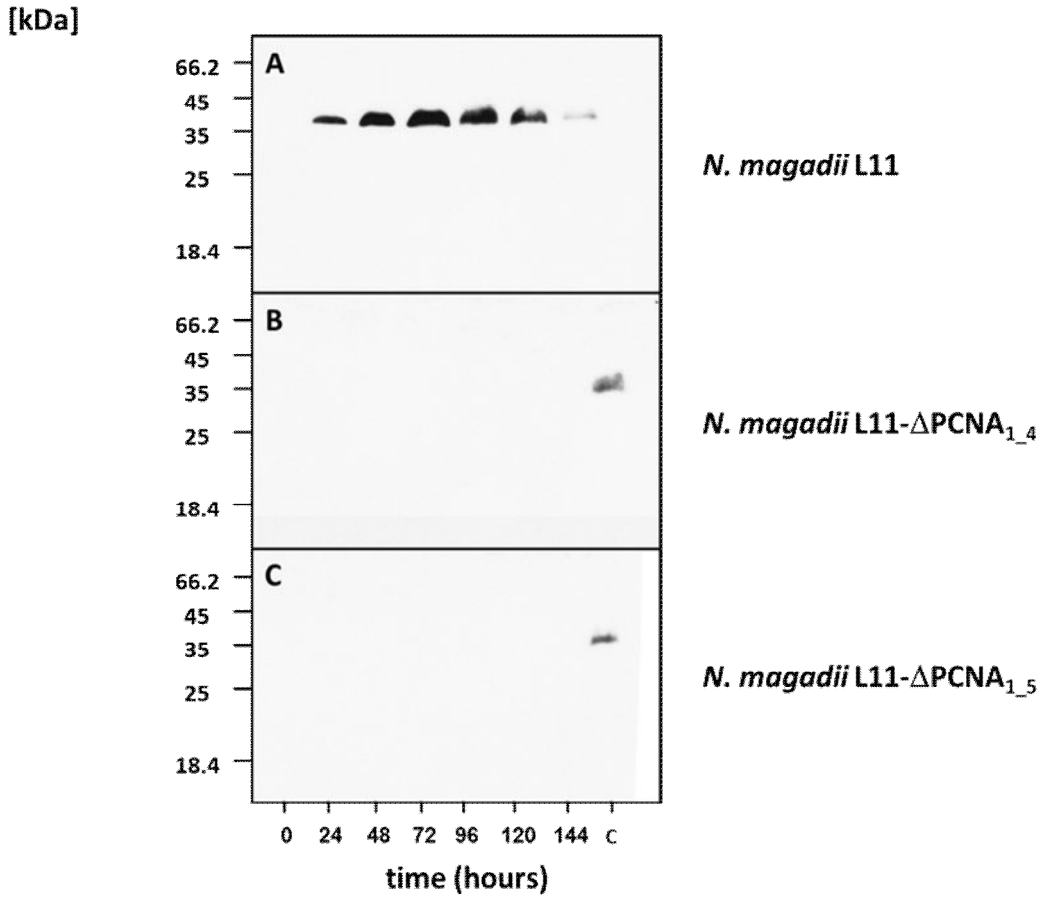


**Figure 27: Expression of ORF59 during the life cycle of the virus ΦCh1.** *N. magadii* L11 as well as the mutant strains were inoculated in rich medium and incubated at 37°C. Crude extracts of *N. magadii* were resolved on a 12% SDS-PAGE and after transferring to nitrocellulose, the PCNA proteins were detected using α-PCNA antibodies. Lane 1: molecular weight marker. B: *Nab. magadii* L11 grown at 37°C (lane 2-9: samples from day 1-7), C: *Nab. magadii* L11 grown at 42°C (lane 2-9: samples from day 1-7), D: *Nab. magadii*

### 3.1.7 Virus particle release of strains *N. magadii* L11 and *N. magadii* L11-ΔPCNA$_{1_4}$.

In order to analyze the consequences that arouse from the earlier onset of lysis of strain *N. magadii* L11-ΔPCNA$_{1_4}$ compared to the wild type strain *N. magadii* L11 the virus titer of supernatants of samples taken at the end of the lysis was determined (Fig. 28).

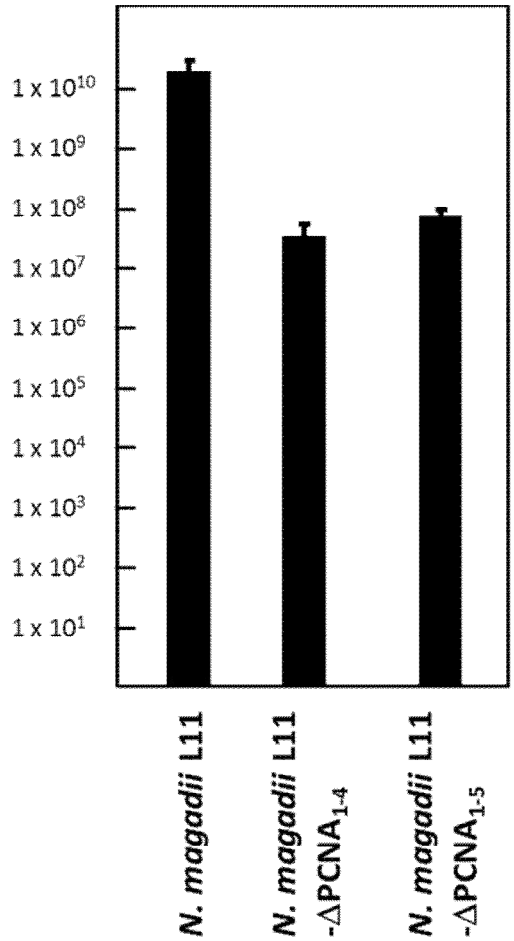


**Figure 28: Virus release of strains *N. magadii* L11, *N. magadii* L11-ΔPCNA1$_{_4}$ and *N. magadii* L11-ΔPCNA$_{1_5}$.** Samples were taken 7 days after inoculation of the strains in rich medium. Arrow bars are indicated

Samples were taken from both cultures at day 7 after inoculation. The supernatants were used to determine the virus titer by plating appropriate dilutions on *N. magadii* L13 agar lawns. *N. magadii* L13 is the cured strain which can be infected by ϕCh1 (8). Plates were incubated for 10 days at 37°C and the titers were calculated. As can be seen in Fig. 28, the virus release of is reduced by a magnitude of 3 for both strains.

### 3.1.8 Expression of ORF11 in *N. magadii* L11 and *N. magadii* L11-ΔPCNA$_{1-4}$

The development of virus particles in strain *N. magadii* L11-ΔPCNA$_{1-4}$ was reduced by 3 magnitudes of order (see above). This would imply a reduction within the replication machinery implying the function of the encoded protein of ORF59 as a PCNA molecule. In order to analyze the implications of this mutation, the expression of the major capsid protein, encoded by ORF11, was studied using western blots.

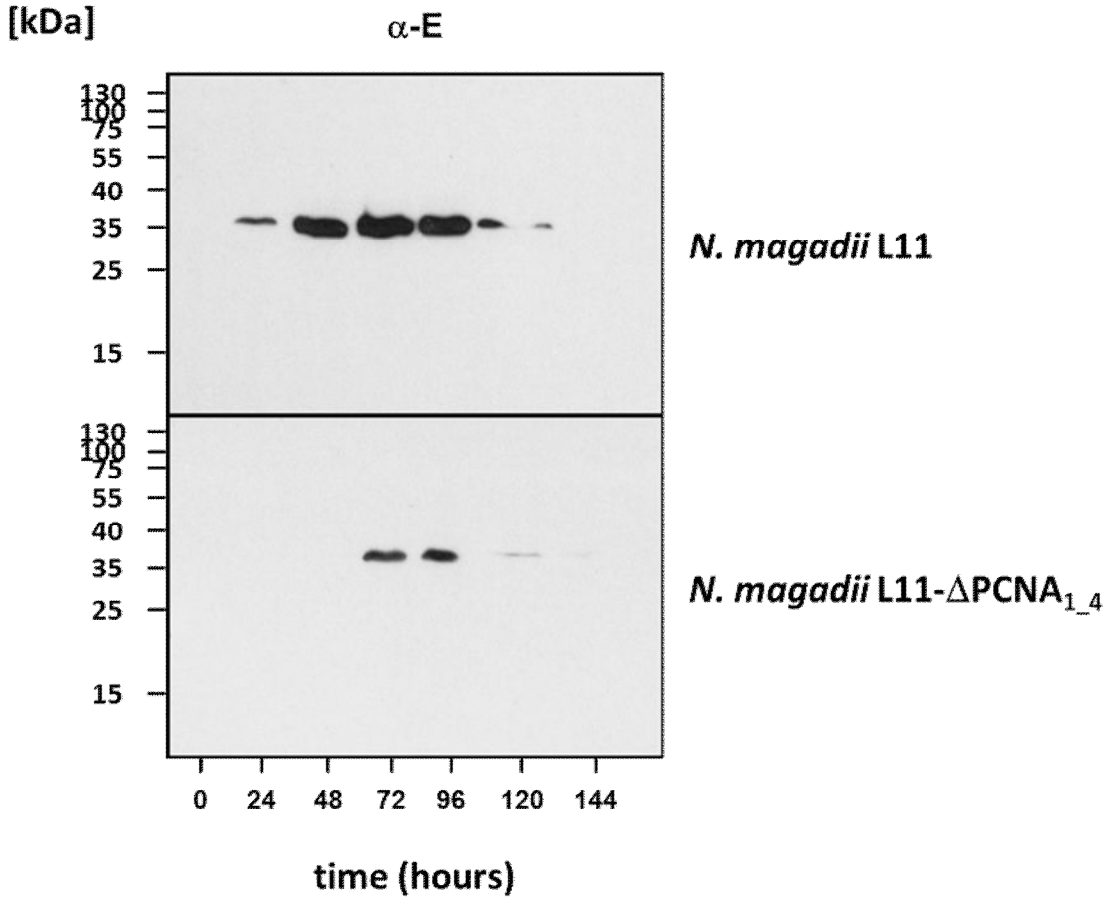


**Figure 29: Expression of ORF11 in *N. magadii* L11 and *N. magadii* L11-ΔPCNA$_{1-4}$.** Samples were taken from strains *N. magadii* L11 (upper panel) and *N. magadii* L11-ΔPCNA$_{1_4}$ (lower panel) at different time points (see Fig. 27 for details). Crude extracts were prepared from these probes according to materials and methods and applied to a 12% SDS-PAGE. After separation, proteins were transferred to a nitrocellulose membrane and incubated with a α-E-antiserum. Detection of protein E was performed as mentioned in materials and methods. The concentrations of the different samples were adjusted using an SDS-PAGE stained with Coomassie.

The expression of ORF11 started 24 h after inoculation of the wild type strain *N. magadii* L11 and increased 24 h later. The level of expression stayed constant the next 48 h and resulted in a decreased status as lysis occurred. This is consistently with the reported expression of ORF11 (40). However, the overall expression of ORF 11 in the mutant strain *N. magadii* L11-ΔPCNA$_{1_4}$ is reduced in general (Fig. 29, lower panel). The first signal could be detected 72 h after inoculation, which means, 48 h later than in the wild type strain. Moreover, the general signal is much brighter than the signal of the wild type strain. In general expression of ORF11 within strain *N. magadii* L11-ΔPCNA$_{1_4}$ is significant reduced compared to strain *N. magadii* L11 reflecting the lower particle number released by the mutant strain.

### 3.1.9 Discussion

PCNA in general is the processivity factor of Polδ and therefore is the functional homologue of other processivity factors, the β subunit of the *E. coli* DNA polymerase III holoenzyme and the product of gene-45 of bacteriophage-T4. Similarities in the function of these proteins gave the first indication of the structure of PCNA.

PCNA has been described as a `sliding clamp'. The first evidence for the sliding clamp structure came from the study of the β subunit of *E. coli* Pol III. In a series of experiments, it was shown that the β subunit bound tightly to nicked circular plasmid but readily dissociated upon linearization of the plasmid via sliding over the ends (74).

The genome of ϕCh1 itself codes for a PCNA gene, ORF59 (40). A second copy of this gene function is located on the chromosome of *N. magadii* (76). Both proteins encoded by these genes have a sequence identity of 65% over 247 AA, as determined using BLAST analysis (Fig. 30).

```
1 ----------------------MFKAIVSAETLTSALDSISVLVDECKIHLEEDGLEIRAVDPANVGMVDLSLDAAAFESY
2 MRYRPGPHDGERSLNQSETTPMFKAIATKSELESFTDPISQLVEECKVNLNEDGLHVRAVDPANVGMVESNAHASGFESY

1 EADGGLIGVDLSRLEDIAGMAESGQLIQFELDEETRKLHIQIDGLEYTLALIDPDSIRQEPDIPELDLPAEVVLEGKDVN
2 EADGGLIGLNVDRFEDVIGMANSGDLVHLELDEETRKLAIQIEGMEFTLALIDPDSIRAEPDIPDLDLTSEIVLEGRHID

1 RSVTAADMVSDHIALGVNDGDDSFYVDAEGDTDDVHLELTEADLIDLQAGPAHSLFSLDYLKDMNKAIPSNTEVTLQLGE
2 RGIKAADMVSDHITLAVDEKEELFEIEAEGDTDDVHIGLERDDLIDLQVGPARSLFSLDYLKDMSKAIDAGDEVTIELGE

1 EFPVKIYFGFAEGQGQVTYMLAPRIQSD
2 EFPVKLAFTTEEGNVDVQYMLSPRIQSN
```

**Figure 30: Protein sequence alignment of PCNA encoded by *N. magadii* and ΦCh1.** One top the sequence of the chromosomal encoded gene pcn1 Nmag_2041(1) is given, below the sequence of ORF59 of ϕCh1 (2). Identical AA are indicated in red, non-identical in black.

As shown in Fig. 30, the N-terminus of the virus encoded protein seems to be prolonged compared to the chromosomal encoded once. Evidence for the longer version of ORF59 and its encoded protein was taken from the analysis of ORF59 expression in strain *N. magadii* L11 using a different antibody than used throughout this study (Fig. 31). Here, two bands were visible in a western blot using α-PCNA_2 antibodies (75).

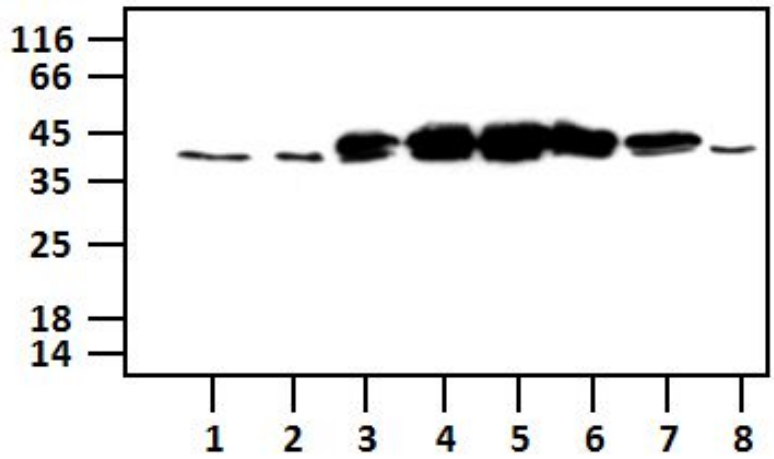


**Figure 31: Expression of ORF59 of ΦCh1 determined by western blot analysis using α-PCNA_2 antibodies (Sabitzer, 2005).** Crude extracts of *N. magadii* L13 (lane 1) and *N. magadii* L11 (lane 2 - 8 taken 24, 48, 72, 96, 120, 144 and 168 h after inoculation) were separated on a 12% SDS-PAGE. Proteins were transferred to a nylon membrane and used for western blot analysis. Size markers are indicated on the left (kDa).

As can be seen in Fig. 31, both copies of PCNA are clearly distinguishable using antibody α-PCNA_2. The annotated sequence encoding the chromosomally encoded PCNA gene has a length of 247 codons given rise to a protein of 26.82 kDa with an isoelectric point (pI) of 3.69 (76). This

gene is quite similar to that of ϕCh1 (247 codons, encoding a protein of 27.18 kDa and a pI of 4.01) (40). However, these two proteins would be hardly separated in a 12% SDS-PAGE assuming a larger protein encoded by ϕCh1. The inspection of the ϕCh1 sequence revealed a possible different start 31 codons upstream of the published one, using GTG as a start codon. This would prolong the protein to a molecular weight of 29.58 kDa and a pI of 4.19.

Being able to distinguish between the two copies of PCNA a deletion mutant of ORF59 was constructed during this master thesis. Here it could be demonstrated that the deletion of ORF59 resulted in an earlier onset of lysis of *N. magadii* L11-ΔPCNA and a reduction in the production of the major capsid protein E resulting in a reduction of progeny virus release of 3 orders of magnitude. Therefore, ORF59 is not essential, however it supports the production of virus particles. Nonetheless the earlier onset of lysis of the mutant strain could also indicate a general influence on regulation of the life cycle of virus ϕCh1. Consequently, the mutant strain *N. magadii* L11-ΔPCNA is an excellent starting point to answer several questions regarding the function of the virus encoded PCNA:

- Complementation of the mutant strain with plasmids carrying the shorter and the longer form of ORF59 should answer the question, which form is able to restore the wild type phenotype.
- In addition, this complementation assay should also answer the question if both the earlier onset of lysis and the reduced production of progeny viruses are based on the deletion of ORF59.
- Does PCNA influence the copy number of ϕCh1 genomes within the cell, being responsible for the reduced synthesis of the major capsid protein E?

Taken everything together this mutant will give a closer look on the replication of ϕCh1 genomes within the lysogenic strain *N. magadii* L11.

# 3.2 Investigation of the function of gp44

## 3.2.1 Aim

The aim of this experiment was to create a homozygous deletion mutant of ORF44 in order to perform further investigation of the function of the gene product gp44. As described above, gp44 is part of a putative toxin-antitoxin system encode by the ORFs 43 and 44. The protein gp44 encoded by ORF44 represents the toxin which showed homology to a VapC PIN domain nuclease. Christoph Hofbauer showed in his Master thesis, that ORF44 has an effect on ORF34 expression. For the experiment three different *N. magadii* L13 strains where constructed. The figure shows the expression of ORF34 in the presence of gp44, gp43 and gp44. Only gp44 has an effect on ORF34. As shown in Fig. 32, strains *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43/44) and *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43) a constant expression of ORF34$_{52}$ using specific antibodies against gp34$_{52}$ was visible (upper and middle panel, marked with ORF43/44 and ORF43 on the left). In contrast strain *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF44) showed a delayed expression of ORF34 starting 72h after inoculation. In addition, the size of the gp34$_{52}$ signal is reduced from 67 kDa of strains *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43/44) and *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43) to app. 50 kDa. Taken into account the possible function of gp44 as a nuclease, the following scenario could be possible: gp44 specifically cleaves the mRNA of ORF34$_{52}$ resulting in a shorter, but precise protein. The expression of ORF34$_{52}$could only be performed at a time point where gp44 is out competed and therefore a complete inhibition of the production of gp34$_{52}$ is not longer possible.

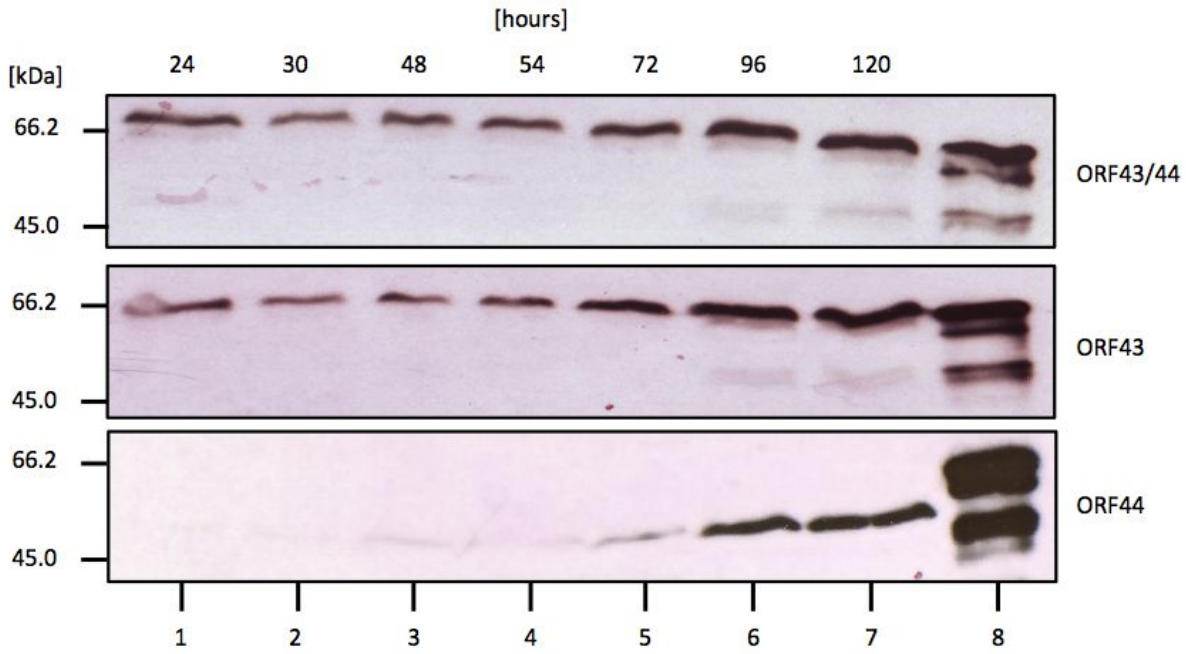


**Figure 32: Western blot analysis of gp34$_{52}$ in the presence of ORF43/44, ORF43 and ORF44.** This western blot shows the expression of ORF34$_{52}$ in the three different *N. magadii* L13 strains. Strain 1: *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43/44) (upper panel), strain 2: *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF43) (middle panel), strain 3: *N. magadii* L13 (pNB102-ORF34$_{52}$, pRo5-p43-ORF44) (lower panel). Samples were taken at different time points and separated on a 12% SDS-PAGE (lanes 1-7). As a quantification control, the crude protein extract of E. coli XL1-Blue (pQE30-ORF34$_{52}$) is shown in lane 8. The signal of pg34$_{52}$ together with gp43/44 or gp43 does not differ from the control. Expression with only gp44 prevents the signal until 48 hours.

### 3.2.2 Experimental setup

For the generation of an ORF44 deletion mutant, the same strategy as for the PCNA experiment was used. The plasmid pBluescript II KS+ included the novobiocin resistance cassette (2,453 bp) which was flanked by the upstream region (480 bp) and the downstream region (659 bp) of ORF44. Again, the novobiocin was cloned in forward and reverse orientation. The plasmids were transformed into *N. magadii* L11 multiple times, but unfortunately no positive clone was found in this master thesis (data not shown)

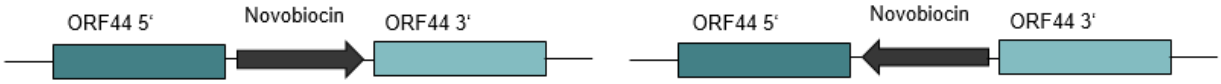


**Figure 33: Schematically presentation of the 2 clones which should be cloned during this thesis.** The novobiocin cassette is indicated by a black arrow, the upstream and downstream sequences of ORF44 with a green or light green box and marked with ORF44 5´and ORF44 3´, respectively. The drawing is not to shale.

### 3.2.3 Conclusion

The cloning of the two suicide plasmids containing the up- and downstream sequences of ORF44 with the novobiocin resistance cassette are a veritable tool to create a deletion mutant of ORF44. The construction of such a mutant will lead to the verification of the function of gp44 in its natural background, meaning embedded in the genome of ϕCh1 integrated into the chromosome of *N. magadii*. Therefore, the consequence of the loss of ORF44 on the development of the virus can be studied in future.

# 4 References


1 **Volker Müller**, Einführung in die Mikrobiologie (WS 04/05), Institut für Molekulare Biowissenschafen, Campus Riedberg

2 **Woese C. R. and Fox G. E**. Phylogenetic structure of the prokaryotic domain: The primary kingdoms. 1977. Proc. Natl. Acad. Sci. **74:** 5088-5090.

3 **Forterre P.** Evolution of the Archaea. 2002. Theoretical Population Biologie. **61**: 409-422.

4 **Pikuta E. V.** Overview of Archaea. 2011. Proceedings of SPIE. 8152

5 **Klingl A.** S-layer and cytoplasmic membrane – exceptions from the typical archaeal cell wall with a focus on double membranes. 2014. Front. Microbiol. **5**: 624. doi: 10.3389/fmicb.2014.00624

6 **Jarrell K. F., Walter A. D., Bochiwal C., Borgia J. M., Dickinson T. and Chong J. P. J.** Major players on the microbial stage: why archaea are important. 2011. *Microbiology* **157**: 919–936

7 **Van de Vossenberg J. L., Driessen A. J. and Konings W. N.** The essence of being extremophilic: the role of the unique archaeal membrane lipids. 1998. *Extremophiles* **2:** 163-170.

8 **Witte A., Baranyi U., Klein R., Sulzner M., Luo C., Wanner G., Krüger D. H. and Lubitz W.** Characterization of *Natronobacterium magadii* phage ϕCh1, a unique archaeal phage containing DNA and RNA. 1997. *Molecular Microbiology.* **23**: 603-616.

9 **Rohwer F. and Barott K.** Viral information. 2012. Biology & Philosophy. **28**: 283-297

10 **Chaib de Mares M**. The evolutionary implications of viral life. 2014. Evo Bites.

11 **Nasir A., Forterre P., Kim M. K. and Caetano-Anolles.** The distribution and impact of viral linages in domains of life. 2014. Front. Microbiol. **5**: 194

12 **Wessner D. R**. The origins of viruses. 2010. Nature Education. **3**: 37

13 **Woese C. R., Kandler O. and Wheelis M. L.** Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. 1990. Proc. Natl. Acad. Sci. **87**: 4576-4579

14 Brock Biology of microorganisms twelfth edtion

15 **Milo R., Phillips R. and Orme N.** Cell Biology by the Numbers Taschenbuch – 2016

16 **Huh, J.W. and Kim, H.S**. SIV and HIV. 2004. AGRIS. **20**: 293-299

17 **Parker J.** Criteria of life: some methods of measuring viability. 1953. American Scientist. **41**: 614-618

18 **Andersson, S. G. E., Zomorodipour A., Andersson J. O., Sicheritz-Poten T., Alsmark U. C. M., Podowski R. M., Näslund A. K., Eriksson A. S., Winkler H. H., and Kurland C. G.** The genome sequence of Rickettsia prowazekii and the origin of mitochondria. 1998. Nature. **396**: 133–143

19 **Lander, E. S. *et al.*** Initial sequencing and analysis of the human genome. 2001. Nature. **409**: 860–921

20 **Ghafourian S., Raftari M., Sadeghifard N. and Sekawi Z.** Toxin-antitoxin Systems: Classification, Biological Function and Application in Biotechnology. 2014. Curr Issues Mol. Biol. **16**: 9-14

21 **Magnuson R. D**. Hypothetical Functions of Toxin-Antitoxin Systems. 2007. Journal of Bacteriology. **189**: 6089-6092

22 **Engelberg-Kulka H., Hazan R. and Amitai S.** mazEF: a chromosomal toxin-antitoxin module that triggers programmed cell death in bacteria. 2005. Journal of Cell Science. **117**: 4327-4332

23 **Wang X. and Wood T. K.** Toxin-Antitoxin Systems Influence Biofilm and Persister Cell Formation and the General Stress Response. 2011. Applied and Environmental Microbiology. **77**: 5577-5583

24 **Zhang Y., Zhang J., Hoeflich K. P., Ikura M., Qing G. and Inouye M.** MazF Cleaves Cellular mRNAs Specifically at ACA to Block Protein Synthesis in Escherichia coli. 2003. Molecular Cell. **12**: 913-923

25 **Vesper O., Amitai S., Belitsky M., Byrgazov K., Kaberdina A. C., Engelber-Kulka H. and Moll I.** Selective Translation of Leaderless mRNAs by Specialized Ribosomes Generated by MazF in Escherichia coli. 2011. Cell. **147**: 147-157

26 **Brock J. E., Pourshahian S., Giliberti J., Limbach P. A. and Janssen G. R.** Ribosomes bind leaderless mRNA in Escherichia coli through recoginition of their 5'-terminal AUG. 2008. RNA. **14**: 2159-2169

27 **Iro M., Galos B., Baranyi U., Rössler N., Witte A.** The lysogenic region of virus ϕCh1: identiication of a repressor-operator system and determination of ist activity in halophilic Archaea. 2007. Extremophiles. **11**: 383-396

28 **Hofbauer C.** The function of gp34 and its regulation by ORF79 of ϕCh1 as well as the influence of other regulation elements. 2015. Master thesis.

29 **McKenzie J. L., Robosn J., Berney M., Smith T. C., Ruthe A., Gardner P. P., Arcus V. L. and Cook G. M.** A VapBC Toxin-Antitoxin Module is a Posttranscriptional

Regulator of Metabolic Flux in Mycobacteria. 2012. Journal of Bacteriology. **194**: 2189-2204

30 **Arcus, V. L., McKenzie J. L., Robson J. and Cook G. M.** The PIN-domain Ribonucleases and the Prokaryotic VapBC Toxin-Antitoxin array. 2011. Protein Eng Des Sel. **24**: 30-40

31 **Maga G. and Hübscher U.** Proliferating cell nuclear antigen (PCNA): a dancer with many partners. 2003. Journal of Cell Science. **116**: 3051-3060

32 **Strzalka W. and Ziemienowicz A.** Proliferating cell nuclear antigen (PCNA): a key factor in DNA replication and cell cycle regulation. 2010. Annals of Botany. **107**: 1127-1140.

33 **O´Donell M., Langston L. and Stillman B.** Principles and Concepts of DNA Replication in Bacteria, Archaea, and Eukarya. 2013. Cold Spring Harb Perspect Biol. **5**: a010108

34 **Kelman, Zvi.** PCNA: structure, functions and interactions. 1997. Oncogene. **14**: 629-640

35 **Williams G. J., Johnson K., Rudolf J., McMahon S. A., Carter L., Oke M., Liu H., Taylor G. L., White M. F. and Naismith J. H.** Structure oft he heterotrimeric PCNA from Sulfolobus solfataricus. 2006. Acta Cryst. **62**: 944-948

36 **Yao N., Turner J., Kelman Z., Stukenberg P. T., Dean F., Shechter D., Pan Z. Q., Hurwitz J. and O'Donnel M.** ***et al.*** Clamp loading, unloading and intrinsic stability oft he PCNA, β and gp45 sliding clamps of human, *E. coli* and T4 replicases. 1996. Genes to Cells **1**: 101-113

37 **Lippke J. A., Gordon L. K., Brash D. E. and Haseltine W. A.** Distribution of UV light-induced damage in a defined sequence of human DNA: Detection of alkaline-sensitive lesions at pyrimidine nucleoside-cytidine sequences. 1981. Proc. Natl. Acad. Sci. **78**: 3388-3392

38 **De Laat W. L., Jaspers N. G. J., Hoeijmakers J. H. J.** Molecular mechanism of nucleotide excision repair. 1999. Genes & Development. **13**: 768-785

39 **Moldovan G. L., Pfander B. and Jentsch S.** PCNA, the Maestro oft he Replication Fork. 2007. Cell. **129**: 665-679

40 **Klein R., Baranyi U., Rässler N., Greineder B., Scholz H and Witte A.** *Natrialba magadii* virus ϕCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. 2002. Molecular Microbilogy. **45**: 851-863

41 **Gribaldo S. and Brochier-Armanet C.** The origin and evolution of Archaea: a state of the art. 2006. Phil. Trans. R. Soc. B. **361**: 1007-1022

42 **Hook S. E., Wright A. D. G. and McBride B. W.** Methanogens: Methane Producers oft he Rumen and Mitigation Strategies. 2010. Archaea, vol. **2010**. 11 pages.

43 **Ochsenreiter T., Selezi D., Quaiser A., Bonch-Osmolovskaya L and Schleper C.** Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. 2003. Envrionmental Microbilogy. **5**: 787-797

44 **Schönfelder K.** Characterization oft he ptnaN promoter and ist use for homologous gene expression in *Natrialba magadii*. 2015. Master thesis

45 **Blöchl E., Rachel R., Burggraf S., Hafenbradl D., Jannasch H. W. and Stetter K. O.** *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113℃. 1997. Extremphiles. **1**: 14-21

46 **Spang A., Hetzenpichler R., Brochier-Armanet C., Rattei R., Tischler P., Spiek E., Streit W., Stahl D. A., Wagner M. and Schleper C.** Distinct genes et in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. 2010. Trends in Microbiology. **18**: 331-340

47 **Brochier-Armanet C., Gribaldo S. and Forterre P.** Spotlight on the Thaumarchaeota. 2012. The ISME Journal. **6**: 227-230

48 **Hatzenpichler R.** Diversity, Physiology, and Niche Differentiation of Ammonia-Oxidizing Archaea. 2012. Applied and Envrionmental Microbiology. **78**: 7501-7510

49 **Elkins J. G., Podar M., Graham D. E., Makarova K. S., Wolf Y., Randau L., Hedlund B. P., Brochler-Armanet C., Kunin V., Anderson I., Lapidus A., Goltsman E., Barry K., Koonin E. V., Hugenholtz P., Kyrpides N., Wanner G., Richardson P., Keller M., and Stetter K. O.** A korarchaeal genome reveals insights into the evolution oft he Archaea. 2008. PNAS. **105**: 8102-8107.

50 **Stetter K. O.** A brief history oft he discovery of hyperthermophilic life. 2013. Biochem. Soc. Trans. **41**: 416-420

51 **Baker B. J., Comolli L. R., Dick G. J., Hauser L.J., Hyatt D., Dill B. D., Land M. L., VerBerkmoes N. C., Hettich R. L. and Banfield J. F.** Enigmatic, ultrasmall, uncultivated Archaea. 2010. PNAS. **107**: 8803-8811

52 **Waters E., Hohn M. J., Ahel I., Graham D. E., Adams M. D., Barnstead M., Beeson K. Y., Bibbs L., Bolanos R., Keller M, Kretz K., Lin X., Mathur E., Ni J., Podar M., Richardson T., Sutton G. G., Simon M., Söll D., Stetter K. O., Short J. M. and Noordewier M.** The genome of Nanoarchaeum equitans: Insights into early archaeal evolution and derived parasitism. 2003. PNAS. **100**: 12948-12988

53 **Wurch L., Giannone R. J., Belisle B. S., Swift C., Utturkar S., Hettich R. L., Reysenbach A. L. and Podar M.** Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. 2016. Nature Communications. **7**: 12115

54 **Albers S. V., van de Vossenberg J. L. C. M., Driessen A. J. M. and Konigs W. N.** Adaptations oft he archaeal cell membrane to heat stress. 2000. Frontiers in Bioscience. **5**: 813-820

55 **Prangishvili D., Forterre P. and Garrett R.** Viruses oft he Archaea: a unifying view. 2006. Natrue Reviews Microbilogy. 4: 837-848

56 **Schleper C. Kubo K. and Zillig W.** The particle SSV1 from the xtremely thermophilic archaeon Sulfolobus is a virus: Demonstration of infectivity and of transfection with viral DNA. 1992. Proc. Natl. Acad. Sci. **89**: 7645-7649

57 **Erdmann S., Scheele U. and Garrett R. A.** AAA ATPase p529 of Acidianus two-tailed virus ATV and host receptor recognition. 2011. Virology. **421**: 61-66

58 **Häring M., Rachel R., Peng X., Garrett R. A. and Prangishvili D.** Viral Diversity in Hot Springs of Pozzuoli, Italy, and Characterization of a Unique Archaeal Virus, Acidianus Bottle-Shaped Virus, from a New Family, the Amullaviridae. 2008. Journal of Virology. **79**: 9904-9911

59 **Arnold P. H., Ziese U. and Zillig W.** SNDV, a Novel Virus oft he Extremely Thermophilic and Acidophilic Archaeon Sulfolobus. 2000. Virology. **272**: 409-416

60 **Prangishvili D., Arnold H. P., Götz D., Ziese U., Holz I., Kristjansson J. K. and Zillig W.** A Novel Virus Family, the Rudiviridae: Structure, Virus-Host Interactions and Genome Variability oft he Sulfolobus Viruses SIRV1 and SIRV2. 1999. Genetics. **152**: 1387-1396

61 **Vestergraad G., Häring M., Peng X., Rachel R., Garrett R. A. and Prangishvili D.** A novel rudivirus, ARV1, oft he hyperthermophilic archaeal genus Acidianus. 2008. Virology. **336**: 83-92

62 **Häring M., Peng X., Brügger K., Rachel R., Stetter K. O., Garrett R. A. and Prangishvili D.** Morphology and genome organization oft he virus PSV oft he hyperthermophilic archaeal genera Pyrobaculum and Thermoproteus: a novel virus family, the Globuloviridae. 2004. Virology. **323**: 233-242

63 **Klein R., Baranyi U., Rösller N., Greineder B., Scholz H. and Witte A.** *Natrialba magadii* virus ϕCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. 2002. Molecular Microbiology. **45**: 851-863

64 **Kogler A.** Characterization of the fdx promoter for establishment of an expression system as well as the construction and analysis of a pibD deletion mutant of *Natrialba magadii*. 2015. Master thesis.

65 **Rössler N., Klein R., Scholz H. and Witte A.** Inversion within the haloalkaliphilic virus ϕCh1 DNA results in differential expression of structural proteins. 2004. Molecular Microbiology. **52**: 413-426

66 **Letarov A. V. and Kirsch H. M**. The episodic evolution of fibritin: traces of ancient global environmental alterations may remain in the genomes of T4-like phages. 2013. Evology and Evolution. **3**: 3628-3635

67 **Klein R., Rössler N., Iro M., Scholz H. and Witte A.** Haloarchaeal myovirus ϕCh1 harbours a phase variation system for the production of protein variants with distinct cell surface adhesion specificities. 2012. Molecular Microbiology. **83**: 137-150

68 **Mochizukli T., Krupovic M., Pehau-Amaudet G., Sako Y. and Forterre P**. Archaeal virus with exceptional virion architectrue and the largest single-stranded DNA genome. 2012. PNAS. **109**: 13386-13391

69 **Abedon S. T. and Murray K. L**. Archaeal Viruses, Not Archaeal Phages: An Archeaological Dig. 2013. Archaea. **2013**: 10 pages

70 **Happonen L. J., Redder P., Peng X., Reigstad L. J., Prangishvili D. and Butcher S. J.** Familial Reletaionships in Hyperthermo- and Acidophilic Archaeal Viruses. 2010. Journal of Virology. **84**: 4747-4754

71 **Allers T. and Mevarech M**. Origins of eukaryotic cells – membrane lipids. 2010. Nature Reviews Genetics. **6**: 58-73

72 **Albers S. V. and Meyer B. H.** The archaeal cell encelope. 2011. Nature Reviews Microbiology. **9**: 414-426

73 **Iro M.** Characterisation of the regulatory region of the virus ϕCh1 infecting *Natrialba magadii*. *Dissertation.* 2006.

74 **Kelman Z and O'Donnell M**. DNA polymerase III holoenzyme: structure and function of a chromosomal replicating machine. 1995. Annu. Rev. Biochem., **64**, 171 - 200.

75 **Sabitzer, S.** PCNA of ϕCh1 and related proteins in halophilic Archaea. Diploma thesis, University of Vienna. 2005.

76 **Siddaramappa S, Challacombe J. F. , Decastro R. E., Pfeiffer F. , Sastre D. E., Giménez M.I., Paggi R. A., Detter J. C., Davenport K. W., Goodwin L. A., Kyrpides N., Tapia R., Pitluck S., Lucas S., Woyke T., Maupin-Furlow J. A**. 2012. A comparative genomics perspective on the genetic content of the alkaliphilic haloarchaeon *Natrialba magadii* ATCC 43099T. BMC Genomics doi: 10.1186/1471-2164-13-165.

# 5 List of Figures

# 6 Abstract

The first part of this thesis focuses on the construction of a ϕCh1::PCNA mutant and its effects on the development of the virus ϕCh1. The proliferating cell nuclear antigen (PCNA) is a processivity factor for Polδ, which is a homologue of the β subunit of the *E. coli* DNA polymerase III holoenzyme. PCNA is encoded by the virus ϕCh1 and also by its only known host *N. magadii.* Both proteins showed to have a sequence identity of 65% over 247 AA, which was determined by BLAST analysis. The fact, that the virus encoded PCNA is 21 AA larger than the one of *N. magadii* makes it possible to distinguish between these two proteins. Why the virus ϕCh1 is not using the PCNA of the host and instead encoding for its own copy is not known yet. It might be possible, that the virus encoded form has a greater affinity to its target. The effects on the virus development upon deleting the ϕCh1 PCNA gene were investigated. In this master thesis it could be demonstrated, that the deletion of ORF59 resulted in an earlier onset of lysis of *N. magadii* L11-ΔPCNA and a reduced production of the major capsid protein E, resulting in a reduction of progeny viruses release. Therefore, ORF59 is not essential, however it support the production of virus particles. In addition lysis of the mutated strain occurred earlier, which could indicate to a general influence on the regulation of the life cycle of virus ϕCh1.

The second part focuses on the construction of a mutant with an ORF44 deletion and the function of its gene product gp44. Gp44 is part of a putative toxin-antitoxin system that is encoded by the ORFs 43 and 44, whereas the latter one represents the toxin. Sequence analysis showed, that ORF44 has homology to a VapC PIN nuclease domain. Previous studies demonstrated, that gp44 probably specifically cleaves mRNA of $ORF34_{52}$ which results in a shorter but precise protein. In course of this study, the construction of these deletion mutants could be obtained, which serve as a basis for future studies.

# 7 Zusammenfassung

Der erste Teil dieser Arbeit beschäftigt sich mit der Konstruktion einer ϕCh1::PCNA Mutante und deren Auswirkung auf die Entwicklung des Viruses ϕCh1. Das proliferating cell nuclear antigen (PCNA) ist ein Prozessivitätsfaktor der Polδ, welche ein Homolog der β Untereinheit von dem DNA Polymerase III Holoenzym von *E. coli* ist. PCNA ist sowohl auf der Virus DNA von ϕCh1 als auch auf der DNA von seinem einzig bekanntem Wirt *N. magadii* kodiert. Beide Proteine zeigen eine Sequenzidentität von 65% auf über 247 AA auf, welche mittels BLAST analysiert wurde. Die Tatsache, dass das Virus kodierte PCNA Protein um 21 AA länger ist als das bei *N. magadii*, macht eine Unterscheidung der zwei Varianten möglich. Aber warum der Virus sein eigenes PCNA Protein exprimmiert und nicht die Wirtseigene ist bis jetzt noch nicht bekannt. Der Grund könnte eine erhöhte Affinität der Virus kodierten Variante zu seinen möglichen Ziel-Sequenzen sein. In dieser Masterarbeit konnte gezeigt warden, dass die Deletion von ORF59 eine frühere Lyse des Wirtes *N. magadii* L11-ΔPCNA zur folge hatte. Desweiteren konnte eine verringerte Produktion des Hauptcapsidproteins E festgestellt werden, was zu einer geringeren Freisetzung von Virusnachkommen zur Folge hatte. Dies führt zu der Schlussfolgerung, dass der ORF59 nicht essenziell für den Virus ist, jedoch unterstzützt er die Produktion von Viruspartikeln. Die frühere Lyse des Mutantenstammes kann auf einen Einfluss auf die Regulation des Lebenszyklus von ϕCh1 deutet.

Im zweiten Teil dieser Arbeit, geht es um die Konstruktion eines Mutanten mit einer ORF44 Deletion, sowie die Funktion dessen Genprodukts gp44. Es wird vermutet, dass gp44 ein Teil eines Toxin-Antitoxin Systems ist, welches von ORF43 und ORF44 kodiert wird. Letzterer representiert das Toxin. Sequenzanalysen haben gezeigt, dass ORF44 Homologien zu einer VapC PIN nuclease Domäne aufweist. Vorangegangene Studien haben gezeigt, dass gp44 vermutlich spezifisch die mRNA von $ORF34_{52}$ schneidet, welches ein verkleinertes Protein ergibt. Im Rahmen dieser Arbeit konnte eine Deletionsmutante hergestellt werden, welche als Basis für zukünftige Studien dient.