



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

„Construction of an ORF34 deletion ϕ Ch1 strain and
secretion of putative tail fibre proteins“

Verfasserin

Dimmel Katharina

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt: A 441

Studienrichtung lt. Studienblatt: Diplomstudium Genetik – Mikrobiologie (Stzw) UniStG

Betreut von: Ao. Univ.-Prof. Dipl.-Biol. Dr. Angela Witte

Table of the contents

1. INTRODUCTION	10
1.1 ARCHAEA – THE THIRD DOMAIN OF LIFE.....	10
1.1.1 <i>Diversity and abundance</i>	11
1.1.2 <i>Archaea a mosaic of Bacteria and Eukarya</i>	13
1.1.3 <i>Archaea, Bacteria and Eukarya in contrast</i>	13
1.1.3.1 Structural aspect	13
1.1.3.2 Genetic aspect.....	14
1.1.3.3 Replication process	14
1.1.3.4 Transcription	16
1.1.3.5 Translation	16
1.1.4 <i>Life in extreme environments</i>	17
1.1.4.1 Physically extremes	18
1.1.4.2 Geochemically extremes:	19
1.1.5 <i>Halo-alkaliphilic Archaea and their adaption to extreme environments</i>	20
1.1.5.1 Cell envelope and osmotic adaptation.....	21
1.1.5.2 Haloalkaliphilic adaptation of proteins	22
1.1.5.3 Adaptation of protein secretion systems of haloalkaliphilic organisms.....	23
1.1.5.4 Metabolic aspects of alkaliphilic <i>Haloarchaea</i>	25
1.1.6 <i>Biotechnical applications of (haloalkaliphilic) Archaea</i>	26
1.2 NATRIALBA MAGDII	27
1.2.1 <i>General characteristics of Natrialba magadii</i>	27
1.2.2 <i>Laboratory strains of Natrialba magadii</i>	27
1.2.3 <i>Working with Natrialba magadii</i>	28
1.2.3.1 Transformation of <i>Natrialba magadii</i>	28
1.2.3.2 Vectors and genetic markers for <i>Natrialba magadii</i>	29
1.2.3.3 <i>Natrialba</i> extracellular protease (Nep).....	30
1.3 ARCHAEL VIRUSES	31
1.3.1 <i>The halophage ϕCh1 infecting Natrialba magadii</i>	32

1.3.2	<i>Morphology and structural components of φCh1</i>	32
1.3.3	<i>Life cycle of the temperate φCh1</i>	34
1.3.4	<i>Genomic organization of φCh1</i>	35
1.3.5	<i>The invertible region of the φCh1 genome</i>	37
1.3.6	<i>Aspects of the putative tail fibre proteins</i>	39
1.3.7	<i>Virus – host interaction</i>	41
2.	MATERIALS AND METHODS	45
2.1	MATERIALS:	45
2.1.1	<i>Strains</i>	45
2.1.2	<i>Growth media</i>	45
2.1.3	<i>Additives</i>	47
2.1.4	<i>Vectors</i>	47
2.1.5	<i>Primer</i>	49
2.1.6	<i>DNA and protein markers</i>	50
2.1.6.1	DNA ladders	50
2.1.6.2	Protein ladders	51
2.1.7	<i>Enzymes and corresponding buffers</i>	51
2.1.7.1	Restriction	51
2.1.7.2	PCR	51
2.1.7.3	Other enzymes	52
2.1.8	<i>Antibodies</i>	52
2.1.8.1	Primary antibodies	52
2.1.8.2	Secondary antibodies	52
2.1.9	<i>KITs</i>	53
2.1.10	<i>Solutions and reagents for DNA methods</i>	53
2.1.10.1	Gelelectrophoresis	53
2.1.10.2	Gel extraction from polyacrylamide gels	54
2.1.10.3	Southern blot analysis	54
2.1.11	<i>Solutions and reagents for proteins methods</i>	55
2.1.11.1	SDS-PAGE and western blot analysis	55

2.1.11.2	Protein purification under denaturing conditions (E.coli)	56
2.1.12	<i>Isolation of chromosomal DNA from Nab. magadii</i>	57
2.1.12.1	Extraction of plasmid DNA from <i>Nab. magadii</i>	57
2.1.12.2	Buffers and solution for phage methods	57
2.1.12.3	Precipitation of phage DNA	57
2.1.12.4	Precipitation of phage proteins	57
2.2	METHODS	58
2.2.1	<i>DNA gel electrophoresis</i>	58
2.2.2	<i>Agarose gel electrophoresis</i>	58
2.2.3	<i>6% Polyacrylamide gel electrophoresis</i>	58
2.2.4	<i>Polymerase chain reaction (PCR)</i>	58
2.2.4.1	Preparative PCR	59
2.2.4.2	Analytical PCR	59
2.2.5	<i>CLONING</i>	60
2.2.5.1	DNA purification	60
2.2.5.2	Restriction of DNA	61
2.2.5.3	Fill-in 5' overhangs	61
2.2.5.4	Ligation	61
2.2.5.5	Transformation of E. coli	62
	Generation of CaCl ₂ competent E. coli cells	62
2.2.5.6	Screening and verification of positive clones – Quick apply	62
2.2.5.7	Plasmid preparation from E. coli	63
2.2.6	<i>Cloning strategies</i>	63
2.2.6.1	Deletion of ORF34 (pΔ34 ₅₂ ::Nov ^R)	63
2.2.6.2	Complementation of φCh1-ΔORF34	64
2.2.7	<i>Construction of nproNep–ORF34₁ and ORF34₅₂</i>	64
2.2.8	<i>Methods for Archaea</i>	65
2.2.8.1	Transformation of <i>Natrialba magadii</i> strains	65
2.2.8.2	Screening for positive clones	66
2.2.8.3	Generation and evidence of a homozygous mutant strain	67
2.2.8.4	Isolation of <i>Nab. magadii</i> chromosomal DNA	68

2.2.8.5	Southern blot procedure	68
2.2.9	<i>Virus methods</i>	70
2.2.9.1	Purification of ϕ Ch1 particles.....	70
2.2.9.2	ϕ Ch1 titer determination by soft agar technique	71
2.2.9.3	Plating efficiency of ϕ Ch1 on <i>Nab.magadii</i> exporting different variants of ORF34	71
2.2.9.4	Isolation of virus DNA.....	71
2.2.9.5	Precipitation of viral proteins.....	72
2.2.10	<i>Protein methods</i>	72
2.2.10.1	Preparation of crude protein extracts.....	72
2.2.10.2	Protein extracts from culture supernatants.....	72
2.2.11	<i>His-tagged protein over-expression and purification in E. coli</i>	73
2.2.11.1	His- tagged protein over-expression in E. coli.....	73
2.2.11.2	His-tagged protein purification under denaturing conditions.....	73
2.2.11.3	Crude quantification of purified proteins.....	74
2.2.12	<i>Protein purification from Nab. magadii</i>	74
2.2.13	<i>SDS-PAGE</i>	74
2.2.13.1	Preparation of protein gels	75
2.2.13.2	Coomassie staining.....	76
2.2.13.3	Western blot analysis.....	76
3.	RESULTS AND DISCUSSION	78
3.1	DELETION MUTANT OF ΦCH1ΔORF34₅₂	78
3.1.1	<i>General aspects of gp34₅₂</i>	78
3.1.2	<i>Construction of a homozygous ΦCh1ΔORF34₅₂-2 deletion mutant</i>	80
3.1.3	<i>Homozygation of the Nab. magadii L11 ΔORF34 strain</i>	82
3.1.4	<i>Expression of the ORF34 gene product in Nab. magadii L11 and L11ΔORF34 strains</i>	83
3.1.5	<i>Growth kinetics of Nab. magadii L11 wild type and L11ΔORF34 strains</i>	85
3.1.6	<i>Expression of the major capsid protein E</i>	86
3.1.7	<i>Complementation of ΦCh1 lacking the ORF34</i>	86
3.1.7.1	Retransformation of pNB102-ORF341 and pNB102-ORF3452 into <i>Nab. magadii</i>	87
3.1.7.2	Quick test appliance for complementation of Φ Ch1 lacking the ORF34	87

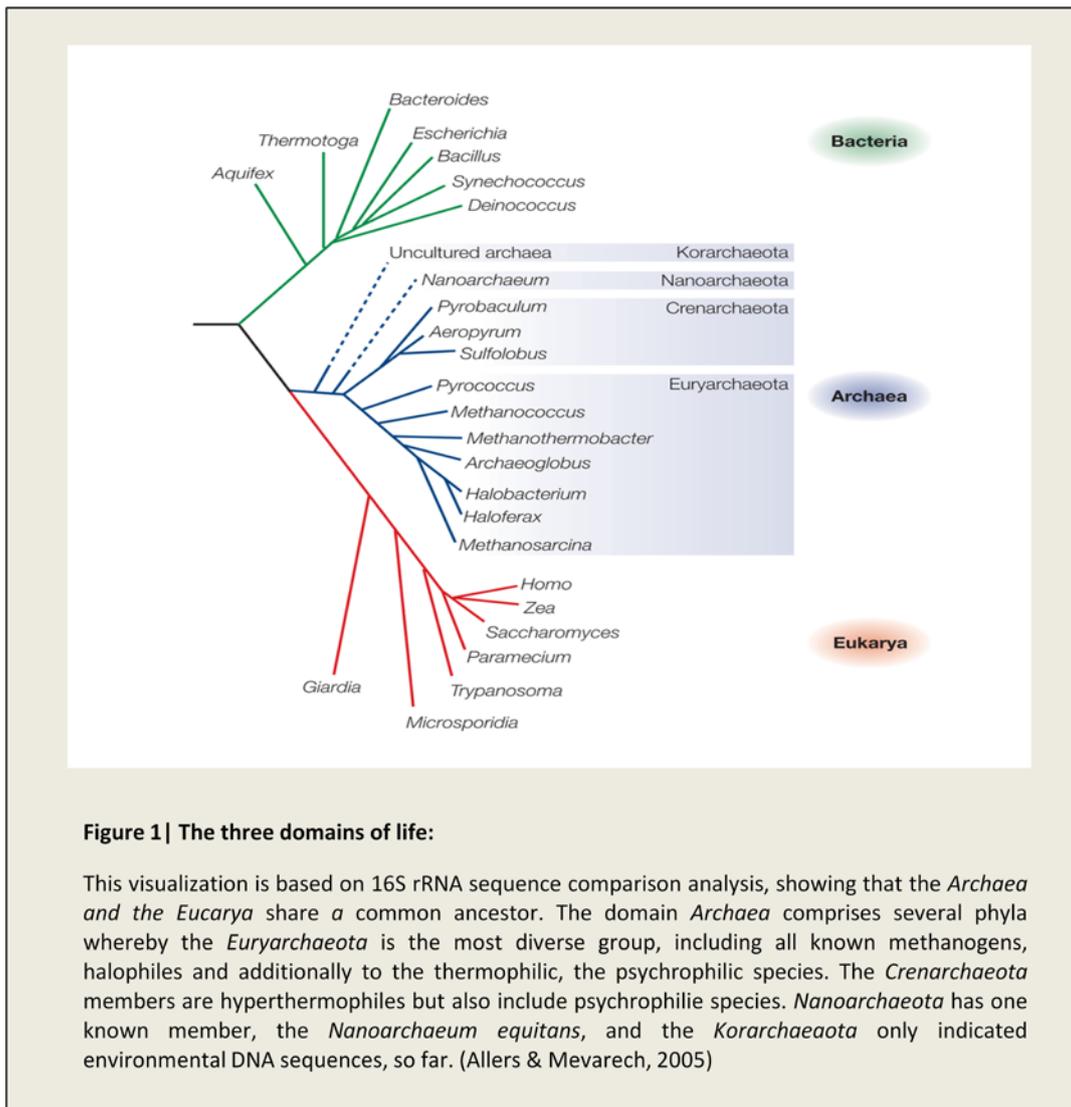
3.1.7.3	Phage titer evaluation	88
3.1.8	Discussion	88
3.2	SECRETION OF RECOMBINANT PROTEINS TO PREVENT ADHESION OF ΦCh1	90
3.2.1	Essential preconditions for exporting recombinant proteins from haloalkaliphilic organism	90
3.2.2	Expression of ORF34 variants in <i>Nab. magadii</i> P3 and verification of secretion	91
3.2.3	Infectivity analysis of Φ Ch1 on gp34 secreting <i>Nab. magadii</i> strains	93
3.2.4	Discussion	93
4.	REFERENCES	97
5.	APPENDIX	105
I.	Index of figures and tables	105
II.	Acknowledgement	106
III.	Abstract	107
IV.	Zusammenfassung	109
V.	Curriculum Vitae	111

1. Introduction

1.1 *Archaea* – the third domain of life

Several billion years ago, the *Archaea* evolved as one of the three primary ancestries and were first described in scientific literatures about 130 years ago, nevertheless the proposal that *Archaea* is the third domain of life was discovered only 20 years ago. The establishment of a meaningful phylogenetic relationship of the living world has created a long debate process in history. In 1866 the phylogenetic tree of Ernst Haeckel had three main branches, evolving from a common origin which represents animals, plants and protists. Prior to this only the dichotomy of animal and plant kingdoms existed. Later Herbert F. Copeland popularized a four-kingdom classification where the bacteria and algae were separated into the kingdom Monera (Prokaryotes). At around the same time, in 1938, Édouard Chatton manifested the theory of two major empires based on the difference between cells without a distinct nucleus, the prokaryotes and those unicellular and multicellular organisms whose cells have a nucleus are called eukaryotes (Chatton, 1938). This view has coexisted with the five-kingdom scheme from Robert Whittaker who created an additional kingdom for the Fungi in 1969. He divided all life forms into Animalia, Plantae, Fungi, Protista (eukaryotes), giving the kingdom Monera (prokaryotes) the same taxonomic rank (Woese *et al.*, 1990). Through the sequencing revolution, Carl R. Woese in 1977 proposed a new system of classifying organisms by analyzing the ribosomal RNA.

“To date, the primary structure of the 16S (18S) ribosomal RNA has been characterized in a moderately large and varied collection of organisms and organelles, and the general phylogenetic structure of the prokaryotic domain is beginning to emerge” (Woese & Fox, 1977). Based on the comparative analysis of molecular sequences a small number of bacteria exhibited significant differences in comparison with other prokaryotes, suggesting that they belong to a distinct domain. Those methanogenic bacteria were tentatively named *archaebacteria* (Woese & Fox, 1977). Further comparative genomics affirmed this theory; in 1990 Woese introduced the “three domain system” dividing the living world in *Eukarya*, *Bacteria* and *Archaea*. The newly established domain *Archaea* shows profound molecular differences to *Eukarya*, so that it also distinguishes from the *Bacteria* which seemed to branch earlier in time as indicated in the picture below (Woese, 1990).



1.1.1 Diversity and abundance

As a result of molecular sequence analysis, which allows comparison of far distant relative organisms and even uncultured microorganisms from environmental studies, a great expanse of phylogenetic diversity was established among the microorganisms. Since the new third domain was discovered further species were included, phylogenetically distinctions into several phyla within the *Archaea* were confirmed and comprised a wide range of variety that had never been predicted. The *Euryarchaeota* which is phenotypically a heterogeneous group comprises of methanogens, extreme halophiles and sulfate-reducing types of hyperthermophiles species. The *Crenarchaeota* contains physiologically relatively homogenous organisms whose niches are entirely thermophilic (Woese *et al.*, 1990). An additional phylum, the *Korarchaeota* was proposed with a pool of sequences isolated from hyperthermophilic

environments, which belong neither to the *Crenarchaeota* nor to *Euryarchaeota* (Barns *et al.*, 1996).

The fourth phylum the *Nanoarchaeota* was proposed when *Nanoarchaeum equitans* was isolated in coculture with a new *Ignicoccus* species (Huber *et al.* 2002; 2003). In 2008 the uncultivated mesophilic *Archaea*, initially classified as *Crenarchaeota* was suggested to form a new phylum with the name *Thaumarchaeota* (Brochier-Armanet *et al.*, 2008). Due to the fact that the first recorded *Archaea* was found living under extreme conditions, biologists even thought that archaeal abundance is predominant over *Bacteria* in extreme environments, like high salt, high or low pH and high pressure but this is only true for high-temperature habitats (Forterre *et al.*, 2002). Whereas in all other environmental conditions *Archaea* are found together with *Bacteria* and *Eukarya* (Rothschild & Mancinelli, 2001). Studies of environmental diversity have shown that *Archaea* are much more widespread as previously thought. They make up about 20% of all microbial cells in the oceans, a habitat where they were not thought to exist 10 years ago (DeLong & Pace, 2001).

This ubiquitous abundance together with diverse functional attributes manifested that those organisms play a crucial importance in global biogeochemical processes. “*The methanogenes are found in anaerobic marine and freshwater environments and in the gastrointestinal tracts of animals, where they participate in the conversion of organic matter by utilizing the metabolic products of bacteria. Haloarchaea reside in hypersaline environments, where they grow as heterotrophs, often in association with phototrophic algae. Thermoacidophiles (including hyperthermophiles, which grow fastest at temperatures above 80°C) colonize volcanic terrestrial environments and deep-sea hydrothermal vents, growing aerobically or anaerobically as heterotrophs or autotrophs*” (Cavicchioli, 2011). *Archaea* and *Bacteria* exhibit a wide range of common metabolic pathways like autotrophy, heterotrophy and photosynthesis; however, only methanogenic organisms have been described within the kingdom *Archaea* (Forterre *et al.*, 2002).

1.1.2 Archaea a mosaic of Bacteria and Eukarya

Genome sequencing analysis have shown that *Archaea* are chimera of bacterial and eukaryotic attributes, whereas the information processing functions mirroring those of *Eukarya* and the core metabolic functions are distinctly bacterial (Allers & Mevarech, 2005). Once *Archaea* were thought to be closely related to *Bacteria*, due to the morphological similarities but now at the molecular level they are considered to be more closely related to the *Eukarya*.

1.1.3 Archaea, Bacteria and Eukarya in contrast

1.1.3.1 Structural aspect

Bacteria and *Archaea* are prokaryotes, typically without a nucleus and complex cell compartments. Similarities concerning morphological characteristics like cell size or shape; this is why they were previously mixed into a common phylogenetic group. Generally archaeal morphologies are more diverse than what is found in *Bacteria*. Accessorily all the metabolic pathways apart from methanogenesis also exist among *Bacteria* (Forterre *et al.*, 2002). Photosynthesis in *Archaea* is based on bacteriorhodopsin not on chlorophyll, which was once thought to be unique to this domain but has been found in planktonic *Bacteria* as well (Béjà *et al.*, 2001).

When distinguishing between *Archaea* and *Bacteria* it is only possible at the molecular level. The archaeal and bacterial ribosomes are quite similar in size and composition, they are composed of 30S (16S molecule) and 50S (23S and 5S molecule) subunits which combine to make a 70S unit, with up to 70 ribosomal proteins depending on the species. However, the primary sequence of archaeal rRNA and ribosomal proteins are closer to those of eukaryotes (Londei, 2010). Beside the specific rRNA *Archaea* are significantly distinguishable from *Bacteria*, as peptidoglycan is lacking in archaeal membrane composition (Forterre *et al.*, 2002). *Archaea* exhibit an enormous diversity of cell envelopes. Instead of bacterial murein archaeal species (e.g. members of *Methanobacteriales*, behaving Gram positive) possess the related pseudopeptidoglycan (pseudomurein) in the cell wall which correlates with the peptidoglycan in function, physiology and morphology. The basic structure components of pseudomurein are N-acetylglucosamine and N-acetyltalosaminuronic acid, which are linked by β -1,3-glycosidic bonds. In contrast to peptidoglycane which consists linkages of β -1,4 N-acetylglucosamine and N-acetylmuramic acid. As a result of the 1-3 linkages *Archaea* are resistant to lysozymes and other bacterial hydrolases (Bullock, 2000). Others like *Halococcus* have rigid cell walls based on heteropolysaccharides but not all archaeal species have cell walls. Species which only possess a

single membrane like *Thermoplasmatales* are covering their cytoplasmic membrane with a surface layer (S-layer) of glycoproteins. An additional peculiarity of archaeal cell envelopes is that most of the extracellular proteins are N- and O-glycosylated. The nature of archaeal glycerolipids in the membrane are ether linkages (2,3-*sn*-glycerol) between glycerol and isoprene side chains. Whereas *Bacteria* and *Eucarya* lipids are ester linkages (1,2-*sn*-glycerol) of glycerol and fatty acids (Ellen 2010, Brown & Doolittle, 1997).

1.1.3.2 Genetic aspect

Archaea and *Bacteria* share similar genomic structures and mechanism of genome duplication, such as fast and bidirectional chromosome replication from a single origin. Metabolic and cell division proteins also resemble those of *Bacteria* (Grabowski & Kelman, 2003). Moreover, *Archaea* are a mosaic of different groups of genes: a part that encodes information processing which complies with eukaryotes in nature, and a lineage that encodes housekeeping functions with bacterial aspects. Comparisons of complete genome sequences of *Archaea* have shown that they are more than a mix of bacterial and eukaryotic parts. As of yet about 50% of archaeal genes have no assigned function and necessity (Allers & Mevarech, 2005).

1.1.3.3 Replication process

Archaea possess a simplified form of the eukaryotic replication machinery which is presumably reflective of a more ancestral state. Comparative genomics has revealed that most of archaeal informational processes show similarities to those in eukaryotes, especially proteins which are involved in DNA replication have eukaryotic homologs or are more related to them than to bacterial ones (Myllykallio, 2000). When origins of replication in archaeal genomes were identified by bioinformatical analysis it turned out that some *Archaea* own a single origin (e.g. *Pyrococcus furiosus*) and others multiple origins (e.g. *Halobacterium*). “The identified origins are rather large, and many are located in close proximity to the genes encoding the archaeal homolog of the eukaryotic initiator protein Cdc6. A similar situation exists in bacteria where genes encoding replication proteins (e.g. *DnaA*) are close to the replication origin” (Kelman & Kelman, 2003). According to *Bacteria* and *Eucarya* archaeal predicted origins are apparently rich on adenine (A) and thymine (T) bases and contain one or more long AT-stretches. They also include inverted repeats of various sizes which also facilitate the binding of initiation proteins (Grabowski & Kelman, 2003). Both *Archaea* and *Bacteria* possess a large circular chromosome and occasionally accompanied by one or more smaller circular DNA plasmids. Many archaeal genes appear to be organized and arranged like bacterial operons (Brown & Doolittle, 1997). So the question arises: how do those organisms with a bacterial-like

chromosome structure and organization, initiate and regulate DNA replication with eukaryal-like proteins? In the archaeal genomes some of homolog proteins with different subsets among *Archaea* have been identified which are involved in the initiation process in *Eukarya*. In *Bacteria* the DnaA protein binds to the origin of replication (*oriC*), more specifically to a 13 bp sequence called DnaA box that is repeated five times in *oriC*. This initiates melting of complementary DNA strands at A-T clusters in order to initiate replication. In *Archaea* the proteins ORC (origin recognition complex) and Cdc6 have been identified as homologues to the eukaryotic initiator proteins.

“The fact that one of ORC/Cdc6 homologs is usually expressed next to the origin (e. g one next to both predicted origins in Halobacterium NCR1) suggests that one functions as the origin-binding protein (OBP), analogous to DnaA in bacteria” (Grabowski & Kelman, 2003). When origins are bound by ORC the Cdc6 protein recruits supposedly with other proteins that cannot be identified. It is thought to facilitate loading of the minichromosome maintenance (MCM) complex which is a helicase responsible for unwinding of the DNA. DnaB the bacterial helicase is loaded with the help of DnaA and DnaC. Primer synthesis is achieved by the monomeric product of the *dnaG* gene in *Bacteria*, whereas archaeal genomes encode homologs of two subunits. The eukaryotic primase complex is made up of a large (PriL) and a small (PriS) subunit, forming a dimer but they do not have homologs of subunit B and $\text{pol}\alpha$, which form a primase complex in *Eukarya* (Bell *et al.*, 2005). In all *Archaea* at least one DNA polymerase of the B family is responsible for strand elongation during replication, which is conserved and homologs have been identified in all three domains of life. In euryarchaeal species an additional class of heterodimeric DNA polymerases, the family D, have been detected (Bell *et al.*, 2005).

Based on biochemically studies it is thought that B polymerases synthesize the leading strand and D polymerases replicate the lagging strand in *Pyrococcus*. Creanarchaeal species only have multiple families of B polymerases and it is suspected that they have distinct roles on the leading and lagging strands. In *Bacteria* the sliding clamp is known as a protein complex which serves as a processivity-promoting factor in DNA replication and is a homodimer. In contrast, in archaeal and eukaryal organisms the sliding clamp, the proliferating cell nuclear antigen (PCNA), is a trimer (Berry & Bell, 2006).

1.1.3.4 Transcription

The archaeal transcription apparatus, including the RNA-polymerase (RNAP) and the general transcription factors, is comparable with the eukaryotic transcription machinery, whereas the transcription regulators are closely related to bacterial factors (Carlo *et al.*, 2010). Bacterial RNAP is a relatively simple complex with 5 subunits, and in addition with a set of up to 20 sigma factors it allows promoter selection in response to changing conditions. Up to five variant RNAP complexes (I-V) are used in *Eukarya* for transcription of distinct genes (Koning *et al.*, 2010). The first step of the initiation process is the recognition of the TATA box that is located about 30 bp upstream of the initiation site. It is made by the pre-initiation complex (PIC) which is relatively simple in construction compared to the eukaryotic ones. In *Archaea* this complex consists the transcription factor B (TFB), an ortholog to the eukaryal TFIIB; the TATA-box-binding protein (TBP) and a 12-subunit RNAP that is highly structurally similar to the eukaryotic RNA polymerase II (Koning *et al.*, 2010). In contrast the eukaryotic PIC includes RNAPII, TBP (present within the multi-subunit TFIID complex) and TFIIB. The latter two are helping to recruit and stabilize RNAPII at the promoter and determine the transcription start site (Carlo *et al.*, 2010). *“The fact that Archaea rely upon TBP and TFB to regulate RNAP-dependent transcription suggests that structural organization of the archaeal PIC may resemble a basic core within the eukaryotic PIC, whereas TBP and TFB orthologs are not found in bacteria”* (Carlo *et al.*, 2010). The subunits of the archaeal RNA polymerase can be assigned to three different functional groups. The “catalytic core” with the active site, the “assembly platform” and the “auxiliary subunits” including the subunit H which is required during promoter opening and the initiation of transcription. The latter subunits undergo structural rearrangement in the transition from initiation to elongation (Koning *et al.*, 2010).

1.1.3.5 Translation

The last phase of gene expression in *Archaea* is initiated by recognition of an initiation codon by the small ribosomal subunit (30S) and the formation of the initiation complex. This complex includes the initiation tRNA (Met-tRNA^{Met}) and mRNA. Like in *Bacteria* archaeal mRNAs are not capped on the 5'-end and some, but not all genes have a shine-Dalgarno-like sequence located upstream of the start codon. After the initiation complex is formed the large ribosomal subunit (50S) joins and the 70S ribosome is fully assembled then elongation of translation can occur in a factor-dependent manner (Bell & Jackson, 1998). Like *Eukarya* and *Bacteria* *Archaea* have IF1/eIF1A and IF2/eIF5B orthologs which ensure that the initiator tRNA gets to the right place on the small ribosomal subunit and that the subunit joining with the start codon happens

at the correct time. The most common start codon used in *Archaea* is AUG as in *Bacteria*, but also GUG and UUG start codons are employed (Bell & Jackson, 1998). The Shine-Dalgarno recognition mechanism is not primarily used by all *Archaea* or *Bacteria*. In *Sulfolobus* and *Pyrobaculum* this mechanism is only used at distal cistrons of polycistronic transcripts. In methanogens and halophiles the gene cluster are larger and bacteria-like, whereas the same genes can be unlinked in thermoacidophile species (Londei *et al.*, 2005).

In *Archaea* the homologue of the eukaryal eEF-1 α recruits the aminoacyl tRNA to the free A-site on the ribosome and the translocation process is mediated by the eukaryotic-like eEF-2 (Bell & Jackson, 1998). Finally the termination step in *Bacteria* is implemented by two releasing factors (RF1 and RF2) that recognize different stop-codon pairs (UAA/UAG and UAA/UGA), whereas in most *Archaea* and *Eukarya* only a single one (aRF1 and eRF1, respectively) exists that recognizes all three stop sequences (Koning *et al.*, 2010).

In *Eukarya* it is different as the 40S subunit is not able to interact directly with the mRNA and needs mediation by the 5'-cap binding complex eIF4F. When the mRNA is bound the complex scans the RNA in 3' direction in order to recognize an initiation codon, after that the 60S joins the complex, the initiation proteins leave and elongation can start. Subsequently leaderless mRNAs are presented in all three domains although quite rare in *Bacteria* and *Eukarya* but they are abundant in many archaeal species (mechanism for monocistronic mRNA). In this case, transcription starts with 5'-terminal initiation codons and can be translated efficiently by all ribosomes (Koning *et al.*, 2010). It is thought that this could be a very old mechanism which is still present in all three domains of the living world (Benelli *et al.*, 2003).

1.1.4 Life in extreme environments

Organisms whose habitats lie outside the range of environmental conditions in which the majority of life exists are called "extremophiles". Those environmental influences include physical and geochemical extremes that are inauspicious to organisms which live in moderate environments are called mesophilic or neutrophilic. Extremophiles are presented in all three domains predominantly microbes, where various special adaptations have evolved to cope with the environmental circumstances. For example cells need precise control over biomolecules, electric currents and the ability for reparation under certain conditions that can damage nucleic acids, proteins and lipids (Rothschild & Mancinelli, 2001).

1.1.4.1 Physically extremes

Temperature:

Temperature influences organisms in different ways, for example structural devastation caused by ice crystals and denaturation of biomolecules. Furthermore it affects the fluidity of the membrane, the solubility of O₂ and CO₂ in aquatic biotopes and with increasing temperature degradation of chlorophyll is induced, this excludes photosynthesis in those environments. Maximum growth of hyperthermophilic organisms occurs at >80°C, psychrophilic at >15°C, and the first isolated hyperthermophilic *Archaea Pyrolobus fumarii* (*Crenarchaeota*) has an optimal growth temperature at 106°C however it will still grow at 113°C (Oarga *et al.*, 2008). In contrast many microbes can be reversibly preserved in liquid nitrogen (-196°C), the lowest recorded temperature for active microbes is at -18°C (Rothschild & Mancinelli, 2001). It is assumed that proteins from cold-adapted *Archaea* consist of a higher content of non-charged polar amino acids (Gln and Thr) and a lower amount of hydrophobic (Leu) one. Membrane lipids tend to become rigid at low temperatures and it was shown that in *Methanococoides burtonii* a growth-temperature-regulated membrane-lipid unsaturation is involved to ensure membrane function (Daniel *et al.*, 2006).

Pressure:

Pressure is ubiquitously present in the environment and increases with ocean depth and underground. Once human beings evolved at an atmospheric pressure of 101 kPa (= 1.013 bar = 1 atm) although aquatic ancestors originated under hydrostatic pressure. Under these conditions life is challenging because pressure compresses packaging of lipids, which results in a decreasing fluidity of the membrane. Although organisms are adapted to the pressure an abrupt change can be lethal. Consequently all environments above 10 MPa are called “high pressure” biotopes, comprising of about 88% of the oceans. The 3800m depth is characterized by a stable average temperature of about 3°C, the lack of sunlight, oxygen and low organic carbon or minerals. 99% of the organic matter which is photosynthetically produced in the surface waters is recycled in the upper 100- 1000m and thereby about 1% reaches the deep-sea region. Most of the isolated organisms are chemolithotrophs and many are aerobes or facultative anaerobes (Rothschild & Mancinelli, 2001).

Radiation:

Radiation is energy which appears like particles (e.g. neutrons, electrons, protons, alpha particles or heavy ions) or electromagnetic waves (UV radiation, microwaves, gamma rays, X-rays or radiowaves). Resistant organisms are characterized by the ability to withstand DNA-damaging agents that causes lethal effects. *Deinococcus radiodurans* was first isolated in canned meat exposed to γ -irradiation in order to achieve sterility. Radioresistant attributes were also observed among hyperthermophilic *Archaea*. For example *Halobacterium salinarum* is an aerobic, mesophilic and obligatory halophilic organism, therefore has a high intracellular KCl concentration and a pigmented membrane with bacterioruberin which provides protection from DNA damaging stressors. It can cope with a narrowed capacity of UV and ionizing radiation. It is suggested that those species combine a variety of physiological applications to ensure radioresistance and have been evolved independently in *Bacteria* and *Archaea* (Confalonieri & Sommer, 2011).

1.1.4.2 Geochemically extremes:

Oxygen:

These days, microbial life exists in environments ranging from strictly anaerobic to aerobic, although the Earth has been anaerobic throughout most of the evolution of life. Generally the aerobic metabolism is more efficient than the anaerobic, but organisms have to defy reduced forms of molecular oxygen. Especially the presence of hydroxyl radicals inside the cell can cause DNA damages and other physiological changes. UV-A radiation (320- 400 nm), metabolic production during aerobic metabolism and photosynthesis are processes where H_2O_2 is presented. Additionally processes where reactive oxygen is presented as an endogenous source such as mitochondrial respiration, cytochrome P450 metabolism of hydroperoxides, production of uric acid and oxidative bursts in case of pathogen defense. The presence of oxygen can support DNA damages and has to involve many repair mechanisms for prevention (Rothschild & Mancinelli, 2001).

Salinity:

Hypersaline environments include natural lakes, salt flats and deep-sea hypersaline niches, which are populated by *Archaea*, or *Bacteria*, green algae, protozoa but also yeasts and fungi. Organisms living in this condition are called "osmophilic" or "halophilic" and can tolerate or require high salt concentrations. Water is essential for all organisms to ensure normal cellular

function. To stabilize cells in response to increased osmotic pressure organisms must be able to adapt physiologically. Therefore compatible solutes such as, K⁺, glycine, proline, glutamate or glutamine accumulates in the cytosol and protects the organism from cytoplasmic dehydration or desiccation. More about adaptation to salinity environments will be mentioned in greater detail in the following chapter 1.1.5 (Oarga, 2009).

pH:

In humans biological processes generally occur towards the neutral range of pH among 5 – 8.5. Higher pH occurring in soda lakes are populated with so called “alkaliphiles” with a pH up to 11, in contrast the “acidophiles” are organisms which are living in extremely acidic biotopes with a pH below 2. Organisms are faced with the problem of protein denaturation and to sustain metabolic processes. Consequently acidophiles use active proton pumps as protection mechanism to maintain their intracellular pH value at almost neutral. The defense mechanism used by “alkaliphiles” will be discussed further in the next chapter (Oarga, 2009).

1.1.5 Halo-alkaliphilic *Archaea* and their adaption to extreme environments

Ecosystems like saltern pond brines and natural salt lakes are often populated with low diversity but high community densities of salt-loving microorganisms. Several of the studied habitats of halophiles are thalassohaline environments that accrued by evaporation from seawater, hence reflecting the ionic composition and nearly neutral or slightly alkaline pH of seawater. Athalassohaline environments such as natural soda lakes or soda deserts are denoted as the most stable natural occurring alkaline hypersaline niches. Organisms living there have the challenge of high salt content combined with the need to thrive at high alkaline pH and raised temperatures (Yanhe Ma *et al.*, 2010). The predominant microbial biomass of hypersaline soda lakes such as Lake Magadi in Kenya, are the aerobic halophilic archaeal family *Halobacteriaceae* of the order *Halobacteriales*. One of the specific characteristic is the mass of C-50 carotenoid pigments in the membrane of the most *Halobacteriaceae* and the cause for the red coloration of such hypersaline lakes.

Additionally halophilic representatives were found within the methanogenic *Euryarchaeota* which are able to carry out methanogenesis in hypersaline environments. However, no halophilic members of the kingdom *Crenarchaeota* have been identified there yet (Oren, 2002). In the course of this thesis the family *Halobacteriaceae* will be elaborated further. The

first representatives of this family were identified about hundred years ago and leading up to the present, 36 genera with 129 species have been discovered (Oren, 2012).

1.1.5.1 Cell envelope and osmotic adaptation

As already mentioned structural diversity of membranes within the family *Halobacteriaceae* are confirmed. In more detail the haloalkaliphilic species of the genus *Natrialba* contain diether core lipids of C₂₀C₂₅ (sn-2-esterpanyl-3-phytanlyglycerol) and C₂₀C₂₀ (dn-2,3-diphytanlyglycerol), polar lipids like PGP-Me (phosphatidylglycerol phosphosphate-methyl ester), PG (phosphatidylglycerol) and PA (phosphatidic acid) as well as glutaminyglycan polymers were reported as characteristic membrane components (Kamekura & Kates, 1999). The increased assemblage of negatively charged lipids upon increasing salt concentrations suggests that membranes of haloalkaliphiles are mainly adapted to the high salt concentration and high pH concentration (Van de Vossenberg, 1999). *“Irrespective of the nature of the lipids, i.e., halophile versus haloalkaliphile, pH does not interfere with membrane integrity, nor does it dramatically affect the proton permeability.”* (Van de Vossenberg *et al.*, 1999)

These particular structural features are an important mechanism of adaptation to their niche since osmotic pressure has to be overcome. Haloalkaliphilic species have evolved additional mechanisms including the intracellular accumulation of inorganic cations and/ or neutral organic compounds. Therefore two different strategies are used to achieve an intracellular high osmotic pressure, keeping the Na⁺ concentration low. The first strategy “high salt-in” involves accumulation of K⁺ and Cl⁻ inside the cell, while pumping Na⁺ out using a variety of cation/proton antiporters.

The second strategy is called “low salt-in” and facilitates osmotic balance in the cytoplasm by biosynthesis, and accumulation of organic osmotic solutes (osmolytes). The latter one is far more widespread and does not require the extensive adaptation of the intracellular enzymes to maintain activity near the saturation point of salt concentrations (Oren, 2008). Energetically less costly to the cell is the “high salt-in” strategy because these antiporters generate a proton motive force based on both, the transmembrane electrical potential ($\Delta\psi$) and the transmembrane pH gradient (ΔpH) (Van de Vossenberg *et al.*, 1999). The genome of *Natrialba magdii* contains an operon encoding a putative pH adaptation K⁺ efflux system. Three other genes are presented encoding putative cation/proton antiporters, additionally genes for biosynthesis of spermine as well as transporters for the uptake of choline/carnitine/betaine, and spermidine/putrescine, supporting the presumption that this species have multiple mechanisms for osmotic adaptation (Siddaramappa *et al.*, 2012).

1.1.5.2 Haloalkaliphilic adaptation of proteins

Proteins synthesized by “salt-in” organisms have various notable changes that allow them to be stable and soluble, which is necessary for a proper folding and activation of the protein in these conditions. One of the earliest properties recognized was the excess of negative charges of the proteins this is clearly reflected in the low values of the isoelectric point (pI) of halophilic proteins. The amino acid compositions of those proteins show an increased number of acidic residues like glutamate (Glu), aspartate (Asp) and also a decreased number of lysines (Lys) which may contribute a low pI value. It was suggested that the acidic residues promotes the binding of a network of hydrated salt ions, prevents aggregation of proteins through repulsive interaction and that it maintains flexibility within the protein. However, some halophilic proteins are active and stable in conditions were no or little salt binds. Other differences between amino acid composition of non-halophilic and halophilic proteins have been reported with an increased amount of serine, threonine and small hydrophobic residues (glycine, alanine, valine) (Shirai *et al.*, 2008). It has been reported that high NaCl concentration enhances hydrophobic interactions, whereby hydrophobicity is the mainly driving force of protein folding (Siglioccolo *et al.*, 2011). Furthermore the enzymatic activity of halophilic enzymes is heavily dependent on the nature of the salt, while the activity in KCl is significantly higher than in NaCl (Madern *et al.*, 2000). To overcome alkaline environments with a pH value of 9 or higher, haloalkaliphilic proteins have to develop strategies for adaptation. Therefore following three major strategies were suggested: a) the pKa modulation, b) the Asp+Glu gain strategy and the c) Asp+Lys loss – Glu+Arg gain strategy.

a) *The pKa modulation strategy*

This strategy means that the proteins are modified by modulation of hydrophobic bonds or by changing the net charge of the protein towards an increasing pH. Adaptation is based on the catalytic activity of a protein which is dependent on pKa values of the catalytic residues at certain pH values. The formation of hydrogen bonds with a catalytic site facilitates the deprotonated conformation of the catalytic residues and as a consequence thereof the pKa is decreased, which can occasionally result in lowering the pH optimum value for an enzyme. As a result the alteration of the pKa to a higher pH optimum often is attained by the use of alternative amino acids, respective of mesophilic homologues, in the proximity of the catalytic site to prevent hydrogen bond formation. Otherwise retaining the protonated catalytic residues is also essential for a proper function since the proton amount is low at higher pH

values. Furthermore, amino acids like Asp and Glu as negatively charged residues can capture protons and raise the pKa of ionizable groups, resulting in a notable increase of the optimum pH of enzymes (Shirai *et al.*, 2008).

b) The Asp+Glu gain strategy

As mentioned previously haloalkaliphilic proteins consists of higher amounts of Asp and Glu than arginine (Arg) and Lys, which causes the negative net charge value of a protein and relates to the theoretical strategy of pKa modulation. A possible application is when the cell wall of haloalkaliphilic organisms where negatively charged polymers are exposed, thus repelling hydroxyl ions (Shirai *et al.*, 2008).

c) The Asp+Lys loss – Glu+Arg gain strategy

This strategy implies the evolutionary adaptation of proteins to alkaliphilic environments by modifications of the amino acid composition, highlighting the decreasing amount of Lys and Asp residues, while the number of Arg, His and Glu is increased. For instance, an exchange of Lys with Arg results in protein which has a higher pKa and pI, therefore it is more adapted for high pH value environments. The “loss and gain” theory is partly based on ASET (ancestral sequence evolutionary trace) analysis where amino acid residue alterations of alkaliphilic proteins are compared to their calculated ancestral sequence. Additional differences in the structure of alkaliphilic enzymes compared to mesophilic enzymes are confirmed by a student’s t-value test (Shirai *et al.*, 2008). Summarizing it is still not completely clear and definable how proteins are active in the intracellular high salt conditions of haloalkaliphilic archaea.

1.1.5.3 Adaptation of protein secretion systems of haloalkaliphilic organisms

All three domains of life exhibit extracytoplasmic proteins which have to master the barricade of hydrophobic and lipid based membranes. Previously noted archaeal protein translocation is a mosaic comprising bacterial, eukaryal and archaeal features. Firstly specific proteins have to be distinguished and targeted from the pool of cytoplasmic proteins. In all domains secretory proteins are synthesized as a pre-protein with an N-terminal signal peptide, which serves to target a special transmembrane export complex and the mature proteins have to be (re)folded outside the cell. *Archaea* as well as *Bacteria* and *Eukarya* use the general secretory (Sec) pathway for protein export. The translocon in *Bacteria* is based on the SecYEG complex and in

Eukarya the complex Sec61 $\alpha\beta\gamma$, embedded in the membrane of endoplasmic reticulum (ER) is competent for translocation (Ring & Eichler, 2004). In *Archaea* homologues of SecY/Sec61 α have been detected in all sequenced archaeal genomes and analysis revealed that this complex is more reminiscent to the eukaryotic than to the bacterial counterparts (Cao & Saier, 2003). The Sec-dependent translocation in *Archaea* and *Eukarya* are suggested to be of similar type since the finding of encoded homologues of eukaryal Sec61 β , as well as Sec61 α and Sec61 γ , were examined in all archaeal genomes (Ring & Eichler, 2004). Additionally archaeal genomes encode for components of the Sec-independent secretion pathway for translocation of already folded proteins. This twin arginine transport (Tat) translocation pathway was also observed in *Bacteria* (Ring & Eichler, 2004). Analysis predicted that this translocation pathway is predominantly used by halophilic *Archaea*, contingently to prevent misfolding and aggregation of proteins which are exposed to the cytoplasmic milieu of high KCl in those organisms (Bolhuis, 2002). The fact of the different natures of cytoplasmic membranes of *Archaea* makes it similar to specific evolutionary adaptations, happened in their proteins translocation pathways. However, Tat-pathway components are similar to those of *Bacteria* and chloroplasts have been detected in *Haloarchaea*, which are suggested to use different sources of energy (Hutcheon & Bolhuis, 2002). Generally archaeal proteins targeted for Tat secretion are antimicrobial enzymes or proteins which are involved in intercellular signaling also include polymer degrading enzymes or membrane-anchored proteins. A wide variety of cellular processes are assigned to the proteins; nutrient uptake, motility and surface attachment (Szabo & Pohlschroder, 2012). Tat-specific proteins are folded prior to translocation, therefore those substrates are too large for the Sec translocon complex and obviously the Tat complex and mechanism is fundamentally different to the Sec pathway. “However, just as Sec substrates are targeted to the Sec pathway by a signal peptide, a Tat signal peptide targets proteins to the Tat pathway” (Gohlke *et al.*, 2005; Robinson *et al.*, 2011). The twin arginine transport signal peptide contains a lower hydrophobic stretch than the Sec signal sequence confirmed by two programs; TatFind and TatP are specialized for Tat substrate identification (Rose *et al.*, 2002; Bendtsen *et al.*, 2005). However, it was established that the Tat signal peptide sequences in bacterial and archaeal species are reasonably similar (Rose *et al.*, 2002). For a functional Tat pathway organisms have to contain at least TatA and TatC, which are playing an essential role with TatC (and in many species with TatB) in the substrate targeting process. Whereas a TatA multimer as a single transmembrane spanning protein allows pore formation of variable sizes to accommodate different masses of secretory proteins (Dilks *et al.*, 2005; Gohlke *et al.*, 2005; Leake *et al.*, 2008; Robinson *et al.*, 2011).

1.1.5.4 Metabolic aspects of alkaliphilic *Haloarchaea*

Although *Haloarchaea* share common features in order to adapt to hypersaline environment, like respiratory chains, rhodopsins and acidic protein machinery, their nutritional demands and metabolic pathways are extremely different (Falb *et al.*, 2008). Comparative analysis, regarding their metabolic properties based on systematic metabolic reconstruction of sequenced halophilic and haloalkaliphilic species show, that they differ not only in their catabolic pathways but also in their nutritional requirements.

Haloarcula marismortui, *Haloferax mediterranei* and *Halococcus saccharolyticus* are carbohydrate-utilizing species which are able to catabolize hexoses, pentoses, sucrose and lactose. Alternatively *Halobacterium salinarum* thrive on amino acids and other typical compounds of highly saline environments, similar to *Natronomonas pharaonis* which is likely incapable of glucose degradation. Glycerol is a highly abundant carbon substrate in hypersaline habitats, catabolized by *Haloferax volcanii* and *Natrialba magadii*. Furthermore in *Nab. magadii* genes encoding putative enzymes which are involved in modification of pathways such as, gluconeogenesis, glycolysis, utilization of glycerol and catabolism of aromatic amino acids have been identified (Siddaramappa *et al.*, 2012).

Although haloalkaliphilic *Archaea* mainly obtain energy sources from degradation of organic nitrogen compounds, many grow proteolytically and therefore secrete extracellular proteases which degrade proteins (De Castro *et al.*, 2008). In regard to the high pH conditions active proton-driven energy conserving ATP synthase is obviously challenging for *Archaea*, therefore in course of adaptation it was proposed that ATP synthase is driven by sodium ions instead of protons (Siddaramappa *et al.*, 2012). Several gene clusters encoding putative ATPase subunits were identified in many halophilic archaeal species. *Nab. magadii* is thought to have a proton-driven ATP synthase since no sequence signature of Na⁺-dependent ATPase was found (Mulkidjanian *et al.*, 2008). In aerobic organisms the respiratory chain (reduction of oxygen and associated proton coupled electron transfer) is the basic source of energy and involves various cytochromes and oxidases. Comparative genomic analysis revealed novel features like a type II NADH dehydrogenase and homologs in several halophilic species. Additionally to the NADH dehydrogenases *Nab. magadii* and further species are predicted to possess a succinate dehydrogenase which may oxidize succinate and is responsible for a reduction of quinones in the electron transport chain (Siddaramappa *et al.*, 2012).

1.1.6 Biotechnical applications of (haloalkaliphilic) *Archaea*

Microbial life is not limited to specific environments and is represented in most diverse conditions, hence unique properties of those biocatalysts are providing possibilities for novel applications in industrial and environmental processes. Halophiles are involved in renowned processes like traditional food fermentation, production of solar salt from seawater and improving cold washing efficiency with protein engineering techniques (Enache & Kamekura, 2010). More modern applications are the treatment of saline wastewaters using halophilic advantages with bacteriorhodopsin as “photoelectric converter” or for the production of biodegradable plastic with properties resembling that of polypropylene (Oren, 2010).

Industrial processes like pickeling industry or the β -carotene production from the halophilic green algae *Dunaliella salina*, is used as an antioxidant and food colouring agent this generates saline wastewater. To degrade organic carbon or toxic compounds in high salinity wastewaters only laboratory-scale models exist that show satisfactory results of up to 6% salt concentrations (Oren, 2010). Another application for wastewater treatment might be the using extracellular enzymes of halolakaliphilic species which are stable at high salt concentrations and high pH values.

Poly- β -hydroxyalkonate production by halophilic organisms is naturally accumulated as storage polymers. The advantage of using halophilic species for bioplastic production is the decreasing danger of contamination in the cultivation medium. Moreover the relatively simple purification of the product since those cells lyse in the absence of salt, resulting in the release of the polymer (Oren, 2010). In *Hbt. salinarum* bacteriorhodopsin serves as a light-driven proton pump and is also stable and active at low salt conditions. It functions within a wide range of temperatures as well as pH-values and it is resistant to digestion by proteases; additionally it is easy to immobilize and maintain photochemical properties over a long time. Possible uses of this molecule have been suggested based on the conversion of light energy into chemical energy. It also may serve as possible application for ultrafast light detection, motion detection, construction of artificial retinas and molecular sensors. To summarize the tremendous diversity of *Archaea* and other extremophile organisms, these may develop novel applications in the future. Research based on available genomics and proteomics of haloalkaliphiles have opened new possibilities for innovative biotechnological processes designed for the current demand of modern society.

1.2 *Natrialba magdii*

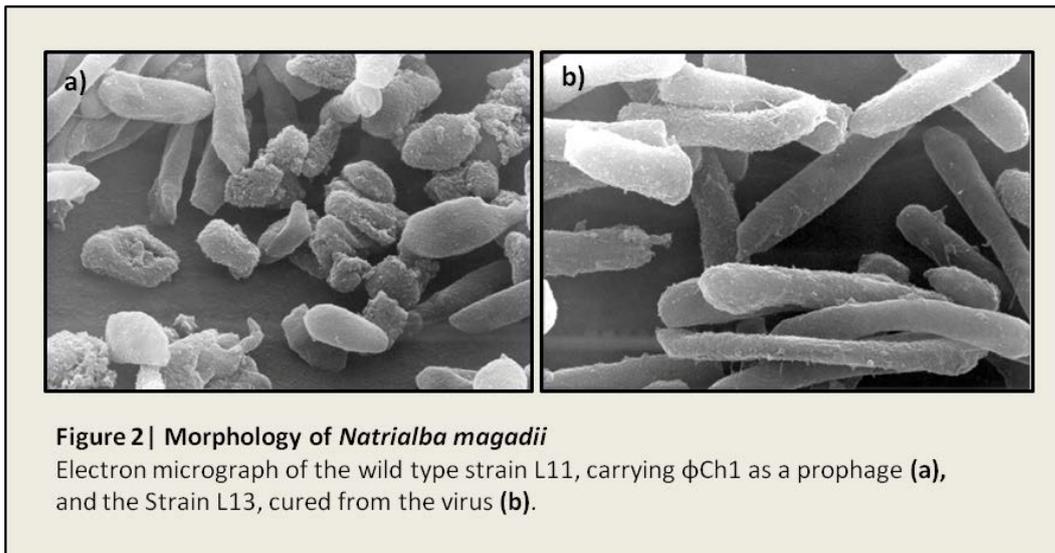
The haloalkaliphilic species *Nab. magdii* was first isolated and described as *Natronobacterium magdii* in 1984 by Tindall *et al.* The initial isolation from alkaline, hypersaline lakes and soils, consisted rods and cocci, therefore were accordingly separated into the genera *Natronobacterium* and *Natronococcus*. In the course of 16S rRNA comparative analysis of sequenced *Halobacteria* have shown that *Natrialba* represent a distinct genus within the family *Halobacteriaceae* (Kamekura *et al.*, 1997). The phylogenetic relationship between *Natrialba magdii* and two *Natrialba asiatica* strains were suggested with a sequence similarity of 93.3 % and 93.7 %.

1.2.1 General characteristics of *Natrialba magdii*

Nab. magdii belongs to the kingdom *Euyarchaeota*, the cells are motile rods with a length of 0.5 – 0.7µm and have a slightly orange/red colour caused by carotinoid pigments which are embedded in the membrane (Tindall *et al.*, 1984). *Nab. magdii*, isolated from Lake Magadi (Kenya) which belongs to the east African Rift valley lakes, requires not only high salt concentrations but also pH values between 8.5 and 11 and Mg²⁺ concentrations below 10mM for optimal growth. Further it prefers temperatures between 37°C and 42°C and is not able to withstand lower temperatures or salt concentration below 2M. However, even at optimal growth conditions the generation time in the logarithmic growth phase takes about nine hours. *Nab. magdii* is a strictly aerobic chemoorganotrophic organism, using amino acids and small peptides as energy and carbon source (Tindall *et al.*, 1984).

1.2.2 Laboratory strains of *Natrialba magdii*

There are two strains of *Nab. magdii* we are working with in the laboratory. In 1997 spontaneous lysis of the wild type strain led to the discovery of the prophage φCh1, which is integrated into the genome of *Nab. magdii* during the lysogenic cycle. The wild type strain L11 carrying the prophage was cured by repeated subculturing revealed a non lysogenic strain of *Nab. magdii*, called L13 (see figure 2) (Witte *et al.*, 1997). *Nab. magdii* is to date the only known host of the halophage φCh1 and the complete nucleotide sequence was first described in the year 2000 (Witte *et al.*, 1997; Klein *et al.*, 2002). The virus-cured strain L13 serves as an indicator strain and is used to demonstrate infectivity, transformations and kinetic studies.



1.2.3 Working with *Natrionalba magadii*

Working with haloalkaliphilic organisms is still a great challenge especially with respect to the application of known basic methods, which are not convertible and efficient enough for the particular conditions. At the start of researching with *Nab. magadii* in the laboratory neither a successful transformation method and vectors were developed, nor selection markers were available for this organism. Several troublesome applications were adapted and established for the high haloalkaliphilic living conditions of *Nab. magadii* and the associated halophage ϕ Ch1.

1.2.3.1 Transformation of *Natrionalba magadii*

The first efficient transformation of a halophilic archaeon was published in 1987, where *Halobacterium salinarum* was transfected with DNA isolated from the virus ϕ H (Cline & Doolittle, 1987). This method is based on a polyethylene glycol (PEG) and EDTA-mediated spheroplast generation, whereby the chelating agent binds Mg^{2+} ions and destabilizes the S-layer which leads to competent cells. The principle of the method was quickly adapted for the transformation of *Haloferax volcanii* and various other archaeal species. Several methods for *Haloarchaea* are published in "The Halohandbook" (Dyall-Smith, 2009). However, this method is not effective for alkaliphilic *Archaea* like *Nab. magadii*, since removing the paracrystalline glycoproteins surface layer with EDTA was unsuccessful. Thereupon spheroplast cells of *Nab. magadii* were successfully achieved by primary treatment with bacitracin which inhibits the glycosylation of the S-layer; followed by partial digestion using the proteinase K. After that procedure the PEG - mediated transformation with foreign DNA was successful. The

transfection rate of bacitracin-proteinase K-treated cells with ϕ Ch1 DNA amounts 10^4 pfu/ μ g and is efficient enough for genetic manipulation of *Nab. magadii* (Mayerhofer-Iro *et al.*, 2013).

1.2.3.2 Vectors and genetic markers for *Natrialba magadii*

After the successful developed transformation method a construction of a shuttle vector for *Nab. magadii* and *E. coli* followed. Therefore sequence similarities of the replication origin (*repH* and AT- rich sequence) of *Nab. magadii* and the ϕ Ch1 genome were analyzed. Sequences (ORF53, ORF54) were discovered with equality to the protein-encoding sequence of a halophilic plasmid (*H. marismortui* plasmid pNRC100) (Iro. *et al.*, in prep). The *repH* gene is known to be a part of the origin of replication of already known plasmids in halophilic *Archaea* (Ng & DasSarma, 1993; Klein *et al.*, 2002). Because selectable markers for *Nab. magadii* were lacking the mutated *gyrB* gene (novobiocin resistance) of *Haloferax alicantei* was isolated from pMDS11 and cloned into the pKSII+ vector, (possessing origin of replication and an ampicillin resistance for *E. coli*) named pNov-1. Subsequently different parts of the regions ORF53 and ORF 54 were cloned into pNov-1 resulting in various plasmids termed pRo-1 to pRo-11. However, the best transformation efficiency result was achieved with the vector pRo-5 and has been successfully employed until present (Mayerhofer-Iro *et al.*, 2013).

For the haloalkaliphilic *Nab. magadii* only a few more shuttle plasmids exist, one of them is the pNB102 which is routinely used in the laboratory. This plasmid construction originates from the plasmid pNB101 which was combined with the pWL102 fragment carrying the ColE1 origin as well as a mevinolin and ampicillin resistance (Zhou *et al.*, 2004). Mevinolin and novobiocin are known as selectable markers which are active in *Nab. magadii* that makes the construction of shuttle plasmids limited. Mevinolin is competitively inhibiting the HMG-CoA reductase (3-Hydroxy-3-Methylglutaryl-Coenzym-A-Reduktase) which is used to synthesize mevalonic acid from acetyl CoA, found in *Archaea*, *Eucarya* and some *Bacteria*. In *Archaea* mevinoline can completely arrest growth as it blocks the production of isoprenoid lipids (Cabrera *et al.*, 1986; Lam & Doolittle, 1989).

Novobiocin is a naturally occurring antibiotic which inhibits the activity by binding the B subunit of the DNA gyrase and blocks the access of ATP to its binding site. This inhibition process is competitively since novobiocin shows little structural similarities to ATP, concluding that the inhibitory affect mechanism is the same in haloarchaeal organisms as in *Bacteria* (Mizuuchi *et al.*, 1978; Holmes & Dyall-Smith 1991).

1.2.3.3 Natrialba extracellular protease (Nep)

Nab. magadii exhibits a serine protease (*Natrialba* extracellular protease Nep) which is secreted at the end of the exponential growth phase. It is able to degrade large proteins such as casein and gelatin however, peptides containing Tyr, Phe and Leu at the carboxyl terminus are degraded more efficiently than those with basic amino acids like Arg or Lys. This secretory haloprotein shows its highest activity for azocasein hydrolysis in 1.5 M NaCl with the temperature of 60°C and a broad range of pH profile (8-10) (Giménez *et al.*, 2000). It is even thought that Nep is a serine protease due to the strong delay to serine protease inhibitors such as DFP (diisopropyl fluorophosphate), PMSF (phenylmethyl sulfonyl fluoride) and chymostatin. To expand, Nep is about 45kDa in size and the putative promoter includes an archaeal TATA-box consensus sequence proximal to a potential transcription start site and downstream from the BRE-like sequence (see figure 3). The protease is translated as a propeptide including a 121 amino acid residue (12,626 Da) which is cleaved to generate a mature 43,828 Da active protease in the environment (see figure 3) (De Castro *et al.*, 2008). Experimental evidence *in vitro* and *in vivo* verified that processing and activation during/after translocation is autocatalytic, with the hypothesis that cleavage of the signal peptide may be required to induce the autocatalytically process of removing the Nep propeptide (Ruiz *et al.*, 2012).

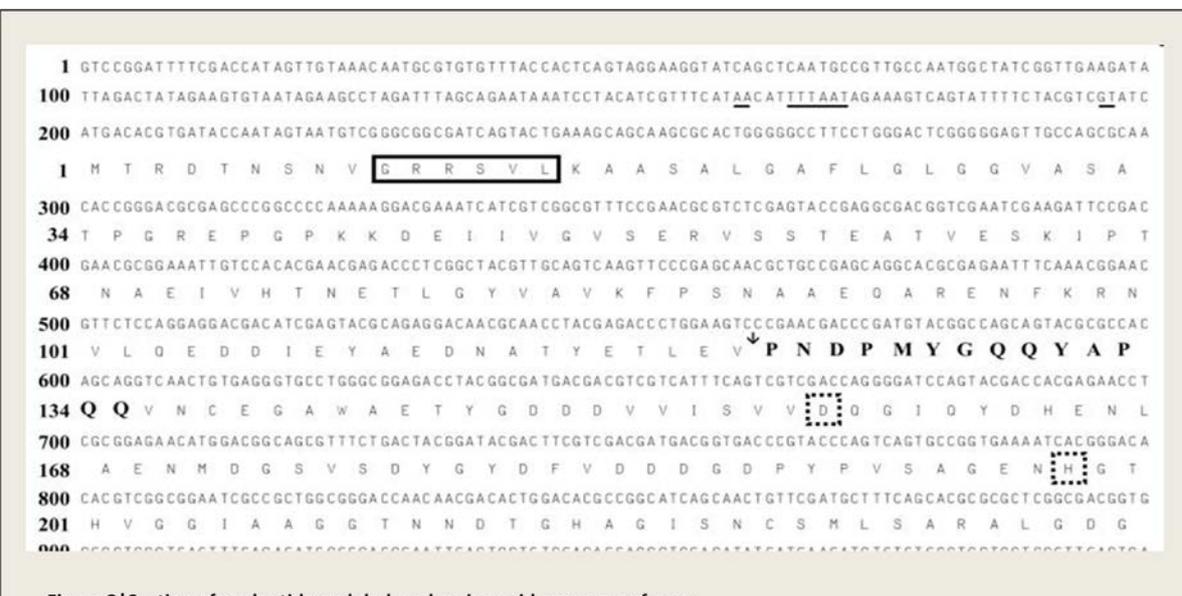


Figure 3 | Section of nucleotide and deduced amino acid sequence of nep :

The nucleotide and amino acid residue positions are numbered on the left and the deduced amino acids in single letter code are written below. The underlined region above indicates the putative TATA-box (-34 to -29) and the BRE-like element (-39 to -38) as well as the potential transcription start region (-5 to -4). The boxed amino acid residues (10-15) with solid lines indicate the predicted signal sequence for the Tat secretion pathway. The protease cleavage site is marked by the vertical arrow whereby the following bold letters indicate the N-terminal amino acid sequence from the mature Nep. The Asp-His-Ser catalytic triad active site are labeled with broken line boxes but serine is not shown yet. (De Castro *et al.*, 2008)

Nep has been detected during the transition to the stationary growth phase which leads to the suggestion that production is stimulated by nutrient scarcity and/or the population density. It was established that the Nep accumulation is increased by nutrients limited conditions and slow growth rate. The activity does not require the presence of protein substrates like casein for induction. However, easily metabolized nitrogen compounds (ammonium salts) inhibits the enzyme activity indicating that repression of Nep may be controlled by nitrogen catabolites. In comparison to this low density cultures starved for carbon and nitrogen sources did not accumulate detectable amounts of extracellular protease activity (D'Alessandro *et al.*, 2007). Hence indicating the possibility of *de novo* synthesis which needs "factors" or autoinducer molecules that are presented in stationary growth phase for Nep formation, due to quorum sensing (Paggi *et al.*, 2010). Bioinformatical analyses of the *Nab. magadii* ATCC43099 genome did not result in identification of proteins similar to those involved in bacterial quorum sensing, and no clear identification of autoinducer molecules in the cultures, since no reproducible results were obtained by the researchers (Paggi *et al.*, 2010).

1.3 Archaeal viruses

In 2003 Dyall Smith *et al.* published an overview about the diversity of haloarchaeal viruses. 15 different viruses were listed infecting species from the family *Halobacteriaceae*, whereby most of them have head-tail morphology. Samples taken from the Dead Sea were investigated by electron microscopy and have predominantly shown a "fusiform" (lemon-shaped) morphology. The density of virus-like particles in hypersaline waters were estimated and amounts at least 10^7 particles/ml (Oren *et al.*, 1997). In 1982 the first halovirus ϕ H was isolated by Schnabel *et al.* and studied in more detail. It is a temperate, head-tail phage infecting *Hbt. salinarum* and it was shown that sequence similarities at the level of replication, morphology and the control of lysogeny resembles those of bacteriophage P1 (Schnabel *et al.*, 1982). Additionally extensive similarity (97 %) of the phage ϕ Ch1 to ϕ H was found. It could be assigned that most of the ϕ Ch1 sequences were homologues of ϕ H, but a few matched sequences were found in *Bacteria* (i.e., *Myobacterium*, *Haemophilus* and *Bacillus*) as well. However full comparison of ϕ Ch1 and ϕ H is not possible yet, as the ϕ H genome sequence is only known for 60 %. Although the phylogenetic distance between their hosts, a close relationship between the phages is affirmed. In 1993 more haloviruses with head-tail morphology were isolated from these natural hypersaline sources. HF1 and HF2 are lytic and by cross-hybridization it turned out that their genomes were at least 80 % the same. The genome of HF2 with 121 predicted ORFs

shows only about 10 % sequence matches including haloarchaeal (*Hfx. Volcanii*; *Hbt. salinarum*), bacterial and bacteriophage lineages (*Listeria*, *Sinorhizobium*, *Bacillus* and the coliphage Mu) (Dyall-Smith et al., 2003) . The viruses His1 and His2 infecting *Haloarcula hispanica* were identified in 1997. They are unique with the classified haloviruses and in respect to their shape morphology it is proposed that His1 was a member of the Fuselloviridae family, which includes also SSV1. However on molecular level His1 and His2 are completely different from SSV1. His1 and His2 replicate by terminal protein-primed DNA synthesis and encode their own DNA polymerase, whereas SSV1 is a temperate phage, enables to integrate into the host genome, has a circular genome and harbors no DNA polymerase (Bath & Dyall-Smith, 1998; Dyall-Smith *et al.*, 2003). In summary the knowledge of the diversity and relationships of halophages is still increasing although in contrast to bacteriophages, the number of known halophages is relatively small. Culturing of all dominant species in salt lakes would explore the true halovirus diversity and informations in respect to their ecology and evolution (Dyall-Smith et al., 2003).

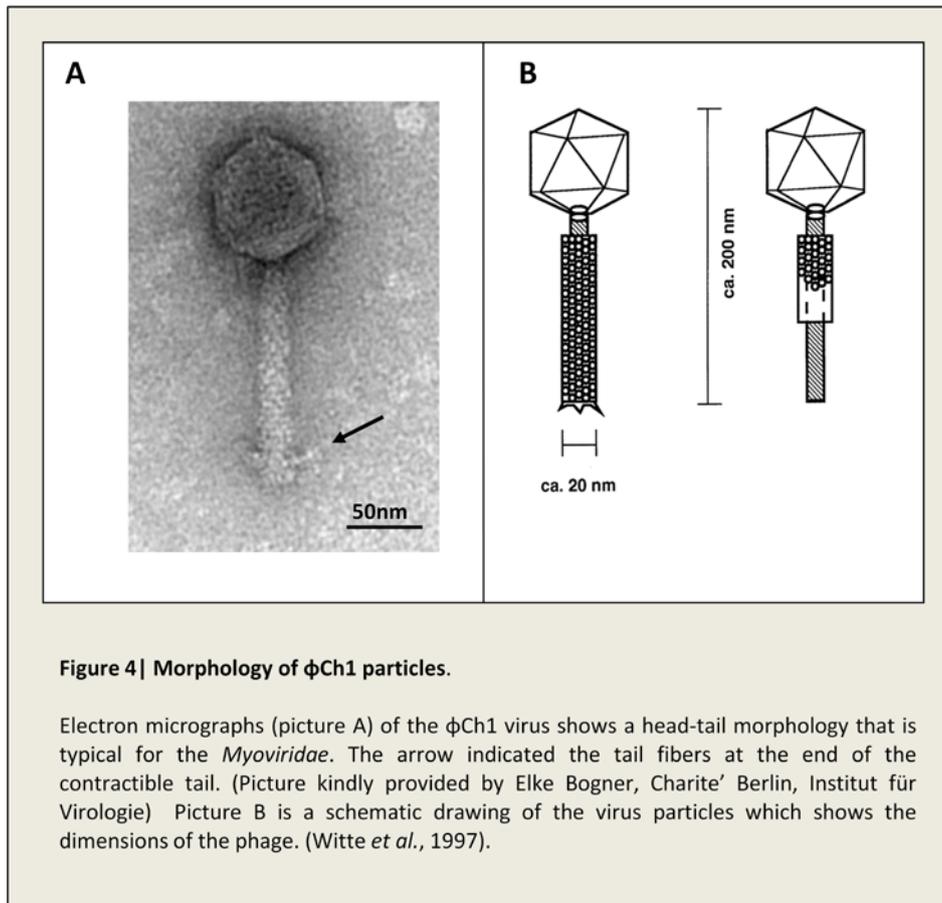
1.3.1 The halophage ϕ Ch1 infecting *Natrialba magadii*

In 1997 the halophage ϕ Ch1 was observed as a consequence of spontaneous lysis of *Nab. magadii* and no other host has been found until now (Witte *et al.*, 1997). The evidence that ϕ Ch1 is a temperate virus was given due to the turbid-plaque morphology and the fact that *Nab. magadii* cells, isolated from plaques, were able to produce viruses. These cells were named L11 and are used for the isolation of ϕ Ch1 (Witte *et al.*, 1997). Superinfection of *Nab. magadii* with the virus is not possible as it is known for the halophilic virus ϕ H infecting *Halobacterium salinarum* (Witte *et al.*, 1997; Stolt & Zillig, 1992). As already mentioned a second *Nab. magadii* strain was isolated which had been cured of ϕ Ch1 and serves as an indicator strain namely L13. This strain is used to demonstrate infectivity (Witte *et al.*, 1997).

1.3.2 Morphology and structural components of ϕ Ch1

The morphology of ϕ Ch1 resembles members of the family *Myoviridae*, characterized by typical head-tail morphology. After inspection of several electron micrographs the size of the phage particles were estimated. Those particles are composed of icosahedral heads and contractile tails with a length of 70nm and 130nm, respectively. Fully assembled virions show a total length of approximately 200nm, as indicated in figure 4 (Witte *et al.*, 1997). The major capsid protein gpE of ϕ Ch1 is encoded by ORF11 of the halophage genome and is 35.8 kDa in

size. Three dimensional structural modeling of gpE resulted in a similar protein fold as the main capsid protein of the bacteriophage HK97, which shows also relatedness at the higher structural level (Krupovič *et al.*, 2010). The ORF11 is expressed during the late phase of the viral life cycle and after translation the protein it gets associated with the host cell membrane. For the assemblage and release of progeny viruses the capsid protein has to be cleaved proteolytically. It is thought that it is also involved in packaging of the DNA, harbored within the virus, or in assembly of coat proteins (Siegel & Schaechter, 1973; Klein *et al.*, 2000). A putative leuzine zipper motif was discovered which might contribute to the multimerization of the structural proteins during virus maturation (Klein *et al.*, 2000). Separation of the complete phage particles on denaturing polyacrylamid gels indicated three further major proteins (A,H and I) and five minor proteins (B,C,D,F and G) of ϕ Ch1 (Witte *et al.*, 1997). Interestingly both protein A (80 kDa) and H (14.4 kDa) are showing the same N-terminal amino acid seuquence, whereby only ORF19 was detected to encode this sequence expressing a 14.4 kDa protein. It is likely that the 80 kDa band represents either homomultimers or heteromultimers consisting of ORF19 protucts and/or other structural proteins of ϕ Ch1 (Klein *et al.*, 2002). The minor proteins B, C and D which are encoded by the ORFs 9, 7 and 8, respectively, are presented in a much lower quantity in mature viral particles compared to the major structural proteins (Klein *et al.*, 2002). Moreover the ϕ Ch1 has contractible tails covering an internal shaft, the determined width amounts approx. 20nm. At the end of the shaft, tail fibre structures are located and the supposed function is the attachment to archaeal cells. It is known that tail fibres of many bacteriophages are involved in the interaction with the host cell, initiating the infection. As it is usually the case for structural proteins, the expression of the tail fibres was detected at the beginning of the stationary growth phase of *Nab. magadii* (Rössler *et al.*, 2004). In the course of this thesis the relevance of the putative tail fibre proteins was studied to evidence a similar function in the infection process. In chapter 1.3.6, the putative tail fibre proteins are described in more detail.



In general all structural proteins of the halophage ϕ Ch1 are characterized by a low isoelectric point, thus enables the exposure to high-salt conditions and alkaline pH. As a consequence of this adaptation the phage particles completely dissociate, or changes their conformation in salt concentrations below 2M and therefore lose infectivity (Witte *et al.*, 1997).

1.3.3 Life cycle of the temperate ϕ Ch1

For the temperate halophage ϕ Ch1 two distinct life forms are known since virus particles were obtained after spontaneous lysis of *Nab. magadii* at stationary growth phase. This suggests a lysogenic state of the virus and the existence of a prophage state within the cell. Virus DNA can either exist as an episomal state, as it has been reported for ϕ H in *Halobacterium salinarum*, or it can integrate into the host chromosome where it is preserved and propagated together with the cellular DNA by cell division. To distinguish between these two possibilities of the physical state of ϕ Ch1 in *Nab. magadii* hybridization experiments were confirmed. It was determined that the prophage of ϕ Ch1 is chromosomally integrated (Witte, *et al.*, 1997). For this lifestyle it is essential to have a precise regulation control of the lysogenic state versus the lytic life cycle. The two open reading frames ORF48 and ORF49 have been identified as putative repressors

which are involved in preventing early onset of the lytic cycle during the lysogenic state (Iro, *et al.*, 2007). ORF48 shows sequence similarities to the known repressor molecule of the related virus ϕ H. Due to the fact that ORF48 is expressed constitutively during the whole life cycle, it cannot be a regulator element defining the time point of lysis. In contrast the expression of ORF49 is steadily increasing over the time during the virus life cycle. Thus the gene product of ORF49 probably acts as a transcriptional repressor, which is involved in triggering the lytic cycle or as a factor that speeds up the process of the lytic life state (Iro, *et al.*, 2007).

1.3.4 Genomic organization of ϕ Ch1

The mature and infectious phage particles of ϕ Ch1 contain both double stranded linear DNA (58 498 base pairs) as well as several RNA species (80 – 700 nucleotides) (Klein *et al.*, 2002; Baranyi *et al.*, 2000). It is suggested that the RNA, ranging in the size of 5S rRNA, is probably host specific and encoded by the genome of *Nab. magadii* (Witte *et al.*, 1997). Restriction analysis by isoschizomeric enzymes (*Sau3A*, *DpnI* and *MboI*) revealed that the viral DNA possesses modified bases. The results show that the DNA of ϕ Ch1 consist Dam-like methylation of adenine residues within the recognition site 5'-GATC-3', but no modifications of cytosine residues were observed. Furthermore it was clearly shown that the phage DNA is only partially methylated and the rest remains unmodified. To estimate the G+C content of the virus DNA a high pressure liquid chromatography (HPLC) was performed and revealed a value of approx. 62% (Witte *et al.*, 1997). The DNA of ϕ Ch1 is both terminally redundant and circularly permuted; therefore it seems that the virions are “stuffed” by the headful packaging mechanism (Klein *et al.*, 2002). When the genome was searched for ORFs consisting at least 30 codons with ATG or GTG start codons, 98 ORFs out of a total of 570 potential ORFs were observed and predicted to be genes. Most of the proteins encoding ORFs, apart four are starting with GTG (ORFs 3, 41, 79 and 83, respectively), were predicted to start with ATG. Comparative analysis of the ϕ Ch1 genome with available sequences in the database and the partially sequenced relative ϕ H results a total of 48 matches. Only 17 ORFs were similar to proteins with a known function. The close arrangement into the same direction or the few nucleotides overlap of many ORFs, suggests that they may form transcriptional units (Klein *et al.*, 2002).

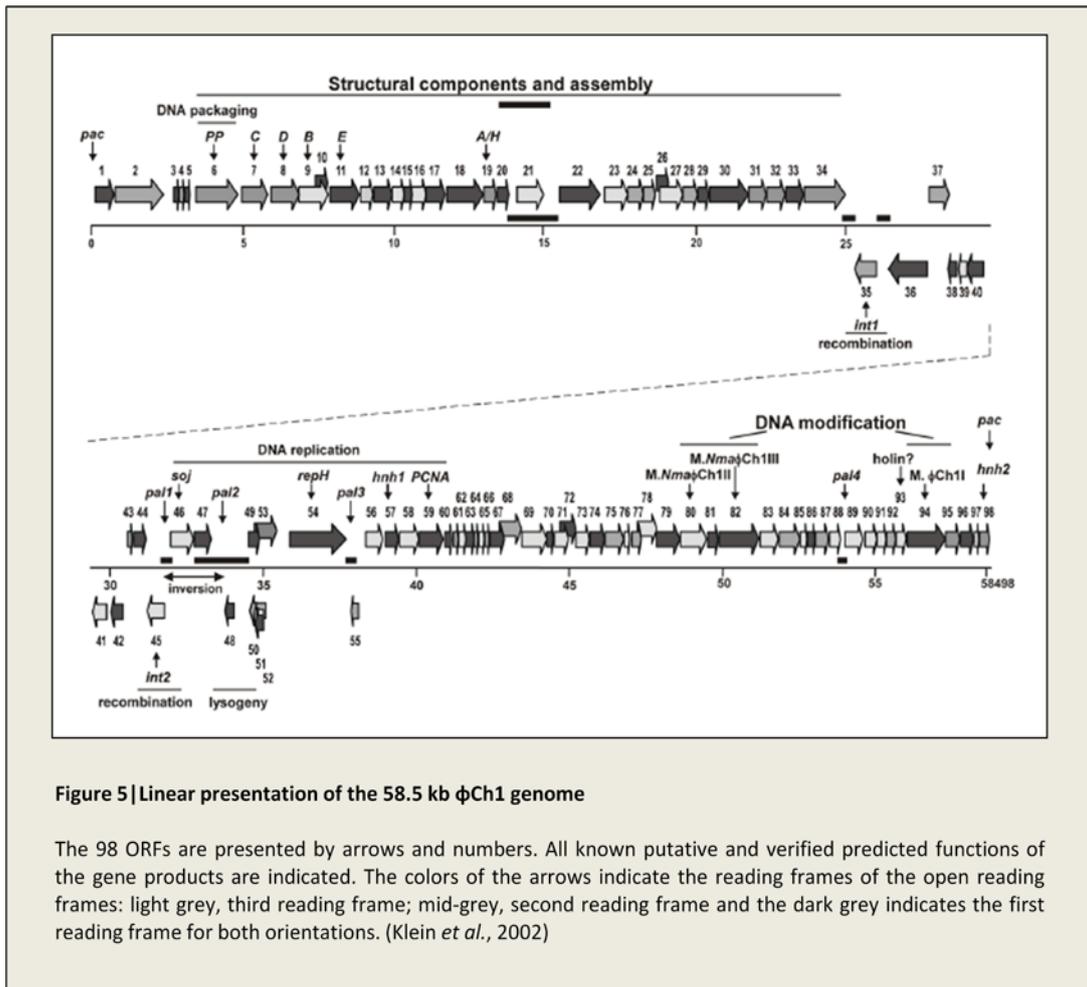


Figure 5 | Linear presentation of the 58.5 kb ϕ Ch1 genome

The 98 ORFs are presented by arrows and numbers. All known putative and verified predicted functions of the gene products are indicated. The colors of the arrows indicate the reading frames of the open reading frames: light grey, third reading frame; mid-grey, second reading frame and the dark grey indicates the first reading frame for both orientations. (Klein *et al.*, 2002)

As indicated in the figure 5 the left part of the genome solely comprises rightward-transcribed genes (ORFs 1-34), whereas the middle part is built up of a mixture of leftward- and rightward-transcribed genes (ORFs 35-55) and the right part again harbors only rightward-transcribed genes (ORFs 56-98).

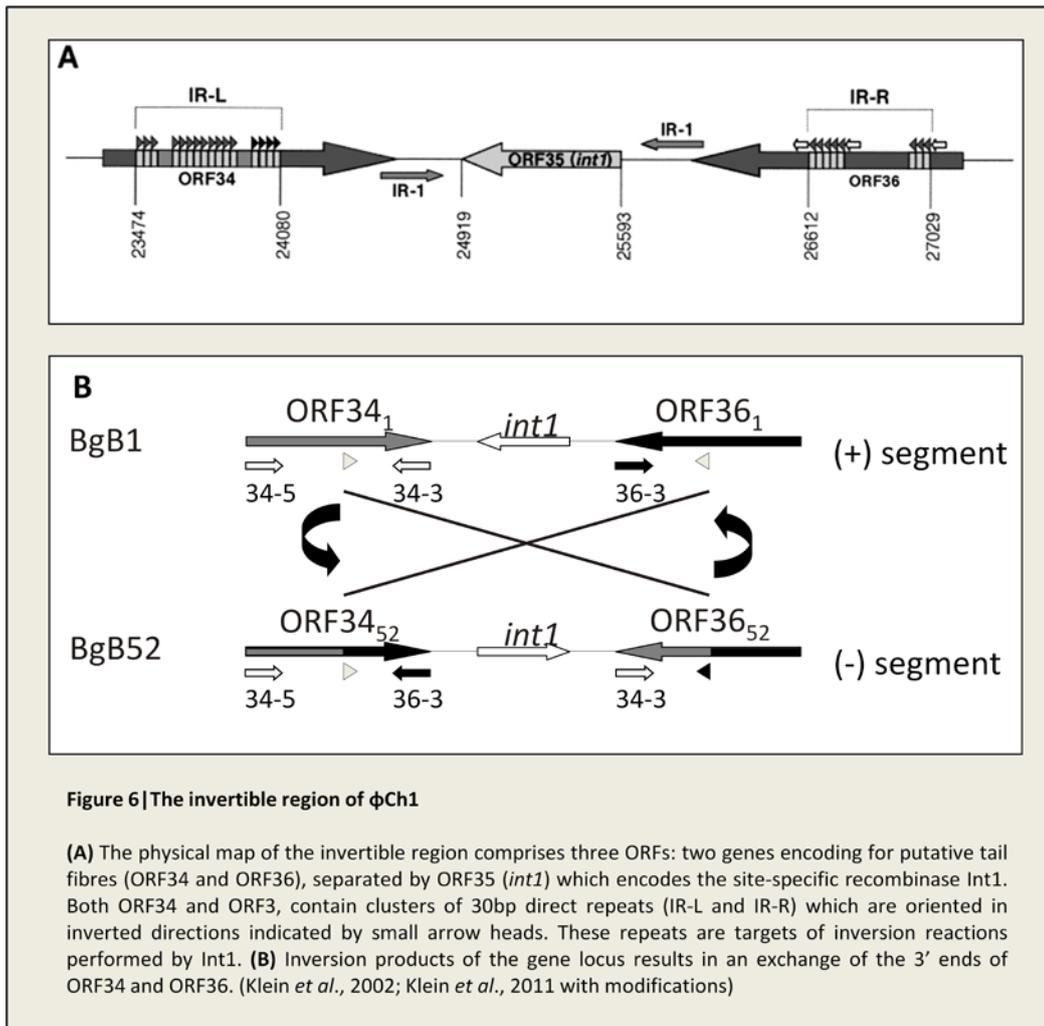
The sequenced genome structured into three parts, mirroring the proposed functions of the encoded proteins. While the left part encodes structural compounds and virion morphogenesis, the right part is predominant by proteins with currently unknown functions and in addition, enzymes performing DNA modification and restriction. The ORFs situated in the central part of the genome include genes for replication, plasmid stabilization and gene regulation.

The central part of the ϕ Ch1 genome resembles the invertible L-segment from ϕ H, which enables circularization and replication as an autosomal plasmid (Gropp *et al.*, 1992). Surprisingly they share a sequence similarity of 50 – 97 %, since the two hosts of ϕ Ch1 and ϕ H

live in distinct habitats with respect to the different pH values required for growth. However, the host range of ϕ Ch1 does not include *Hbt. salinarum* but it could be suggested, that the virus either loss infectivity or the capability to produce progeny viruses in *Hbt. salinarum*. The segmented composition of the genome of the head tail virus ϕ Ch1 is not only shared with ϕ H but also with a number of other bacteriophages and may come as a result of “module exchange” during the course of phage evolution (Hendrix *et al.*, 2000). Moreover putative site-specific recombinases (*int1* and *int2*) of the λ integrase type were discovered by sequencing analysis of the ϕ Ch1 genome (Klein *et al.*, 2000). The position of the ORFs within the genome and the repetitious amino acid stretches are similar to tail fibre proteins of the bacteriophage λ . This crucial sequence segment is known to perform inversion reactions leading to a variation of structural proteins. In the following chapter this region will be described more precise.

1.3.5 The invertible region of the ϕ Ch1 genome

The invertible region located in the middle part of the ϕ Ch1 genome contains the recombinase encoded by the ORF 35 and two flanking ORFs (34 and 36) (Rössler *et al.*, 2004). Another site-specific recombinase of the same type is encoded by the ORF45 (*Int2*) (Klein *et al.*, 2002). The integrases belong to the λ integrase family and are involved in processes such as transposition, integration of viral or plasmid DNA into the host genome, resolution of circular forms as well as DNA excision or inversion due recombination (Hallet and Sherratt, 1997). Sequence similarities to known integrases include the four high conserved amino acids R-H-R-Y (Arg-His-Arg-Tyr), forming a tetrad at the C-terminal region (Argos *et al.*, 1986; Abremski and Hoess, 1992; Blakely and Sherratt, 1996). As shown for other members of the λ -*int* family, the sequence is structured in three separate regions on the polypeptide strand. The BoxA segment has conserved arginine and the Box B contains the highly conserved motif His-X-X-Arg, they are intermitted with a 123 amino acids spacer. The catalytic tyrosine which is involved in strand cleavage and rejoining, is found at the distance of 25 amino acids from BoxB; in the so called BoxC (Hallet and Sherratt, 1997). All boxes come close together during folding of the protein, providing the four certain amino acids R-H-R-Y (Rössler *et al.*, 2004). The suggested function is that the integrase 1 is involved in phase variation reactions by performing inversion within the neighboring ORFs flanking its gene region. It is known that the flanking ORFs, orientated in opposite directions, encode structural proteins. As implied in the figure 6 the ORF34 is transcribed rightward in contrast to the ORF36 and the *int1* gene, which are transcribed leftward (Klein *et al.*, 2002).



In more detail the ORFs 34 and 36 are built up of cluster composed of various 30bp indirect repeats and one final inverted repeat: IR-L (within ORF34) and IR-R (within ORF36) (Klein *et al.*, 2002). Only ORF34 possesses a functional promoter and expression was detected in the lysogenic strain *N. magadii* L11 at the end of the logarithmic growth phase, about 98 h after inoculation (Klein *et al.*, 2012; Rössler *et al.*, 2004). This was confirmed by reverse transcription PCR using specific primers for ORF34 and ORF36. ORF36 could not be detected. The recombination events performed by the Int1 results in inversion of the gene region, which leads to the exchange of the 3'-end of ORF 34 and ORF36. Therefore a pool of heterogenous proteins of gp34 and gp36 with variable length and/or distinct C-termini are presented (Rössler *et al.*, 2004). Moreover this is a reason why different kinds of this gene region, differing in their orientation, are found in the isolated DNA of the mature phages. As already mentioned inversion leads to different versions of gene products and was worked out more in detail in 2004. Analyses were performed by cloning different versions of the *Bgl*II-B fragments in *E. coli*. This *Bgl*II-B fragment comprises 5995bp of the ϕ Ch1 genome and includes the region from

ORF31 to ORF36. About 60 different clones were found in *E.coli* and five of them were analyzed in detail. The BgB1, BgB5 and BgB51 share an (+) orientation of the *int1* gene and resembles the non-inverted sequence. Both products BgB43 and BgB52 show an exchange of the 3' ends of ORF34 and ORF36 and an (-) orientated *int1* gene. As representatives a (+) and (-) orientated variant of the invertible region was selected for further experiments (BgB1 and BgB52). However, all observations lead to the assumption that ORF34 encodes tail fibre proteins, which determines the host range of tailed phages such as observed in P1 and Mu (Iida, 1984; Sandmeier, 1994). PCR analysis show that the non-inverted (+) fragment yielded a product in both clones at all time points of the growth phase. Due to the fact that the PCR product resulting from inversion was first detected after a growth time of 78 hours, it is supposed that inversion happens in the stationary growth phase of *Nab. magadii* L13 (Ladurner, 2008). This correlates to the observations in the wild type strain L11, which shows expression of the integrase during the early logarithmic growth phase (Rössler *et al.*, 2004). The demonstration that inversion occurred and that *int1* is expressed early, leads to the assumption that another proteins is involved as activator in the lysogenic strain L11; but has not been confirmed so far.

1.3.6 Aspects of the putative tail fibre proteins

The putative tail fibre proteins result from an exchange of the 3' ends of ORFs 34 and 36 and are translated as proteins with a range of 41-54 kDa. Both proteins are characterized by a low isoelectric point of 3.7 and by an expression about 98h after inoculation of the strain *Nab. magadii* L11, similar to other halophilic structural proteins. Detection by western blot analysis using an antiserum against mature ϕ Ch1 particles (α - ϕ Ch1) and an antibody specifically raised against gp36 (α -gp36) confirmed the assumption that both proteins are part of the mature virus particle. The latter detection strategy show signals for both proteins (ORF 34 and 36) at 66 kDa, explained by the occurrence of extensive clusters of 30bp direct repeats (10 amino acid) found in both ORFs and with an identical range of 33% at the DNA and amino acid level (E-value: 1×10^{-42}) (Rössler *et al.*, 2004; Klein *et al.*, 2002). Those specific sequences are highly reminiscent of tail fibre proteins of certain bacteriophages. For example, the central shaft domain of the bacteriophages λ and T4 frequently contain repetitive sequence motifs with invariant residues and conserved pattern of hydrophobic amino acids, which results in formation of β -structural elements (Klein *et al.*, 2012; Mitraki *et al.*, 2002; Sandmeier *et al.*, 1994). As already mentioned, the invertible region is built up of two ORFs (34 and 35) orientated in an inverse direction with their 3' end facing each other, separated by ORF35

which encodes the putative site specific recombinase. This entire region tends to allow rearrangement which obviously occurs through the highly similar repeat clusters within the ORF34 and ORF36. As a result of those recombination events, variants of ORF34 and ORF36 containing different numbers of repeats are arising. Accessorily, these recombination occurrences can lead to an inversion of the intervening fragment that consequently causes not only an inversion of the *int1* gene (ORF35) but also can produce two different variants of gp34; the unmodified gp34₁ and one version carrying the C-terminus of gp36 called gp34₅₂ (Klein *et al.*, 2012). According to that, a heterogenous pool of gp34 and gp36 proteins with variable length and/or distinct C-termini is presented. In several bacteriophages like Mu and P1, this feature has been shown to function as a host - range switch, since tail fibres are widely used for interaction with the host cell receptor to enable attachment for following infection (Sandmeier *et al.*, 1994). After sequencing analysis, which indicates the pool of different DNA fragments of ORF34 and ORF36 from a ϕ Ch1 population, two orientations were selected for further investigations (BgB1 and BgB52) (see figure 6). To determine the time point of expression in *Nab. magadii* L11 specific primers for the BgB1-type ORF34 and ORF36 (presented in [+] orientation) were designed and used for reverse transcription-PCR. During the early and mid -logarithmic growth phase no transcript was detected, however, in the transition of the late growth phase and after the onset of lysis the specific ORF34 fragment could be investigated. Although no specific ORF36 fragment appeared at this time point. The same procedure were done to investigate transcripts originating from the inverted ORF34 and ORF36 in BgB52 ([-] orientation) and confirmed that only ORF34 transcription was detectable. These investigations lead to the assumption that only the transcribed ORF34 possesses a promoter sequence, which was further investigated due expression analysis in *Hfx. Volcanii* (a model system for extreme halophilic *Archaea*) and primer extension analysis.

To verify the function of the putative tail fibres and their ability to bind to the cell surface of *Nab. magadii* in vitro binding assays were performed. Therefore gp34₁ and gp34₅₂ expressing *Hfx. volcanii* extracts were incubated with *Nab. magadii* cells. The components bound to the cell surface were separated from the unbound proteins by centrifugation, and the presence of gp34₁ and gp34₅₂ in the resulting cell fraction as well as in the supernatant was detected by western blot analysis using the antibody α -gp36. It was shown that gp34₅₂, but not gp34₁, was able to attach to specific cell surface structures of *Nab. magadii* (Klein *et al.*, 2012). Also the binding assays of the same proteins which were expressed and purified from *E. coli*, supported this discovery. Besides the findings further experiments indicated that the C-terminal but not the N-terminal truncation of gp34₅₂ abolished the attachment to the host cell and represents

the functional binding domain. While the N-terminal part is supposed to be associated with the viral capsid and the shaft of the tail it seems to be negligible for the actual binding process. Moreover it was found that both gp34₁ and gp34₅₂ contain predicted galactose-binding domains within their C-termini. This suggests that a pretended galactose residue is located on the *Nab. magadii* cell surface and therefore involved in binding. It could be shown that α -D-galactose dramatically inhibited the attachment of viral particles to the cell surface, hence preventing infectivity of ϕ Ch1. Assumedly the galactose is essential for ejection of viral DNA and requires an interaction between additional viral or host factors, or the galactose-containing moiety on the cell surface conduces as a co-receptor and is only required for the initial attachment to the cell surface (Klein *et al.*, 2012). This view also correlates with the fact that bacteriophages of gram-positive bacteria use cell surface carbonates as receptors for attachment (Lindberg, 1973). The carbohydrates are found in the peptidoglycan, teichoic acid and teichuronic acid, associated with the surface layer (S-layer) as part of the cell wall. *Nab. magadii* does not own a peptidoglycan layer but possesses S-layers consisting glycoproteins on the cell surface (Eichler, 2003). Alternatively, it was shown that two out of four flagella proteins (FM2 and FM4) from *Nab. magadii* are glycosylated and may also mediate the binding of ϕ Ch1 (Fedorov *et al.*, 1994; Serganova *et al.*, 2002). In this manner the crucial role of the ϕ Ch1 proteins, involved in infection of *Nab. magadii* is very likely, but has to be further confirmed. Regarding to this a mutant of ϕ Ch1 lacking the protein gp34₅₂ was constructed in course of this diploma thesis.

1.3.7 Virus – host interaction

The best studied phage-host interactions were carried out on bacteriophages infecting *E. coli* such as coliphage λ and T-even viruses. Bacteriophages show structure components which are unique among viruses in that most of them have specialized tails for host cell attachment and are generally called *Caudovirale*. Accordingly to their tail morphology this family is divided into three sub-families: *Myoviridae* (T-even bacteriophages with long contractile tails), *Siphoviridae* (lambdoid phage with a long non-contractile tail), and the *Podoviridae* (P22 phage with a short non-contractile tail). The most complex attachment structures with the greatest number of proteins which are involved in the tail constitute and function are represented by the *Myoviridae* phages, including the halophage ϕ Ch1 (Leimann *et al.*, 2010). However, infection processes are highly diverse and rather little is known about the way of infection, performed by archaeal viruses. The initialization of the adsorption process is characterized by the recognition of some structural components on the host cell surface through virus tail fibre

proteins or analogous structures, for viruses lacking these specified structures. For an irreversible binding of phage particles to the receptor on the host cell, it is essential that a sufficient amount of long tail fibres binds to the base plate, which is composed of base-plate parts (Iwashita & Kanegasaki, 1976). Cascades of molecular changes are responsible for the following introduction of DNA from the virus into the microbial cell.

As a representative of the family *Myoviridae*, the T4-type phage which was classified on basis of morphological criteria, will be used for explanation of adhesion. All of them consist of an elongated head structure and the complex tail is built up by a sheath, an internal tail tube and a baseplate which is situated at the distal end of the tail. Two different types of fibers, the long tail fibers and the short tail fibers, which are attached to the baseplate, are responsible for host cell recognition and binding. In the act of attachment to the host cell, the tail undergoes a large conformational change where the baseplate opens up, similar to a flower, while the shaft contracts and the internal tube is pushed through the baseplate, penetrating the outer membrane of the host, creating a channel for DNA ejection from the phage capsid into the host cell (Leimann *et al.*, 2010). The T4 baseplate is significantly more complex than other well studied phages with contractible tails, like phage Mu or P2, so I will go through the infection process in a little bit more detail. During infection the phage attaches the receptor pattern on the host cell, using the long tail fibres (LTFs) which are connected to the baseplate. The recognition signal is transmitted through the LTFs to the baseplate attachment protein (gp9) and then to the baseplate itself. Thereafter the baseplate anchors to the lipopolysaccharide cell surface receptors with help of the short tail fibres (STF) which are initially assembled under the baseplate. This event leads to the conformational change in the baseplate, which is coupled to that of the sheath. This strength, irreversible interaction is a striking feature of tailed bacteriophages and bound phages can not be separated from their host, generally (Grahm *et al.*, 1999). The event described before, triggers a conformation change from hexagonal to star-shaped conformation and drives the head closer to the cell surface, caused by contraction of the tail sheath. This exerts a force, transmitted through gp27 cylinder and the N-terminal domain of gp5 to the β -helix needle, onto the tail tube toward the cell membrane; thereby the latter two puncture the outer cell membrane of the cell. When the β -helix needle enlarges the pore in the membrane and comes into contact with the periplasmic peptidoglycan layer it dissociates from the tip of the tube and activates the lysozyme domain gp5. This lysozyme digests the cell wall and allows the tail tube to penetrate the inner membrane. A trimer formation of gp27, forming the tip of the tail tube, is thought to interact

with the specific receptor molecule on the cytoplasmic membrane and initiates the release of the DNA from the phage capsid into the host. Nevertheless, the DNA is not released until the tail tube tip binds to a cytoplasmic membrane receptor, suggesting that tail contraction does not cause the release of DNA. During that transfer process the membrane remains virtually undamaged, since the proton motive force is required for the transfer of the DNA.

The exact formation and structure of the long tail fibres are built up as described in the following way. Each of them consists of the rigid proximal and distal part which is connected by a hinge region, whereby the gp34 trimer forms the proximal part and the distal region is composed of the trimeric gp36 and gp37 together with the monomeric gp35. The baseplate binding bulge is formed by the N-termini of the gp34 trimer and its C-termini, gp35, and parts of the gp36 trimer are components of the hinge formation. The already mentioned proteins are the major player in the host range determination. The fibre is based on a repeating motif involving the proteins gp34, gp37 and gp12 (STFs) and additive intervening parts of different length. However, although a broad sequence diversity of T-even tail fibre loci are presented, functionally analogous proteins in T2, T4 and T6 are discovered. Generally, viral tail fibre genes are mosaic structured and are composed of elements that may be exchanged among related, as well as unrelated phages. As a consequent result the adhesions have a high degree of sequence divergence and so the individual phages vary greatly in the bacterial recognition receptors. Bacteriophages have to transport their genetic material across the bacterial cell envelope to ensure a proper infection. In case of infecting Gram-negative bacteria, phages must penetrate through the outer membrane which is composed of lipopolysaccharide, phospholipids and several proteins. In some cases phages require molecules of both types for adhesion (Lindberg, 1973). When phages infect Gram-positive bacteria, they have to overcome a thick peptidoglycan layer and a cytoplasmic membrane. All these structural components can act as phage receptors. In the interests of easy accessibility, bacteriophages tend to use structures exposed on the outer membrane of the host *Bacteria* as a receptor. Though the fact, those archaeal viruses clearly differ from bacterial phages, some viruses are associated with members of the *Euarchaeota* that belong to the myovirus and siphovirus families. Since myoviruses and siphoviruses can infect *Bacteria* and some existing mesophilic members of the *Euryarchaeota* can colonize humans and other animals, more closely investigations are necessary for investigation of potential similarities between archaeal viruses and bacterial phages. To date, there are no reports of any archaeal virus that can infect *Bacteria*, or conversely, no report of bacterial phages that are able to infect *Archaea*. Receptors for archaeal viruses still have not been well studied and characterized. Despite the differences in

cell wall and membranes of *Archaea* and *Bacteria*, the archaeal flagella also appear to have a different ancestry and possess many divergent characteristics. Therefore, the structures that bacterial phages recognize and bind to are largely absent from archaeal cells, it seems unlikely that bacterial phages would be able to attach archaeal cells. However, due to the fact that phages can expand their host range by modification of the phage anti-receptor, it is unclear whether anti-receptor mutation would allow phages to bind to an archaeal receptor. It is also possible that bacterial phages which are attached to the host cell, are incapable of transferring their DNA into archaeal cells because of intrinsic differences in DNA processing/translocation machinery (Gill & Brinkman, 2011).

2. Materials and methods

2.1 Materials:

2.1.1 Strains

Domain	Strain	Genotype	Reference
Bacteria	<i>E. coli</i> XL-1-Blue	<i>endA1, gyrA96, hsdR17 (r_k-m_k+), lac, recA1, relA1, supE44, thi,</i>	Stratagene
	<i>E. coli</i> GM48F'	<i>F', thr, leu, thi-1, lacY, galK, galT, ara, tonA, tsx, dam, dcm, supE44</i>	Palmer and Marinus, 1994
Archaea	<i>Nab. magadii</i> L11	wild type, integrated prophage ϕ Ch1	Witte <i>et al.</i> , 1997
	<i>Nab. magadii</i> L13	ϕ Ch1 cured derivate of L11	Witte <i>et al.</i> , 1997
	<i>Nab. magadii</i> L11 Δ 34	ORF34 deficient derivate of L11	this thesis
	<i>Nab. magadii</i> P3	NEP deficient derivate of L13	Derntl, 2009

2.1.2 Growth media

LB medium

Peptone 10 g
Yeast extract 5 g
NaCl 5 g

pH 7.0

add dH₂O to a final volume of 1 l

15 g/l agar were added for plates

NVM+ (rich medium for haloalkaliphilic *Archaea*)

Casamino acids	8.8 g
Yeast extract	11.7 g
Tri-Na citrate	0.8 g
KCl	2.35 g
NaCl	235 g

pH 9.0

add dH₂O to a final volume of 934ml (8 g/l agar were added for plates, 4 g/l agar were added for soft agar)

After autoclaving the medium or the agar were complemented with:

0.57 M Na ₂ CO ₃ (dissolved in steril dH ₂ O)	65 ml
1 M MgSO ₄ (autoclaved)	1 ml
20 mM FeSO ₄ (dissolved in steril dH ₂ O)	1 ml

NMMb+ (mineral medium for haloalkaliphilic *Archaea*)

Na ₂ HPO ₄	0.28 g
NaH ₂ PO ₄	0.28 g
Alanine	2.23 g
Leucine	0.66 g
Arginine	0.81 g
Histidine	0.778 g
Lysine	0.731 g
Na ₃ -citrate dihydrate	0.8 g
Na acetate	2.23 g
KCl	1.66 g
Na pyruvate	1.1 g
NaCl	205 g

pH 9.0 – 9.5

add dH₂O to a final volume of 900ml

8 g/l agar were added for plates

After autoclaving the medium or the agar were complemented with:

1.75 M Na ₂ CO ₃ (dissolved in steril dH ₂ O)	100 ml
1 M MgSO ₄ (autoclaved)	1 ml
20 mM FeSO ₄ (dissolved in steril dH ₂ O)	0.25 ml
1000 x trace elements	1 ml

(trace elements: MnCl₃ 4 mM, CaCl₂ 3 mM, CuSO₄ 4 mM, ZnSO₄ 3 mM add dH₂O to 100 ml final volume)

2.1.3 Additives

<i>Domain</i>	<i>Additive</i>	<i>Stock concentration</i>	<i>Final concentration</i>	<i>Annotations</i>
<i>Bacteria</i>	Ampicillin	20 mg/ml	100 µg/ml	dissolved in ddH ₂ O, sterile filtered, stored at 4°C
	Tetracycline	10 mg/ml	10 µg/ml	dissolved in ½ Vol. ddH ₂ O, completed by ½ Vol. 96 % EtOH, stored at -20°C (light protected)
<i>Archaea</i>	Bacitracin	7 mg/ml	7 µg/ml	dissolved in ddH ₂ O, sterile filtered, stored at 4°C
	Mevinolin	10 mg/ml	7,5 µg/ml	isolated from pulverized tablets, dissolved in 96 % EtOH, stored at -20°C
	Novobiocin	3 mg/ml	3 µg/ml	dissolved in ddH ₂ O, sterile filtered, stored at -20°C (light protected)

2.1.4 Vectors

<i>Plasmid</i>	<i>Features</i>	<i>Reference</i>
pUC19	bla, pMB1ori, lacZa, mcs	Yanisch-Perron <i>et al.</i> , 1985
pBluescript II KS+	mcs, bla, ColE1 ori, lacZa	Stratagene
pBAD24	bla, araC, rrnB, mcs, PBAD promoter, pBR322 ori	Guzman <i>et al.</i> , 1995
pRo-5	bla, ColE1 ori, gyrB (NovR), ϕCh1 derived ori	Mayrhofer-Iro <i>et al.</i> , 2013
pNB102	bla, ColE1 ori, hmg (MevR), pNB101 ori	Zhou <i>et al.</i> , 2004

pMDS11	bla, f1 ori, ColE1 ori, gyrB (NovR), pHK2 ori	Holmes <i>et al.</i> , 1991
pBGB1	pKS _{II} + (mcs, bla, ColE1 ori, lacZa)with ϕ Ch1 BglII-B fragment (ORF31-ORF36) in (+) orientation	Rössler <i>et al.</i> , 2004
pBGB52	pKS _{II} + (mcs, bla, ColE1 ori, lacZa)with ϕ Ch1 BglII-B fragment (ORF31-ORF36) in (-) orientation	Rössler <i>et al.</i> , 2004
pKS _{II} + Δ ORF34-fragment 1	pKS _{II} + (mcs, bla, ColE1 ori, lacZa) with upstream region of ϕ Ch1 ORF34	this thesis
pKS _{II} + Δ ORF34-fragment 1+2	pKS _{II} + (mcs, bla, ColE1 ori, lacZa) with upstream and downstream regions of ϕ Ch1 ORF34	this thesis
pKS _{II} + Δ ORF34-Novobiocin deletion cassette	pKS _{II} + (mcs, bla, ColE1 ori, lacZa) with upstream and downstream regions of ϕ Ch1 ORF34, disrupted by gyrB (NovR, derived from pMDS11)	this thesis
pNB102-ORF34 ₁	pNB102 (bla, ColE1 ori, hmg (MevR), pNB101 ori) with ϕ Ch1 ORF34 ₁	Petra Till, 2011
pNB102-ORF34 ₅₂	pNB102 (bla, ColE1 ori, hmg (MevR), pNB101 ori) with ϕ Ch1 ORF34 ₅₂	Petra Till, 2011
pUC19-nproNEP	(bla, pMB1ori, lacZa, mcs) promoter and N-terminus of Natrialba extracellular protease	this thesis
pUC19-nproNEP-ORF34 ₁	(bla, pMB1ori, lacZa, mcs) promoter and N-terminus of Natrialba extracellular protease with pBGB34 ₁	this thesis
pUC19-nproNEP-ORF34 ₅₂	(bla, pMB1ori, lacZa, mcs) promoter and N-terminus of Natrialba extracellular protease with pBGB34 ₅₂	this thesis
pNB102-nproNEP-ORF34 ₁	pNB102 (bla, ColE1 ori, hmg (MevR), pNB101 ori) with promoter and N-terminus of Natrialba extracellular protease with pBGB34 ₁	this thesis
pNB102-nproNEP-ORF34 ₅₂	pNB102 (bla, ColE1 ori, hmg (MevR), pNB101 ori) promoter and N-terminus of Natrialba extracellular protease with pBGB34 ₅₂	this thesis

2.1.5 Primer

<i>Name</i>	<i>Sequence*</i>	<i>Restriction site</i>
34-5	5'- CAG CAG <u>AGA TCT</u> ATG AGT AAA ATC TGG GAA CCG AG - 3'	<i>Bgl</i> II
36-3	5'- CAG CAG <u>AAG CTT</u> ATT CAG GTT TCA TGT CGC TG - 3'	<i>Hind</i> III
49-Hind3-3	5'- CAG <u>CAA GCT</u> TTC GAG GCG TCA TCC T - 3'	<i>Bgl</i> II
NB-1	5'- TCTACCGGGTGCTGAACG - 3'	
NB-2	5'- CGCTGATGTACGAACCGAG - 3'	
79-Pst	5'- GATA <u>CTGCAGCT</u> CTTTGTACCGATGCGTC - 3'	<i>Pst</i> I
79-Nae	5'- GATAGCCGGCGACTCTCACAAGATCTC - 3'	<i>Nae</i>
Nov-11	5'- GCA TGT CGT GGC TGT TCG - 3'	
pKS-5	5'- GTA GCT CTT GAT CCG GCA - 3'	
pKS-3	5'- GCC CTC CCG TAT CGT AGT - 3'	
34-XbaI-a	5'- GAC <u>GTC TAG</u> ACT CCG ATG AAC ACG ACA CTC - 3'	<i>Xba</i> I
Δ34-1	5'- GAC <u>CTC TAG</u> AGT ACC GAA CGC ATC TCG - 3'	<i>Xba</i> I
Δ 34-2	5'- GCA ACC CGG <u>GAA GCT</u> TCT CGT AGC TCT GGT TTT CCT - 3'	<i>Hind</i> III
Δ 34-3	5'- GAC <u>CAA GCT</u> TAA CTG ATC TTC ACA CCG GAT - 3'	<i>Hind</i> III
Δ 34-4	5'- GAA <u>AGG TAC</u> CGA GAG ACA TGC CCA CGA - 3'	<i>Kpn</i> I
Δ 34-2-Sma	5'- GCA <u>ACC CGG</u> GCT CGT AGC TCT GGT TTT CCT - 3'	<i>Sma</i> I
36-3X	5'- GCA <u>GTC TAG</u> ACC ATC GGT TAT TCG AGT TTC - 3'	<i>Xba</i> I
RSET-Bgl	5'- GAC <u>GAG ATC</u> TAT GCG GGG TTC TCA TCA TC - 3'	<i>Bgl</i> II
Npro-Bam	5'- GAT <u>CGG ATC</u> CGA CTT CCA GGG TCT CGT - 3'	<i>Bam</i> HI
Npro-Kpn	5'- GAT <u>TGG TAC</u> CGG ATT TTC GAC CAT AGT TG - 3'	<i>Kpn</i> I

* underlined sequences represent recognition sites for used restriction endonucleases

2.1.6 DNA and protein markers

2.1.6.1 DNA ladders

<i>Application</i>	<i>Marker</i>	<i>Fragments</i>	<i>Reference / Annotations</i>
Agarose gel (0.8%)	λ - <i>Bst</i> EII (<i>Eco</i> 91I)	8453, 7242, 6369, 5687, 4822, 4324, 3675, 2323, 1929, 1371, 1264, 702 [bp]	λ DNA and reaction buffer O was obtained from Fermentas; final concentration of digested DNA in a 1x loading dye was 50 ng/ μ l (Denaturation of cos- fragments at 65° C for 20 min)
	λ - <i>Pst</i> I	11501, 5077, 4749, 2838, 2556, 2443, 2140, 1986, 1700, 1159, 1093, 805, 514, 468, 448, 339, 264, 247, 216, 211, 200, 164, 150, 94, 87, 72, 15 [bp]	preparation as before
PAA gel (6%)	pUC19 <i>Hae</i> III (<i>Bsu</i> R1)	587, 458, 434, 298, 257, 174, 102, 80, 18, 11 [bp]	pUC19 plasmid DNA digested in buffer R (Fermentas); final concentration in a 1x loading dye was 50 ng/ μ l
	pUC19 <i>Sau</i> 3AI	955, 585, 341, 258, 141, 105, 78/75, 46, 36, 18/17, 12/11.8 [bp]	pUC19 plasmid DNA, digested in NEBuffer 1 (New England BioLabs #R0169S)
Southern blot	Biotinylated 2-Log DNA Ladder	10, 8, 6, 5, 4, 3, 2, 1.5, 1.2, 1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, 0.1 [kbp]	New England BioLabs #N7554S

2.1.6.2 Protein ladders

Marker	Fragments	Reference
Unstained Protein Molecular Weight Marker	116, 66.2, 45, 35, 25, 18.4, 14.4 [kDa]	Fermentas #SM0431
PageRuler™ Plus Prestained Protein Ladder	250, 130, 100, 70, 55, 35, 25, 15, 10 [kDa]	Fermentas #SM1811
Spectra™ Multicolor Broad Range Protein Ladder	260, 140, 100, 70, 50, 35, 25, 15, 10 [kDa]	Fermentas #SM1849

2.1.7 Enzymes and corresponding buffers

2.1.7.1 Restriction

All restriction enzymes were obtained from Fermentas and used with the provided buffers. Double digestions were performed in Tango buffer whenever possible or else with the recommended buffer as suggested on the Fermentas DoubleDigest™ webpage (<http://www.fermentas.com/en/tools/doubledigest>).

2.1.7.2 PCR

Polymerase	Reference
<i>Pwo</i> (from <i>Pyrococcus woessii</i>)	Peqlab (Cat. 01-5020)
<i>Pfu</i> (from <i>Pyrococcus furiosus</i>)	Promega (Cat. M776A)
<i>GoTaq</i> ®	Promega (#M31745)

2.1.7.3 Other enzymes

<i>Enzyme</i>	<i>Reference</i>
Klenow Fragment	Fermentas (#EP0051)*
Proteinase K	Qiagen® (Cat.No.19133)**
T4 DNA Ligase	Fermentas (#EL0011)*

*enzymes were used with the appropriate buffers and/or in concentrations recommended by the manufacturer

** Proteinase K was used in concentrations experimentally determined to be most suitable for spheroblasts formation (see section: 2.2.8.1)

2.1.8 Antibodies

2.1.8.1 Primary antibodies

<i>Antibody</i>	<i>Application</i>	<i>Reference</i>	<i>Remarks</i>
α -gp34 ₅₂ (rabbit)	antiserum rabbit 17 diluted 1:5000 in TBS, 0.3 % BSA, 0.02 % NaN ₃	Moravian-Biotechnology	purification of ϕ Ch1 gp34 (Petra Till, 2011) and immunization of rabbits
RGS°-His™ antibody BSA-free; Isotype: mouse IgG ₁	dilution varying from 1:500 to 1:5000 in TBS, 0.3 % BSA	Quiagen (Cat. No.34650)	

2.1.8.2 Secondary antibodies

<i>Antibody</i>	<i>Application</i>	<i>Reference</i>	<i>Annotations</i>
α -rabbit	antibody diluted 1:5000 in TBS	GE Healthcare (Cat. NA934V)	conjugated with horseradish peroxidase (HRP)
α -mouse Ig	antibody diluted 1:5000 in TBS	GE Healthcare (#NXA931)	HRP linked hole antibody (from sheep) GPR

2.1.9 KITS

All Kits were used with supplied buffers and solutions according to manufacturer's instructions.

<i>Application</i>	<i>Product</i>	<i>Reference</i>
DNA purification	QIA [®] Gel Extraction Kit: Buffers PB and PE	Qiagen (Cat. 28706)
DNA purification – elution from gel	QIA [®] Gel Extraction Kit: QG, PE and Elution Buffer	Qiagen (Cat. 28706)
Miniprep	Gene JET [™] Plasmid Miniprep Kit: Resuspension Solution, Lysis Solution, Neutralisation Solution, Wash Solution	Fermentas (#K0503)
Southern blot – labeling of the probe	NEBlot [®] Phototope [®] Kit	New England Biolabs, (#N7550S)
Southern blot – development of the blot	Phototope [™] -Star Detection Kit for Nucleic Acids	New England Biolabs (#N7020S)
Western blot – development	Super Signal [®] West Pico Chemiluminescent Substrate	Thermo Scientific (#34080)

2.1.10 Solutions and reagents for DNA methods

2.1.10.1 Gelelectrophoresis

50% TAE

2 M Tris base
1 M Acetic Acid
0.1 M EDTA
adjust pH 8.2 with HCl

5x DNA loading buffer

50 mM Tris base
0.1% SDS
0.05% (w/v) Bromphenol blue
0.05% (w/v) Xyclene cyanol
pH 8.2
After autoclaving sucrose is added to a final concentration of 25% (v/v).

Agarose gel

x % solid agarose
1 x TAE buffer

10x TBE

108 g Tris base
55 g boracic acid
0.5 M EDTA pH 8
adjust pH 8.0 with boracic acid

6 % Polyacrylamid gel

30 % Polyacrylamid
1 x TBE

30 % Polyacrylamid solution

29 % Acrylamid	1.2 ml
1 % N,N'- Bisacrylamid	4.8ml
10 % APS	60µl
TEMED	6µl

2.1.10.2 Gel extraction from polyacrylamide gels

Elutionbuffer

10 mM EDTA pH 8
10 mM MgAc.4H₂O
0.5 M NH₄OAc

2.1.10.3 Southern blot analysis

20x SSC

3 M NaCl
0.3 M Sodium citrate
adjust pH 7.2 with HCl

Blocking solution

125mM NaCl
17mM Na₂HPO₄
8mM NaH₂PO₄
173mM SDS
pH 7.2

10x Washing solution II

12.1 g Tris base
5.85 g NaCl
2.03 g MgCl₂
adjust pH 9.5

1x Washing solution I

1:10 dilution of Blocking solution

50x Denhardt' solution

1 g Ficoll 400
1 g Polyvinylpyrrolidone
1 g BSA
add ddH₂O to a final volume of 100 ml

Hybridisation buffer (store at -20°C)

5 ml ddH₂O
25 ml 20x SSC
10 ml 50x Denhardt's sol.
5 ml 10 % BSA
5 ml 1 M Na₂HPO₄
500 µl 20 % SDS
200 µl 0.5 M EDTA

Denaturing solution before blotting; (after transfer)

0.4 M NaOH / 0.6 M NaCl ; (0.4 M NaOH)

Neutralizing solution before blotting; (after transfer)

1.5 M NaCl / 0.5 M Tris-HCl (pH 7.5) ; (0.2 M Tris / HCl pH 7.5)

Prehybridization solution

100 µg /ml herring sperm in Hybridization buffer

Wash solutions after hybridization

1) 2x SSC / 0.1 % SDS (room temperature)

2) 0.1x SSC / 0.1 % SDS (65°C)

2.1.11 Solutions and reagents for proteins methods

2.1.11.1 SDS-PAGE and western blot analysis

2 x Protein sample buffer (Laemmli buffer)

0.12 mM Tris-HCl pH 6.8

17.4 % Glycerol

4 % SDS

2 % β-mercaptoethanol

0.02 % Bromphenol blue

10 x SDS running buffer

0.25 M Tris base

1.92 M Glycine

1 % SDS

30 % Polyacrylamid (29:1)

29 % Acryl amide

1 % N,N₁-methylene bisacrylamide

10 % APS

0.5 g Ammonium persulfate

add ddH₂O to a final volume of 5 ml

4 x Separation gel buffer

1.5 M Tris-HCl pH 8.8

0.4 % SDS

4 x Stacking gel buffer

0.5 M Tris-HCl pH 6.8

0.4 % SDS

Coomassie staining solution

25 % Methanol

10 % Acetic acid

0.15 % Coomassie Brilliant Blue R-250

Transblotbuffer

48 mM Tris base

39 mM Glycine

0.037 % SDS

20 % MetOH

10x TBS / TBS-T

8 % NaCl
3 % Tris base
0.2 % KCl
(5ml Tween 20)
adjust pH 8.0

Ponceau S staining solution

1 % Acetic acid
0.5 % Ponceau S

Harsh stripping solution

10 ml SDS 20 %
6,25 ml 1M Tris pH 6.8
0,8 ml β - mercaptoethanol
add ddH₂O to final volume of 100 ml

Destaining solution

10 % Acetic acid

2.1.11.2 Protein purification under denaturing conditions (*E.coli*)

Buffer B (lysis buffer)

100 mM NaH₂PO₄
10 mM Tris-HCl
8 M urea
adjust pH 8.0 with NaOH

Buffer C (washing buffer)

100 mM NaH₂PO₄
10 mM Tris-HCl
8 M urea
adjust pH 6.3 with HCl

Buffer D (Elution buffer)

100 mM NaH₂PO₄
10 mM Tris-HCl
8 M Urea
adjust pH 5.9 with HCl
(immediately before use)

Buffer E (Elution buffer)

100 mM NaH₂PO₄
10 mM Tris-HCl
8 M Urea
adjust pH 4.5 with HCl
(immediately before use)

10 x PBS

8% (w/v) NaCl
0.2 % (w/v) KCl
1.44 % (w/v) NaH₂O₄
0.24 % (w/v) NaH₂PO₄
pH 7.5

Dialysis buffer I

4M urea
2M NaCl
50mM Tris-HCl
pH 9.5

Dialysis buffer II

2M NaCl
50mM Tris-HCl
pH 9.5

2.1.12 Isolation of chromosomal DNA from *Nab. magadii*

required materials:

14mM Desoxycholat 1:1 Phenol / Chloroform Isopropanol EDTA (final concentration 10mM) Solid CsCl 10mg /ml Ethidium bromide Butanol	High alkaline salt solution 4M NaCl 50mM Tris base
--	--

2.1.12.1 Extraction of plasmid DNA from *Nab. magadii*

<u>Resuspension solution</u> 2M NaCl	<u>Lysis solution</u> 1 % SDS 0.2M NaOH	<u>Neutralization solution</u> 5M Potassium acetate
---	---	--

additional materials:

Isopropanol
70 % EtOH

2.1.12.2 Buffers and solution for phage methods

Isolation of phage particles

<u>High alkaline salt solution</u> 4 M NaCl 50 mM Tris base	<u>Salt buffer solution</u> 2M NaCl 50mM Tris base pH: 8.5 - 9	PEG 6000
<u>Solution 1.1</u> 20 g CsCl 200 ml salt buffer	<u>Solution 1.3</u> 90 g CsCl 200ml salt buffer	<u>Solution 1.5</u> 135 g CsCl 200ml salt buffer

2.1.12.3 Precipitation of phage DNA

required materials:

1:1 Phenol Chloroform
96 % EtOH
70 % EtOH

2.1.12.4 Precipitation of phage proteins

required materials

50 % TCA
ammonia

2.2 Methods

2.2.1 DNA gel electrophoresis

Generally for analysing DNA probes, concerning fragments > 600 bp, an agarose concentration of 0.8 % was chosen whereas smaller fragments were usually separated on a 6% PAA gel in electric fields.

2.2.2 Agarose gel electrophoresis

The agarose in 1 x TAE buffer the gel was completely covered with 1x TAE buffer after solidification and the fitted in combs were removed. Separation of the DNA fragments according to size was performed by applying the sample and the migration behaviour of bromphenol blue and xylene cyanol of the 5x DNA loading dye was used to estimate DNA localisation. An electric potential of about 6V/cm of gel was applied. Afterwards the gel was stained in an ethidium bromide bath (10µg/ml) and visualized under a UV – transilluminator.

2.2.3 6% Polyacrylamide gel electrophoresis

PAA gel electrophoresis was used for DNA fragments <600 bp. For casting of a gel the equipment of the company Bio-Rad (Mini-PROTEAN[®]) was applied and used as suggested in the protocol. A 6 % PAA gel was mixed as described in section 1.1.10.1. The gel was poured between two glass plates fixed in the casting system and covered with 1 x TBE buffer. DNA samples were mixed with 5 x DNA loading buffer in the appropriate ratio and loaded onto the gel. A current strength of 20 mA/ gel was applied for about 20 minutes. Afterwards, the gel was stained in an ethidium bromide solution and visualised under UV light.

2.2.4 Polymerase chain reaction (PCR)

This method allows the amplification of a definite DNA sequence, of which the 5' and 3' ends have to be known. This information is essential for the design of primers. It was performed to amplify DNA to increase material for cloning strategies and for certification of the presence of certain sequences in probes, isolates from living cells. Applying to the purpose, different templates and polymerases were used. The primers define the sequence to be amplified by binding to the 3' and 5' end of the target region of interest, therefore acting as the starting

point for the DNA polymerase. This oligonucleotides were obtained from VBC genomics, delivered in a lyophilized mode and were solved in ddH₂O gaining 1µg/µl stocks, which were diluted 1:10 for final usage. The melting temperatures (T_M), based on the G+C content of the primers, were calculated using the program Gen Runner Version 3.01 from Hastings Software and the primer annealing temperature was usually set 4°C lower than the T_M. The values of probe concentration and salt concentration were adjusted to 25nM and 50mM, accordingly. Additional PCR conditions such as the elongation temperature and time were adapted to the specific activity and efficiency of the used polymerases, and the length of amplified sequences.

2.2.4.1 Preparative PCR:

In this case recombinant polymerases with proofreading activity, as characteristics of *Pwo* or *Pfu*, were applied in the recommended and supplied complete buffers containing KCl and MgSO₄. As a template, either purified plasmid DNA from *E. coli*, chromosomal DNA or φCh1-DNA, diluted 1:30 was used. Raw extracts of archaeal strains were generated by centrifugation of 25µl of the culture, removing the supernatant and lyse the cells by adding 100µl ddH₂O and briefly vortexing.

PCR batch *Pwo/Pfu* (100µl):

68µl ddH₂O
 10µl 10x *Pwo/Pfu* buffer
 10µl dNTPs (2mM)
 5µl primer forward (0.1µg/µl)
 5µl primer backward (0.1µg/µl)
 1µl template DNA
 2µl polymerase (*Pwo/Pfu*)

PCR programm:

1. 95°C 5'
2. 95°C 1' denaturation
3. X°C* 30'' - 1' primer annealing
4. Y°C* 2' per 1kb product elongation)
5. Y°C* 2- fold time of elongation
6. 4°C infinite

} 33 cycles

*X : individually calculated annealing temperature of the primer with a lower T_M

*Y: The elongation temperature, corresponding to the optimal activity of the *Pwo* polymerase (68°C) and the *Pfu* polymerase (75°C), was chosen.

2.2.4.2 Analytical PCR:

Analytical PCRs were performed to investigate positive clones consequential from transformation of certain plasmids into *Nab. magadii*, and other examinations where proofreading activity was irrelevant. For this purposes the recombinant *GoTaq* polymerase of Promega was used which was established to be a good option, especially for long DNA fragments regarding it elongates 1000bp per minute.

PCR batch *GoTag* (50µl):

28.75µl ddH₂O
10µl 5x *GoTag* buffer
5µl dNTPs (2mM)
2.5µl primer forward (0.1µg/µl)
2.5µl primer backward (0.1µg/µl)
1µl template DNA
0.25µl *GoTag* polymerase (5u/µl)

PCR programm:

- | | | |
|----------|--------------------------------|-------------|
| 1. 95°C | 5' | } 20 cycles |
| 2. 95°C | 1' denaturation | |
| 3. X°C* | 30''- 1' primer annealing | |
| 4. 72°C* | 1' per 1kb product elongation) | |
| 5. 72°C* | 2- fold time of elongation | |
| 6. 4°C | infinite | |

*principally a master mix for several approaches was prepared

*X : individually calculated annealing temperature of the primer with a lower T_M

The resulting products of both preparative and analytical PCRs, were controlled by gel electrophoresis.

2.2.5 CLONING

2.2.5.1 DNA purification:

Agarose gel

Generally purification of DNA was performed to get rid of unspecific DNA fragments, dNTPs, primers, enzymes, salt or buffer systems after PCR or restriction cuts, which possibly could impact further reactions. Simple DNA purification was executed with the QIAquick PCR purification kit (Qiagen) according to manufacturer's instructions. For purification of DNA, to remove unspecific fragments after PCR or plasmid backbone after restriction cut, the QIAquick gel extraction kit (Qiagen) was used. The purified DNA sample was eluted in an appropriate amount of ddH₂O, depending on the initial sample and the final concentration desired.

Elution from an agarose and 6 % PAA gel

In the case of unwanted PCR or restriction reaction site products, the total sample was applied to a gel which was chosen depending on the size of the fragments. The DNA fragment of interest was isolated by cutting out the corresponding band under 70 % UV-light, after short ethidium bromide staining. DNA isolated from agarose gels was simply purified using the QIAquick Gel extraction kit (Qiagen). All steps were accomplished as suggested by the company. For the extraction from 6 % PAA gels it had to be mashed after cutting out and incubated in 300µl elution buffer, shaking at 37° C at least four hours. Finally the supernatant containing the DNA was transferred into a new tube after centrifugation (13.2krpm, 5 min) and purified using the QIAquick gel extraction kit (Qiagen) again. Purity was checked using the appropriate gel.

2.2.5.2 Restriction of DNA

Approximately all DNA restriction endonucleases and appropriate buffers were applied of the company Fermentas. Occasionally restrictions were performed by supplies of the companies New England Biolabs (*Sau3a*) and Promega (*KpnI*). All restriction cuts involved in cloning processes were achieved using the optimal conditions recommended in the supplied protocols. Double digests were implemented following the buffer suggestions obtained from Double Digest™ tool from the Fermentas homepage (<http://www.fermentas.com/en/tools/doubledigest>). To obtain best restriction results the batches were incubated at least 3 h or even over night if no star activity was referred. For strictly analytical purposes the duration of the restriction was reduced to 1 h.

2.2.5.3 Fill-in 5' overhangs

Blunt-end ligations, filling-in reactions of 5' overhangs were performed by the Klenow Fragment, obtained from the company Fermentas. The reaction was incubated with the provided unique buffer or the Fastdigest® buffer from Fermentas, and the dNTPs mixture (0.05 mM final concentration) for 30 minutes at 37 °C. Inactivation of the enzyme was accomplished by incubation at 75° C for 10 minutes.

2.2.5.4 Ligation

Ligation reactions were performed to link linear vector DNA with a DNA fragment, therefore resulting in a plasmid with DNA insert. Both linearised vector and DNA insert were analyzed on a gel with the purpose to estimate the concentrations. The amount of the vector was varied by dilution, in order to obtain a vector: insert ratio of 1:1 to 1:5. The batches were done with T4 DNA ligase and the corresponding buffer, both provided by Fermentas, and were incubated either at room temperature for 3 h or at 16° C overnight.

Following batches were composed:

11.5 µl DNA fragment of interest

1 µl corresponding vector

1.5 µl T4 ligase buffer (Fermentas)

1 µl T4 ligase (Fermentas)

After the reaction the total volume of 15 µl were deployed for transformation of *E. coli* (see section 2.2.5.5)

2.2.5.5 Transformation of *E. coli*

Generation of CaCl₂ competent *E. coli* cells

To generate competent cells from *E. coli* (XL-1 –Blue; Rosetta), 100ml LB were inoculated with a well grown overnight culture to an initial optical density (OD₆₀₀) of 0.1. Incubation takes place at 37° C, shaking at 160 rpm until an OD₆₀₀ of about 0.6 was reached. Those logarithmic growing cells were centrifuged at 6krpm for 15 minutes at 4° C. The harvested cells were resuspended in 40 ml cooled MOPS I solution and incubated on ice for 10 minutes. Again cells were collected by centrifugation at the same conditions like before, and the received pellet was resolved in 40 ml MOPS II solution. After another incubation step of 30 minutes on ice, the cells were centrifuged one last time and resolved in 2 ml MOPS IIa solution and afterwards aliquots of 100µl were taken and stored at – 80° C, or were used immediately.

Transformation of competent *E. coli* cells

Stored frozen aliquots of competent cells were thawed on ice for 10 minutes. In case of ligation product the complete batch was used, while for transformation of an existing plasmid just 1µl, approximately 100 ng, was added to the competent cells. After addition of DNA the cells were incubated on ice for 30 minutes. After that a heat shock at 42° C were performed for 2 minutes, followed by the addition of 300µl LB medium to regenerate the stressed cells at 37° C without shaking for 30 minutes. Finally the transformation batch was triplicate plated on selective LB agar plates containing adequate additives and were incubated at 37° C overnight.

2.2.5.6 Screening and verification of positive clones – Quick apply

This method allows pre-screening of potential positive candidates by quick plasmid preparations via comparison of different migration of plasmids with and without insert. Therefore numerous single colonies were picked from the selective agar plates after transformation and inoculated in 5 ml fluid LB medium; supplemented with the respective antibiotics and incubated, shaking at 37° C overnight. 300µl of the well grown cells were centrifuged in an eppendorf tube at 13.2 krpm for 5 minutes and the supernatant was removed. The pellet was resuspended in 30 µl 5x DNA loading dye and then 15µl phenol/chloroform (1:1) were added. After that the tube was vortexed for about 10 seconds and finally centrifuged again at 13.2 krpm for 5 minutes. Finally 12µl of the resulting aqueous phase were applied to a 0.8 % agarose gel to analyze the rough extracted nucleic acids from the transformants. Possibly positive clones were tested for the right insertion

by matching restriction cuts (see section 2.2.5.2) and/or analytical PCR (see section 2.2.4.2), using plasmid DNA templates purified as described in the following section. Cultures of positively verified clones, possessing the right plasmid, were mixed with 50 % sterile glycerol (1 ml of culture, 800 µl of glycerol) and stored at -80° C.

2.2.5.7 Plasmid preparation from *E. coli*

This method was based on the protocol of the GeneJET™ Plasmid Miniprep Kit from Fermentas which is designed for rapid preparation of plasmid DNA from recombinant *E. coli* cultures. Here DNase free water was used to elute the plasmid DNA from the silica based membrane in the spin column, instead of elution buffer.

Quantification and lyophilization of DNA

Semiquantitative estimation of DNA concentration was carried out by analysis on agarose gel or 6 % PAA gel, respectively. UV/Vis-spectrophotometer NanoDrop ND-1000 from PeqLab was used to measure DNA concentration based on OD_{260/280} and to check purity quality. Necessarily increasing of DNA concentration for some procedures, the Speed Vac UNIVAPO 100H from UniEquip was used to lyophilize the sample by reducing the volume of the batch.

2.2.6 Cloning strategies

Following abbreviations apply:

Nab. magadii L11 for *Nab. magadii* containing the prophage, *Nab. magadii* L13 for the cured strain, *Nab. magadii* P3 for the protease deficient strain and NovR for novobiocin resistance cassette.

2.2.6.1 Deletion of ORF34 (pΔ34₅₂::Nov^R)

The proposed tail fibre gene ORF34 of φCh1 was deleted by cloning the upstream and downstream regions of ORF34; interrupted by a novobiocin resistance cassette. The initial vector pKSII+ does not exhibit an origin of replication for *Nab. magadii*. The construction was performed in a suicide plasmid, aiming that homologous recombination of the described φCh1 loci into the chromosomal DNA of *Nab. magadii* L11, occurs.

The first cloning step was performed by PCR amplification of the upstream region (nu. 22252 – 23235) using the primers Δ34-1 (*Xba*I) and Δ34-2 (*Sma*I), which resulted in a 1004bp fragment. This sequence was digested with the restriction enzymes *Xba*I and *Sma*I, reintegrated with the

primers and the vector was approached equally. After verification of a positive recombinant plasmid, amplification of the 980bp downstream region (nu. 24625 – 25584) of ORF34 with the primers Δ 34-3 (*Hind*III) and Δ 34-4 (*Kpn*I) followed. Both vector and insert were digested with the enzymes *Hind*III and *Kpn*I. In any case the flanking regions were obtained by PCR amplification of purified ϕ Ch1 DNA. The yielding plasmid with the upstream and downstream region of ORF34 was digested with *Sma*I and *Hind*III, and the novobiocin cassette was inserted.

The Nov^R (gyrase B) resistance cassette harboring a promoter sequence was obtained from plasmid pMDS11 by digesting the plasmid with the restriction enzyme *Hind*III and *Sma*I, resulting in excision of the Nov^R fragment from the rest vector. The fragment was isolated by gel extraction. The final construct, amplified in *E. coli*, was purified to a final concentration of ≥ 500 ng/ μ l for transformation into *Nab. magadii* L11, followed by screening for positive clones resulting from a proper recombination event.

2.2.6.2 Complementation of ϕ Ch1- Δ ORF34

To verify the effect caused by deletion of ϕ Ch1 ORF34, the mutation was complemented by providing a functional ORF34 on a plasmid. For this purpose two different variants were used, the non-inverted fragment ORF34₁ and the inverted fragment ORF34₅₂, the latter is carrying the 3' end of ORF36 with its own promoter. The constructs of pNB102 with ORF34₁ and ORF34₅₂ were done by Petra Till (2011), however transformation into *Nab. magadii* L13, screening for positive clones and verification of well expressing strains (see section 2.2.8.) were performed in the course of my work.

2.2.7 Construction of nproNep–ORF34₁ and ORF34₅₂

In this case secretion of gp34₁ and gp34₅₂ was performed to show prevention of adhesion of ϕ Ch1 wild type with the latter, inverted product of ORF34. Therefore *Natrialba* extracellular protease shortly named Nep, was used to facilitate the export.

The 5'-end of Nep was amplified first using *Nab. magadii* L13 as template for PCR amplification with the primers Npro-Kpn and Npro-Bam, resulting a 559bp fragment (nu.712748 – 713306 of the *Nab. magadii* ATCC 43099 genome). The product and the vector pUC19 were digested with *Bam*HI and *Kpn*I and ligated. After screening of positive clones in *E. coli* the accruing plasmid was digested with *Bam*HI and *Xba*I. Additionally the variant sequences of ORF34 were amplified and the accepting fragments about 1400bp in size were restricted with *Bgl*II and *Xba*I for each ligation. For the non inverted fragment 34₁ (nu.23246 - 24660) a PCR from plasmid pBgB1 sequence with the primers 34-5 and 34-3x, was performed. The inverted variant 34₅₂ (nu. 24625 - 25584)

carrying the 3' end of ORF36 was amplified with the oligonucleotides 34-5 and 36-3x, using the vector pBgB52 as template. Due the identical sticky overhangs resulting from the restriction with *Bam*HI and *Bgl*II, respectively, ligation of vector and fragment was warranted. After verification of successful clones carrying the 5'-end of Nep fused with each variants of ORF34, were isolated by restriction with *Kpn*I and *Xba*I and purified from an agarose gel. The shuttle vector pNB102 containing an origin of replication for halophile organism was digested with the same endonuclease enzymes *Kpn*I and *Xba*I. The finally yielded constructs pNB102-nproNep-34₁ and pNB102-nproNep-34₅₂ were transformed into *Nab. magadii* P3, a protease deficient strain.

2.2.8 Methods for Archaea

2.2.8.1 Transformation of *Natrialba magadii* strains

Generation of competent *Natrialba magadii* cells

Preparing of spheroblast cells are necessary for generation of competent cells and transformation of *Nab. magadii*. Therefore a fresh grown culture was splitted into three buffled 500ml flasks containing 60ml NVM+, supplemented with a final concentration of 70µg/ml bacitracin and different amounts of the well grown pre-culture; 4ml, 6ml and 8ml respectively. The batches were incubated at 37° C shaking at about 160 rpm until one of them reached an OD₆₀₀ of 0.5-0.6. The cells were harvested by centrifugation at 6 krpm for 15 minutes at room temperature and were resuspended in half of the volume from the original culture of high salt (HS) buffered spheroblasting solution containing glycerol. Finally proteinase K was added with a concentration of 20µg/ml and the culture was further incubated, shaking in a 100 ml Erlenmeyer flask at 42° C until spheroblast formation could be observed under the light microscope. Due the treatment with bacitracin and proteinase K the archaeal S-layer is destabilized, additionally EDTA causes the surface layer to fall apart and the naturally rod shaped cells changes into well-rounded cells. In this stage the cells can be stored at – 80° C when 1.5 ml aliquots are centrifuged and resuspended in 150 µl HS buffered spherobalsting solution with 15 % glycerol, however, maximum storage time should not exceed one week because transformation efficiency is decreasing with every day of storage.

PEG-600 based transformation of competent *Natrialba magadii* cells

Preliminary the PEG-600 aliquot (stored at -80° C) was defrosted at 65° C and blended with HS unbuffered spheroblasting solution (w/o glycerol) (60 % PEG-600; 40 % HS solution). The buffer was kept warm at 37° C until use.

- centrifugation of 1.5 ml competent cells (incl. negative control) at RT, 10 krpm for 3'
- addition 15 µl of 0.5 M EDTA pH 8.0 and incubate at RT for 10 minutes
- addition of DNA (generally 5 µg)
- incubation on RT for 5 minutes
- addition of 150 µl PEG-600 (60% in HS unbuffered spheroblasting solution w/o glycerol)
- incubation on RT for 30 minutes
- addition 1 ml of rich medium and mix carefully using a blue tip
- centrifugation for 3 minutes, 11 krpm
- resolution of the pellet in 1 ml rich medium
- regeneration at 37°C shaking at 800 rpm affixed on the shaker, until the cells are rod shaped again
- plating 100 µl of the total transformation batch on selective medium plates (undiluted, 1:10 dilutions, one-time neg. control) ; positive control was plated on agar medium w/o antibiotics
- plates sealed in plastic bags were incubated at 37° C or 43° C respectively, until colonies were visible

2.2.8.2 Screening for positive clones

For pre-investigation of chromosomal gene replacement in *Nab. magadii* L11, single colonies were picked and inoculated in 1 up to 5 ml liquid medium containing the selecting antibiotics. Incubation took place shaking at 37° C or 42° C and airing the batches at least every other day to ensure aerobic conditions, grown up to an optical density which allows analytical PCR analysis. For examination by PCR 25 µl of the turbid cultures were centrifuges for 3 minutes at 13 krpm and the resulting pellet was lysed with 100 µl ddH₂O. Up to 10 transformants were pooled per PCR approach and tested, which allows screening of a considerable number of single colonies. When a positive signal was obtained the PCR was repeated, whereby each of the mixed cultures was tested individually. Without a proper homologous recombination there are two possibilities of DNA integration by single homologous recombination, assuming that only one cross-over event

happened; leading to integration of vector DNA into the chromosome of *Nab. magadii*. The composition depends on between which sequences the homolog recombination event occurred. In order to obtain the absence of plasmid DNA a PCR strategy with the primers pKS-3 and pKS-5, was performed.

Colony PCR analysis were performed for verification of up taken plasmids in *Nab. magadii*. For multiple samples an analytical PCR master mix was prepared and 25 µl were aliquoted into each, serially numbered reaction tube. Then a single colony was picked, stroken out on a numbered sectorred selective agar plate and inoculated into the corresponding numbered PCR tube to comprehend the putative positive clone which gives a signal in the following PCR reaction. Additionally in both applications a negative (strain without plasmid) and positive control (plasmid construct) were executed, that would give false positive signal.

2.2.8.3 Generation and evidence of a homozygous mutant strain

As already mentioned before *Natrialba magadii* possesses up to 50 copies of its genome per cell. For survival of a recombinant cell only one transcript of the *gyrB* (novobiocin resistance) gene is necessary. The applied method described before does not yield a homozygous strain which is completely lacking the ORF34. Primary, occurrence of a recombination event where the novobiocin cassette has disrupted the ORF34, was tested by analytical PCR analysis with the primers Δ34-1 and Nov-9. The positive clones were passaged in order to achieve a homozygous mutant strain. In this case 20 ml rich medium containing novobiocin was inoculated with the 100µl of the previous culture, consistently growing them until the stationary phase was reached. Simultaneously dilutions of the fresh grown passages were plated on selective medium plates and single colonies were picked again. This procedure was continued until no wild type signal was detectable anymore. Therefore, primers 34-5 and 36-3 were used to investigate the wild type gene by analytical PCR, employing 50 cycles of synthesis to ensure that all ORF34 wild type genes are eliminated. In addition to PCR analysis a southern blot was performed in order to verify the homozygous deletion mutant of ORF34. This is a method which allows the detection of DNA sequences, routinely used to check for the presence of DNA fragment in a DNA sample. In this case the method was applied to visualize an insertion or deletion event due to a size shift of the fragment.

2.2.8.4 Isolation of *Nab. magadii* chromosomal DNA

The culture of interest was grown in a 500 ml batch until stationary phase was reached and the cells were harvested by centrifugation at 8 krpm for 20 minutes at room temperature. The pellet was completely dissolved in 5 ml high alkaline salt buffer (4M NaCl/50mM Tris-HCl pH 9.5) and transferred into two SS34 centrifugation tubes. To each tube 2.5 ml 14mM Desoxycholat were added and after homogenous distribution 7.5 ml sterile ddH₂O were admitted; rolling the tubes until slimy white threads were visible. For completely lysis of the cells incubation on ice for about one hour was performed. For extraction the cell lysate was mixed with 12.5 ml phenol/chloroform 1:1 in each tube, and further centrifuged at 10 krpm for 30 minutes at 4° C. The uppermost aqueous phase was transferred into a 100 ml Erlenmeyer flask and overlaid with approximately 0.6x volume isopropanol for precipitation of DNA. A hook-shaped glass Pasteur pipette was used to wind up the DNA, to dissolve it afterwards in 5 ml sterile ddH₂O. After that sterile EDTA with a final concentration of 10mM was added, followed by supplementation of 7g CsCl per 5 ml DNA solution and ethidium bromide to get a rosé-colored solution with a final concentration of about 10µg/ml. The pinkish solution was transferred to a Quick-seal ultracentrifuge tube, sealed and centrifuged at 60 krpm at room temperature for 20 hours. The yielding DNA band (upper pink band) was precisely removed with a pipette and transferred into an Eppendorf tube; an additional lower band is representing the RNA fraction. For removal of ethidium bromide from the chromosomal DNA, extraction with water saturated butanol was performed. Therefore about a half milliliter of the solvent was added to the DNA solution and mixed by inverting the tube then, the upper phase of the yielding 2-phase system were removed. This procedure was repeated until the batch was completely decolorized. To get rid of the CsCl the purified chromosomal DNA was transferred into a dialysis tube, which was thoughtfully washed with sterile ddH₂O for following restriction approaches, letting dialyze the DNA against sterile ddH₂O. The quality and amount of the purified DNA was checked on an agarose gel and the DNA solution was stored at 4° C for further approaches.

2.2.8.5 Southern blot procedure

For this purpose isolated chromosomal DNA of the wild type strain *Nab. magadii* L11 and of the putative homozygous *Nab. magadii* L11-ΔORF34 (see section 2.2.8.4) was digested with the restriction endonuclease *Bam*HI. The resulting DNA fragments were separated on a 0.8 % agarose gel, applied with the marker λ-*Bst*EII and a biotinylated 2-log DNA ladder. After standard ethidium bromide staining a picture was taken under UV-light and the gel was incubated for 10 minutes in 0.25 M HCl, allowing fragmentation of DNA for an easier transfer of the nucleic acids to the

membrane. After that the gel was briefly rinsed with ddH₂O. To obtain and retain single DNA strands a bath with the solution 0.4 M NaOH / 0.6 M NaCl, following with incubation in 1.5 M NaCl/ 0.5 M Tris-HCl (pH 7.5) for neutralization was performed, in each case for 30 minutes. DNA transfer from the gel to the nylon membrane was achieved by capillary blotting procedure in 10x SSC buffer over night. Success of transfer was checked by visualizing the remaining DNA in the gel after an ethidium bromide bath under UV-light. The membrane was incubated in 0.4 M NaOH and 0.2 M Tris-HCl for 1 minute each. Shortly after the membrane became touch dry, cross – linking procedure by UV-light was realized to fix the nucleic acids covalently to the membrane, using the Stratalink cross - linker from Stratagene.

Preparation of the biotinylated DNA probe

The upstream region of ORF34 was amplified by preparative PCR with the primers Δ34-1 and Δ34-2 using ϕCh1 DNA as template, after purification and labeling procedure, intended to conduce as probe. For biotinylating the probe approximately 500 ng of DNA was denatured by heating at 95° C for 10 minutes, followed by incubation on ice for 5 minutes. Afterwards 10 µl 5x labeling mix (biotinylated random octamers in 5X labeling buffer), 5 µl dNTPs (containing Biotin-14-dATP) and 1 µl Klenow Fragment (3' → 5' exonuclease activity, 5.000 units/ml) were added to the batch. All components were provided from NEBlot[®] Phototope[®] Kit of the company New England BioLabs. The labeling reaction was achieved at 37° C for 3h and stopped by addition of 5 µl 0.2 M EDTA (pH 8.0). Precipitation of biotinylated DNA was performed by adding 10 µl 4 M LiCl and 200µl 96 % ethanol to the reaction batch and incubation followed at -20° C, at least for 20 minutes. After a centrifugation step carried out at 16.4 krpm at 4° C for 30 minutes, the resulting pellet was washed with 1 ml 70 % EtOH, dried at 65° C and resolved in 20 µl nuclease-free ddH₂O.

Blocking and hybridization

For blocking the nylon membrane with the immobilized DNA was transferred into the southern blot tube and pre-hybridized in 20 ml hybridization buffer containing 100 µg/ml salmon sperm DNA at 65°C for at least 3 hours. For hybridization the total batch of biotinylated DNA probe was added to the hybridization buffer and incubation was continued over night.

Detection of the southern blot

The membrane was washed twice with 2x SSC, 0.1 % SDS for 5 minutes at room temperature, then with 0.1x SSC / 0.1 % SDS at 65° C for 15 minutes, also done twice. Subsequently the membrane was transferred into a small basin for development, using the Phototope[®]-Star

detection Kit from New England BioLabs. At this time the membrane was incubated in the blocking solution for 5 minutes and then incubated in 5 ml blocking solution containing 5 μ l streptavidin solution for 5 minutes. Further the membrane was washed thrice with washing solution I (1:10 dilution of blocking solution) and then incubated with 5 ml blocking solution, containing 5 μ l biotinylated alkaline phosphates solution of the kit. Before visualization the membrane was washed once with blocking solution for 5 minutes and 3 times with washing solution II. At last the membrane was incubated with 3 ml 1x CDP-Star[®] dilution buffer as well as 6 μ l of CDP-Star[®] Reagent. For recognition of the signal the biotinylated DNA probe binds to the target sequence on the membrane via base pairing. The alkaline phosphatase which is conjugated to biotin, binds to streptavidin and catalyzes the removal of the phosphate from the CDP-Star[®] (phenylphosphate substituted 1,2 dioxetane) yielding in a more or less stable intermediate which then spontaneously decays and emits light that is detectable by exposing an X-ray film. The signal enhancement is warranted, as several biotin molecules can bind strong but non-covalently to one streptavidin molecule. The exposure time had to be determined experimentally but usually one to ten minutes were sufficient for detection.

2.2.9 Virus methods

2.2.9.1 Purification of ϕ Ch1 particles

For generation a fresh lysogenic strain of *Nab. magadii* L11 a plaque of ϕ Ch1 was inoculated into rich medium and grown to an optical density of about 0.6. This culture was used to inoculate a 5 l batch of rich medium. The cultures were grown until complete cell lysis was observed. The virus particles were isolated from the supernatant of the lysed culture by centrifugation (20 minutes, room temperature and 6 krpm). The pooled supernatant, transferred into new flasks, were supplemented with 10 % (w/v) PEG 6000 and gently stirred overnight to coat the virus particles. At this time the virus particles could be harvested by another centrifugation step with equal conditions as before. The resulting pellets were pooled and resolved in as little as possible high alkaline salt buffer (4M NaCl/50mM Tris-HCl pH 9.5), to concentrate the batch for further purification steps.

CsCl density centrifugation

The ϕ Ch1 particles were purified using an isopycnal density centrifugation method. The first discontinuous CsCl gradient was prepared by overlaying 4 ml CsCl solution 1.5 with 4 ml solution 1.3 followed by 6 ml of the solution containing ϕ Ch1 phage particles. Finally the Beckmann Ultracentrifuge tubes were exactly balanced (CsCl solution 1.1) and centrifuged for 20 hours, at

room temperature with 30 krpm, using the swinging bucket rotor SW40Ti. The accumulated particles are visible as a defined band were carefully removed from the tubes and further purified and concentrated by an additional continuous CsCl centrifugation step. Therefore 5ml of the ϕ Ch1 particles were mixed with an equivalent volume of CsCl solution 1.3 and centrifuged again at room temperature at 30 krpm for 20 hours. Routinely the quality and amount of the purified phage particles was checked on an agarose gel and a titer determination by infecting the cured strain *Nab. magadii* L13 was performed to calculate the plaque forming unit per milliliter.

2.2.9.2 ϕ Ch1 titer determination by soft agar technique

This method was used to determine the pfu/ml of purified viruses as described above, but also for further infection experiments as described below (see section 2.2.9.3). Therefore 100 μ l of different virus dilutions (10^{-2} , 10^{-4} , 10^{-6} , 10^{-8} , 10^{-10} dilutions in rich medium) were mixed with 400 μ l of *Nab. magadii* L13 culture at early stationary phase in 5 ml rich medium soft agar and poured on a rich medium plate. Each infection was performed in triplets. Plates were drawn-in on room temperature over night, sealed and incubated at 37°C until plaques were visible (usually 10 days).

2.2.9.3 Plating efficiency of ϕ Ch1 on *Nab.magadii* exporting different variants of ORF34

This experiment was performed in order to investigate a suspected reduction or prevention of ϕ Ch1 infection by export of gp34₁ and the putative tail fibre protein gp34₅₂ variant. Therefore the strain *Nab.magadii* P3, a Nep deficient strain, was used for secretion of both protein variants. The plating efficiency was executed by addition of 800 μ l culture (*Nab. magadii* P3 pNB102-npro-NEP-34₁ and P3 pNB102-npro-NEP-34₅₂) to 100 μ l wild type ϕ Ch1 dilutions. The procedure was carried out according to the protocol described above (see section 2.2.9.2.).

2.2.9.4 Isolation of virus DNA

100 μ l of purified viruses were mixed with 300 μ l sterile ddH₂O in order to lyse phage particles. For extraction of nucleotides 200 μ l phenol/ chloroform (1:1) were added and mixed by vortexing, followed by centrifugation at 13.2 krpm for 3 minutes at room temperature. The upper aqueous phase was carefully transferred into a new Eppendorf tube. This procedure was repeated until no white borders were visible anymore. After that, precipitation was archived by addition of twofold volume of 96 % EtOH and centrifugation at 16.4 krpm at 4° C for 30 minutes. The supernatant was removed and the pellet was washed twice with 1 ml 70 % EtOH and dried at 65° C. Finally the DNA

was dissolved in 20 µl sterile ddH₂O and analysis on agarose gel was performed to check the quality and concentration.

2.2.9.5 Precipitation of viral proteins

To precipitate viral proteins 200µl of purified virus particles were mixed with 5 % TCA (final concentration). The batch was incubated for 1 hour on ice; afterwards centrifugation at 16 krpm for 30 minutes at 4° C was carried out. The pellet was resuspended with 25 µl 5mM sodium phosphate buffer (Na₂HPO₄ and NaH₂PO₄, pH 6.8) and completed with same volume of 2x Laemmli buffer (protein sample buffer) which contains a pH indicator. Due to the application of TCA, the pH of the sample became very low (yellow colored) and had to be neutralized by treatment with gaseous ammoniac (without changing the volume) until the color changes into blue again.

2.2.10 Protein methods

2.2.10.1 Preparation of crude protein extracts

For preparation 1.5 ml of a growing culture with known optical density was centrifuged at 13.2 krpm for 3 minutes. The supernatant was removed and the pellet was resuspended in the volume calculated with the formula: $OD_{600} \times 75 \mu\text{l} = x \mu\text{l}$ sodium phosphate buffer. The same volume of 2x protein sample buffer (Laemmli) was added. To denature proteins the extracts were heated up to 95 ° C for 10 minutes, before the sample was applied on a protein gel. In some cases the proteins of *Nab. magadii* raw extracts were incubated at 37° C over night to ensure less viscosity for superiorly application onto protein gel.

2.2.10.2 Protein extracts from culture supernatants

The protease deficient strain *Nab. magadii* P3 was transformed with the two constructs, nproNep–ORF34₁ and nproNep–ORF34₅₂, differing in the ORF34 sequence. Additionally the strain was transformed with the empty vector pNB102, serving as negative control. Those cultures were incubated until the stationary phase was reached (about an OD₆₀₀ of 1.3) because the expression of protease starts at the end of the logarithmic growth phase and is predominant in the stationary phase. For preparation of proteins from the culture supernatant 1.5 ml of the culture was centrifuged as described above in section 2.2.10.2, with the difference that the supernatant was transferred into a new Eppendorf tube and a second centrifugation step was performed. To

ensure that no cells are left anymore the supernatant were filtered through a sterile 0.2 μm membrane. For precipitation 5 %TCA (final concentration) was added into the sample and incubated on ice at least for one hour, following by centrifugation at 4° C for 30 minutes at 16 krpm. The resulting pellet was resuspended with the calculated volume 5mM sodium phosphate buffer and 2x Laemmli buffer, according to the optical density of the culture. Exposure to gaseous ammoniac was necessary to neutralize the pH of the samples described in section 2.2.10.2.

2.2.11 His-tagged protein over-expression and purification in *E. coli*

2.2.11.1 His- tagged protein over-expression in *E. coli*

For expression of recombinant protein in *E. coli* the respective strain was transformed with the desired gene region on an expression vector (see section 2.2.5.5) and inoculation of a one liter (containing adequate antibiotics) batch to an OD_{600} of 0.05- 0.1 was performed, using a well grown overnight culture. The culture was incubated, shaking at 37° C until an optical density of approximately 0.3 was reached. At this time point an aliquot for crude extract of the culture was taken, serving as a negative control and to compare expression with the induced culture. The over-expression was induced with IPTG (Isopropyl- β -D-thiogalactopyranosid) to a final concentration of 1mM and the culture was grown at 37° C for another 3 hours. Harvesting of the cells were performed by centrifugation at 7 krpm for 15 minutes at 4° C and followed by removing the supernatant. The pellet could be stored at -20° C until final use.

2.2.11.2 His-tagged protein purification under denaturing conditions

In the case of a frozen pellet it was thawing on ice, then solved in approximately 50 ml buffer B and for lysing the cells the batch was stirred gently (preventing foam formation) over night at room temperature. For complete cell lysis the cell suspension was sonicated 4 times for 3 minutes on ice. To remove cell remnants a centrifugation step was performed for 20 minutes at 7 krpm at 4° C. After that the supernatant was supplemented with 500 μl Ni-NTA (Qiagen) and stirred again over night at room temperature, in order to enable binding of the His-tagged proteins to the Ni^{2+} ions. The cell lysate containing Ni-NTA was applied to a column from the QIAexpressionist™ Kit (Qiagen) and the flow-through was collected. The nickel-beads are accumulated on the filter of the column. The procedure was continued with two wash steps, each with 4 ml buffer C, collected as well in sterile tubes. The elution, caused by decreasing pH values, of the Ni-NTA bound proteins starts by addition of 6x 0.5 ml buffer E. Those fractions were collected in Eppendorf tubes,

respectively. To analyze the presence of the protein of interest, small aliquots (30 μ l) of the collected fractions were prepared for SDS-PAGE by adding the same amount of 2x protein sample buffer (Laemmli) and denaturing at 95° C for 5 minutes before applying on the gel. Visualization of proteins was performed by coomassie staining (see section 2.1.13.2).

Dialysis of purified proteins

For finalization the purified protein fractions containing uncorrupted amounts of the protein of interest were pooled and transferred into a dialysis tube, in order to remove the urea. Proteins from ϕ Ch1 were adjusted to high salt and pH conditions by dialysis against buffer I (4M urea/ 2M NaCl/ 50mM Tris-HCl pH 9.5) for one hour, additionally with dialysis buffer II (4M NaCl/ 50mM Tris-HCl pH 9.5) over night. In other applications the dialysis was carried out against 1x PBS.

2.2.11.3 Crude quantification of purified proteins

The dialyzed protein was applied on SDS-PAGE with particular standards of BSA (bovine serum albumin). Dilutions of BSA (0.025 μ g/ μ l, 0.05 μ g/ μ l, 0.1 μ g/ μ l, 0.2 μ g/ μ l, 0.4 μ g/ μ l, 0.8 μ g/ μ l) were used to determine the protein concentration for further applications.

2.2.12 Protein purification from *Nab. magadii*

This procedure is quite similar to the protein purification protocol from *E. coli* (see section 1.4.2.1), with the difference in buffer composition, which was adapted to the haloalkaliphile conditions. Another contrast is that all centrifugation steps were performed at room temperature. First, after harvesting the cells, the pellet was resuspended in 180 - 200ml resuspension buffer (3M KCl/10mM Tris-HCl (pH 8)/20mM Na₂PO₄). After sonification and centrifugation (10krpm, 15 minutes) the supernatant was incubated with 1 ml Ni-NTA over night. On the next day the whole lysate was applied on a column. Two wash steps, each with 4 ml were performed and elution buffers 1 to 5 were eluted with 500 μ l. The protein samples were mixed 1:1 with 2x Laemmli buffer and loaded on a SDS-Page gel, having heated at 95°C. After running the gel, Coomassie staining was done to visualize the eluted proteins.

2.2.13 SDS-PAGE

SDS-PAGE, sodiumdodecylsulfate polyacrylamide gel electrophoresis, was performed in order to analyze proteins. This method separates proteins according to their molecular weight under

denaturing conditions, caused by heating, β -mercaptoethanol and SDS. A polypeptide is binding amounts of SDS, an anionic detergent, in proportion to its relative molecular mass. The negatively charged SDS linearizes the complex structure and through applying electricity they are strongly attracted towards a positively charged electrode. This is valid for proteins which have an equalize percentage of charged amino acids. In case of halophilic proteins, which have predominantly acidic residues, the migration behavior is much slower compared to mesophilic proteins of the same size. The matrix for SDS-PAGE polymerizes by a source of free radicals and addition of the stabilizer APS (ammonium persulfate) and TEMED (Tetramethylethylenediamine), which triggers a radical chain reaction of acrylamide and N, N-methylene bisacrylamide. Varying the ratio of acrylamide to bisacrylamide can be adjusted according to protein sizes. Lower percentage gels are supposed to use for high molecular weight proteins and vice versa.

2.2.13.1 Preparation of protein gels

For casting a discontinuous SDS-polyacrylamid gel a vertical system provided from the manufacture Bio-Rad (Mini-PROTEAN[®]) according to the instructions of the company. At the beginning the reagents for the separation gel (according table below) were mixed together on ice and immediately after addition of APS and TEMED it was poured between two glass plates about 6 cm high. To achieve a perfectly horizontal surface without air bubbles, the separation gel was overlaid with a thin layer isopropanol, instantly. The isopropanol was removed after polymerization of the gel. The solutions for the stacking gel (according table below) were mixed on ice again, casted above the separation gel and the comp was fitted in. In general 8 μ l of a prepared protein sample were applied onto the gel, having them heated at 95° C for 10 minutes. First an electric potential of 40V was set and after protein accumulation between the gel border, the voltage was increased to 60V. The run was performed until the bromphenol blue was run out of the gel matrix to ensure sufficient separation.

required materials	12 % Seperation gel	4 % stacking gel
ddH ₂ O	1750 μ l	1233 μ l
Separation gel buffer	1250 μ l	
Stacking gel buffer		500 μ l
30 % AA-solution	2000 μ l	267 μ l
10 % APS	60 μ l	20 μ l
TEMED	10 μ l	5 μ l

2.2.13.2 Coomassie staining

Unspecific visualization of the proteins was achieved by staining the gel in a Coomassie brilliant blue staining solution, length depending on the quality of the solution, followed by destaining with 10 % acetic acid solution. This staining solution binds to aromatic amino acids as well as histidine and arginine, making the proteins visible on the gel but also cross-links them to the gel matrix. This procedure was done at once to the western blots, in order to document the consistent protein loading across the samples.

2.2.13.3 Western blot analysis

This method allows detection of specific proteins which were transferred from a gel on a nitrocellulose membrane (Protran from Whatman) and the recognition is based on a specific antibody or antiserum. Usually a second antibody, coupled to a horseradish peroxidase (HRP), binds to the primary antibody and in presence of luminol and hydrogen peroxide it is able to detect by emission of light.

Blotting procedure:

For the transfer procedure from a SDS-Page gel onto the membrane the semi-dry blotting apparatus were chosen. Therefore one nitrocellulose membrane and six Whatman filter papers, about the same size of the gel, were cut and a stack composed of a double-layer (membrane-polyacrylamidgel) flanked by three layers of filter papers on each side, were prepared. Transfer was performed by applying voltage of 20 V for 20 minutes for one gel and for two gels 30 minutes were chosen. To analyse successful blotting, the membrane was stained in Ponceau S shortly to visualize all unspecific proteins reversibly and to mark the bands of the unstained protein ladder.

Blocking:

After the membrane was completely destained with water the blot were blocked with 50 ml 1x TBS -T containing 5 % skim milk powder, shaking over night at 4° C or for 1 hour at 37° C. This was performed to prevent unspecific binding of the antibodies to the membrane.

Western blot development and visualization:

After blocking procedure the membrane was washed once by shaking in 1x TBS-T for 10 minutes, followed by incubation with the first antibody solution (see section 2.1.8.1) for 1 hour shaking at room temperature (reusable antibody). Again the blot was washed 3 times for 10 minutes with 1x TBS-T. The second antibody solution (α -rabbit, 1:5000 in TBS-T) was applied and the blot was incubated, again for 1 hour. Solutions and conditions for the development of western blots were

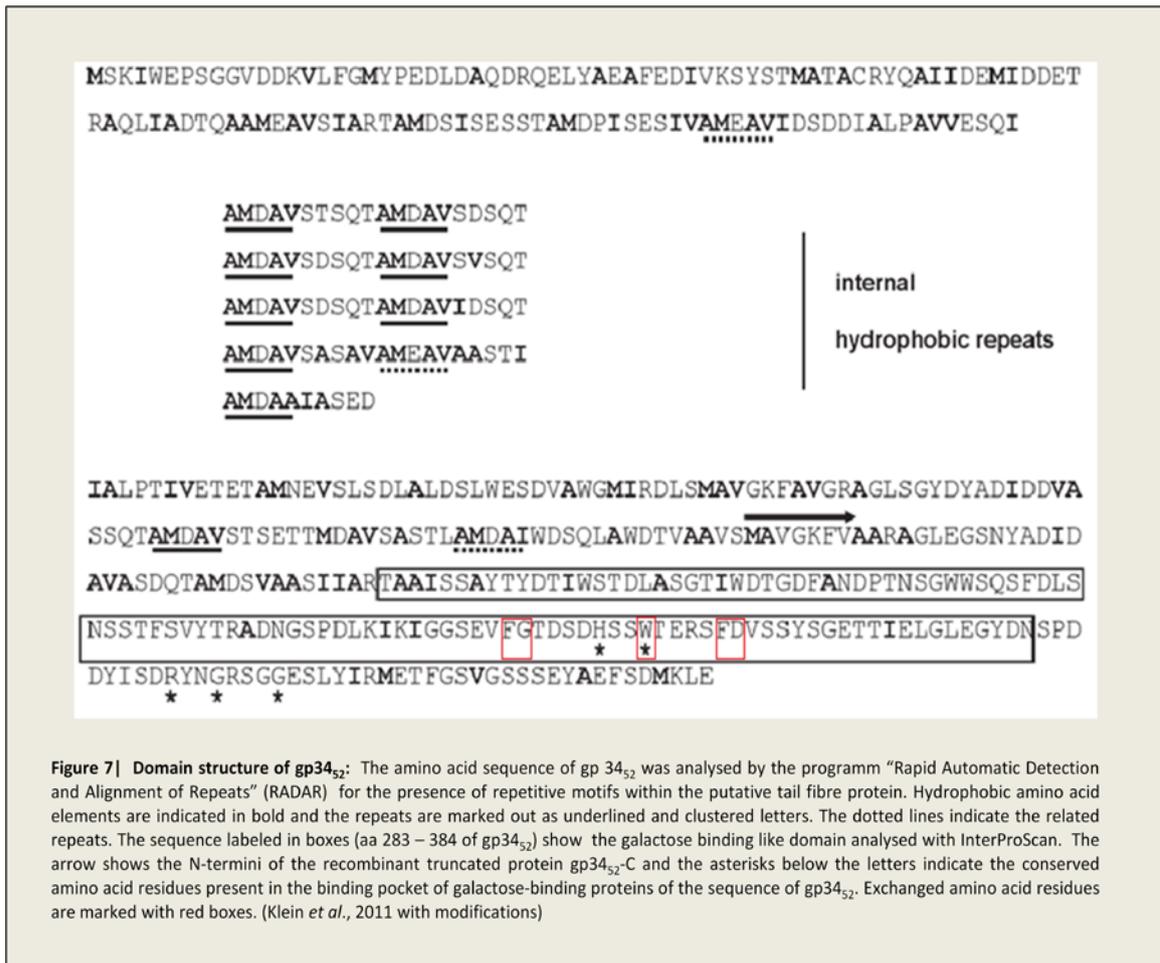
adapted to the properties of the antibodies. In course of this diploma work, the primary antibodies α -gp34, α -gpE and α -His-tag were used. Optionally, in case of application of the primary antibody α -His-tag, the α -mouse antibody (1: 5000 in TBS-T) was chosen as secondary antibody. After abolishing the antibody the blot was finally washed three times with 1x TBS-T. For detection of the now HRP-coupled proteins, the SuperSignal[®] West Pico Chemiluminescent solutions, applied from Thermo Scientific, were mixed 1:1 and the blot was incubated 5 minutes. Subsequently the membrane was placed into the developing cassette and the signal was detected on the X-ray hyper film, provided from Amersham Bioscience, by adequate exposure time in the dark.

3. Results and Discussion

3.1 Deletion mutant of Φ Ch1 Δ ORF34₅₂

3.1.1 General aspects of gp34₅₂

Project outcomes so far lead to the assumption that the corresponding proteins gp34₁ and gp34₅₂ of the invertible region may act as tail fibre proteins. Both proteins share high sequence similarities at the amino acid and the DNA level; the containing extensive clusters resemble in other tail fibre proteins of certain phages, like bacteriophage T4 and λ . The C-terminal part of those proteins exhibited similarities to galactose-binding domains elaborated by InterProScan analysis. However, only the C-terminus of the gp34₅₂ comprises a conserved amino acid sequence within the galactose-binding domain, which is known to form the binding pocket for α -D-galactose, gp34₁ lacks this conserved aa sequence (see Fig. 7, indicated by asterisks below the letters)(Klein *et al.*, 2011). A common function of these domains was found in galactose binding enzymes of *Escherichia coli* which are able to bind specific cell surface-attached carbohydrates ligands (Iro *et al.*, 1991). It can therefore be concluded that the C-terminal domain of gp34₅₂ seems to be responsible for adsorption of Φ Ch1 to the receptor on the cell surface of *Nab. magadii*. In order to verify this assumption, the potential interaction of gp34₅₂ with *Nab. magadii* was analyzed. Therefore, ORF34₁ and ORF34₅₂ were individually expressed in the haloalkaliphilic archaeon *Hfx. volcanii* and the crude extracts in its native form were used for incubation with *Nab. magadii* L13 cells. Western Blot analysis showed that gp34₁ was exclusively found in the cell supernatant, while gp34₅₂ was only found in the pellet fraction, signifying that the inverted ORF34 product (gp34₅₂) was able to bind to the cell surface of *Nab. magadii* L13. Further approaches with truncated versions of the proteins (either the C-terminal (gp34₁-C and gp34₅₂-C) or the N-terminal part (gp34₁-N and gp34₅₂-N) were detected), showed as expected, that exclusively the gp34₅₂-C without the N-terminus was able to bind to the cell surface. Additionally, an exchange of one of the conserved amino acid residues (see Fig. 7; point mutations are marked with red boxes) within that galactose-binding domain were performed, in order to investigate if the predicted galactose-binding domain in gp34₅₂ is responsible for the adhesion of the protein to *Nab. magadii* cell surface. This mutated gp34₅₂ was not able to bind to *Nab. magadii* L13 cells, in contrast to the control where the amino acid residue exchange occurred close but outside of the predicted galactose-binding domain.

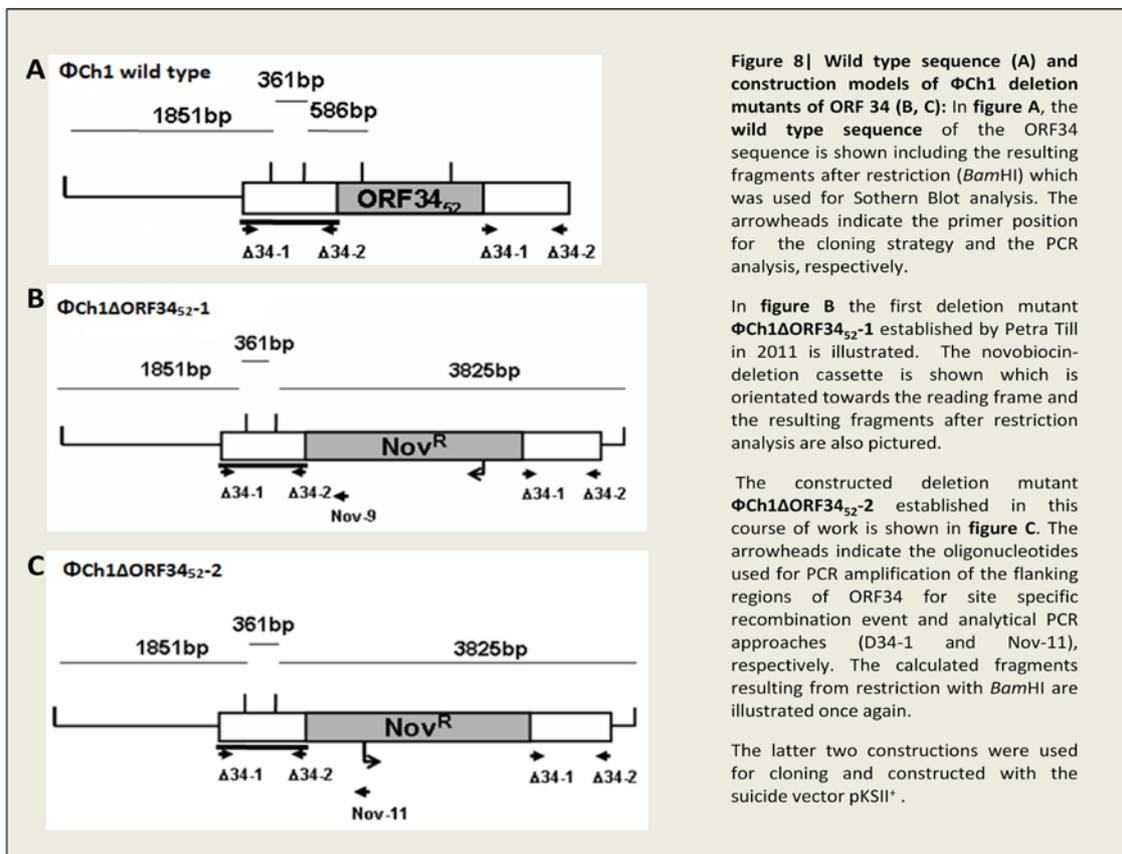


To visualize the binding of gp34₅₂ to *Nab. magadii* L13 cells as well as the glycosylation of the cell wall surface layer proteins of *Nab. magadii* was shown with the periodic acid-Schiff staining (PAS-staining) method. To illustrate the necessity of carbohydrates involved in the infection process of *Nab. magadii* with ΦCh1, a competition assay with increasing α-D-galactose concentrations were performed. An inhibitory effect on infection was observed upon the addition of 50mM α-D-galactose to the incubation medium (Klein *et al.*, 2011). Anyhow, to present evidence for the putative tail fibre protein gp34₅₂ and its importance in the infectivity process a deletion mutant of ΦCh1, namely ΦCh1ΔORF34₅₂-1 lacking the ORF34 was constructed by Petra Till in 2011. The result of the non-directional cloning strategy caused an inverse direction of the novobiocin-cassette into the genome of ΦCh1 which induces a polar effect within the operon with an impact on transcription of following capsid encoding genes. Electron micrographs of purified phage particles ΦCh1ΔORF34₅₂-1 has shown significant differences compared to the ΦCh1 wild type. The heads seem to be empty although viral DNA could be isolated successfully. Furthermore the mutated phage particles exhibited extended structured tails where the start- and end points were

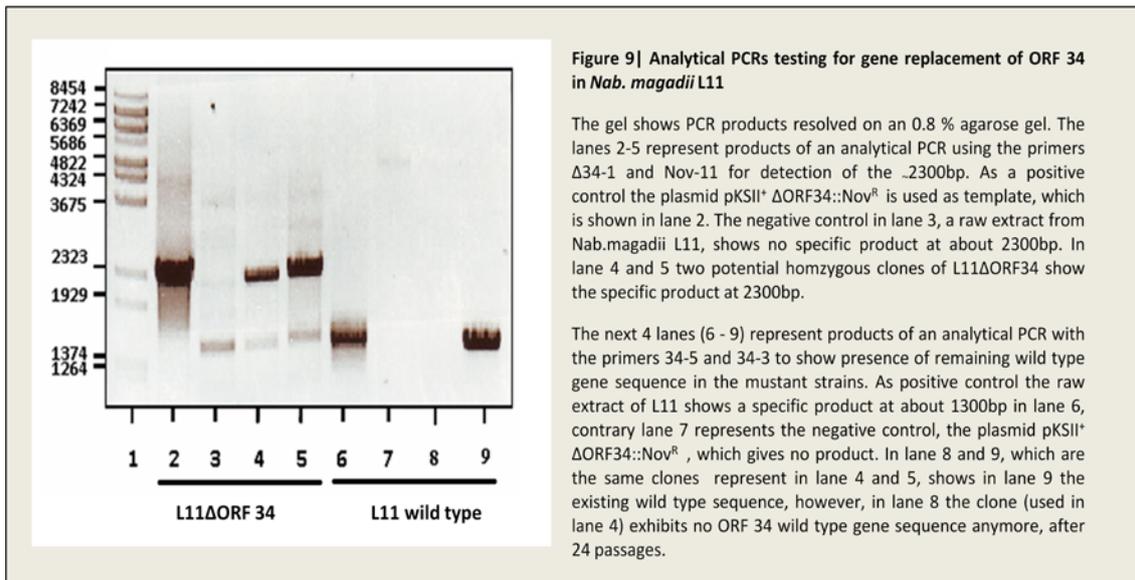
not visible in the electron micrographs. The nature of the filamentous structures was not clear, though finally considered as tail particles of the virus $\Phi\text{Ch1}\Delta\text{ORF34}_{52-1}$, which might be caused by the absence of the termination signal during tail synthesis thereby, became more and more extended. In this course of work a new knock out mutant of ΦCh1 was established according to the well-tried strategy described in the section methods and materials 2.2.8.3. This mutant virus was termed as $\Phi\text{Ch1}\Delta\text{ORF34}_{52-2}$, since Petra Till created the first deletion mutant of ORF34, named as $\Phi\text{Ch1}\Delta\text{ORF34}_{52-1}$. The principle aim of the project was the confirmation of the importance of the putative tail fibre protein gp34₅₂ within the infection process. The resulting deletion mutant $\Phi\text{Ch1}\Delta\text{ORF34}_{52-2}$ was constructed with the expected feature of loss of infectivity in contrast to the ΦCh1 wild type. Supposed phenotypically characteristics were intended to be analysed and evaluated.

3.1.2 Construction of a homozygous $\Phi\text{Ch1}\Delta\text{ORF34}_{52-2}$ deletion mutant

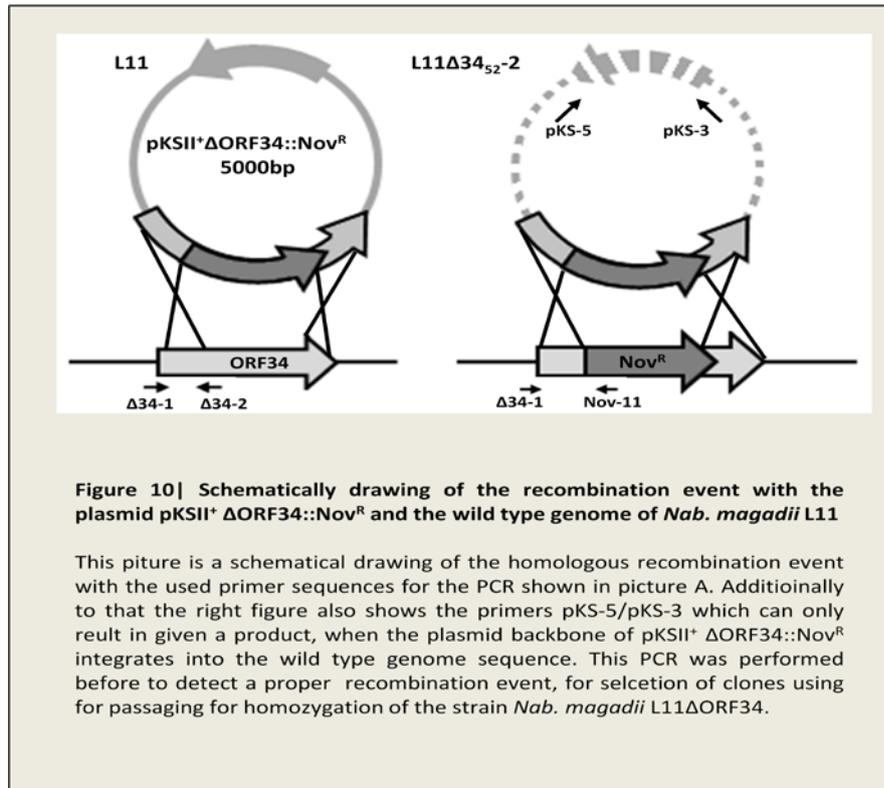
For the construction of ΦCh1 lacking the ORF34 a new directional cloning strategy was established to ensure an insertion arranged within the transcriptional reading frame activity as pictured in figure 8.



The construct contains a centered novobiocin-cassette, acting as a deletion- and selection tool with two flanking regions (upstream and downstream regions) of the ORF34 for homolog recombination. For creation of the ORF34 deficient Φ Ch1, the suicide vector pKSII⁺ was used for transformation of the *Nab. magadii* L11 strain which harbors the prophage Φ Ch1 integrated in the genome. As a result of the absence of an origin of replication which is working in Archaea, the vector carrying the deletion construct is not able to persist within the cells. The construct was transformed into *Nab. magadii* L11. The transformants were selected for novobiocin (*gyrB*) resistant clones and tested for the presence of *gyrB* within the ORF34 gene by analytical PCR with the primers D34-1 and Nov-11. This PCR results in a 2300 bp long product, if the deletion-cassette inserted in at least one of the ORF34 gene independent of localization or existence of remaining plasmid DNA in the *Nab. magadii* chromosome. As shown in figure 9, for two of the potential clones, a positive PCR product of about 2300bp could be detected (Fig. 9, lanes 4 and 5).



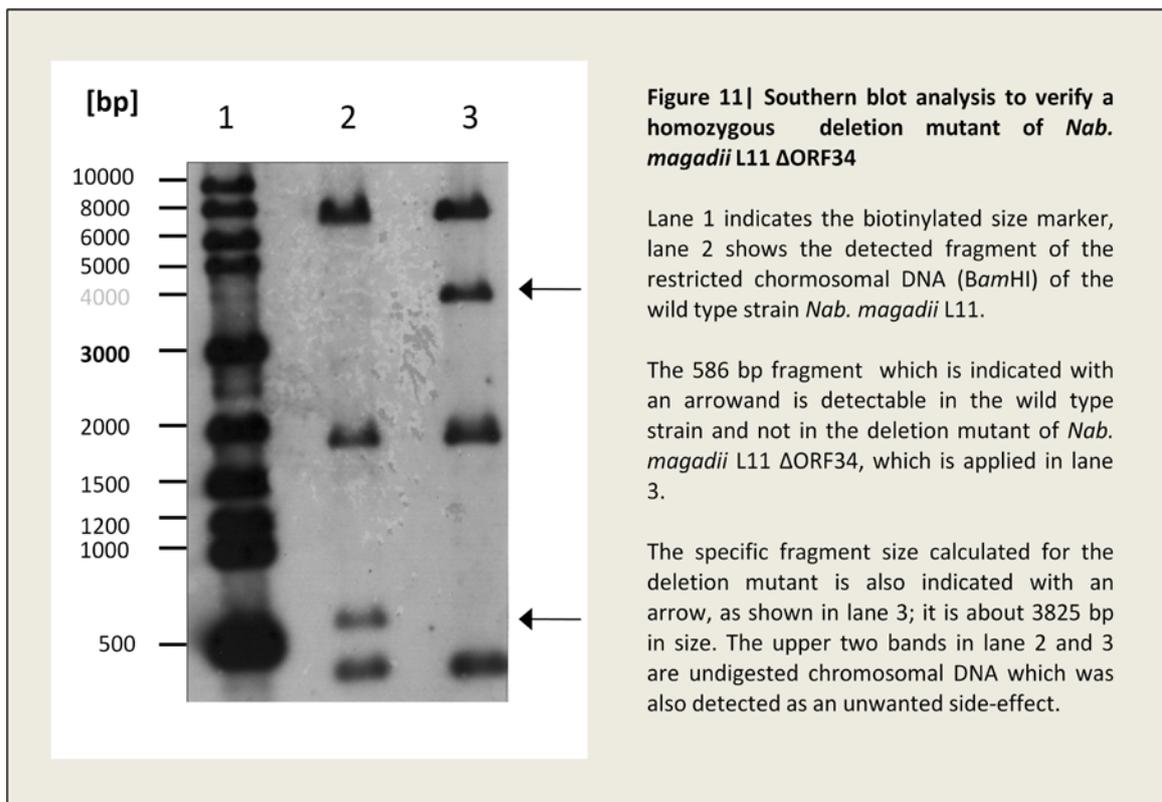
This event happened in consequence of the homologous recombination between the up- and downstream regions of ORF34 on the vector and the corresponding wild type sequences in the Φ Ch1 genome which is shown in a schematically drawing in figure 10. In this way the replacement of the wild type ORF34 gene was successful, respecting the lysogenic cycle of the virus. When a proper recombination event occurred, no PCR product of the plasmid pKSII⁺ Δ ORF34::Nov^R backbone is detectable with the primers pKS-5 and pKS-3 (data not shown; schematically drawing in figure 10).



3.1.3 Homozygation of the *Nab. magadii* L11 ΔORF34 strain

As described before, haloalkaliphilic *Archaea* were shown to represent polyploidy with up to 50 copies of their chromosome, a frequently occurring feature within *Archaea* living extreme conditions. Due to this fact the established deletion mutant was passaged and each time dilutions of the cultures were plated on selective media. The resulting colonies were tested of remaining wild type sequences, as shown in figure 9. After 5 passages the plated culture revealed one potential homozygous strain of *Nab. magadii* L11 ΔORF34 with no detectable wild type signal in the analytical PCR approach (figure 9). The PCR was performed according to the protocol in section 2.2.4.2, using the primers 34-5 and 34-3 for amplification of the wild type ORF 34 gene sequence, yielding a ~1390 bp fragment. For detection of deletion of the ORF 34 sequence caused by the novobiocin cassette, the primers D34-1 and Nov-11 were used. In both cases a positive and negative control was carried out and illustrated in figure 9. As a result the wild type sequence of ORF34 had been completely removed from *Nab. magadii* L11ΔORF34, confirmed by PCR analysis. In order to confirm this homozygous deletion mutant an additional method was performed, the southern blot analysis. For the southern blot analysis the chromosomal DNA from the *Nab. magadii* L11 wild type strain and the putative mutant strain L11ΔORF34 was isolated and followed up as described in section 2.2.8.4. The chromosomal DNA of wt and mutant strain was restricted

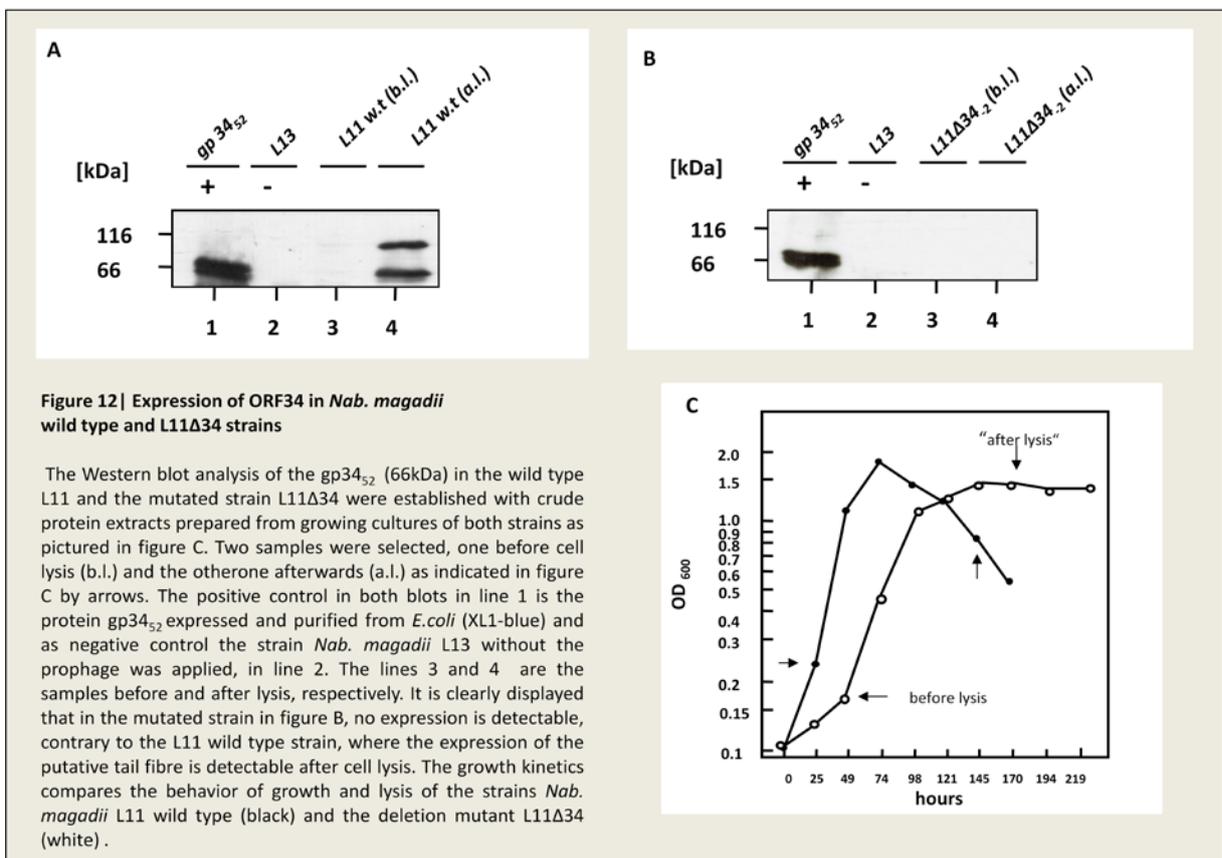
with *Bam*HI and separated on a 0.8 % gel. After blotting onto the nylon membrane, a biotinylated DNA probe of 983 bp in size specific for the upstream region of ORF 34 sequence was used for hybridization. The southern blot shows the expected fragments calculated for the homozygous deletion mutant, the 361 bp, 1847 bp and the 3825 bp fragment specific for the mutant strain (Fig. 11, lane 3), whereas the wild type specific fragment of 586 bp could not be detected in the chromosomal DNA of *Nab. magadii* L11 Δ ORF34 (Fig. 11, lane 2). This leads to the conclusion that the homozygation of the strain *Nab. magadii* L11 lacking the wild type sequence ORF34 was successful, as shown in figure 11.



3.1.4 Expression of the ORF34 gene product in *Nab. magadii* L11 and L11 Δ ORF34 strains

Yet another method was applied to verify the protein expression of the deleted gene region ORF34 additionally to the evidence on DNA level. Therefore a western blot analysis was performed, using an antibody raised against the putative tail fibre protein gp34₅₂. For this purpose a growth curve was performed and the optical density was measured, at the same time crude protein extracts were prepared as described in section 2.2.10. The protein samples were separated on an 8 % PAA gel by SDS PAGE and analysed by an antigen-antibody detection

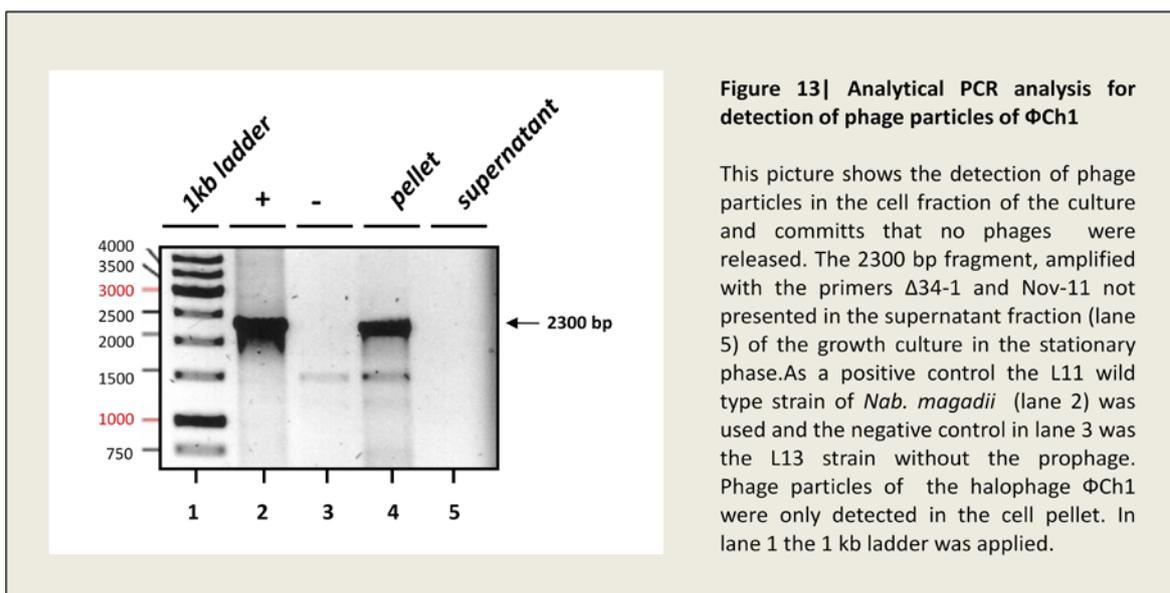
method, carried out according to the instructions described in section 2.2.13.3. For the western blot analysis two comparable time points were selected, one before and the other after onset of cell lysis, as illustrated in figure 12 C (indicated by arrows). It was demonstrated that the *Nab. magadii* L11 wild type strain expresses the gp34₅₂ after cell lysis but not earlier; as it is shown in figure 12, lane 4). A protein with a molecular mass of about 66kDa was detectable. First expression of the putative tail fibre protein was verifiable approximately 98h after inoculation (Rössler et al., 2004). Hence as an unwanted side effect, the band above is considered as an unspecific binding of the antibody α -gp34₅₂. The deletion mutant L11 Δ ORF34 gives no signal of a protein at all. The signal of the positive control (gp34₅₂ expressed and purified in *E.coli* -XL1-blue) in both Western blot analysis is reproducible just like the negative control *Nab. magadii* L13, sample taken form the stationary growth phase (Fig. 12, lanes 1 and 2, respectively).



Therefore it can be concluded, that the gene sequence ORF34 had successful been deleted, hence the putative tail fibre protein is not expressed anymore in the strain *Nab. magadii* L11 Δ ORF34. Therefore it can be assumed that the phage particles lack the terminal structural elements of their contractile tails which are probably responsible for adhesion, consequently combined with the loss of infectivity.

3.1.5 Growth kinetics of *Nab. magadii* L11 wild type and L11ΔORF34 strains

One of the first phenotypically characteristics which had to be observed was the comparison of the growth and lysis behavior of the *Nab. magadii* L11 wild type and the L11ΔORF34 strains. For this analytical purpose the cultures were grown until the late stationary growth phase was reached. As in picture 12 C shown, the graph illustrates a relevant difference of *Nab. magadii* L11 wild type and L11ΔORF34 concerning growth and lysis behavior. It is shown in figure 12 C that the wild type cells L11 (black) began to lyse about 74h after inoculation but the mutated strain L11ΔORF34 did not show a decrease of optical density and therefore no cell lysis. It is supposed that this behavior is combined with no release of mature phage particles. To confirm this assumption a PCR analysis, illustrated in picture 13, was performed to find out whether the phage particles are detectable in the cells or in the supernatant of the growth culture.



The result of the PCR analysis clearly demonstrates that the phage particles of ΦCh1 are enclosed in the *Nab. magadii* L11ΔORF34 cells and are unable to enter the lytic life cycle. The trial to isolate ΦCh1 virus particles from cells of *Nab. magadii* L11ΔORF34 cells was not successful, thus made it not possible to characterize the protein pattern compared to the ΦCh1 wild type protein composition (data not shown).

3.1.6 Expression of the major capsid protein E

Based on the fact that the viruses are not able to lyse the host cells, a further approach was performed in order to verify if the major capsid gene *E* is still expressed in the *Nab. magadii* L11ΔORF34 strain after deletion of ORF34. The western blot analysis, shown in figure 12, was also performed with the *Nab. magadii* L11 wild type strain serving as control. As in figure 14 A and B illustrated, expression of gpE is detectable after onset of lysis in the wild type strain. Opposite to that, no expression of the major capsid protein occurred in the deletion mutant L11ΔORF34. The appearance of the two bands in the positive control (lane 1) of the blot A in picture 14 shows the protein E in a cleaved or degraded form. In contrast the detected gpE in *Nab. magadii* L11 wild type raw extracts in lane 4 (figure 14 A) gave only one signal according to the mature, processed form of the major capsid protein E.

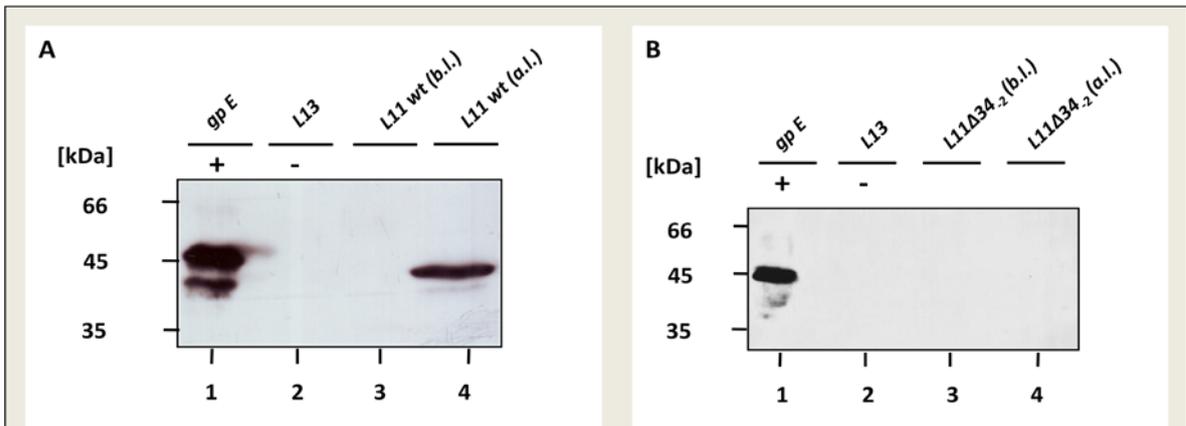


Figure 14 | Western blot analysis of Capsid protein gpE expression

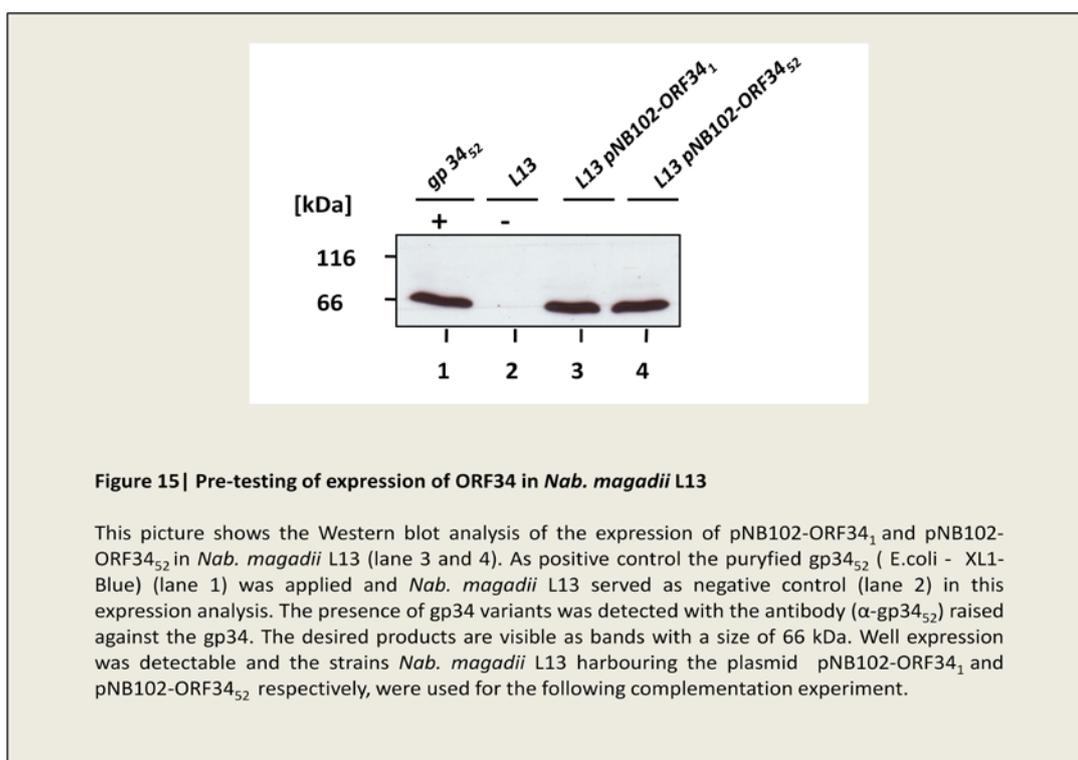
The samples (before lysis and after lysis), selected for this western blot analysis, were the same as those used for the gp34₅₂ expression analysis, shown in figure 12. In lane 1 of both blots, an overexpressed gpE in *E. coli* was purified and used as positive control. The negative control in lane 2 was performed with the *Nab. magadii* L13 strain in late stationary growth phase. In picture A the blot was performed with the *Nab. magadii* L11 wild type strain using an antibody raised against the gpE and in picture B the mutant L11Δ34 vstrain is presented. As illustrated, expression of the major capsid protein E is expressed shortly after cell lysis, whereas the deletion mutant in picture B is not able to express gpE, 170h after inoculation.

3.1.7 Complementation of ΦCh1 lacking the ORF34

In order to verify the effect caused by deletion mutation of the ORF34 sequence in ΦCh1, the mutation was aimed to be complemented by introduction of a functional ORF34 on a plasmid.

3.1.7.1 Retransformation of pNB102-ORF34₁ and pNB102-ORF34₅₂ into *Nab. magadii*

For this purpose, Petra Till constructed vectors with two variants of ORF34, the non-inverted fragment ORF34₁ and the inversion product ORF34₅₂, in 2011. In this course of work the plasmids pNB102-ORF34₁, pNB102-ORF34₅₂ and the empty plasmid pNB102 were transformed into the strain *Nab. magadii* L13. Furthermore crude protein extracts were prepared from those three different clones and pre-selected for well expression, by western blot analysis as illustrated in figure 15.



3.1.7.2 Quick test appliance for complementation of Φ Ch1 lacking the ORF34

After selection of well expressing *Nab. magadii* L13 strains carrying the variants of ORF34 on a shuttle vector (pNB102), the cultures have been prepared for further work. Therefore, competent cells were transformed with isolated chromosomal DNA of *Nab. magadii* L11 Δ ORF34, since the isolation of L11 Φ Ch1 Δ ORF34 was not feasible. The procedure of transformation and isolation of the chromosomal DNA has been carried out according the instruction described in section 2.2.8.4. The quick test was performed for determination of the phage titer without screening for positive transformants and following isolation of the virus. After the procedure of transformation the

phage titer with four different dilution batches (undiluted, 10^{-2} , 10^{-4} , 10^{-6}) was performed as described in section 2.2.9.2.

3.1.7.3 Phage titer evaluation

The phage titer was incubated for 9 days at 37°C. After this time no plaque formation, resulting from infection of *Nab. magadii* L13 with re-established, intact wild type viruses, was observed on the plates. Extended incubation time did not yield another outcome of the titer analysis. No plaques were visible in all executed approaches, performed with the strains *Nab. magadii* L13 (pNB102) as control, L13 (pNB102-ORF34₁) and L13 (pNB102-ORF34₅₂) transformed with chDNA of *Nab. magadii* L11ΔORF34.

3.1.8 Discussion

In the course of this diploma work a deletion mutant of the halophage ΦCh1, lacking the ORF34, was performed by homolog recombination with a deletion cassette on a vector and the genome of *Nab. magadii* L11, which has the integrated ΦCh1 DNA in its genome. The homozygotation of the strain *Nab. magadii* L11ΔORF34 was verified both, by PCR and Southern blot analysis. Moreover, a Western blot analysis was performed for the confirmation of the absence of the deleted gene product. The deployed antibody was directed against the gene product gp34₅₂ (see figure 12) and the fail of expression of the desired gene was definitely approved in the strain *Nab. magadii* L11ΔORF34. However, the elimination of ORF34, which is supposed to represent the sequence of encoded putative tail fibre proteins, yielded no mature phage particles. Due to the fact that the *Nab. magadii* L11ΔORF34 strain did not enter the lytic cycle, no ΦCh1ΔORF34 phage particles could be isolated for observations under the electron microscope to visualize the absence of the putative tail fibre proteins. Furthermore, the presence of the major capsid protein E was analysed, by Western blot detection (see figure 14). The appearance of the two bands in the positive control (lane 1) of the blot A in picture 14 showed the protein E in a cleaved or degraded form. In contrast the detected gpE in *Nab. magadii* L11 wild type raw extracts in lane 4 (figure 14 A), gave only one signal according to the mature, processed form of the major capsid protein E. Cleavage of viral head proteins is thought to be a demand for release of progeny precursors from the host membrane, and therefore it is feasible that also the protein E of ΦCh1 is involved in DNA packaging (Siegel and Schaechter, 1973; Klein *et al.*, 2010). Due the provided evidence of the absence of the major capsid protein E in *Nab. magadii* L11ΔORF34, it is conceivable that protein E is involved in virus assembly of structural proteins or responsible for

virus release from the host cell. This would also explain why the deletion mutant *Nab. magadii* L11ΔORF34 is not able to enter into the lytic life cycle. The halophage ΦH is one of the most intensively studied viruses within the group which also includes ΦCh1 and is infecting *Halobacterium salinarum*. It has been revealed that the similarity of the capsid protein Hp32, from ΦH and the major capsid protein E from ΦCh1 have an 80% identity (Klein *et al.*, 2010). Further studies included analysis of lytic control which seem to be controlled by action of transcriptional repressor as well as antisense RNA (Stolt *et al.*, 1992, Klein *et al.*, 2010). There is the possibility that the very late expressed (mRNA detectable after 74h) capsid protein E underlies the control of other structural component proteins or factors. The onset of transcription of the ORF11 (gpE) takes place within a very short period of time, hence resulting in a nearly immediate high level of transcription (Klein *et al.*, 2010). It is assumed that the protein gpE might be associated with the host membrane. Typically features for virus proteins that are coupled with the host membrane are, the involvement in DNA packaging and/or progeny assembly, like it is believed in bacteriophages T4 or P22 respectively (Klein *et al.*, 2010). It is evident that due to the failure of expression of the capsid protein E no assembly of mature phage particle is practicable.

Additionally, the deletion of ORF34 was attempted to recover its functionality by introducing a functional sequence, expressing the gp34₅₂. The complementation application was performed with investigation of plaque formation, caused by retrieval of infectivity of the mutant virus. Yet, performance of complementation with variants of ORF34 encoded products (gp34₁ and gp34₅₂) did not yield any effect on plaque formation. It is presumably that the induced deletion mutation affects transcription whereby, subsequently expression performance (capsid protein E) is dysfunction.

Based on the present results, the key role of the gene product of ORF34 in the initiation of infection cannot really be supposed with established outcome of experiments, in this case of course. For further proposal, to achieve a proper transcription, an archaeal termination sequence integrated in the deletion construction sequence might be an approach to solve the problem.

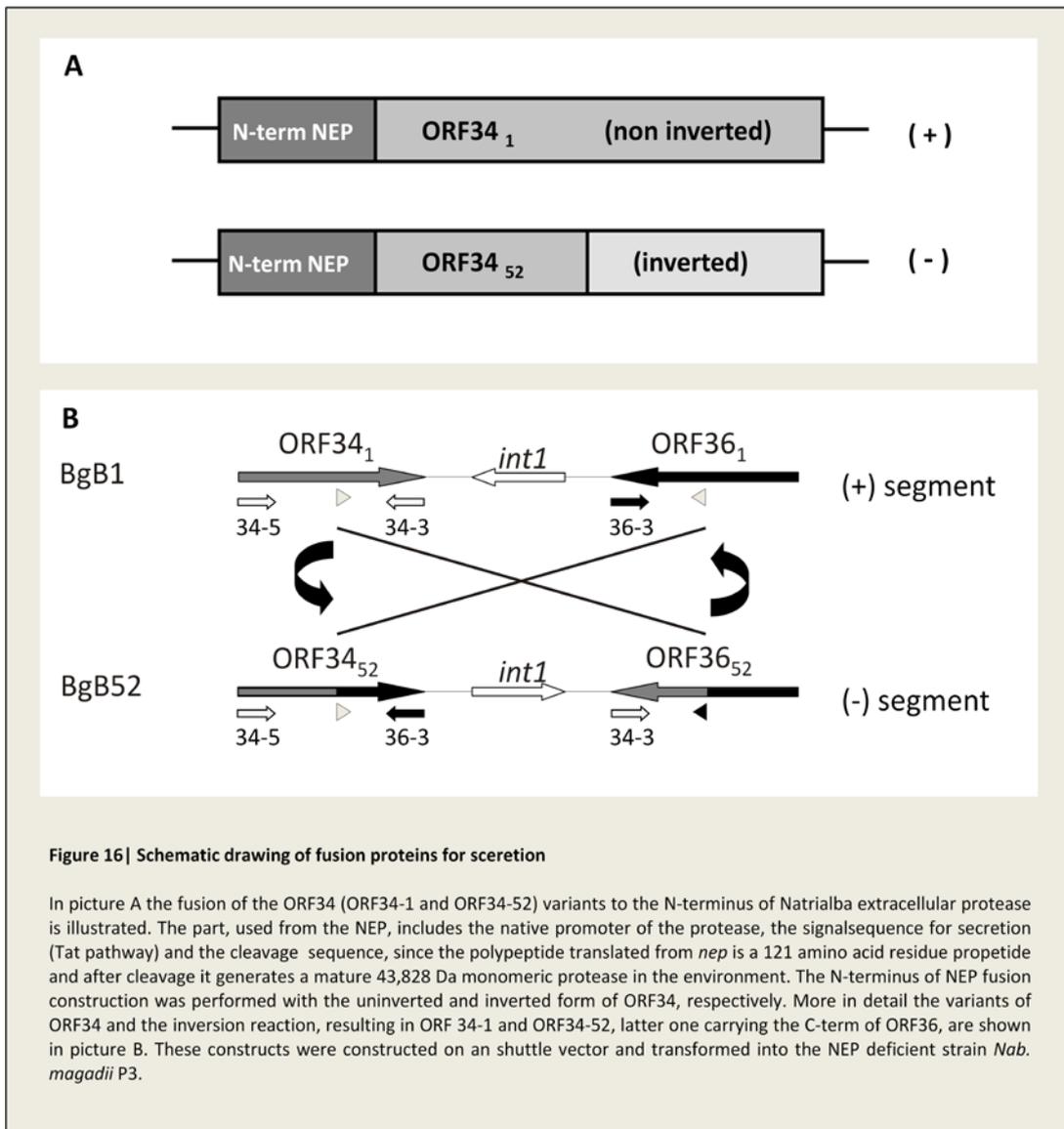
3.2 Secretion of recombinant proteins to prevent adhesion of Φ Ch1

In this course of work a further experiment was performed to confirm that the product of ORF34 acts as a putative tail fibre protein. In 2011, Petra Till attempted to prevent infectivity of the virus Φ Ch1 by addition of purified gp34₅₂ (overexpressed in *E. coli*) into the growth culture of *Nab. magadii*. This experiment showed limited success since the protein precipitated, which was possibly affected by different proteins folding and/ or modification in *E. coli*, thus making it not compatible with the translated protein in its host strain *Nab. magadii*. This evolved into the idea, of exporting the protein variants of gp34 (gp34₁ and gp34₅₂) to prevent infection of the virus Φ Ch1 with the purpose of occupying the receptors on the *Nab. magadii* cell surface; which are involved in viral tail fibre interaction.

3.2.1 Essential preconditions for exporting recombinant proteins from haloalkaliphilic organism

A crucial requirement to perform secretion of recombinant fusion proteins was the knowledge of the fact that *Nab. magadii* possesses an extracellular protease, since it is a proteolytically growing organism. In 2009, Christian Derntl successfully created a *Natrialba* extracellular protease (Nep) deficient strain, named *Nab. magadii* P3 (Derntl, 2009).

Therefore, Nep was used as facility for secretion of the putative tail fibre protein in this application by fusion of the putative tail fibre proteins with the N-terminus of the protease, as shown in figure 16. The construct, cloned into a shuttle vector according to the description in section 2.2.7, including the native promoter of the protease, the signal sequence for export (Tat-pathway) and the cleavage sequence at which the latter one is essential to ensure, that the gp34 variants are presented in the supernatant.



Generally, the expression of Nep starts in the late exponential growth phase and is translated as a polypeptide including a 121 amino acid residue propeptide (De Castro *et al.*, 2008). Hence, after cleavage it generates a mature 43,828 Da monomeric protease in the supernatant, which was replaced in this case by the gp34₁ and gp34₅₂, respectively.

3.2.2 Expression of ORF34 variants in *Nab. magadii* P3 and verification of secretion

The shuttle vector pNB102 carrying the N-terminus of Nep with the fused ORF34 (inverted and non inverted) variants were transformed into the Nep deficient strain *Nab. magadii* P3. Additionally the empty shuttle vector was transformed, serving as negative control. Cultures of all three strains (*Nab. magadii* P3 (pNB102), *Nab. magadii* P3 pNPro-34₁ and *Nab. magadii* P3 pNPro-

34₅₂) were grown until the stationary phase was reached and crude protein extracts were prepared. For the Western blot analysis, the pellet fraction and the proteins precipitated from the supernatant were used, respectively. The supernatant fractions were carefully prepared by separating cells by centrifugation with a following second centrifugation step the supernatant was cleared and additionally filtration (0.2µm) was carried out to ensure that no cells were remaining in the samples. Then the proteins were precipitated with TCA and further preparations, needed for western blot analysis. The following detection was performed with an antibody raised against the gp34₅₂ protein. As illustrated in figure 17 the samples of all growing *Nab. magadii* P3 transformants were taken at the stationary phase were expression of Nep already started. The western blot analysis detected the gp34 variants in the pellet fractions and in the supernatant samples, which were applied from the strains *Nab. magadii* P3 pNPro-34₁ and *Nab. magadii* P3 pNPro-34₅₂. In contrast, no detection of ORF34 products is visible in the pellet or supernatant fraction of the strain *Nab. magadii* P3 carrying the empty pNB102 shuttle vector. As positive control the expressed and purified gp34₅₂ from *E. coli* was applied. For comparison all constructs were transformed into the wild type *Nab. magadii* L13 strain and western blot analysis was performed (data not shown).

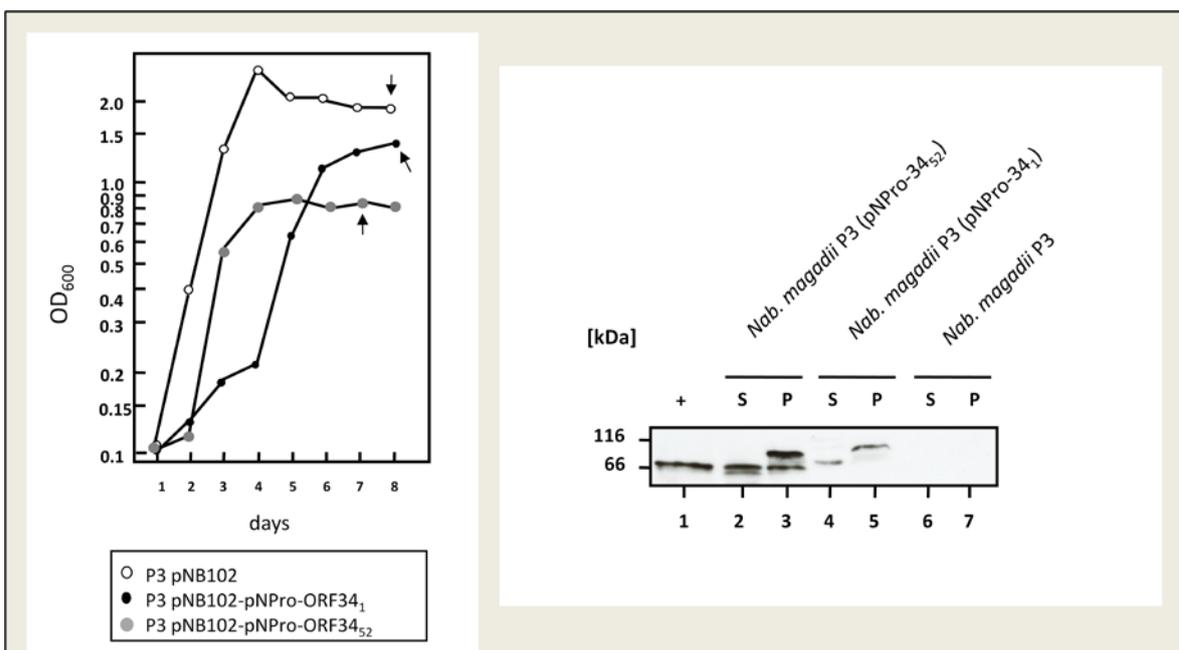
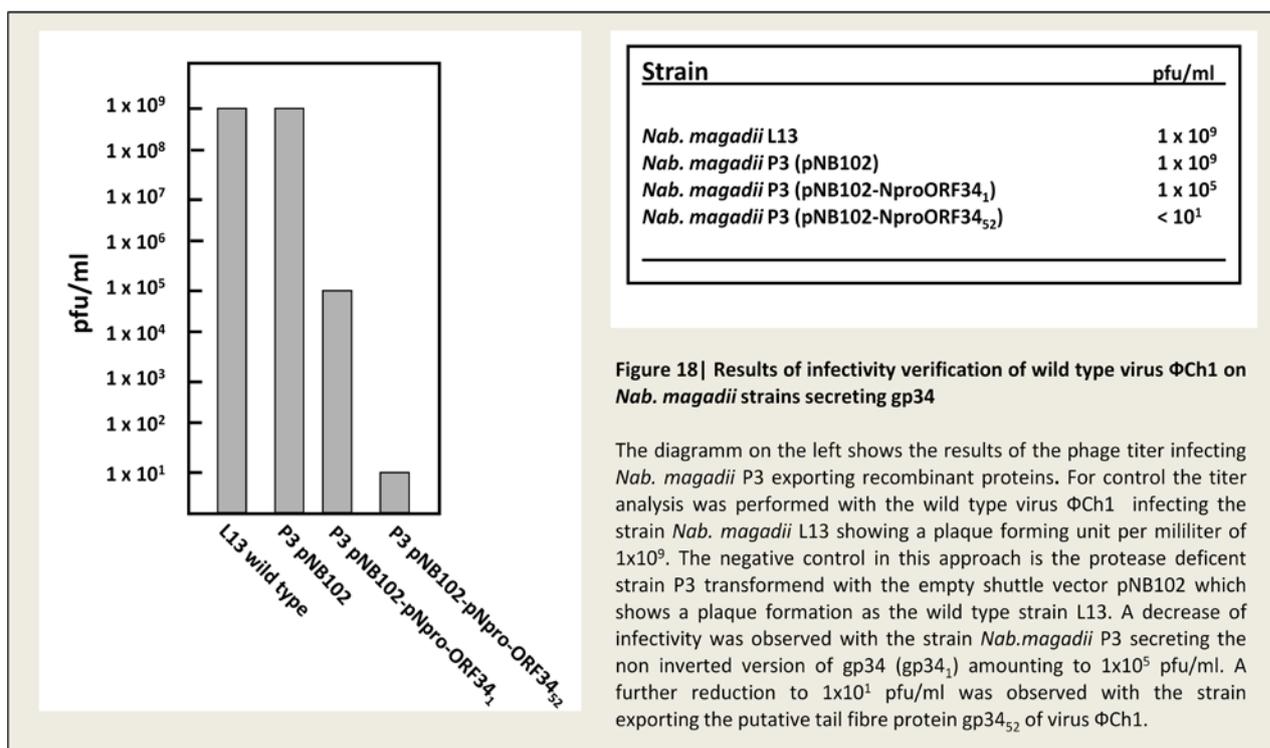


Figure 17 | Growth curve of *Nab. magadii* P3 transformants and verification of secretion of recombinant proteins

The growth curve shows three *Nab. magadii* P3 transformants (*Nab. magadii* P3 (pNB102), *Nab. magadii* P3 pNPro-34₁ and *Nab. magadii* P3 pNPro-34₅₂) with different ORF34 variants. The errors are labeling the time points (stationary growth phase) which were taken for the western blot analysis, which is shown on the right part of the figure. The positive control in lane 1 is gp34₅₂ (expressed and purified from *E. coli*), in lane 6 and 7 the negative control strain was applied including both, pellet and supernatant fraction. In lane 2 and 3 the putative tial fibre protein in its inverted version was detected in the pellet fraction as well as in the supernatant. Lane 4 and 5 shows the detection of the gp34₁ variant, which is not inverted and is also presented in both fractions. The different in size of the bands mirrors the unprocessed and processes or cleaved protein in the supernatant and pellet fraction, respectively.

3.2.3 Infectivity analysis of Φ Ch1 on gp34 secreting *Nab. magadii* strains

For this purpose a phage titer was performed to observe if the virus Φ Ch1 is still able to adhere on gp34 exporting *Nab. magadii* strains. The procedure was carried out as described in section 2.2.9.2. All *Nab. magadii* P3 strains transformed with the respective constructs (pNB102 empty, pNB102 pNPro-34₁ and pNB102 pNPro-34₅₂) were grown until the stationary phase was reached; infected with the wild type virus Φ Ch1 and incubated at 37°C. Plaque formation was observed as illustrated in the diagram of figure 18, after 8 days of incubation. This infectivity analysis was conducted more often (six times) to be absolutely sure that the application worked.



3.2.4 Discussion

Within this course of diploma work a further experiment was performed in order to confirm gp34₅₂ as a tail fibre protein, since the deletion of the respective gene sequence executed no affirmation. Among others, the origin of the idea to export the putative tail fibre protein into the environment in order to prevent adhesion of the virus Φ Ch1 on *Nab. magadii* cells arose, since addition of purified gp34₅₂ from *E. coli* into the *Nab. magadii* L13 culture did not yield any success.

The expression of gp34 variants was thought to be performed in its host conditions to ensure proper folding and possible modification of the putative tail fibre protein. The putative tail fibre proteins result from an exchange of the 3' ends of ORFs 34 and 36, yielding, among others, in gp34₁ (non-inverted) and gp34₅₂ (inverted) proteins, as illustrated in figure 6. Experiments indicated that just gp34₅₂ but not gp34₁ is responsible for the binding to the host cell. Further observations showed that the C-terminal but not the N-terminal truncation of gp34₅₂ enabled the attachment to the host cell and represents the functional binding domain. While the N-terminal part is supposed to be associated with the viral capsid and the shaft of the tail it seems to be negligible for the actual binding process. Both, gp34₁ and gp34₅₂, contain predicted galactose-binding domain within their C-termini, suggesting that a pretended galactose residue is located on the *Nab. magadii* cell surface and therefore is involved in binding. It could be shown that α -D-galactose dramatically inhibited the attachment of viral particles to the cell surface, hence preventing infectivity of ϕ Ch1 (Klein *et al.*, 2013). As vehicle for the export of those two selected gp34 variants (gp34₁ and gp34₅₂), the *Natrialba* extracellular protease was chosen. This protease is translated as a propeptide including a 121 amino acid residue (12,626 Da) which is cleaved to generate a mature 43,828 Da active protease in the environment (De Castro *et al.*, 2008). Experimental evidences verified that processing and activation during/after translocation is autocatalytic, with the hypothesis that cleavage of the signal peptide may be required to induce the autocatalytically process of removing the Nep propeptide (Ruiz *et al.*, 2012). Detection of Nep was observed during the transition to the stationary growth phase; which leads to the suggestion that production is stimulated by nutrient scarcity and/or the population density. In figure 16 a schematically drawing show the fusion construction of the N-terminus of Nep with the gene sequences for the products gp34₁ and gp34₅₂, respectively. This in frame fusion constructions were transformed by a shuttle vector into *Nab. magadii* P3, a protease deficient strain, created in 2009 by Christian Derntl. The autocatalytically features of Nep enables to export the putative tail fibre proteins without residues from the N-terminus of the protease because of the cleavage before export occurs. The confirmation of a successful secretion was performed by western blot analysis and is shown in figure 17. The selection of the timepoints for the samples was set in the stationary growth phase of all three cultures, followed by a carefully protein preparation of the supernatant to get rid of remaining cells, to overcome false positive results (see section 2.2.10.2). The detection of the putative tail fibre proteins were carried out with the polyclonal antibody against the gp34₅₂ protein. The detection strategy show signals for both proteins (gp 34₁ and gp34₅₂), explained by the occurrence of similar epitops in both variants which are recognized from α -gp34₅₂. The effect of a more width band in lane 1 with 66kDa in size is, which is indicating the

positive control, is caused by a different preparation method with different conditions (expressed and purified from *E. coli*). In figure 17, lane 6 and 7 the samples for the negative control (*Nab. magadii* P3 pNB102 empty) was applied and no detection was observed. This indicates that no phage particle, encoding a putative tail fibre protein, is presented in the culture of the negative control. Whereas in the figure 17, lanes 2, 3, 4 and 5, protein detection of two different sized bands were observed. The explanation of the varying sizes of the bands is given due to the fact, that the N-terminus of the protease is still coupled to the fused viral proteins during the process of export. This is clearly visible in the cell fraction, in lane 3, where the strain *Nab. magadii* P3 is secreting gp34₅₂. This agrees with the about 13 kDa protease residue, which is cleaved after secretion and therefore not presented in the supernatant fraction in lane 2. However the detection of gp34₁ in both applied fractions (pellet and supernatant), gives a weaker signal as the other variant does; maybe because of a lowered expression rate of the vector and/or the recognition of the antibody is decreased.

Hence with these results a successful confirmation of secreting recombinant proteins from a haloalkaliphilic system could be observed. This reveals a large number as variety of features and applications in the area of biotechnological research, for instance towards the water treatment where harsh conditions have to be overcome; obviously with tough archaeal systems.

Stated previously, the most complex attachment structures with the greatest number of proteins which are involved in the tail constitute and function are represented by the *Myoviridae* viruses, including the halophage ϕ Ch1 (Leimann *et al.*, 2010). Infection processes are highly diverse and rather little is known about the way of infection, towards archaeal viruses. The initialization of the adsorption process is characterized by the recognition of some structural components on the host cell surface through viral tail fibre proteins or analogous structures. Receptors for archaeal viruses still have not been well studied and characterized.

Relating to the infectivity analyses it is expected that, the secreted, inverted version gp34₅₂ is able to bind the receptors on the cell surface of *Nab. magadii* P3, hence preventing the adhesion of the halophage ϕ Ch1; consequentially leading to no, or much less plaque formation on the plates.

Apart from unusual statistic outliers, explained by handling failure during preparation of the phage titer, the means of the method does render reliable results and can be reproduced. The infectivity analyses were performed in triplicates with 400 μ l of the respective strains exporting gp34 variants from stationary growth phase which were infected with 100 μ l of ϕ Ch1 dilutions (10^4 , 10^6 and 10^8). The strain *Nab. magadii* P3, carrying the empty shuttle vector pNB102, shows as

much plaque formation as the wild type strain *Nab.magadii* L13 after infection with the wild type virus ϕ Ch1. In both cases, a plaque forming unit per milliliter of 1×10^9 was observed showing turbid plaques with an average of 2–5 mm in size. However, fewer plaques were observed on infectivity analyses performed with the strain *Nab. magadii* P3 secreting the non inverted product of ORF34 (gp34₁), although it is assumed not to be involved in the infectivity process of ϕ Ch1. The reduction to a value of 10^5 pfu /ml could be a result of a worse growth behavior of this strain.

As shown in figure 17, *Nab.magadii* P3 (pNB102 pNPro-34₁) reached an optical density (OD₆₀₀) of 0.8, whereas strain *Nab.magadii* P3 (pNB102 pNPro-34₅₂) grows up to an OD₆₀₀ of 1.3. Therefore it is supposed that the reduced growth has an direct influence on the infectivity of ϕ Ch1. The infectivity analyses of the strain *Nab. magadii* P3 (pNB102 pNPro-34₅₂) exporting the putative tail fibre protein gp34₅₂ yielded a clearly decrease in the plaque forming unit with a rate less than 1×10^1 pfu/ml. This is a reduction of infectivity of at least 8 orders of magnitude compared to the control strain *Nab. magadii* P3 (pNB102) and 5 orders of magnitude compared to *Nab. magadii* P3 (pNB102 pNPro-34₁), indicating that infectivity of ϕ Ch1 can be hindered by masking the receptor with gp34₅₂. These results gave another hint, that gp34 is a tail fibre protein.

4. References

- A Oren, 2002: **Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications.** J Ind Microbiol Biotechnol. 28(1):56-63.
- Abremski, K.E. and Hoess, R.H., 1992: **Evidence for a second conserved arginine residue in the integrase family of recombination proteins.** Protein engineering 5: 87-91
- Allers T, Mevarech M., 2005: **Archaeal genetics - the third way.** Nature Reviews Genetics. 6(1):58-73.
- Andreea Oarga, 2009: **Life in extreme environments.** Revista de Biologia e Ciências da Terra ISSN 1519-5228 Volume 9 – Nr. 1 2009
- Argos, P. et al., 1986: **The integrase family of site-specific recombinases: regional similarities and global diversity.** EMBO journal 5: 433-40
- Armen Y Mulkidjanian, Michael Y Galperin, Kira S Makarova, Yuri I Wolf and Eugene V Koonin, 2008: **Evolutionary primacy of sodium bioenergetics.** Biology Direct 3:13
- Baranyi U, Klein R, Lubitz W, Krüger DH, Witte A., 2000: **The archaeal halophilic virus-encoded Dam-like methyltransferase M.φCh1-I methylates adenine residues and complements dam mutants in the low salt environment of Escherichia coli.** Molecular Microbiology. 35(5):1168-79.
- Barns SM, Delwiche CF, Palmer JD, Pace NR., 1996. **Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences.** Proc. Natl. Acad. Sci. USA. 93:9188-93.
- Barry Elizabeth R. and Bell Stephen D., 2006: **DNA Replication in the Archaea.** Microbiol. Mol. Biol. Rev. 70(4):876. DOI:
- Barry ER, Bell SD., 2006: **DNA Replication in the Archaea.** Microbiology and Molecular Biology Reviews. 70(4):876-87.
- Bart de Koning, Fabian Blombach, Stan J. J. Brouns, and John van der Oost, 2010: **Fidelity in Archaeal Information Processing.** Nucleic Acids Res. 38(6):1950-63.
- Bath C, Dyall-Smith ML, 1998: **His1, an archaeal virus of the Fuselloviridae family that infects Haloarcula hispanica.** J Virol. 72:9392-9395
- Béjà O, Spudich EN, Spudich JL, Leclerc M, DeLong EF., 2001: **Proteorhodopsin phototrophy in the ocean.** Nature. 2001 Jun 14;411(6839):786-9.
- Bell SD, Jackson SP., 1998: **Transcription and Translation in Archaea: a mosaic of eukaryal and bacterial features.** Trends in Microbiology. 6(6):222-8.

- Bell Stephen D, Dionne Isabelle, Si-houy Lao-Sirieix, Victoria L. Marsh, Adam McGeoch, Nicholas P. Robinson, 2005: **Archaeal DNA Replication:A Robust Model for Eukaryotes**. Geothermal Biology and Geochemistry in YNP 1:277-288
- Bendtsen J. D., Nielsen H., Widdick D., Palmer T., Brunak S., 2005: **Prediction of twin-arginine signal peptides**. BMC Bioinformatics 6, 167, 1471-2105-6-167.
- Benelli D, Maone E, Londei P., 2003: **Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation**. Mol Microbiol. 2003 Oct;50(2):635-43.
- Blakely, G.W. and Sherratt, D.J., 1996: **Cis and trans in site-specific recombination**. Molecular Microbiology 20: 233-8
- Bolhuis A., 2002: **Protein transport in the halophilic archaeon Halobacterium sp. NRC-1: a major role for the twin-arginine translocation pathway?** Microbiology.;148 (Pt 11):3335-46.
- Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P., 2008: **Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota**. Nature Reviews Microbiology 6, 245-252
- Bullock, C. ,2000: **The Archaea- a biochemical perspective**. Biochem. Mol. Biol. Edu. **28** 186-191.
- Cabrera, J. A., J. Bolds, P. E. Shields, C. M. Havel, and J. A. Watson., 1986: **Isoprenoid synthesis in Halobacterium halobium. Modulation of 3-hydroxy-3-methylglutaryl coenzyme A concentration in response to mevalonate availability**. J. Biol. Chem. 261:3578-3583.
- Cao TB, Saier MH Jr., 2003: **The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions**. Biochim Biophys Acta. 1609(1):115-25.
- Chatton, E., 1938: **Titre et travaux scientifique (1906-1937)** de Edouard Chatton. E. Sottano, Sète, France
- Cavicchioli Ricardo, 2011: **Archaea – timeline of the third domain**. Nat Rev Microbiol. Jan;9(1):51-61.
- Cline SW, Doolittle WF., 1987: **Efficient transfection of the archaeobacterium Halobacterium halobium**. J. Bacteriol. 169:1341-1314.
- D’Alessandro C.P., De Castro R.E., Giménez M.I. and Paggi R.A., 2007: **Effect of nutritional conditions on extracellular protease production by the haloalkaliphilic archaeon Natrionalba magadii**. Lett Appl Microbiol. 44(6):637-42.
- Daniel I, Oger P, Winter R., 2006: **Origins of life and biochemistry under high-pressure conditions**. Chem Soc Rev. 2006 Oct;35(10):858-75.
- De Carlo Sacha, Shih-Chieh Lin, Dylan J Taatjes, and Andreas Hoenger, 2010: **Molecular basis of transcription initiation in Archaea**. Transcription 1(2): 103–111.

De Castro RE, Maupin-Furlow JA, Giménez MI, Herrera Seitz MK, Sánchez JJ, 2006: **Haloarchaeal proteases and proteolytic systems**. FEMS Microbiol Rev. 30:17-35

De Castro Rosana E., Ruiz Diego M., Giménez María I., Silveyra María X., Paggi Roberto A., Maupin-Furlow Julie A., 2008: **Gene cloning and heterologous synthesis of a haloalkaliphilic extracellular protease of *Natrialba magadii* (Nep)**. Extremophiles 12:677–687

DeLong EF, Pace NR., 2001: **Environmental Diversity of Bacteria and Archaea**. Syst Biol. Aug;50(4):470-8.

Derntl, C., 2009: **Construction of the first deletion mutant of a haloalkaliphilic archaeon and analysis of gene expression of the methyltransferase M.Nma ϕ Ch1I of the halophage ϕ Ch1**. Diploma thesis, University of Vienna

Dilks K., Gimenez M. I., Pohlschroder M., 2005: **Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea**. J. Bacteriol. 187, 8104–8113.

Dyall-Smith ML, Tang SL, Bath C, 2003: **Haloarchaeal viruses: how diverse are they?** Res Micro-biol. 154:309-313

Eichler J., 2003: **Facing extremes: archaeal surface-layer (glyco)proteins**. Microbiology. 149(Pt 12):3347-51.

Enache M., Kamekura M., 2010: **Hydrolytic enzymes of halophilic microorganisms and their economic values**. Rom. J. Biochem., 47(1): 47 – 59.

F Confalonieri and S Sommer, 2011: **Bacterial and archaeal resistance to ionizing radiation**. J. Phys.: Conf. Ser. 261 012005

Falb Michaela, Müller Kerstin, Königsmaier Lisa, Oberwinkler Tanja, Horn Patrick, von Gronau Susanne, Gonzalez Orland, Pfeiffe Friedhelm, Bornberg-Bauer Erich, Oesterhelt Dieter, 2008: **Metabolism of halophilic archaea**. Extremophiles 12:177–196

Forterre P, Brochier C, Philippe H., 2002: **Evolution of the Archaea**. Theoretical Population Biology. 61(4):409-22.

Gill Erin E and Brinkman Fiona S L., 2011: **The proportional lack of archaeal pathogens: Do viruses/phages hold the key?** Bioessays. 33(4): 248–254.

Giménez MI, Studdert CA, Sánchez JJ, De Castro RE, 2000: **Extracellular protease of *Natrialba magadii*: purification and biochemical characterization**. Extremophiles 4:181-188

Gohlke U., Pullan L., Mcdevitt C. A., Porcelli I., De Leeuw E., Palmer T., Saibil H. R., Berks B. C., 2005: **The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter**. Proc. Natl. Acad. Sci. U.S.A. 102, 10482–10486.

Grabowski B, Kelman Z., 2003: **Archeal DNA replication: eukaryal proteins in a bacterial context**. Annu Rev Microbiol.;57:487-516.

Grabowski B, Kelman Z., 2003: **Archeal DNA replication: eukaryal proteins in a bacterial context.** Annu Rev Microbiol.;57:487-516.

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

Gropp F, Grampp B, Stolt P, Palm P, Zillig W., 1992: **The immunity-conferring plasmid p ϕ HL from the Halobacterium salinarium phage ϕ H: Nucleotide sequence and transcription.** Virology. 190(1):45-54.

Grünberg S, Reich C, Zeller ME, Bartlett MS, Thomm M., 2010 : **Rearrangement of the RNA polymerase subunit H and the lower jaw in archaeal elongation complexes.** Nucleic Acids Res. 38(6): 1950–1963.

Hallet, B., and Sherratt, D.J., 1997: **Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements.** FEMS microbiology reviews 21: 157-178

Hendrix RW, Lawrence JG, Hatfull GF, Casjens S., 2000: **The origins and ongoing evolution of viruses.** Trends in Microbiology. 8(11):504-8.

Holmes ML, Nuttall SD, Dyll-Smith ML., 1991: **Construction and use of halobacterial shuttle vectors and further studies on Haloferax DNA gyrase.** J. Bacteriol. 173:3807-3813.

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO., 2002: **A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont.** Nature. 417:63-7.

Huber H, Hohn MJ, Stetter KO, Rachel R., 2003: **The phylum Nanoarchaeota: Present knowledge and future perspectives of a unique form of life.** Research in Microbiology. 154(3):165-71.

Hutcheon GW, Bolhuis A., 2003: **The archaeal twin-arginine translocation pathway.** Biochem Soc Trans.31(Pt 3):686-9.

Iro M, Klein R, Gálos B, Baranyi U, Rössler N, Witte A., 2007: **The lysogenic region of virus ϕ Ch1: identification of a repressor-operator system and determination of its activity in halophilic Archaea.** Extremophiles. 11(2):383-96.

Iwashita Shintaro and Kanegasaki Shiro, 1976: **Enzymic and Molecular Properties of Base-Plate Parts of Bacteriophage P22.** Eur. J. Biochem. 65, 87-94

Kamekura M and Kates M, 1999: **Structural diversity of membrane lipids in members of halobacteriaceae.** Bioscience, Biotechnology and Biochemistry, 63:969_972

Kamekura, M., Dyll-Smith, M. L., Upasani, V., Ventosa, A. & Kates, M., 1997: **Diversity of alkaliphilic halobacteria: proposals for transfer of *Natronobacterium vacuolatum*, *Natronobacterium magadii*, and *Natronobacterium pharaonis* to *Halorubrum*, *Natrialba* and *Natronomonas* gen. nov., respectively, as *Halorubrum vacuolatum* comb. nov., *Natrialba magadii* comb. nov., and *Natronomonaspharaonis* com. nov., respectively.** Znt J Syst Bacteriol47, 853-857

- Kelman LM, Kelman Z., 2004: **Multiple origins of replication in archaea.** Trends Microbiol. 2004 Sep;12(9):399-401.
- Kelman, L.M., and Z. Kelman., 2003. **Archaea: an archetype for replication initiation studies?** Mol Microbiol 48:605-16.
- Klein R, Baranyi U, Rössler N, Greineder B, Scholz H, Witte A., 2002: **Natrialba magadii virus ϕ Ch1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon.** Molecular Microbiology. 45(3):851-63
- Klein R, Greineder B, Baranyi U, Witte A., 2000: **The Structural Protein E of the Archaeal Virus ϕ Ch1: Evidence for Processing in Natrialba magadii during Virus Maturation.** Virology. 276(2):376-87.
- Krupovič Mart, Forterre Patrick, Bamford Dennis H.,2010: **Comparative Analysis of the Mosaic Genomes of Tailed Archaeal Viruses and Proviruses Suggests Common Themes for Virion Architecture and Assembly with Tailed Viruses of Bacteria.** J Mol Biol. 19; 397(1):144-60.
- Lam, W. L., and W. F. Doolittle., 1989: **Shuttle vectors for the archaeobacterium Halobacterium volcanii.** Proc. Natl. Acad. Sci. USA 86:5478-5482.
- Leake M. C., Greene N. P., Godun R. M., Granjon T., Buchanan G., Chen S., Berry R. M., Palmer T., Berks B. C., 2008: **Variable stoichiometry of the TatA component of the twin-arginine protein transport system observed by in vivo single-molecule imaging.** Proc. Natl. Acad. Sci. U.S.A. 105, 15376–15381.
- Leiman Petr G, Arisaka Fumio, van Raaij Mark J, Kostyuchenko Victor A, Aksyuk Anastasia A, Kanamaru Shuji, and Rossmann Michael G., 2010: **Morphogenesis of the T4 tail and tail fibers.** Virol J. 7: 355.
- Lindberg Ali A., 1973: **Bacteriophage receptors.** Annual Review of Microbiology Vol. 27: 205-241
- Lynn J. Rothschild & Rocco L. Mancinelli, 2001: **Life in extreme environments.** Nature 409, 1092-1101
- M L Holmes and M L Dyall-Smith, 1991: **Mutations in DNA gyrase result in novobiocin resistance in halophilic archaeobacteria.** J Bacteriol. 173(2): 642–648.
- Madern D, Ebel C, Zaccai G, 2000: **Halophilic adaptation of enzymes.** Extremophiles 4:91-98
- Mayrhofer-Iro M. , Ladurner A., Meissner C., Derntl C., Reiter M. , Haider F., Dimmel K., Rössler N., Klein R., Baranyi U. , Scholz H., and Witte A. , 2013: **Transformation of Natrialba magadii: application of ϕ Ch1 elements for a shuttle vector system for haloalkaliphilic Archaea.** Appl. Environ. Microbiol. AEM.03287-12
- Meixian Zhou, Hua Xiang, Chaomin Sun & Huarong Tan, 2004: **Construction of a novel shuttle vector based on an RCR-plasmid from a haloalkaliphilic archaeon and transformation into other haloarchaea.** Biotechnology Letters 26: 1107–1113.

Mitraki A, Miller S, van Raaij MJ., 2002: **Review: conformation and folding of novel beta-structural elements in viral fiber proteins: the triple beta-spiral and triple beta-helix.** J Struct Biol. 137(1-2):236-47. Review.

Mizuuchi, K., M. H. O'Dea, and M. Gellert., 1978: **DNA gyrase: subunit structure and ATPase activity of the purified enzyme.** Proc. Natl. Acad. Sci. USA 75:5690-5963.

Myllykallio H, Lopez P, López-García P, Heilig R, Saurin W, Zivanovic Y, Philippe H, Forterre P., 2000: **Bacterial mode of replication with eukaryotic-like machinery in a hyperthermophilic archaeon.** Science.;288 (5474):2212-5.

Ng WL, DasSarma S., 1993: **Minimal replication origin of the 200-kilobase Halobacterium plasmid pNRC100.** J. Bacteriol. 175:4584-4596.

Oren A, Bratbak G, Heldal M, 1997: **Occurrence of virus-like particles in the Dead Sea.** Extre-mophiles 1:143-149

Oren A., 2010: **Industrial and environmental applications of halophilic microorganisms.** Environ Technol. 31(8-9):825-34.

Oren Aharon, 2012: **Taxonomy of the family Halobacteriaceae: a paradigm for changing concepts in prokaryote systematic.** Int. J. of Systematic and Evolutionary Microbiology 62, 263–271

Paola Londei, 2010: **Archaeal Ribosomes.** Based in part on the previous version of this Encyclopedia of Life Sciences (ELS) article Archaeal Ribosomes, by George Harauz and Abdiwahab Musse

Ring Gabriela and Eichler Jerry, 2004: **Extreme Secretion: Protein Translocation Across the Archaeal Plasma Membrane.** Journal of Bioenergetics and Biomembranes, Vol. 36, No. 1,

Robinson C., Matos C. F., Beck D., Ren C., Lawrence J., Vasisht N., Mendel S., 2011: **Transport and proofreading of proteins by the twin-arginine translocation (Tat) system in bacteria.** Biochim. Biophys. Acta 1808, 876–884.

Rose R. W., Bruser T., Kissinger J. C., Pohlschroder M., 2002: **Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway.** Mol. Microbiol. 45, 943–950.

Rössler N, Klein R, Scholz H, Witte A., 2004: **Inversion within the haloalkaliphilic virus ϕ Ch1 DNA results in differential expression of structural proteins.** Molecular Microbiology. 52(2):413-26.

Ruiz DM, Paggi RA, Giménez MI, De Castro RE., 2012: **Autocatalytic maturation of the Tat-dependent halophilic subtilase Nep produced by the archaeon Natrionalba magadii.** J Bacteriol. 194(14):3700-7.

Sandmeler Heinrich, 1994: **Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail fibres.** Mol Microbiol. 12(3):343-50.

Schnabel H, Zillig W, Pfäffle M, Schnabel R, Michel H, Delius H, 1982: **Halobacterium halobium phage Φ H**. EMBO J. 1:87-92

Shigeru Iida, 1984: **Bacteriophage P1 carries two related sets of genes determining its host range in the invertible c segment of its genome**. Virology, Volume 134, Issue 2, Pages 421–434

Shirai T, Kobayashi T, Ito S, Horikoshi K., 2008: **Alkaline Adaptation of Proteins**. In: Siddiqui KS, Thomas T, editors. Protein Adaptation in Extremophiles (Molecular Anatomy and Physiology of Proteins). Nova Publishers. p 105-141.

Siddaramappa S, Challacombe JF, Decastro RE, Pfeiffer F, Sastre DE, Giménez MI, Paggi RA, Detter JC, Davenport KW, Goodwin LA, Kyrpides N, Tapia R, Pitluck S, Lucas S, Woyke T, Maupin-Furlow JA., 2012: **A comparative genomics perspective on the genetic content of the alkaliphilic haloarchaeon *Natrialba magadii* ATCC 43099T**. BMC Genomics. 4;13:165.

Siegel P J, and Schaechter M., 1973: **The Role of the Host Cell Membrane in the Replication and Morphogenesis of Bacteriophages**. Annual Review of Microbiology Vol. 27: 261-282

Siglioccolo Alessandro, Paiardini Alessandro, Piscitelli Maria and Pascarella Stefano, 2011: **Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface**. BMC Structural Biology 11:50

Soyer-Gobillar Marie-Odile, 2006: **Edouard Chatton (1883–1947) and the dinoflagellate protists: concepts and models**, International microbiology 9:173-177

Stolt P, Zillig W., 1992: **In vivo studies on the effects of immunity genes on early lytic transcription in the *Halobacterium salinarum* phage Φ H**. Mol. Gen. Genet. 235(2-3):197-204.

Szabo Zalan and Pohlschroder Mechthild, 2012: **Diversity and Subcellular Distribution of Archaeal Secreted Proteins**. Front Microbiol. 3, 207.

Till P., 2011: **Deletion of the tail fibre protein of Φ Ch1 and further characterization of the inversion within its gene locus**. Diploma thesis, University of Vienna

Tindall B.J., Ross H.N.M., Grant W.D., 1984: **Natronobacterium gen. nov. and Natronococcus gen. nov., Two New Genera of Haloalkaliphilic Archaeobacteria**. Systematic and Applied Microbiology Volume 5, Issue 1, Pages 41–57

van de Vossenberg JL, Driessen AJ, Grant WD, Konings WN, 1999: **Lipid membranes from halophilic and alkali-halophilic archaea have low H⁺ and Na⁺ permeability at high salt concentration**. Extremophiles, 3:253_257

Wendoloski, D., C. Ferrer, and M. L. Dyal-Smith., 2001: **A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2**. Microbiology 147:959-964.

Witte A, Baranyi U, Klein R, Sulzner M, Luo C, Wanner G, Krüger DH, Lubitz W., 1997: **Characterization of *Natronobacterium magadii* phage ϕ Ch1, a unique archaeal phage containing DNA and RNA.** *Molecular Microbiology*. 23(3):603-16.

Woese CR, Fox GE., 1977: **Phylogenetic structure of the prokaryotic domain: The primary kingdoms.** *Proc. Natl. Acad. Sci. USA*. 74(11):5088-90.

Woese CR, Kandler O, Wheelis ML., 1990: **Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya.** *Proc. Natl. Acad. Sci. USA*. 87(12):4576-9.

Yanhe Ma, Erwin A. Galinski, William D. Grant, Aharon Oren, and Antonio Ventosa, 2010: **Halophiles 2010: Life in Saline Environments.** *Appl. Environ. Microbiol. November 2010 vol. 76 no. 21 6971-6981*

Yosuke Koga, 2012: **Thermal Adaptation of the Archaeal and Bacterial Lipid Membranes.** *Archaea* Volume 12, Article ID 789652

5. Appendix

I. Index of figures and tables

Figure 1 The three domains of life.	11
Figure 2 Morphology of <i>Natrialba magadii</i>	28
Figure 3 Section of nucleotide and deduced amino acid sequence of <i>nep</i>	30
Figure 4 Morphology of ϕ Ch1 particles.	34
Figure 5 Linear presentation of the 58.5 kb ϕ Ch1 genome.	36
Figure 6 The invertible region of ϕ Ch1.	38
Figure 7 Domain structure of gp34 ₅₂	79
Figure 8 Wild type sequence and construction models of Φ Ch1 deletion mutants of ORF34. ...	80
Figure 9 Analytical PCRs testing for gene replacement of ORF 34 in <i>Nab. magadii</i> L11.	81
Figure 10 Schematically drawing of the recombination event with the plasmid pKSII ⁺ Δ ORF34::Nov ^R and the wild type genome of <i>Nab. magadii</i> L11.	82
Figure 11 Southern blot analysis to verify a homozygous deletion mutant of <i>Nab. magadii</i> L11 Δ ORF34.	83
Figure 12 Expression of ORF34 in <i>Nab. magadii</i> wild type and L11 Δ 34 strains.	84
Figure 13 Analytical PCR analysis for detection of phage particles of Φ Ch1.	85
Figure 14 Western blot analysis of Capsid protein gpE expression.	86
Figure 15 Pre-testing of expression of ORF34 in <i>Nab. magadii</i> L13.	87
Figure 16 Schematic drawing of fusion proteins for secretion.	91
Figure 17 Growth curve of <i>Nab. magadii</i> P3 transformants and verification of secretion of recombinant proteins	92
Figure 18 Results of infectivity verification of wild type virus ϕ Ch1 on <i>Nab. magadii</i> strains secreting gp34.....	93

II. Acknowledgement

First of all, I want to thank my supervisor Prof. Dr. Angela Witte for giving me the opportunity to graduate in her lab and for supporting me to write this thesis. Her patience, enthusiasm, motivation and wealth of knowledge as well as breadth of ideas opened a time where I have learnt much about the exciting topic *Archaea* and more generally, how microbiology works routinely. Furthermore, I enjoyed the pleasant working environment which facilitates enjoyable conversations not only about work-flow, but also privately.

Next, my thanks go to Judith Beraha and Lèa Schöner for being my cooperative and helpful colleagues but mostly for the new friendship which I don't want to miss anymore. In addition my thank goes to Petra Till, Beatrix Alte and Tatjana Svoboda for introducing and supporting me during the beginning of my work and for the further working time.

I want to express my gratitude especially to my mother Eugenia, because of her believe in me, her incredible encouragement and that she is always there for me.

Finally I thank my partner Michael and all my friends, specific Marion, Nicole, Andi and Rene.

III. Abstract

The virus ϕ Ch1, which infects the haloalkaliphilic archaeon *Natrialba magadii* contains an invertible genomic region where a recombination event leads to an exchange of the 3' ends of the convergent open reading frames (ORFs) 34 and 36. During the lysogenic life cycle, this inversion reaction yields in various gene products coding for putative tail fibre proteins. Increased emphasis has been put on two selected products of ORF34, the non-inverted variant gp34₁ and the variant comprising the C-terminus of gp36 gp34₅₂. Previous studies identified that just gp34₅₂ is able to bind to the cell surface of *Nab.magadii*, the only known host of ϕ Ch1, yet (Klein *et al.*, 2012). The aim of this work was another confirmation of the protein gp34₅₂ as a putative tail fibre protein. Therefore an ORF34 deletion mutant was created by homologous recombination, using a selection cassette (Nov^R) for gene replacement. Per Western blot and Southern blot the absence of the respective gene product and homozygation of the strain L11 Δ ORF34 respectively, were demonstrated. However, the elimination of ORF34, which is supposed to represent the sequence of encoded putative tail fibre proteins, yielded no mature phage particles. Furthermore, the presence of the major capsid protein E, which is thought to be essential for release of progeny precursors from the host membrane, could not be detected by Western blot analysis. This result may explain why the deletion mutant *Nab. magadii* L11 Δ ORF34 is not able to enter the lytic life cycle. By introducing a sequence expressing ORF34₅₂ it was tested if the deletion mutant is able to recover wild type behavior. The complementation application was performed with investigation of potential plaque formation, which would be due to regained infectivity of the mutant virus. However, complementation of the mutant strain resulted in no plaque formation. It is presumably that the introduced deletion mutation affects transcription. In order to achieve a proper transcription, an archaeal termination sequence integrated in the deletion construction sequence might be an approach.

In the second part of this work, an experiment was performed in order to prevent infection of the virus Φ Ch1 with the purpose of occupying the receptors on the *Nab. magadii* cell surface, by exporting the protein variants of gp34 (gp34₁ and gp34₅₂). Therefore, Nep (*Natrialba* extracellular protease) was used as facility for secretion of the putative tail fibre protein by fusion of the respective proteins with the N-terminus of the protease. Those in frame fusion constructions were transformed by a shuttle vector into *Nab. magadii* P3, a protease deficient strain. The autocatalytically features of Nep enables to export the putative tail fibre proteins without

residues from the N-terminus of the protease because of the cleavage, before export occurs. The confirmation of a successful secretion was performed by western blot analysis which was followed with an infectivity analyses with the wild type ϕ Ch1. The results showed that the strain *Nab. magadii* P3 (pNB102 pNPro-34₅₂) exporting the putative tail fibre protein gp34₅₂ yielded a reduction of infectivity of at least 8 orders of magnitude compared to the control strain *Nab. magadii* P3 (pNB102) and 5 orders of magnitude compared to *Nab. magadii* P3 (pNB102 pNPro-34₁), indicating that infectivity of ϕ Ch1 can be hindered by occupying the receptor with gp34₅₂. These results gave another hint, that gp34 acts as a tail fibre protein.

IV. Zusammenfassung

Der Virus ϕ Ch1, welcher das haloalkaliphile Archaeon *Natrialba magadii* infiziert, besitzt eine spezialisierte Gensequenz mit einem invertierbaren Bereich. Dieser umfasst die mutmaßliche Invertase *int1*, flankiert von den beiden offenen Leserahmen 34 und 36 welche nach einer Inversionsreaktion zu unterschiedlichen Genprodukten führt, die als putative „tail fibre“ Proteine dienen sollen und durch die entstehende Vielfalt ein erweitertes Wirtsspektrum sicherstellen sollen. Dieses Ereignis führt unter anderem zu einem Austausch der 3'Enden von ORF34 und ORF36 wobei das Genprodukt 34₅₂ entsteht und frühere Studien konnten die Interaktion mit der Zelloberfläche von *Nab. magadii* nachweisen (Klein *et al.*, 2010). In diesem Zusammenhang wurde im Zuge dieser Arbeit eine Deletionsmutante dieser fraglichen Region ORF34 hergestellt, wobei die Abwesenheit des vermeintlichen „Infektionswerkzeuges“ des Virus ϕ Ch1 einen Verlust der Infektion zur Folge haben sollte. Dieser *Nab. magadii* L11- Δ ORF34 Stamm wurde mittels homologer Rekombination mit einer Novobiocin Kasette hergestellt. Die Abwesenheit dieses Produktes wurde anhand Western Blot nachgewiesen und die Bestätigung eines homozygoten Stammes wurde mit einem Southern Blot durchgeführt. Nachdem festgestellt wurde, dass keine intakten Virus Partikel freigesetzt werden konnten, wurde auf das Vorhandensein des Hauptcapsidprotein E getestet. Dieses Protein ist unter anderem für das Entlassen von reifen Viren verantwortlich. Es konnte keine Expression nachgewiesen werden und lässt daraus schließen, dass die eingefügte Mutation den Leserahmen der Transkription beeinflusst und somit keine gänzlich funktionierende Virus Partikel produziert. Im Anschluss wurde versucht die ursprüngliche Funktion dieser „Knock out“ Mutante L11- Δ ORF34 durch das Einbringen einer funktionellen Sequenz, welche für das Genprodukt 34₅₂ kodiert, wiederherzustellen. Für diese Komplementationsanalyse wurden die Transformanten unter Durchführung eines Phagen Titer untersucht wobei resultierende Plaques auf wieder intakte wild typ Viren schließen lässt. Es konnte keine Plaque Formation nachgewiesen werden und somit keine funktionstüchtigen Viren. Um diesen Einfluss der Deletion auf nachfolgende Genexpression unter Kontrolle zu halten, soll ein zusätzliches Stopp-codon, spezifisch für *Archaea*, eingefügt werden und weiter Erfolge versprechen.

Der zweite Teil dieser Arbeit umfasst ein weiteres Experiment welches die Funktion des Genproduktes 34₅₂ als „Infektionswerkzeug“ nachweisen soll. Dafür wurde ein weiteres Produkt dieser bereits erwähnten Inversionsreaktion ausgewählt (Genprodukt 34₁) von dem bekannt ist, dass es nicht an die Zelloberfläche von *Nab. magadii* bindet. Das Ziel war es, durch Sekretion

beider Genprodukte 34₁ und 34₅₂ aus *Nab. magadii* die Infektion des wildtyp Virus ϕ Ch1 zu verhindern oder einzuschränken. Dies wurde mithilfe der Natrialba Extrazellulären Protease (Nep) durchgeführt. Es ist bekannt, dass dieses proteolytische Enzym zu Beginn der Stationären Phase autokatalytisch expremiert und prozessiert wird, wobei der N-Terminus während des Exportes abgespalten wird um die Protease in den Extrazellulären Raum zu entlassen (De Castro *et al.*, 2008). Diese Eigenschaften wurden als Hilfsmittel zum Export der vermeintlichen „tail fibre“ Proteinen genutzt und Konstrukte hergestellt, welche den N-Terminus der Protease fusioniert mit den jeweiligen Varianten der Inversionsprodukte von ORF34/36 tragen. Diese Konstrukte wurden in den Stamm *Nab. magadii* P3 transformiert welcher defizient für diese Protease ist und mittels Western Blot konnte die erfolgreiche Sekretion dieser Proteine nachgewiesen werden. Die anschließende Durchführung eines Phagen Titers, wobei die exportierenden Stämme *Nab. magadii* P3 (pNB102 pNPro-34₁), *Nab. magadii* P3 (pNB102 pNPro-34₅₂) und *Nab. magadii* P3 (pNB102, leerer Vektor) mit dem wild typ Virus infiziert und mit der Infektionsrate des Kontrollstammes *Nab. magadii* P3 (pNB102) verglichen wurden. Die Ergebnisse zeigten eine Reduktion der Infektionsrate des Stammes *Nab. magadii* P3 (pNB102 pNPro-34₅₂). Dieser Stamm exportiert das vermutete „tail fibre“ Protein welches an die Zelloberfläche binden kann und lässt die Infektionsrate im Vergleich zum Kontrollstamm *Nab. magadii* P3 (pNB102) mindestens um ein Achtfaches sinken und etwa eine Fünffache Abnahme der Infektion konnte bei einer Gegenüberstellung mit dem Stamm *Nab. magadii* P3 (pNB102 pNPro-34₁) festgestellt werden. Diese Ergebnisse unterstreichen deutlich die bereits angestellte Vermutung, dass das Genprodukt 34₅₂ eine essentielle Rolle im Infektionsprozess des Virus ϕ Ch1 mit *Natrialba magdii* spielt.

V. Curriculum Vitae

Personal details:

Name: Katharina Dimmel
Date and place of birth: July, 22th, 1983 in Vienna, Austria
Nationality: Austria

Education:

1989 – 1993 Volksschule – Leitzersdorf / Bruderndorf, Austria
1993 -1997 Hauptschule – Ost - Stockerau , Austria
1997 – 2003 BAKIP Maria Regina - Vienna, Austria
2004 – 2005 Medical technologists - AKH Vienna, Austria
2006 –2013 Study Biology “Microbiology and Genetics” (Diploma course), University of Vienna, Austria
2011 –2012 Diploma thesis: Laboratory of Prof. Dr. Angela Witte; Department of Microbiology, Immunobiology and Genetics; University of Vienna, Austria

Working experience:

2007 – 2010 *KSG Medical Consultig GmbH, Austria: Temporary employment*
2012/2013 *University of Vienna, Austria: Tutor for student practical trainings: “Übung in Molekularer Mikrobiologie” and “Übungen III A – Molekularbiologische Laborarbeiten”*
2011 – 2013 *FH Campus Wien (University of Applied Sciences, Vienna, Austria): Tutor of students for practical trainings of “Molekularbiologische Übungen I and II” in the course of the studies in “BioEngineering” and “Biotechnologie”*