universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Gene regulation of ϕCh1: characterization of ORF49

and

further characterization of the origin of replication of the halophage ϕCh1

angestrebter akademischer Grad

Magister der Naturwissenschaften (Mag. rer. nat.)

Verfasser:	Michael Reiter
Matrikel-Nummer:	0004564
Studienrichtung /Studienzweig (lt. Studienblatt):	A 441 Diplomstudium Genetik – Mikrobiologie (Stzw) UniStG
Betreuerin:	Ao. Univ.-Prof. Dipl.-Biol. Dr. Angela Witte

Wien, im Mai 2010

Table of Contents

1. INTRODUCTION 11

1.1. *Archaea* 11

1.1.1. The third domain of life 11

1.1.2. Archaeal diversity 12

1.1.3. Molecular characteristics of the *Archaea* 14

1.1.3.1. Structural features 14

1.1.3.2. Genomic features 15

1.1.3.2.1. DNA replication 15

1.1.3.2.2. Transcription 16

1.1.3.2.3. Translation 17

1.1.4. Haloalkaliphilic *Archaea* 18

1.1.4.1. Diversity of halophilic *Archaea* 18

1.1.4.2. Halophily – living under high salt conditions 18

1.1.4.2.1. Adaptations to high salt concentrations 19

1.1.4.2.1.1. Properties of halophilic proteins 20

1.1.4.3. (Halo-) alkaliphilic *Archaea* 20

1.1.4.3.1. Adaptations to high pH 21

1.1.4.3.1.1. Properties of alkaliphilic proteins 22

1.1.5. Benefits of doing research with (haloalkaliphilic) *Archaea* 24

1.1.6. *Natrialba magadii* 25

1.1.6.1. *Nab. magadii* in the laboratory 26

1.1.6.1.1. Transformation of *Nab. magadii* 27

1.1.6.1.2. Genetic markers in *Nab. magadii* 27

1.2. Viruses of the *Archaea* 28

1.2.1. Diversity of archaeal viruses 28

1.2.1.1. Haloviruses 29

1.2.2. The Halovirus ϕCh1 – a general view 31

1.2.2.1. Morphology of ϕCh1 31

1.2.2.2. Protein composition of ϕCh1 particles 32

1.2.2.3. ϕCh1 is a temperate phage 32

1.2.2.4. Genetic organization of ϕCh1 33

1.2.2.5. The ϕCh1 replication region and development of a shuttle vector 36

1.2.2.6. Gene regulation in ϕCh1 37

2. MATERIALS & METHODS 41

2.1. Materials 41

2.1.1. Strains 41

2.1.1.1. Bacterial strains 41

2.1.1.2. Archaeal strains 42

2.1.2. Growth media 42

2.1.3. Antibiotics 43

2.1.3.1. for *E. coli* 43

2.1.3.2. for *Nab. magadii* 43

2.1.4. Vectors 43

2.1.5. Primer 46

2.1.6. DNA and protein markers 47

2.1.6.1. DNA ladders 47

2.1.6.2. Protein ladders 48

2.1.7. Enzymes and corresponding buffers 48

2.1.7.1. Restriction 48

2.1.7.2. PCR 48
2.1.7.3. Other enzymes 49
2.1.8. Antibodies 49
2.1.9. Buffers and solutions 49
2.1.9.1. Generation of $CaCl_2$ competent *E. coli* cells 49
2.1.9.2. Transformation of *Nab. magadii* 50
2.1.9.3. Isolation of ϕCh1 virus particles 50
2.1.9.4. DNA methods 51
2.1.9.4.1. Electrophoresis 51
2.1.9.4.2. Gel extraction from polyacrylamide gels 51
2.1.9.5. Protein methods 52
2.1.9.5.1. SDS-PAGE and Western blot 52
2.1.9.5.2. Protein purification under native conditions 53
2.1.9.5.3. Protein purification under denaturing conditions 53
2.1.9.5.4. Inclusion body purification from *E. coli* 54
2.1.9.5.5. Protein renaturation 54
2.1.10. Bandshift assays 55

2.2. Methods 55

2.2.1. Preparation and transformation of $CaCl_2$ competent *E. coli* cells 55
2.2.2. Preparation and transformation of competent *Nab. magadii* cells 56
2.2.3. Phage methods 57
2.2.3.1. Isolation of ϕCh1 virus particles 57
2.2.3.2. Determination of phage titers by soft plating technique 58
2.2.3.3. Isolation of ϕCh1 virus DNA 58
2.2.4. DNA Methods 59
2.2.4.1. Agarose gel electrophoresis 59
2.2.4.2. Polyacrylamide gel electrophoresis 59

2.2.4.3. Polymerase chain reaction (PCR) 60
2.2.4.3.1. Primers 60
2.2.4.3.2. Template DNA 60
2.2.4.3.3. PCR batches and -programs 61
2.2.4.3.4. Confirmation and purification of PCR products 62
2.2.4.4. Plasmid isolation from *E. coli* 62
2.2.4.5. Gel extraction of DNA fragments from agarose gels 62
2.2.4.6. Gel extraction of DNA fragments from polyacrylamide gels 63
2.2.4.7. Cloning in *E. coli* strains 64
2.2.4.7.1. Restriction of DNA 64
2.2.4.7.2. Purification of DNA 64
2.2.4.7.4. Blunting DNA with the *Klenow* fragment 64
2.2.4.7.5. Ligation of respective DNA fragments 64
2.2.4.7.6. Pre-screening of candidates by quick plasmid preparations 65
2.2.4.7.7. Verification of positive clones 65
2.2.4.7.8. Storage of verified clones 66
2.2.4.7.9. Cloning strategies 66
2.2.5. Protein methods 69
2.2.5.1. Overexpression of recombinant protein and screening of cultures for optimal time point for cell harvesting 69
2.2.5.2. Protein purification under native conditions 69
2.2.5.2.1. Purification of His-tagged proteins 69
2.2.5.2.2. Purification of Maltose-tagged proteins 70
2.2.5.3. Purification of His-tagged proteins under denaturing conditions 71
2.2.5.4. Purification of inclusion bodies from *E. coli* 71
2.2.5.5. Protein renaturation 72
2.2.5.5.2. by dialysis 72
2.2.5.5.3. by solubilization of inclusion bodies and rapid dilution 72

2.2.5.5.4. by size exclusion chromatography 72
2.2.5.5.5. Gel extraction from polyacrylamide gels 72
2.2.5.6. SDS-PAGE 73
2.2.5.6.1. Preparation of a discontinuous polyacrylamide gel 73
2.2.5.6.2. Sample preparation and running conditions 74
2.2.5.6.3. Staining of separated proteins 74
2.2.5.7. Western blot 74
2.2.5.7.1. The blotting procedure 75
2.2.5.7.2. Detection of the protein of interest 75
2.2.6. Bandshift assays 75
2.2.7. Structure prediction of gp49 76

3. RESULTS & DISCUSSION 77

3.1. Further analysis of the ϕCh1 replication region 77

3.1.1. Introduction of +1 frameshift mutations into ORF53 and ORF54 77
3.1.2. Deletion of up- and downstream palindromic sequences 78
3.1.3. Transformation of pRo-5 derivates into *Nab. magadii* 79
3.1.4. Discussion 79

3.2. Characterization of ORF49 82

3.2.1. Gp49 acts as ϕCh1 repressor 82
3.2.2. Putative 3D structure of gp49 – an attempt in silico 84
3.2.3. Overexpression of ORF49 in *E. coli* sp 87
3.2.4. Purification of gp49 88
3.2.5. Bandshifit assays (EMSA) 90
3.2.5.1. Bandshifting of total ϕCh1 DNA and fragments 90
3.2.5.2. Bandshifting selected ϕCh1 ORFs and DNA fragments 93
3.2.5.3. Comparison of gp49 wt and 3´ deletion mutants 95

3.2.5.4. Comparison gp49 wt and N-terminal deletion 97

3.2.6. Construction and analysis of 5 different deletion mutants of ORF49 98

3.2.7. Discussion 99

3.2.8. Closing words 101

4. REFERENCES 103

5. APPENDIX 113

Index of Figures 113

Acknowledgements 115

Abstract 117

Zusammenfassung 119

Curriculum Vitae 121

1. Introduction

1.1. *Archaea*

1.1.1. The third domain of life

In the early 1990s the common view upon phylogeny and taxonomy dramatically changed. Up to that point in time two major models for the evolutional relationships between all organisms were conventionally accepted. On the one hand the five kingdom model, which had been developed from the mid 1860s, introduced by Ernst Haeckel (Haeckel, 1866), throughout the twentieth century during which this theory was rounded up by Copeland (Copeland, 1938) and Whittaker (Whittaker, 1959). In its latest form, which also became known as the "Whittaker Scheme", it categorizes all life forms either as Animalia, Plantae, Fungi, Protista or Monera (Whittaker & Margulis, 1978). On the other hand there stood a far more radical theory introduced by Eduard Chatton. He proposed only two major categories of life based on the absence or presence of a cellular nucleus – the prokaryotes and the eukaryotes (Chatton, 1938). This dichotomy later on became more and more supported by the accumulation of evidence gathered by the utilization of the ever higher developing methods of biochemistry and molecular microbiology. *"However, the eukaryote prokaryote concept itself has been seriously misunderstood and, consequently, wrongly interpreted."* (Woese *et al.*, 1990). A major mistake was buried in the assumption, that all the prokaryotes constitute a monophyletic group (Woese *et al.*, 1990). By analysis of ribosomal RNA sequences of various "bacterial strains" it was found that amongst this domain a small group of bacteria displayed major differences in comparison with other prokaryotes. This group by that time comprising only a small number of methanogenic bacteria was termed *Archaebacteria* (Woese & Fox, 1977). Further comparison of the small subunit rRNAs of eukaryotes, prokaryotes and archaebacteria eventually led to the rearrangement of the universal phylogenetic tree and to the introduction of a whole new three domain theory dividing all

organisms into *Bacteria*, *Archaea* and *Eukarya* (Woese *et al.*, 1990). Amazingly the newly introduced domain of *Archaea* seemed to share with the *Eukarya* a common ancestor while the *Eubacteria* seemed to branch earlier in time thus making *Eukarya* and *Archaea*, though distant, specific relatives (Woese *et al.*, 1990).

1.1.2. Archaeal diversity

Originally it was believed that the domain of *Archaea* comprises a rather exclusive group of microorganisms, which mainly inhabit extreme environments but it soon became clear that *Archaea* are in fact widespread (Olsen, 1994). For example a recent study of archaeal and bacterial abundance in the Pacific Ocean suggests that the global oceans harbor approximately 1.3×10^{28} archaeal cells and 3.1×10^{28} bacterial cells and that the phylum *Crenarchaeota* alone accounts for 20 % of all the picoplankton cells in the oceans (Karner *et al.*, 2001). Thus the so called "third domain of life" contributes a significant, not to say crucial amount of earth´s biomass.

Based upon the 16S rRNA analysis the domain *Archaea* currently consists of two phyla (according to Bergey´s manual (Garrity *et al.* (eds), 2001)) that are commonly accepted – the *Euryarchaeota* and the *Crenarchaeota* (Woese *et al.*, 1990). While the *Euryarchaeota "…encompass the greatest phenotypic diversity among known cultivable species, with the halophiles, the methanogens, some thermoacidophiles and some hyperthermophiles"* (Forterre *et al.*, 2002) the *Crenarchaeota* so far comprise only hyperthermophilic species (Forterre *et al.*, 2002). However additional archaeal phyla like the *Korarchaeota* (Barns *et al.*, 1996) and the *Nanoarchaeota* (Huber *et al.*, 2002; Huber *et al.*, 2003) have also been suggested indicating that the current classification of the domain of *Archaea* is certainly not the end of the tunnel. Eventually this classification could end up looking much more similar to that of the *Bacteria* in terms of the diversity of phyla (Forterre *et al.*, 2002).

Although a common view of archaeal abundance is that they are in fact predominant in extreme habitats (e.g. high or low pH, high salt, high pressure) this assumption cannot be accounted fully true since *Bacteria* (and even some *Eukarya*) have also been found residing side by side with members of *Archaea* even under extreme conditions – so far only to one exception – hyperthermic environments above 95° C (Rothschild & Mancinelli, 2001; Forterre *et al.*, 2002). Another

widespread opinion is that *Archaea* are mostly extremophiles which is due to the fact that the first characterized *Archaea* were found to live under extreme conditions (Valentine, 2007). However, as stated above, the findings of Karner *et al.* and others changed the view of *Archaea* being restricted only to extreme habitats. Yet, the altering view does not diminish the interest in this domain, as in many extreme habitats *Archaea* play the dominant role and are still the record holders for growth at the highest temperature, lowest pH, and highest NaCl concentration (Chaban *et al.*, 2006). Thus the importance of archaeal life in these environments remains unquestioned.

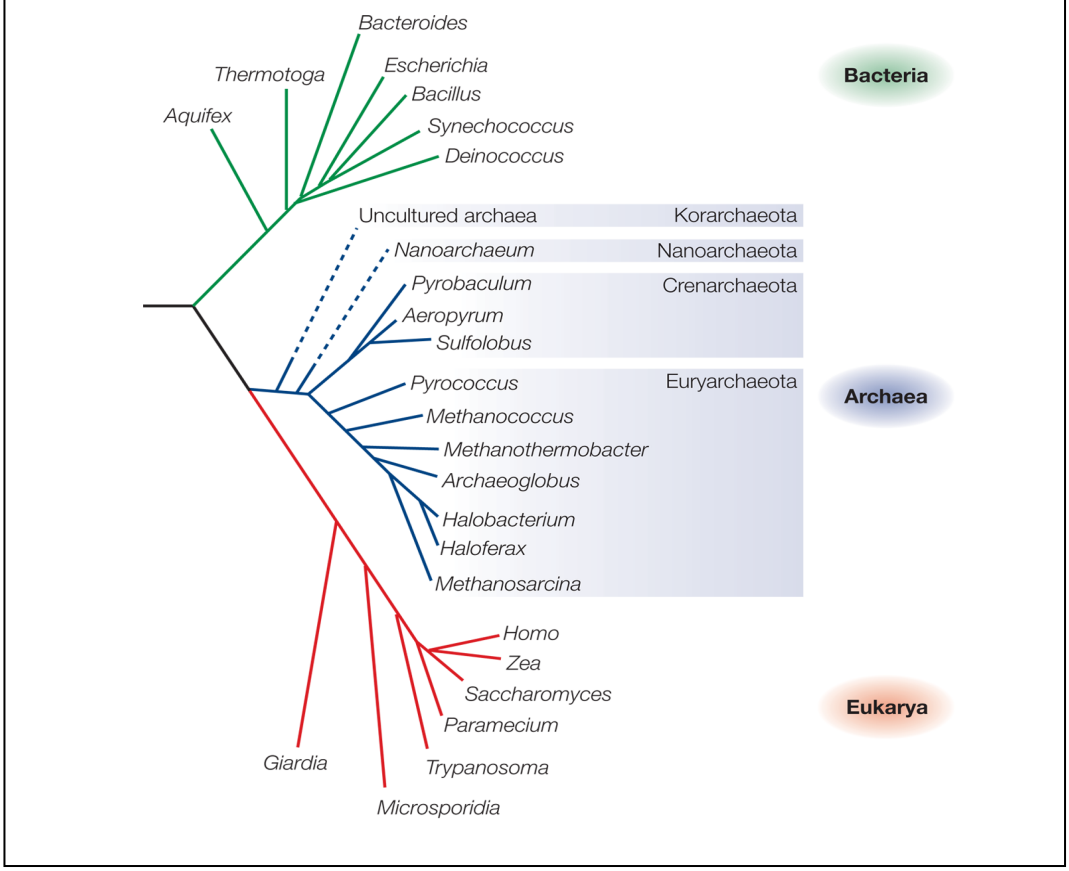

Figure 1| **The three domains of life.** One of many visualizations of the phylogenetic tree of life based on 16S rRNA sequence comparison showing the *Archaea* and the *Eukarya* sharing a common ancestor. Among the *Archaea* the *Euryarchaeota* and the *Crenarchaeota* are commonly accepted according to Bergey´s manual (Garrity *et al.* (eds), 2001) with the *Euryarchaeota* exhibiting the greatest diversity. Currently all cultivable *Archaea* belong to these two phyla. However, other phyla have been suggested. The phylum *Korarchaeota* has been proposed only on the basis of 16S rRNA analysis (Barns *et al.*, 1996) whereas the *Nanoarchaeota* so far comprise only one genus, *Nanoarchaeum equitans*, which exhibits extraordinary small cell sizes of approx 0.5 µm (Huber *et al.*, 2003).

(Allers and Mevarech, 2005)

As far as metabolic diversity is concerned *Archaea* and *Bacteria* seem to match each other – with one exception: methanogenic organisms have so far only been described within the domain of *Archaea* (Forterre *et al.*, 2002). All other metabolic pathways comprising heterotrophy, autotrophy and photosynthesis have been found in both domains *Archaea* and *Bacteria* (Forterre *et al.*, 2002).

1.1.3. Molecular characteristics of the *Archaea*

Archaea are not only distinguished from *Bacteria* and *Eukarya* by their 16S rRNA sequence. They also exhibit certain unique features which cannot be found in the two other domains. On the other hand some of their features are more closely related to the *Eukarya* whereas others seem to be reminiscent of *Bacteria.*

1.1.3.1. Structural features

In terms of cell sizes both *Archaea* and *Bacteria* are relatively similar. However, the membrane lipids of *Archaea* provide a substantial criterion for their distinction from the other two domains of life. Their glycerolipids are ethers of glycerol and isoprenol, while in *Bacteria* and *Eukarya* lipids exhibit esters of glycerol and fatty acids (Forterre *et al.*, 2002). These glycerol ethers contain 2,3-sn-glycerol instead of 1,2-sn-glycerols found in *Bacteria* and *Eukarya* (Brown & Doolittle, 1997). In addition to that *"…archaea have highly methylbranched isopranyl chains, while hydrocarbons in bacteria and eukaryotes are predominantly straight-chain fatty acyl chains."* (Brown & Doolittle, 1997). Furthermore the domain of *Archaea* is characterized by the total lack of peptidoglycan, which is compensated in some species by the presence of pseudopeptidoglycan (e.g. members of the *Methanobacteriales*) or heteropolysaccharides (e.g. *Halococcus*) (Brown & Doolittle, 1997). The pseudopeptidoglycan shows fundamental differences to the bacterial peptidoglycan. It contains L-talosaminuronic acid instead of muramic acid and D-amino acids are not found in the peptide moiety (Kandler & König, 1998). Besides that many members of the *Archaea* possess only proteinaceous or glycoproteinaceous cell envelopes (S-layers) or only have a cytoplasmic membrane containing glycoproteins (e.g. *Thermoplasmatales*) (Kandler & König, 1998; Forterre *et al.*, 2002).

The vast differences between bacterial and archaeal cell wall structure and synthesis are also reflected by their different susceptibility to certain antibiotics. So it seems only natural that classic antibiotics that are directed against bacterial cell wall synthesis like the ß-lactams (e.g. penicillin) have virtually no effect on archaeal growth (Kandler & König, 1998). Thus the number of antibiotics effective against members of the *Archaea* is limited with certain antibiotics being more restricted to smaller groups of them.

1.1.3.2. Genomic features

At the first glance the genomic organization of *Archaea* resembles that of *Bacteria*. The DNA is organized into a, sometimes large, circular chromosome accompanied by one or more smaller circular DNA plasmids (Brown & Doolittle, 1997). Yet the packaging of the DNA in the archaeal cell is more reminiscent of *Eukarya* in a variety of ways as they possess histones. So far, however, these DNA binding proteins have only been found in the *Euryarchaeota*. In *Archaea* the histones are shorter than in *Eukarya* and lack the N- and C-terminal tail extensions (White & Bell, 2002). Another difference to eukaryotic histones is that in *Archaea* they can not only form heterodimers (like in eukaryotes) but also homodimers. In addition to that, *Archaea* also have DNA compacting proteins, similar to bacterial HU (White & Bell, 2002; Brown & Doolittle, 1997).

Archaeal genes are often organized into operons like in *Bacteria* (Brown & Doolittle, 1997). Considering these similarities it was initially believed that like in *Bacteria* archaeal chromosomes also contain only one origin of replication. However, recently *Sulfolobus Solfataricus* has been shown to possess three origins of replication (Lundgren *et al.*, 2004).

1.1.3.2.1. DNA replication

In general DNA replication in *Archaea* seems to be more related to that of *Eukarya* than to that of *Bacteria*. In *Bacteria* the protein DnaA binds to the origin of replication and initiates melting of the complementary DNA strands, in order to initiate replication. In *Archaea* homologues to the eukaryotic initiator proteins Orc and Cdc6 have been identified in almost all of the currently sequenced archaeal genomes (Barry & Bell, 2006). Helicases, proteins responsible for unwinding of the DNA, which is done in *Bacteria* by a protein called DnaB, are found to be

homologous in *Archaea* and *Eukarya*. Furthermore while Primer synthesis in *Bacteria* is achieved by a protein called DnaG, in *Eukarya* and *Archaea* the same is done by family B polymerases (Edgell & Doolittle, 1997). Also archaeal and eukaryal DNA polymerases are homologues and not related to any bacterial DNA polymerase except that of *E.coli* (Brown & Doolittle, 1997). However, in addition to family B DNA polymerases which are responsible for strand elongation in *Archaea* and *Eukarya* a novel family of DNA polymerases, subsequently named family D DNA polymerases, has been found in euryarchaeal genomes. These new polymerases so far seem to be unique to the *Euryarchaea*, though (Barry & Bell, 2006).

1.1.3.2.2. Transcription

The transcription of archaeal genes constitutes a mosaic of eukaryal and bacterial features. Although *Archaea*, like *Bacteria* seem to have only one RNA polymerase the subunit complexity of this enzyme is similar to that of *Eukarya* (Bell & Jackson, 1998). Latter feature three different RNA polymerases which transcribe different sets of genes. However, all three of these show significant structural similarities to archaeal RNA polymerases (Brown & Doolittle, 1997). Furthermore phylogenetic studies show that the large subunits of the archaeal enzyme are closely related to the eukaryal RNA polymerase II. Archaeal promoter elements are also reminiscent of *Eukarya* as they contain TATA box like binding sites, usually situated about 30 bp upstream of the transcription initiation (Bell & Jackson, 1998). The recognition of the TATA like element occurs by a homologue to the eukaryal TATA box binding protein (TBP). Other eukaryotic transcription factors, like TFB or TFIIS are also found to have homologues in *Archaea* (Bell & Jackson, 1998). Interestingly archaeal RNA polymerase requires only two transcription factors (TBP, TFB) for initiation in vitro, in contrast to eukaryal RNA polymerase II, which additionally requires TFIIE, TFIIF and TFIIH (Bartlett, 2005).

Transcriptional elongation in *Archaea* however is less well understood, than initiation since as stated above, eukaryotic TFIIS homologues have been found in *Archaea*. It exhibits homologies to eukaryotic TFIIS as well as to eukaryotic RPB9 (a subunit of eukaryal RNA polymerase II, with homologues in the other two eukaryal RNA polymerases) and has been shown to be able to induce a cleavage

activity in the RNA polymerase, as TFIIS does in *Eukarya* (Lange & Hausner, 2004). This is important to improve the fidelity of transcription and to rescue and prevent arrested elongation (Lange & Hausner, 2004). Along with TFIIS also homologues of bacterial elongation factors NusA and NusG have been identified, leaving the possibility that transcriptional elongation in *Archaea* may be also closely related to *Bacteria* (Bell & Jackson, 1998).

1.1.3.2.3. Translation

In *Archaea* the processes involved in the translation of mRNA for protein biosynthesis seem also to be a mixture of eukaryal and bacterial features. As in *Bacteria* archaeal genes do not contain introns but since lower eukaryotes also do not necessarily seem to have introns as well the absence or presence of spliceable elements should not be regarded as a typically prokaryotic or eukaryotic feature, respectively (Brown & Doolittle, 1997). On the other hand archaeal mRNAs are not 5´ capped like eukaryal ones and in many cases (but not always) they possess a Shine Dalgarno (SD) like sequence, similar to that of *Bacteria*. Furthermore *Archaea* exhibit 70S ribosomes, which contain ribosomal RNA (rRNA) components reminiscent to that of *Bacteria* in number and sizes (23S rRNA, 16S rRNA, 5S rRNA) (Brown & Doolittle, 1997). However, while *Bacteria* are susceptible to streptomycin, an anti 70S ribosomal inhibitor, *Archaea* as well as *Eukarya* are unaffected by such antibiotics, whereas susceptible to certain anti 80S ribosomal inhibitors (like anisomycin) (Brown & Doolittle, 1997). Initiation of translation in *Archaea* is performed by factors that are homologous to eukaryal ones, like eIF-1A, eIF2, eIF2B, eIF-4A (though for eIF-2B the eukaryotic version is more complex). Factors involved in the recognition of the 5´ cap of eukaryal mRNAs are missing in *Archaea*, which is consistent with the fact that *Archaea* lack the 5´ caps (Bell & Jackson, 1998). Another noteworthy characteristic is that *Archaea*, like *Eukarya* are using initiator tRNAs carrying methionine and not N-formylmethionine as is known for *Bacteria* (Keeling & Doolittle, 1995). On the other hand the recognition of the translational start sites resembles that of *Bacteria*. The Shine Dalgarno sequence is situated about 3 – 10 nucleotides upstream of the start codon (which is most frequently AUG, as in *Bacteria*) and exhibits a sequence complementary to the 3´ end of the 16S rRNA. Thus the start codon comes close to the anticodon of the initiator tRNA (Bell & Jackson, 1998). Elongation of the polypeptide chain is mediated also by factors homologous to eukaryal elongation factors such as eEF-1α,

which is responsible for recruitment of the aminoacyl tRNAs to the A-site of the ribosome, or eEF-2, which is involved in translocation of the ribosome. Recognition of stop codons is then, yet again, achieved by a single factor that shares similarities to eukaryotic release factors (Bell & Jackson, 1998).

As a conclusion, considering all these characteristics, *Archaea* in many ways seem to exhibit simplified versions of eukaryal features. On the other side they share significant similarities with *Bacteria*. This underlines their fundamental importance for understanding evolutionary relationships between the domains of life, as well as evolution as a whole.

1.1.4. Haloalkaliphilic *Archaea*

Haloalkaliphilic organisms face both high salt concentrations as well as a high pH. These organisms are found in all three domains of life. Herein only the properties of the *Archaea* and their adaptation to high salt and high pH conditions will be considered subsequently.

1.1.4.1. Diversity of halophilic *Archaea*

The *Halobacteriaceae* of the order *Halobacteriales* constitute the model halophilic microorganisms and contribute the largest part of microbial biomass in habitats as the Dead Sea, hypersaline soda lakes (such as Lake Magadi, Kenya) and saltern crystallizer ponds. Furthermore these microbes are almost solely responsible for the reddish color of such lakes since the membranes of many halophilic species hold large concentrations of C-50 carotenoid pigments, like bacterioruberin and its derivatives (Oren, 2002). In addition to that also the methanogens of the *Euryarchaeota* contain halophilic species and methanogenesis was found to take place even at nearly saturated salt concentrations. However within the kingdom *Crenarchaeota* no halophiles have been identified so far (Oren, 2002).

1.1.4.2. Halophily – living under high salt conditions

There is currently no unifying definition of where halophily starts. All (micro-) organisms have a minimum amount of NaCl that they require for growth, as well as an optimum, where they grow best and a maximum of NaCl that they can

tolerate. Based upon this fact Donn Kushner established three different categories for halophiles: "*extreme halophiles (growing best in media containing 2.5–5.2 M salt), borderline extreme halophiles (growing best in media containing 1.5–4.0 M salt), moderate halophiles (growing best in media containing 0.5–2.5 M salt), and halotolerant microorganisms that do not show an absolute requirement for salt for growth but grow well up to often very high salt concentrations (considered extremely halotolerant if the growth range extends above 2.5 M salt)"* (Kushner, 1978; Oren, 2008).

The majority of the halophilic *Archaea* belongs to the order of *Halobacteriales*. Within this order most of the species could be classified as extremely halophilic, according to Kushner´s definition. In addition to that also within the *Methanosarcinales* halophilic or at least highly halotolerant organisms have been found. It is also worth mentioning, that all archaeal species categorized as halophiles so far belong to the phylum *Euryarchaeota* (Oren, 2008).

1.1.4.2.1. Adaptations to high salt concentrations

In principle adaption to high salinity implies that the cytoplasm is at least isoomotic with the surrounding medium. There are two fundamental strategies to achieve this: (i) Accumulation of high molar concentrations of salt, usually KCl, inside the cell in order to maintain intracellular water activity. This however also requires the adaption of the complete proteome to these conditions, leaving no possibilities for surviving under low salt conditions. This technique is also referred to as "high-salt-in strategy" (Oren, 1999; Oren, 2008). (ii) Biosynthesis and/or uptake of organic osmotic solutes also allows for coping with high extracellular salt concentrations. This strategy is also called "compatible-solute strategy" and does not necessarily involve specially adapted proteins. There is a large variety of such compatible solutes, which ranges from polyols like glycerol, over certain sugars to amino acids and quaternary amines (Oren, 1999; Oren, 2008). Whatever strategy is used, there have to be potent mechanisms for extruding Na^+ from the interior of the cell. All halophilic microorganisms seem to have such a mechanism usually based upon Na^+/H^+ antiporters (Oren, 2002).

Though the compatible-solute strategy seems to be the most widespread in nature, the high-salt in strategy is most common in *Archaea*, especially in the extremely halophilic family of the *Halobacteriacae* (Oren, 1999; Oren, 2008).

1.1.4.2.1.1. Properties of halophilic proteins

Many halophilic proteins exhibit a large excess of acidic amino acid residues. This fact is important since electrostatic effects are significantly contributing to protein folding and stability (Lanyi, 1974). Normally one would assume, that a large number of similar charges in macromolecules, favors protein unfolding due to electrostatic repulsion yet it has been shown that residue linkages as well as disulfide interactions are indeed capable of overcoming the electrostatic force. On the other hand it has been suggested that the need of halophilic proteins for high concentrations of cations is due to the screening of negatively charged residues (Lanyi, 1974). This theory has been at least partially supported by experiments in which the high concentrations of monovalent cations, usually found within the cells of extreme halophiles, were exchanged by lower concentrations of divalent cations, without losing significant protein stability and activity. However, the requirement of high salt concentrations of halophilic proteins cannot be exclusively explained by this theory, since halophilic enzymes also exhibit specificity for certain anions (Lanyi, 1974). For example it has been observed that the activity of 3-hydroxy-3-methylglutaryl-coenzyme A reductase from *Haloferax volcanii* increases with rising concentrations of KCl, while it decreases when the concentration of NaCl is raised (Madern *et al.*, 2000).

There are also some other effects of high salinity that have to be taken into consideration, in order to understand why halophilic proteins are able to function in such extreme environments. With increasing concentrations of salt, new hydrophobic interactions are beginning to form, which were not stable enough before. Thus the protein becomes more tightly folded (Lanyi, 1974). Therefore halophilic proteins contain a low amount of hydrophobic amino acids, resulting in the need for high salinity in order to maintain these hydrophobic interactions (Fendrihan *et al.*, 2006). Furthermore local residues play a significant role, as they seem to be influenced under halophile conditions, (i) through structural changes or (ii) through direct effects on the residues themselves (Lanyi, 1974).

1.1.4.3. (Halo-) alkaliphilic *Archaea*

As it is for halophilism, there is also no common definition for alkaliphily. Usually organisms that have their growth optimum at around or above pH 9 are considered to be alkaliphile. Alkaliphiles can be further subdivided into alkaliphiles and the

haloalkaliphiles (Horikoshi, 1999). The latter have been observed to live in high salt and high pH environments like the lakes situated along the east African rift valley (e.g. Lake Elmenteita or Lake Magadi) and the western soda lakes of the United States. Hypersaline soda lakes like Lake Magadi (Kenya) or the Wadi Natrun (Egypt) are the natural habitat for the alkaliphilic members of the halophilic *Archaea*, where they show titers up to 10^7 – 10^8 cells/ml (Horikoshi, 1999).

1.1.4.3.1. Adaptations to high pH

Although alkaliphilic microorganisms live in environments that exhibit a high pH, the intracellular pH of these organisms has not necessarily to be high as well. There are various ways to measure the cytoplasmic pH inside the cell. For the α-galactosidase of alkaliphile *Micrococcus sp.* strain 31-2 (a representative of the *Bacteria*) the optimal pH for activity of this enzyme has been shown to be 7.5 suggesting that the internal pH of the cell is around neutral (Horikoshi, 1999). In addition to that in vitro protein synthesis systems from alkaliphiles show best activity at pH 8.2 – 8.5, which is only marginally higher than in *B. subtilis* (Horikoshi, 1999). Furthermore it is possible to measure the distribution of weak bases (which are not actively transported by the cell) within and outside the cell. In doing so it has also been found that the internal pH of the observed organisms lays around 8, although outside in the surrounding media it is far higher (Horikoshi, 1999). These facts lead to the conclusion, that the cell walls and cell membranes of alkaliphilic organisms are crucial for their survival in such extreme environments, as they separate the rather mesophile pH of the interior of the cell from the highly alkaline pH in the surrounding media (Horikoshi, 1999).

The cell walls of alkaliphilic microorganisms often contain acidic polymers. They provide additional negative charges that are capable of adsorbing Na^+ and H_3O^+ ions but repulsing OH^- ions. Thus the pH on the very surface of the cell may be considerably lower than in the surrounding media (Horikoshi, 1999). The cell wall of the highly alkaliphilic archaeon *Natronococcus occultus* contains glutaminylglycan, a glycoconjugate of two oligosaccharides that are linked to a backbone of poly-γ-L-glutamine via their α-amide group. One of which consisting of GalNAc and Glc, the other consisting of GlcNAc and GalA (Kandler & König, 1998).

To regulate intracellular pH aerobic alkaliphiles use Na^+/H^+ antiporters in addition to H^+ -coupled respiration. These antiporters generate a proton motive force based on both the transmembrane pH gradient (ΔpH) and the transmembrane electrical potential (Δψ) (van de Vossenberg *et al.*, 1999). For the majority of haloalkaliphiles the sodium concentration in the surrounding medium is high, while the concentration of H^+ is low. The Na^+/H^+ antiport reaction is responsible for both the extrusion of sodium ions from and the uptake of H^+ into the cytoplasm (van de Vossenberg *et al.*, 1999). Obviously this mechanism would be negatively influenced by an unidirectional influx of sodium or efflux of H^+. Thus the membrane of haloalkaliphiles should per se not be permeable for Na^+ and H^+ (van de Vossenberg *et al.*, 1999).

1.1.4.3.1.1. Properties of alkaliphilic proteins

Microorganisms do not only have proteins and enzymes inside the cell, but also enzymes that are secreted to the extracellular medium. These proteins have to be stable at a high pH in the case of alkaliphiles but also at a high salinity as it is the case for haloalkaliphiles. Neutrophile proteins usually denature when the pH is raised to a certain extent (Shirai *et al.*, 2008). Thus these proteins exhibit a number of interesting modifications in order to maintain functionality under these extreme conditions. The mechanisms involved in high pH stability of alkaliphilic proteins can be currently categorized into three different strategies: (a) the pKa modulation strategy, (b) the Asp+Glu gain (+DE) strategy and (c) the Asp+Lys loss – Glu+Arg gain (-DK+ER) strategy (Shirai *et al.*, 2008).

a) *The pKa modulation strategy*

   The catalytic activity of an enzyme at a certain pH is highly dependent on the pKa values of its catalytic residues. When the surrounding media exhibit (highly) alkaline conditions, modifying the pKa value towards a higher pH provides a possible solution (Shirai *et al.*, 2008). This is done by either modulation of the hydrophobic bonds, shielding catalytic residues from the solvent, or changing of the net charge of the molecule. (i) Hydrogen bond formation with a catalytic site often (but not always) results in lowering the pH optimum for an enzyme, since hydrogen bonds favor the deprotonated conformation of certain catalytic residues thus lowering the pKa. Therefore modification of the pKa towards a higher pH optimum often is achieved by the use of different amino acids (if compared to mesophilic homologues) in the proximity of the catalytic residues in order to avoid

formation of hydrogen bonds (Shirai *et al.*, 2008). (ii) To function properly in an alkaline environment it is also vital to keep the catalytic residues protonated, as the proton density is relatively low at high pH values. Amino acid residues with large side chains, though are capable of shielding the catalytic residues from the solvent molecules surrounding the protein (Shirai *et al.*, 2008). (iii) Negatively charged amino acid residues (like Asp and Glu) can attract protons thus raising the pKa of ionizable groups and resulting in considerably raising the pH optimum of the enzyme (Shirai *et al.*, 2008).

b) *The Asp+Glu gain (+DE) strategy*

It has been shown that alkaliphilic proteins exhibit a higher aspartic acid (Asp, D) and Glutamic Acid (Glu, D) over Arginine (Arg, R) and Lysine (Lys, K) ratio (Shirai *et al.*, 2008). This alters the net charge of alkaliphilic proteins towards negative values and goes along with the pKa modulation theory (see a). In reminiscence to the alkaliphile cell walls, where acidic polymers provide negative charges on the surface, proteins are suggested to do so by exposing negatively charged amino acid residues thus repelling hydroxyl ions (Shirai *et al.*, 2008).

In addition to that this strategy seems to be similar to the adaptation of proteins to high salt conditions and therefore accounts for both alkaliphilic and halophilic proteins (Shirai *et al.*, 2008).

c) *The Asp+Lys loss – Glu+Arg gain (-DK+ER) strategy*

Two analyses found the basics of this strategy. Together with a crystal structure analysis, a ancestral sequence evolutionary trace (ASET) analysis of the adaptation process of several proteins, like the alkaline α-amylase AmyK or the alkaline cellulose CelK was done (Shirai *et al.*, 2008). By the ASET analysis the amino acid residue changes of alkaline proteins were compared to their calculated ancestral sequence. In summery it was found that during evolution alkaliphilic proteins underwent an alteration in their amino acid composition pointing at decreases in the number of Lys and Asp residues, while the number of Arg, His and Glu increased (Shirai *et al.*, 2008). This was confirmed in a second analysis, where differences in the structure of alkaline enzymes were compared to those of (recent) non-alkaline enzymes and assessed with a student´s t-value. In addition to that

it could be shown, that the numbers of Lys-Asp/Glu ion pairs decreased, while that of Arg-Glu increased (Shirai *et al.*, 2008). Both analyses, the ASET, as well as the student´s t-value test led to the conclusion, that ion pair remodeling in the above described manner was necessary for alkaline adaptation and might be important for protein stability under alkaline conditions (Shirai *et al.*, 2008).

1.1.5. Benefits of doing research with (haloalkaliphilic) *Archaea*

In general extremophilic microorganisms, especially the relatively less explored *Archaea*, bury great potential of biotechnological applications let alone the potential of elucidating profound evolutional questions, thereby contributing to help satisfying man´s desire to explain how humanity itself evolved.

As the diversity of extremophiles is high, comprising thermophiles, psychrophiles, acidophiles, alkaliphiles, halophiles, barophiles and others, the range of biotechnological applications is wide as well, considering the large variety of different "extremozymes" (Eichler, 2001), that have been discovered in extremophiles, as well as considering the various types of enzymes that have not yet been discovered.

A common example of the use of halophilic enzymes is bacteriorhodopsin, which is responsible for the reddish color of many members of the *Halobacteriacae*, where it mediates photosynthesis. It has found its way into several light sensitive or bioelectrical applications, one of which being holography (Oesterhelt *et al.*, 1991; Eichler, 2001). Halophilic biopolymers also are valuable for industrial purposes. Biosurfactants that can be produced under high saline conditions could prove invaluable for the remediation of oil contaminated soil and water. Since many petroleum reservoirs exhibit high salinity, exopolysaccharides from halophilic *Archaea* (such as *Halobacterium salinarium*, *Haloferax volcanii* and *Halobacterium distributum*) could be applied in microbial enhanced oil recovery, as emulsifiers and mobility controllers (Margesin & Schinner, 2001). Because of their distinct lipid composition liposomes of *Haloarchaea* could contribute to medical and cosmetic applications. There liposomes are utilized as vessels for the transport of certain active agents to specific target sites in the body. The unique ether linked lipids of *Haloarchaea* (e.g. *Halobacterium cutirubrum*) exhibit a higher stability

and are far more resistant to esterases than their mesophilic ester linked counterparts (Margesin & Schinner, 2001). Other important products of *Haloarchaea* may be polyhydroxyalkanoates (PHA), some of which exhibiting similar properties as polyethylene and polypropylene, while being biodegradable at the same time. One example is *Haloferax mediterranei* that produces large quantities of poly ß-hydroxy butyric acid (PHB). In addition to that the extraction of PHB is very easy, since haloarchaeal cells simply lyse when the salt concentration is lowered to a certain point (Rodriguez-Valera & Lillo, 1992; Margesin & Schinner, 2001; Ventosa & Nieto, 1995). Furthermore enzymes that work under highly alkaline conditions are very important for industry. Alkaline proteases are mainly used as detergent additives, as well as in hide – dehairing processes (Horikoshi, 1999). From a biotechnological point of view haloalkaliphilic *Archaea* constitute an important group of microorganisms, as their proteins and enzymes combine halophilic and alkaliphilic features, both being valuable for industrial purposes.

However, the benefit of studying extremophiles lies not only in the finding of new applications for their products and enzymes but also has it greatly contributed to our understanding of protein folding, stability, structure and function (Gomes & Steiner, 2004).

1.1.6. *Natrialba magadii*

The archaeon *Natrialba magadii* was initially described as *Natronobacterium magadii* by Tindall *et. al.* in 1984 (Tindall *et al.*, 1984) where it was isolated from Lake Magadi (Kenya), which belongs to the east African Rift valley lakes and provides a high saline as well as a high alkaline environment thus being suitable for haloalkaliphilic microorganisms. Originally the discovered microbes where put into two generas – the rod shaped *Natronobacteria* and the cocci shaped *Natronococci* (Tindall *et al.*, 1984). However, in 1997 a 16S and 23S rRNA analysis performed by Kamekura *et al.* led to the introduction of new genera, one of which being the genus *Natrialba*, where *Natronobacterium magadii* was transferred to and since then is classified as *Natrialba magadii* (Kamekura *et al.*, 1997).

Nab. magadii belongs to the kingdom *Euryarchaeota* and as an alkaliphilic member of the family *Halobacteriaceae*, it requires a high sodium chloride

concentration, a high pH, as well as low Mg^{2+} concentrations (below 10 mM) (Kamekura *et al.*, 1997). The cells are motile and have an orange to red color, due to carotenoids stored in their membrane. *Nab. magadii* exhibits a rod shaped morphology, the cells measuring 5 – 7 µm in length and grows aerobically. It exerts optimal growth at a temperature from 37° C – 42° C, a pH of 8.5 – 10.5 and a sodium chloride concentration of 4 M. In order to prevent cell lysis, a minimum of 2 M NaCl is required (Tindall *et al.*, 1984). But even under optimal conditions growth of *Nab. magadii* in comparison to *E. coli* is very slow. The equivalent to an *E. coli* over night liquid culture can easily take up to seven days, and streaks on plates containing *Nab. magadii* rich medium can take even longer to show colonies. Incubation for 14 days or longer is no exception.

1.1.6.1. *Nab. magadii* in the laboratory

In our laboratory currently two strains of *Nab. magadii* are available – the wt L11 strain, containing the lysogenic halovirus ϕCh1 (see chapter 1.2.2.) and the cured strain L13, which does no longer contain the virus. *Nab. magadii* L13 has been obtained by repeated subculturing and testing of colonies by infection with ϕCh1 (Witte *et al.*, 1997). L13 serves as an indicator strain, as it can be infected with ϕCh1, in contrast to L11, where a super infection is impossible.

We are also able to transform *Nab. magadii* with a shuttle vector that has been developed in this laboratory (Iro *et al.*, in prep.).

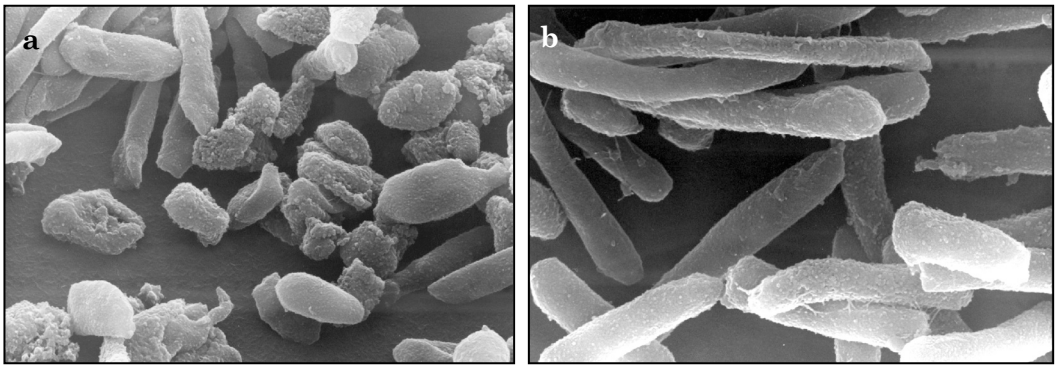

Figure 2| **Morphology of *Nab. magadii*.** Electron micrographs showing *Nab. magadii* cells. Belonging to the family of *Halobacteriacae Nab. magadii* was originally isolated from Lake Magadi (Kenya), a soda lake situated along the North African Rift valley where it faces both a high saline as well as a high alkaline environment. For optimal growth a sodium chloride concentration of 4 M, as well as a pH of 9.5 – 11 is required. **a**|The wild type strain L11 carries ϕCh1 as a prophage. **b**| The strain L13 has been cured of the phage and serves as indicator strain since it can be infected with ϕCh1.

1.1.6.1.1. Transformation of *Nab. magadii*

The first reported event of the transformation (transfection) of an archaeon dates back to 1987 where Cline and Doolittle managed to transfect *Halobacterium halobium*[*] with DNA obtained from the phage ϕH. The success of their method was measured by performing plaque assays. The method is based on the removal of the S-layer by EDTA, a chelating agent, in the absence of Mg^{2+} resulting in the formation of round shaped spheroblasts. These can be subsequently transformed with DNA with the aid of polyethylene glycol (PEG 600) (Cline & Doolittle, 1987; Charlebois *et al.*, 1987). The principles for transformation of *Haloarchaea* that were discovered then are still valid today. To the present day several other *Haloarchaea* could be transformed with DNA and the current knowledge in doing so is presently thoroughly collected in Michael L. Dyall-Smith´s manuscript "The Halohandbook" (Dyall-Smith, 2008), a compilation of various methods for working with *Haloarchaea*. However, for the transformation of *Nab. magadii* some alterations to the standard protocols had to be made. As EDTA has no effect on *Nab. magadii´s* S-layer our protocol involves the growth of the cells in media containing bacitracin, in order to weaken the glycolysation of the S-layer glycoproteins, subsequently followed by an enzymatic digest mediated by *Tritirachium album* proteinase K (Moens & Vanderleyden, 1997; Mescher & Strominger, 1976). After this treatment the resulting spheroblasts can be transformed according to standard protocols (for detailed information of the transformation protocol see chapter 2.2.2.) (Iro *et al.*, in prep.).

1.1.6.1.2. Genetic markers in *Nab. magadii*

Back in 1987 when Cline and Doolittle started doing transformation of *Haloarchaea* their work was exacerbated by lack of genetic loci that would work as selectable markers. Today, although the list of antibiotics affecting *Archaea* is still not very long, in our laboratory we take advantage of two genetic markers, one of which being a novobiocin resistance, the other being a mevinolin resistance. Novobiocin inhibits DNA gyrase in *Bacteria*, by binding to the *gyr*B subunit thereby preventing binding of ATP which results in a strong growth inhibitory effect (Holmes & Dyall-Smith, 1991). It has been shown that novobiocin also targets DNA gyrase of *Haloarchaea* and that resistance to this antibiotic in

[*] later termed *Halobacterium salinarium* (Ventosa & Oren, 1996)

Haloferax sp. is due to three point mutations in the gyrB homologue, all of which possibly affecting the region of the ATP binding site (Holmes & Dyall-Smith, 1991). Mevinolin inhibits the 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase, which in eukaryotes and *Archaea* is participated in the synthesis of mevalonic acid from coenzyme-A (Lam & Doolittle, 1992). Mevalonate in *Archaea* is used for the production of isoprenoid side chains for their unique lipids. Two different events independently lead to mevinolin resistance, being either the introduction of a tandem repeat of the HMG-CoA gene or a point mutation upstream of the gene, resulting in an up-regulation, and thereby producing an excess of the HMG-CoA gene product (Lam & Doolittle, 1992).

1.2. Viruses of the *Archaea*

In the context of *Archaea* both terms "virus" and "phage" are used synonymously. The term "virus" is derived from the Latin word for toxin or poison whereas "phage" is Greek for eating. In a microbial context "phage", simply is an abbreviation for "bacteriophage" ("bacterium eater"), which after the findings of Woese *et al.* could no longer be used as a valid term with respect to *Archaea*. However, many archaeal viruses have been discovered before the domain *Archaea* was split from the *Bacteria* and thus have been called "phages", though nowadays the term "virus" would scientifically be more correct.

1.2.1. Diversity of archaeal viruses

The first archaeviruses to be discovered were viruses infecting *Halobacterium salinarium* and *Halobacterium cutirubrum*, respectively (Torsvik T, 1974; Wais *et al.*, 1975). They were reminiscent to the bacteriophages of the family *Myoviridae*. Later viruses resembling phages of the family *Siphoviridae* were found, leading to the sophism of archaeviruses being just a variety of the common head tail type morphology bacteriophages (Prangishvili *et al.*, 2006a). This view however changed dramatically since the early days of the discovery of archaeviruses. Though, until now all of the described viruses of the *Archaea* contain linear or circular dsDNA, they exhibit an overwhelming morphological diversity with unique morphotypes

that are currently unknown for eukaryal viruses, as well as bacteriophages (Prangishvili *et al.*, 2006a). The viruses of the crenarchaeal *Sulfolobus neozealandicus* and *Acidianus gen.* display so exceptional shapes that upon discovery they were assigned to two totally new families. While the *Acidianus* bottle shaped virus (ABV) and the *Sulfolobus neozealandicus* droplet shaped virus (SNDV) constitute members of the family *Ampullaviridae*, the Acidianus two-tailed virus (ATV) was assigned to the family *Guttaviridae* (Prangishvili *et al.*, 2006a). Among the main archaeal kingdoms of the *Euryarchaeota* and the *Crenarchaeota* the morphologies of the viruses differ dramatically, with only two types of virions occurring in both kingdoms: *(i) spindle-shaped, enveloped virions with a short tail at one pointed end and (ii) spherical, lipid-containing virions with layered shell appearance and no discernible tail* (Prangishvili *et al.*, 2006a; Prangishvili *et al.*, 2006b). Otherwise both kingdoms differ greatly, with the kingdom of the *Euryarchaeota* currently displaying mostly tailed dsDNA viruses (Prangishvili *et al.*, 2006b). So far head tail viruses in *Archaea* have only been found to exclusively infect extreme halophiles or methanogens. Among these are the closely related *Halobacterium salinarium* virus ϕH and *Nab. magadii* virus ϕCh1, both of which being reminiscent of bacteriophages in their genome content, as well as in containing mosaic genomes due to extensive genetic exchange during the course of evolution (Prangishvili, Forterre and Garrett, 2006a).

1.2.1.1. Haloviruses

The first haloviruses found happened also to be the first archaeal viruses ever to be described (Torsvik T, 1974; Wais *et al.*, 1975). Given the numerous species of *Haloarchaea* that have been discovered since then, the number of described haloviruses is surprisingly low, also with respect to the fact, that the viral titer of hypersaline environments can be as high as 10^7 pfu/ml. *Halobacterium salinarium* virus ϕH currently is probably the best described halovirus and like many of the euryarchaeal viruses it exhibits a head-tail morphology (Dyall-Smith *et al.*, 2003). After discovery by Torsvik *et al.* in 1974 the lab of Wolfram Zillig at the Max-Planck Institute in Munich was mostly responsible for elucidating the characteristics of ϕH being a temperate virus possessing a genome size of 58.9 kbp of dsDNA and except from some methylase genes showing little sequence

similarities with known bacteriophages, though replication as well as control of lysogeny resemble the P1 bacteriophage (Schnabel *et al.*, 1982; Dyall-Smith *et al.*, 2003).

In 1993 Nuttal and Dyall-Smith at the University of Melbourne, Australia described two new head-tail viruses, HF1 and HF2. They showed identical morphologies and about 80 % genomic identity, but surprisingly had a completely separate host range, although isolated from the same lake, both being able to infect more than one host and both viruses being lytic (Nuttall & Dyall-Smith, 1993; Dyall-Smith *et al.*, 2003). This led to the question, whether temperate phages like ϕH and the very closely related ϕCh1 really are the dominant form of viruses among the *Haloarchaea*. Furthermore Bath and Dyall-Smith discovered two more lytic viruses, one of which being the *Haloarcula hispanica* infecting phage His1, similar in shape to the fusiform SSV1 virus particle of *Sulfolobus solfataricus*. The other one, His2, also using *Haloarcula hispanica* as a host, though did not show a morphological resemblance of SSV1. However, on a molecular level both His1 and His2 differ fundamentally from SSV1. The latter one being a temperate virus, exhibiting a circular genome, whereas His1 and His2 are lytic, display a linear genome and have their own DNA polymerase (Bath & Dyall-Smith, 1998; Bath *et al.*, 2006; Dyall-Smith *et al.*, 2003). In addition to that after the analysis of Dead Sea water by Oren *et al.* in 1997 the perception began to emerge, that head-tail viruses as well might not even be the most abundant archaeviral morphotypes (Oren *et al.*, 1997; Dyall-Smith *et al.*, 2003).

Though since the description of the first haloviruses major advances on this field of research have been made, the number of described haloarchaeal viruses is still relatively low. In their 2003 paper on haloviral diversity Dyall-Smith, Tang and Bath spot the major reason for this fact in the lack of cultivable hosts that are the dominant microorganism in their respective environment. Thus being able to cultivate the dominant microorganisms in high saline environments will ultimately lead also to the identification of novel, as well as dominant haloarchaeal virus species (Dyall-Smith *et al.*, 2003).

1.2.2. The Halovirus ϕCh1 – a general view

In 1997 the first phage ever detected to infect a haloalkaliphilic archaeon was ϕCh1 (Witte *et al.*, 1997). It is a temperate phage and infects *Nab. magadii*, an archaeon belonging to the kingdom *Euryarchaeota*, requiring both high salinity and an alkaline pH. ϕCh1 was discovered upon spontaneous cell lysis of *Nab. magadii* batch cultures. Lysis only occurred after cultures were grown to stationary phase, suggesting a growth phase dependent lysis behavior (Witte *et al.*, 1997). The virus causes turbid plaques when infected cells are plated upon a cell lawn. Phages were isolated from a single colony obtained from the edge of a plaque that contained vital cells which just then had not lysed. These cells were given the name "L11" and were used for all subsequent isolations of ϕCh1 (Witte *et al.*, 1997). However, superinfection of *Nab. magadii* host cells with the virus is not possible, as is known for other (bacterio-) phages like the *E. coli* λ phage or ϕH of *Halobacterium salinarium* (Witte *et al.*, 1997; Stolt & Zillig, 1992). Thus, a second *Nab. magadii* strain was isolated, by repeated subculturing and searching for altered lysis behavior. In this manner, a strain could be isolated, which had been cured of ϕCh1 and would serve as indicator strain, as it could be infected with the virus. This strain was termed *Nab. magadii* L13 (Witte *et al.*, 1997).

1.2.2.1. Morphology of ϕCh1

ϕCh1 exhibits a morphology typical for head-tail phages of the family *Myoviridae,* resembling *E. coli* T4 phage or ϕCh1 close relative ϕH of *Halobacterium salinarium*. The overall length of the phages lies around 200 nm, with icosahedrical heads approx. 70 nm and tails approx. 130 nm in length . The tail has an internal shaft covered by the contractible tail, exhibiting a total width of about 20 nm (Witte *et al.*, 1997). Electron micrographs show structures on the end of the tails, which are likely to be responsible for phage adsorption to the host cell (see Fig. 3a). Sodium chloride concentrations lower than 2 M result in the loss of infectivity, suggesting either phage disassembly under these conditions or significant conformational changes of structures participating in phage adsorption. Thus ϕCh1 seems to be perfectly adapted to the haloalkaline conditions its only known host, *Nab. magadii*, requires (Witte *et al.*, 1997).

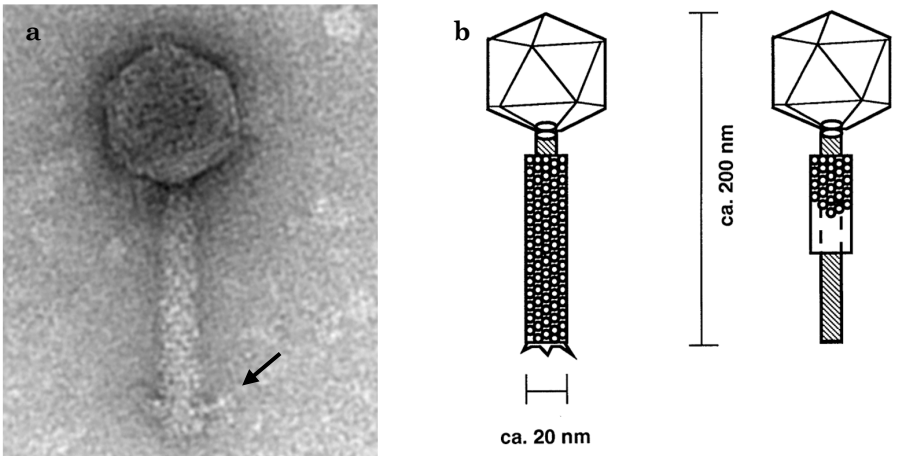

Figure 3| **Morphology of ϕCh1 particles**. **a**| The electron micrograph of a ϕCh1 virus particle shows a head tail morphology that is typical for the Myoviridae. At the bottom of the phage tail fibers are visible. (Photo kindly provided by Elke Bogner, Charite' Berlin, Institut für Virologie) **b**|Schematic representation of the virus particles (Witte *et al.*, 1997).

1.2.2.2. Protein composition of ϕCh1 particles

Early analysis of the virus revealed, that it consists of nine proteins termed A – I, respectively. Concerning the quantity by which they occur in the mature virus there are four major (A, E, H, I) and five minor (B, C, D, F, G) proteins (Witte *et al.*, 1997). Protein E has been shown to be the major capsid protein of ϕCh1. It is expressed in the late phase of virus development and during virus maturation undergoes proteolytic cleavage within the host *Nab. magadii* (Klein *et al.*, 2000). The majority of these nine proteins is acidic and displays isoelectric points between pH 3.3 and pH 5.2 thus rendering these proteins typical for halophiles (Witte *et al.*, 1997).

1.2.2.3. ϕCh1 is a temperate phage

Within the host cell temperate phages can have two distinct forms of existence. Phage DNA can either exist as an episomal element being ready for the lytic phage cycle, or it can integrate into the bacterial/archaeal chromosome where it is preserved and propagated together with the cellular DNA via cell division.

The fact that ϕCh1 lyses *Nab. magadii* only after it has been grown to stationary phase already suggests a lysogenic state of the virus. By hybridization of phage DNA with *Nab. magadii* L11 DNA the location where ϕCh1 is integrated into the chromosome could be determined (Witte *et al.*, 1997). It could also be shown, that extrachromosomal phage DNA does not show up until one day before onset of lysis. Furthermore, as assumed, no hybridization occurred when DNA from the non lysogenic *Natrialba magadii* L13 was used, thus proving that L13 indeed is cured from the phage (Witte *et al.*, 1997).

1.2.2.4. Genetic organization of ϕCh1

The genome of ϕCh1 consists of linear dsDNA. It has been sequenced in 2002 by Klein *et al.* and has a size of 58498 bp. Furthermore it was shown to be circularly permuted and terminally redundant, a feature that ϕCh1 shares with a number of bacteriophages (Klein *et al.*, 2002). By restriction analysis it could also be shown that aside from the linear form in the mature viral capsid, the viral DNA also must have a circular replicative form inside the host cell and that packaging of DNA occurs via the well known "head-full" mechanism, which explains the terminal redundance (Klein *et al.*, 2002). Prior to sequencing high pressure liquid chromatography (HPLC) was performed and revealed a G+C content of approx. 62 % (Witte *et al.*, 1997). Interestingly, in addition to that mature ϕCh1 particles also contain RNA. In a series of hybridization experiments this RNA was shown to be host derived, i.e. encoded by *Nab. magadii* chromosomal DNA. The DNA to RNA ratio is about 1:5 and at least eight different species of ϕCh1 packaged RNA have been identified, all of which ranging in the size of 5S rRNA. However, their function is still unknown (Witte *et al.*, 1997). Participation in the packaging process of DNA into the phage particles, like for example the eubacterial phage ϕ29 (Guo *et al.*, 1987a; Guo *et al.*, 1987b), remains a possibility. However, while ϕ29 RNA has been shown to be associated with the prohead of the phage particle and thus is susceptible to RNAse treatment (Guo *et al.*, 1987a; Guo *et al.*, 1987b), RNA of ϕCh1, being protected by the mature capsid, is totally unaffected by such a treatment (Witte *et al.*, 1997).

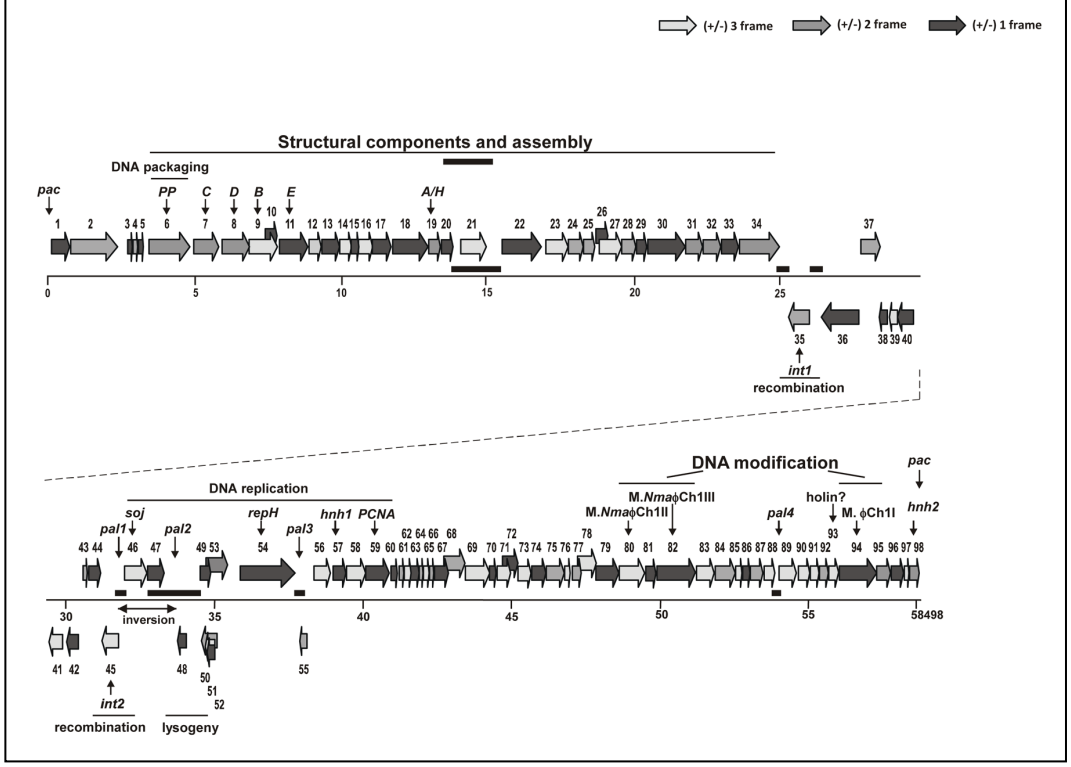

Figure 4| **Genetic organization of virus ϕCh1.** Sequencing the 58498 bp genome by Klein *et al.* in 2002 revealed a total of 98 predicted open reading frames (if the minimum length for an ORF was considered to be at least 30 codons). The genome shows a modular organization with the left part coding for structural components and assembly, the central part coding for genes important for DNA replication and the right part containing genes responsible for DNA modification.

(Klein *et al.*, 2002; with modifications)

Another interesting fact is that ϕCh1 DNA is partially methylated (Witte *et al.*, 1997). Restriction analyses with isoschizomeric enzymes *Sau*3A, *Dpn*I, *Mbo*I, respectively revealed a Dam-like adenine methylation of the cognate restriction site (5´-GATC-3´), as is known for *Enterobacteriaceae*. These three enzymes are either completely unaffected by adenine methylation (*Sau*3A), dependent on this methylation (*Dpn*I) or inhibited by it (*Mbo*I). It is also apparent that this Dam-like modification occurs on both strands as *Dpn*I and *Mbo*I only digest Dam sites if they are methylated or non-methylated on both strands, respectively (Witte *et al.*, 1997). On the other hand modification of cytosine residues could not be observed, upon cleavage with enzymes corresponding to cytosine methylation. Furthermore, restriction analyses clearly show that only a fraction of ϕCh1 DNA is methylated, whereas the rest remains unmodified. However, if methylated, this affects the whole genome, as there were no intermediate DNA fragments observed after restriction analysis (Witte *et al.*, 1997). After all Dam-like methylation of ϕCh1

DNA is rather surprising, given that such a modification in the DNA of the host, *Nab. magadii* is not observed (Lodwick *et al.*, 1986). Thus it was suspected that either ϕCh1 encodes for its own methyltransferase (Mtase), or that an otherwise silent host encoded Mtase is activated during virus maturation (Witte *et al.*, 1997). The former was confirmed when a virus encoded Mtase homolog could be identified, that is expressed in the late phase of virus development, and thus only a small fraction of viral DNA is becoming methylated, since packaging of DNA then is already in progress (Baranyi *et al.*, 2000). Yet, the purpose of ϕCh1 DNA methylation is still unclear. One option is that it represents an adaptation to a restriction/modification system of a so far unknown host (Baranyi *et al.*, 2000).

As mentioned above, the ϕCh1 genome has already been sequenced. A total of 98 predicted open reading frames (ORFs) could be identified, when ATG and GTG were considered as start codons and an ORF length of at least 30 codons was assumed (see Figure 4). However, only four of the predicted ORFs were found to actually start with a GTG. The tight arrangement of many ORFs pointing in the same direction over large areas of the genome indicates an organization into transcriptional units (Klein *et al.*, 2002). Comparison to available sequences in databases and to the only partially sequenced close relative ϕH gives a total of 48 matches, with only 17 matching to proteins with known functions and 31 matching to conserved hypothetical proteins of unknown function. However, the majority of the latter show high similarities to *Halobacterium salinarium* virus ϕH, only (Klein *et al.*, 2002). The ϕCh1 genome is organized into three major parts: (i) the left part, harboring genes mainly responsible for structural proteins and virion morphogenesis, (ii) the central part, where genes for replication, plasmid stabilization and gene regulation are situated and (iii) the right part with most genes of currently unknown functions in addition to genes involved in methylation and restriction (Klein *et al.*, 2002). The central part strikingly resembles the L-part of ϕH, which can circularize and replicate as an autonomous plasmid (Gropp *et al.*, 1992). The two fragments share sequence identities between 50 – 97 %. This is surprising, since the host range of ϕCh1 and ϕH is completely different with hosts that do not even belong to the same genus (Klein *et al.*, 2002). This segmented composition could be the result of the acquisition, i.e. exchange, of such modules among head tail viruses beyond the three domains during the course of evolution. Such a mechanism could be the main driving force of viral evolution. It explains

the similarities that ϕCh1 shares not only with ϕH but also with other bacteriophages (Hendrix *et al.*, 2000).

Sequencing also revealed another interesting feature in the ϕCh1 genome. It also encodes for two site specific recombinases (*int1 and int2*) of the λ-integrase type (Klein *et al.*, 2002). Int1, encoded by ORF35 could be shown to be participated in the inversion of the neighboring ORFs 34 and 36, respectively (Rössler *et al.*, 2004). The gene products of these two ORFs are likely to be part of the mature virus particle, as they are both detectable in an α-ϕCh1 antiserum. The inversion reaction produces a rearrangement of the C-terminal ends of both ORFs among each other, thereby possibly resulting in a structural change of the mature virus particle (Rössler *et al.*, 2004). As is known for bacteriophages Mu and P1 (Iida, 1984; Sandmeier, 1994) this could eventually lead to a shift in the host specificity, though to this date no hosts other than *Nab. magadii* have been found (Rössler *et al.*, 2004).

1.2.2.5. The ϕCh1 replication region and development of a shuttle vector

Sequencing of the ϕCh1 genome revealed homologies to a gene called *repH*. *RepH* as well as an AT-rich region 5´ to the gene were shown to be required for the replication of plasmids in halophilic *Archaea* (Ng & DasSarma, 1993). However, in ϕCh1 two ORFs, show a remarkable similarity to *repH* of *Haloarcula marismortui* plasmid pNRC100. (i) ORF54 contains GTG as a start-codon and encodes a protein of 581 amino acids in length, weighing approx. 65 kDa. This protein possibly contains a coiled coil structure in its center and the highest similarities to RepH were found to be located in the C-terminal end. (ii) ORF53 which lies upstream of ORF54 has a lower, yet significant homology to *repH* than ORF54. However, it resembles RepH only with its N-terminal end. In addition to that, an AT-rich region lies upstream of ORF53 and another one downstream of ORF54. These facts suggest that the ϕCh1 origin of replication is comprised of two rather than one open reading frames (Klein *et al.*, 2002; Iro *et al.*, in prep.).

In order to further investigate the origin of replication of ϕCh1, as well as to construct a functional vector system for haloalkaliphilic *Archaea*, the above described region was used to develop a shuttle vector for *E. coli* and *Nab. magadii*. The basis for this system was provided by a pKSII+ vector. Onto this plasmid the

mutated *gyrB* gene of *Haloferax alicantei* (Holmes & Dyall-Smith, 1990), which confers novobiocin resistance, was cloned. The resulting vector, pNov-1, was used for a cloning series during which different fragments of the putative ϕCh1 origin of replication were analyzed (Iro *et al.*, in prep.). This was done by determining the transformation efficiency of *Nab. magadii* with pNov-1 plasmids containing the respective fragments. These vectors were termed "pRo" and given a number (see Figure 6) (Iro *et al.*, in prep.). Plasmids pRo-1 and pRo-2, which lack the promoter in front of ORF49 as well as the 5´ AT-rich region, did not give any transformants. Plasmid pRo-4, containing the whole ϕCh1 replication region as depicted in Figure 6, gave very low transformation efficiencies. Only pRo-3, pRo-5 and pRo-6 could be transformed with significant efficiencies with pRo-5 performing slightly better. The latter no longer contains the start codons of two ORFs suspected to be involved in transcriptional repression of ϕCh1 – ORF49 and ORF55. (Iro *et al.* in prep).

The present work also participates in the investigation of the ϕCh1 replication region, as it remained to be proved, whether ORF53 or ORF54 are both necessary for replication. In addition to that the role of the ORF53/54 5´ and 3´ palindromic flanking regions is further characterized. These regions constitute putative binding sites for the gene products (gp) of ORF53 or ORF54, respectively and as stated above such interactions have been reported for *repH* by Ng and DaSarma (Ng & DasSarma, 1993). However, for ϕCh1 their role is still not fully elucidated as well.

1.2.2.6. Gene regulation in ϕCh1

The ϕCh1 genome contains two putative repressor genes, ORF48 and ORF49, respectively. ORF48 shows significant sequence similarities to the known repressor of the ϕCh1 close relative ϕH (Klein *et al.*, 2002). It contains an amino acid pair, Ala^{73}-Gly^{74}, which is also known to be conserved in repressors of various other bacteriophages (Raymond-Denise & Guillen, 1991; Perry *et al.*, 1985; Iro *et al.*, 2007). For these reasons ORF48 was named *rep*. It belongs to the winged helix-turn-helix DNA binding proteins, a family where the loops are small beta sheets. However, in contrast to many other known repressors *rep* was shown to be constitutively expressed during the virus maturation, suggesting an additional mechanism for ϕCh1 gene regulation (Iro *et al.*, 2007).

The repressor function of the gene product of ORF49 (gp49) was suspected after the analysis of the lysis behavior of several ϕCh1 infected *Nab. magadii* cultures derived from single plaques. One virus variant, termed ϕCh1-1, showed an increase in plaque formation as well as an earlier onset of lysis.

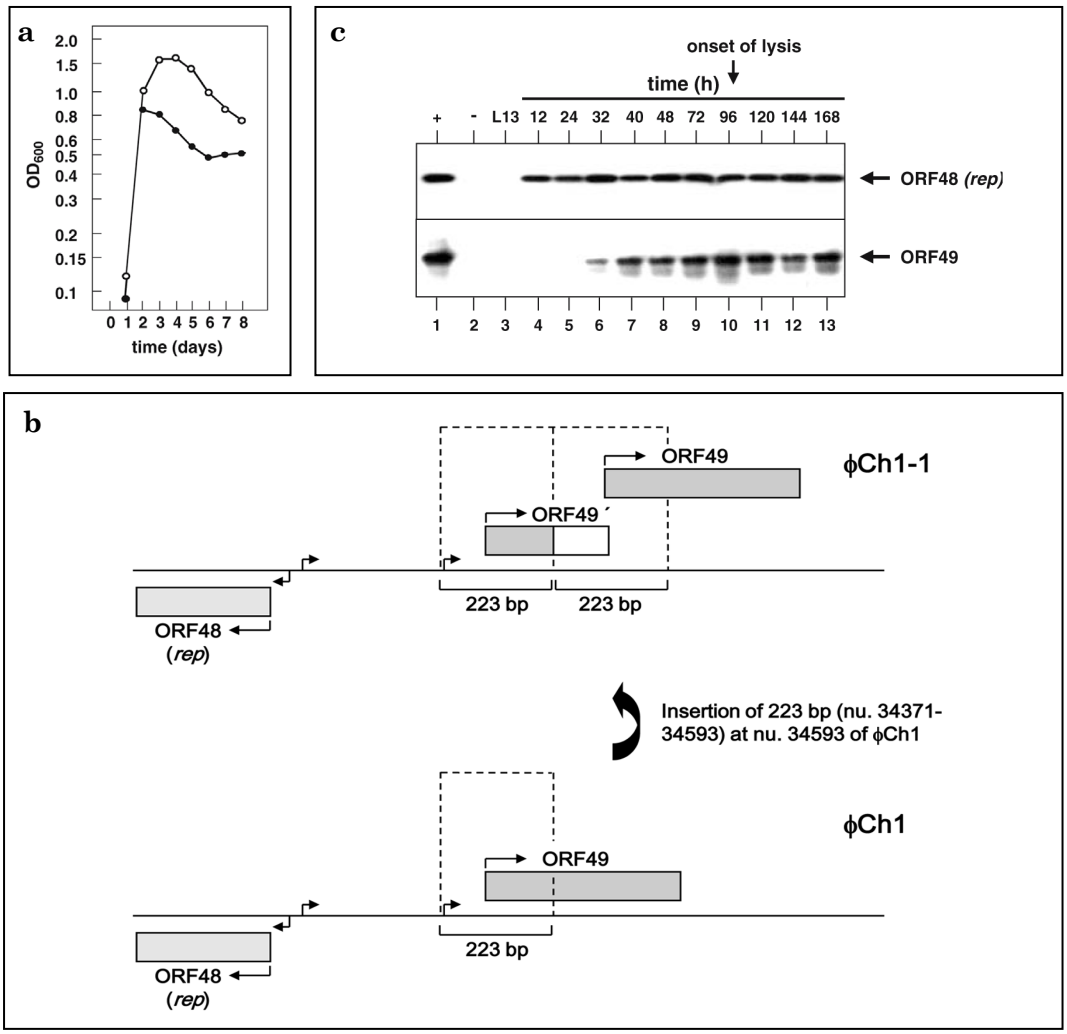

Figure 5| **Gene regulation in ϕCh1 by ORF48 (*rep*) and ORF49.** Studying the lysis behavior of ϕCh1 infected *Nab. magadii* cells obtained from different plaques revealed a virus variation that showed an earlier onset of lysis. It was termed ϕCh1-1. **a**| Growth curves of different *Nab. magadii* cultures infected with ϕCh1 wt (○) and ϕCh1-1 (●). **b**| Sequence analysis of ϕCh1-1 revealed a duplication of a 223 bp fragment resulting in an additional open reading frame – ORF49´ in front of the original ORF49 with a few overlapping nucleotides. **c**| RT-PCR showing that while ORF48 (*rep*) is constitutively expressed during the ϕCh1 life cycle expression of ORF49 starts in the logarithmic growth phase of the cells and increases during virus development.

(a, c: Iro *et al.* 2007; b: personal communication)

Sequence analysis revealed that ϕCh1-1 carries a duplicated 223 bp fragment comprising a part of ORF49 and a few nucleotides of its upstream region (nu. 34371–34593) (Iro *et al.*, 2007; see also Figure 5). This duplication produces an additional open reading frame, ORF49´, which lies just in front of the original

ORF49 with only a few overlapping nucleotides. Thus a co-transcription and -translation of both ORF49 and ORF49´ seems to be likely. The importance of ORF49 for the ϕCh1 life cycle is also underlined by the fact, that the mutation that leads to ORF49´ is not stable and reverts into the wild type relatively quickly (Iro *et al.*, 2007).

In contrast to ORF48, ORF49 is not constitutively expressed. ORF49 expression starts in the logarithmic phase and increases during virus development (Iro *et al.*, 2007). However, the mechanism that triggers its expression is so far unknown. Both, *rep* and ORF49 are situated head to head on the ϕCh1 genome with their intergenic region containing promoter consensus sequences. Reporter gene assays showed that ORF48 is responsible for a dramatic decrease of *bga*H activity, when *bga*H is transcribed from the ORF49 promoter (Iro *et al.*, 2007). This indicates the role of Rep as a transcriptional repressor. The fact, that *rep* is expressed constitutively while ORF49 is not points at other factors involved in the regulation of the ORF48/ORF49 system, where the expression of ORF48 has to be overcome at a certain point in time to allow for ORF49 to be expressed (Iro *et al.*, 2007).

The present work contributes in the further elucidation of the role of ORF49 in the regulation of the ϕCh1 life cycle and the characterization of the activity of the repressor.

2. Materials & Methods

2.1. Materials

2.1.1. Strains

2.1.1.1. Bacterial strains

Strain	***Genotype***	***Reference***
E. coli XL-1-Blue	endA1, gyrA96, hsdR17 (r_{k}-m_{K+}), lac, recA1, relA1, supE44, thi, (F´, $lacI^{q}$, lacZDM15, $proAB^{+}$, tet)	Stratagene
E. coli BL21(DE3) pLysE	F^{-}, ompT, $hsdS_{B}$ ($r_{B}^{-}m_{B}^{-}$), gal, dcm, (DE3), pLysE (Cm^{R})	Novagen
E. coli Rossetta	F^{-}, ompT, $hsdS_{B}$ ($r_{B}^{-}m_{B}^{-}$), gal, dcm, lacY1, (DE3), $pRARE^{6}$, (Cm^{R})	Novagen
E.coli Tuner^{Tm}	F^{-}, ompT, $hsdS_{B}$ ($r_{B}^{-}m_{B}^{-}$), gal, dcm, lacY1	Novagen
E. coli C41 pLysS	F^{-}, ompT, $hsdS_{B}$ ($r_{B}^{-}m_{B}^{-}$), gal, dcm, (DE3), pLysS (Cm^{R})	(Miroux & Walker, 1996)
E. coli C43 pLysS	F^{-}, ompT, $hsdS_{B}$ ($r_{B}^{-}m_{B}^{-}$), gal, dcm, (DE3), pLysS (Cm^{R})	(Miroux & Walker, 1996)
E.coli GM48F´	F^{-}, thr, leu, thi-1, lacY, galK, galT, ara, tonA, tsx, dam, dcm, supE44	(Palmer & Marinus, 1994)

2.1.1.2. Archaeal strains

Strain	*Genotype*	*Reference*
Nab. magadii L11	wt., ϕCh1 as prophage	(Witte *et al.*, 1997)
Nab. magadii L13	ϕCh1 cured derivate of L11	(Witte *et al.*, 1997)

2.1.2. Growth media

LB and LB + 0.2 % Glucose (rich media for *E. coli*)

Peptone	10 g
Yeast Extract	5 g
NaCl	5 g
(Glucose	2 g)*
pH 7.0	
add dH2O to 1 l	

for plates 15 g/l Agar were added

*LB + 0.2% Glucose was used with cultures containing the pMal-c2X vector

NVM+ (rich medium for *Nab. magadii*)

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 933 ml	
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 by adding the following solutions (to obtain 1 l):

0.57 M	Na_2CO_3 (dissolved in sterile dH_2O)	65 ml
1 M	$MgSO_4$ (autoclaved)	1 ml
20 mM	$FeSO_4$ (dissolved in sterile dH_2O)	1 ml

2.1.3. Antibiotics

2.1.3.1. for *E. coli*

Antibiotic	*Final concentration*	*General remarks*
ampicilin	100 µg/ml	dissolved in sterile ddH_2O, sterile filtered, stored at +4° C
tetracycline	10 µg/ml	dissolved in 70 % EtOH, stored at -20° C, light sensitive
chloramphenicol	20 µg/ml	dissolved in 96 % EtOH, stored at -20° C

2.1.3.2. for *Nab. magadii*

Antibiotic	*Final concentration*	*General remarks*
novobiocin	3 µg/ml	Dissolved in sterile ddH_2O, sterile filtered, stored at -20° C
bacitracin	70 µg/ml	dissolved in sterile ddH_2O, sterile filtered, stored at +4° C

2.1.4. Vectors

Plasmid	*Features*	*Source*
pUC19	bla, pMB1ori, lacZa, mcs	(Yanisch-Perron *et al.*, 1985)
pKSII+	mcs, bla, ColE1 ori, lacZa	Stratagene
pRSETA	mcs, bla, EK, PT7, RBS, His-tag, pUC ori, f1 ori	Invitrogen
pRSETC	mcs, bla, EK, PT7, RBS, His-tag, pUC ori, f1 ori	Invitrogen

pMal-c2X	lacI^{q}, Ptac, malE, pMAL-p4X polylinker, lacZa, rrnb terminator, Ampr, M13 ori, pbr322 ori	New England BioLabs
pKSII+ - TR4/Ro6-TR3/Ro7	mcs, bla, ColE1 ori, lacZa; carrying ϕCh1 fragments TR4/Ro6 and TR3/Ro7	(Ladurner, 2008)
pNov-1	*bla*, ColE1 ori, *gyrB*	(Ladurner, 2008)
pRo-4	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	(Iro *et al.*, in prep.)
pRo-5	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	(Iro *et al.*, in prep.)
pRo-6	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	(Iro *et al.*, in prep.)
pRo-7	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	this thesis
pRo-8	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	this thesis
pRo-9	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	this thesis
pRo-10	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	this thesis
pRo-11	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori	this thesis
pRo-5-ORF49	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 (nu. 34480-34833)	(Meissner, 2008)
pRo-5-ORF49Δ1	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 Δ1 (nu. 34321-34795)	this thesis
pRo-5-ORF49Δ2	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 Δ2 (nu. 34321-34760)	this thesis
pRo-5-ORF49Δ3	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 Δ3 (nu. 34321-34723)	this thesis

pRo-5-ORF49Δ4	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 Δ4 (nu. 34321-34686)	this thesis
pRo-5-ORF49Δ5	*bla*, ColE1 ori, *gyrB*, ϕCh1 derived ori, ϕCh1 ORF49 Δ5 (nu. 34321-34650)	this thesis
pMal-c2X-ORF49	$lacI^q$, P_{tac}, *malE*, pMAL-p4X polylinker, *lacZa*, *rrnb* terminator, Amp^r, M13 ori, pBR322 ori, ϕCh1 ORF49 (nu. 34480-34833)	this thesis
pRSETA-ORF49	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49 (nu. 34480-34833)	this thesis
pRSETA-ORF49Δ1	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49Δ1 (nu. 34480-34795)	this thesis
pRSETA-ORF49Δ2	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49Δ2 (nu. 34480-34760)	this thesis
pRSETA-ORF49Δ3	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49Δ3 (nu. 34480-34723)	this thesis
pRSETA-ORF49Δ4	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49Δ4 (nu. 34480-34686)	this thesis
pRSETA-ORF49Δ5	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49Δ5 (nu. 34480-34650)	this thesis
pRSETA-ORF49ΔN	mcs, *bla*, EK, P_{T7}, RBS, His-tag, pUC ori, f1 ori, ϕCh1 ORF49ΔN (nu. 34636-34833)	this thesis
pKSII+ promoter 43 + ORF44 / 3	mcs, bla, ColE1 ori, lacZa, P_{ORF43}, ORF44 of ϕCh1	(Iro *et al.*, 2007)

2.1.5. Primer

Name	***Sequence***	***Restriction site***
TR-1	AATTGCGGCCGCCGCGTTGAAGGCA	*Not*I
TR-2	AATTTCTAGATCCTGGGCCTCTTTGAA	*Xba*I
TR-31	AATTTCTAGACCATCGTGATAACGTTTGCG	*Xba*I
Ro-7	AATTCCCGGGGGCCGTG	*Sma*I
TR-4	GCAGAAGCTTCGGCGTGATCGCGAGTAA	*Hind*III
Δ53-1	GACCGAATTCGGATGCAAGCTGCTCGTGG	*EcoR*I
Δ53-2	GACCGAATTCCGTTGGATGGAGTCTACCAGTCC	*EcoR*I
Δ54-1	GACCGAATTCGCGGCGTCACTCAGCAAC	*EcoR*I
Δ54-2	GACCGAATTCGCGAGATCTTCACCGTTGAAGC	*EcoR*I
12-7-3	CAGCAGAAGCTTTCATCCTGCGGTTTC	*Hind*III
12-7-5	CAGCAGAAGCTTTCATCCTGCGGTTTC	*Hind*III
12-7-5C	AATTGGATCCATGAGAAAAATCAACGC CG	*Bam*HI
49-blunt	ATGAACACCCCCAATAGC	n/a
49-HindIII	CAGCAAGCTTTCGAGGCGTCATCCT	*Hind*III
49-Kpn	CAGCGGTACCTTGCGTTCAGTTCCG	*Kpn*I
49D1	CAGCAAGCTTTCAGCCATTGGTCCGCGAGC	*Hind*III
49D2	CAGCAAGCTTTCAGCCCGGAAAGGACGACA	*Hind*III
49D3	CAGCAAGCTTTCAGCCTCTCACCGAGGCGC	*Hind*III
49D4	CAGCAAGCTTTCACAAGAACAGGAGAGTGTCCA	*Hind*III
49D5	CAGCAAGCTTTCACCGGCGTTGATTTTTCG	*Hind*III
p17-5	GATCGGATCCATGCTCACAACTGCTGAAG	*Bam*HI
21-Hind	CAGGAAGCTTCAACCCCGAACTTCTACTC	*Hind*III
F1-F	GACCAGATCTGATCTGGATCGATCACTACGACC	*Bgl*II
F1-R	GACCAAGCTTCCGATGTGGCACCGTT	*Hind*III
43-3	CAGCAGTCTAGACGTGTCGACGAACAGC	*Xba*I
43-5	CAGCAGTCTAGACGTTGTGCCAGCCGT	*Xba*I
prom-1	GACGACGAATTCGTCCGACAACACAATTCC	*Eco*RI
prom-2	GACGACGGATCCTCCTGGGCCTCTTTG	*Bam*HI
prom-3	GACGACGAATTCGATGCGATCTCCTCTGG	*Eco*RI

N6-1b	CAGGTGATCATGCAACTTGAAGAACTACCAACA CG	*Bcl*I
MT-RT-I1	GGACGAGGTCAACCGAGTCACC	n/a
T7CRISPR_F	GAAATTAATACGACTCACTATAGGGCTCTTAAT AAATGCAGTAATACAGGG	n/a
CRISPR6R_R	CATGGAACTCAACAAGTCTCAGTGTGC	n/a

2.1.6. DNA and protein markers

2.1.6.1. DNA ladders

DNA ladder	***Fragments***	***Source/general remarks***
λ:*Bst*EII	8454, 7242, 6369, 5687, 4822, 4324, 3675, 2323, 1929, 1371, 1264, 702 bp	λ DNA was obtained from New England BioLabs or Fermentas; final concentration of digested DNA in a 1x loading dye was 50 ng/µl (Denaturation of cos-fragments at 65° C for 10 min when necessary)
pUC19:*Hae*III	587, 458, 434, 298, 257, 174, 102, 80, 18, 11 bp	final concentration of digested DNA in a 1x loading dye was 50 ng/µl
MassRuler™ DNA Ladder Mix, ready-to-use	80, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1031, 1500, 2000, 2500, 3000, 4000, 5000, 6000, 8000, 10000 bp	Fermentas; #SM0403
GeneRuler™ 1kb DNA Ladder	250, 500, 750, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 8000, 10000 bp	Fermentas; #SM0311

2.1.6.2. Protein ladders

Protein Ladder	*Fragment Sizes*	*Source*
Unstained Protein Molecular Weight Marker	116, 66.2, 45, 35, 25, 18.4, 14.4 kDa	Fermentas; #SM0431
PageRuler™ Unstained Protein Ladder	10, 15, 20, 25, 30, 40, 50, 60, 70, 85, 100, 120, 150, 200 kDa	Fermentas; #SM0661
PageRuler™ Plus Prestained Protein Ladder	10, 15, 25, 35, 55, 70, 100, 130, 250 kDa	Fermentas; #SM1811

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 *Bam*H1 buffer whenever possible or else with the recommended buffer according to the Fermentas DoubleDigest™ webpage (http://www.fermentas.com/en/tools/doubledigest).

2.1.7.2. PCR

Polymerase	**Source**
Pwo (from *Pyrococcus woesii*)	*Peqlab*
CrimsonTaq	New England BioLabs
GoTaq	Promega
Pfu (from *Pyrococcus furiosus*)	Promega

The listed enzymes were used with the provided buffers.

2.1.7.3. Other enzymes

Enzyme	Source
T4 Ligase	Fermentas*
Klenow Fragment	Fermentas*
DNAse I	Fermentas*
Proteinase K	Roche**

*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 2.2.2.)

2.1.8. Antibodies

For detection of His-tagged proteins an α-His-tag antibody conjugated with horseradish peroxidase (HRP) was used. This type of antibody renders a second antibody unnecessary and accelerates the detection procedure. The conjugated α-His-tag antibody was used in a 1:5000 dilution in TBS-T buffer containing 0.6 % Bovine Serum Albumin (BSA).

2.1.9. Buffers and solutions

2.1.9.1. Generation of $CaCl_2$ competent *E. coli* cells

MOPS I

100 mM	MOPS
10 mM	$CaCl_2$
10 mM	$RbCl_2$

pH 7.0 (KOH)

MOPS II

100 mM	MOPS
70 mM	$CaCl_2$
10 mM	$RbCl_2$

pH 6.5 (KOH)

MOPS IIa

100 mM	MOPS
70 mM	$CaCl_2$
10 mM	$RbCl_2$
15 %	glycerol
pH 6.5 (KOH)	

2.1.9.2. Transformation of *Nab. magadii*

Buffered High Salt Spheroblasting Solution (with or without glycerol)

2 M	NaCl
27 mM	KCl
50 mM	Tris-HCl pH 8.2
15 %	sucrose
(15 %	glycerol)

Unbuffered High Salt Spheroblasting Solution

2 M	NaCl
27 mM	KCl
15 %	sucrose

2.1.9.3. Isolation of ϕCh1 virus particles

Solution 1.1

2 M	NaCl
50 mM	Tris-HCl pH 9.5
0.6 M	CsCl

Solution 1.3

2 M	NaCl
50 mM	Tris-HCl pH 9.5
2.7 M	CsCl

Solution 1.5

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

High salt alkaline solution

4 M	NaCl
50 mM	Tris-HCl pH 9.5

2.1.9.4. DNA methods

2.1.9.4.1. Electrophoresis

50x TAE

2 M	Tris-HCl
1 M	Acetic Acid
0.1 M	EDTA
pH 8.2	

5x DNA sample buffer (with or without XC / Orange G)

50 mM	Tris-HCl
0.1 %	SDS
0.05 %	bromphenol blue (BPB)
(0.05 %	Xylene Cyanol (XC))
(0.05 %	Orange G)

pH 8.2

after autoclaving add sucrose to a concentration of 25 %

10x TBE

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

add ddH_2O to 1 l and adjust pH 8.0 (boracic acid)

6 % Polyacrylamide gel

1.2 ml	30 % PAA (29:1)
4.8 ml	1x TBE
60 µl	10 % APS
6 µl	TEMED

2.1.9.4.2. Gel extraction from polyacrylamide gels

Elutionbuffer

10 mM EDTA pH 8

10 mM MgAc.4H_2O

0.5 M NH_4OAc

2.1.9.5. Protein methods

2.1.9.5.1. SDS-PAGE and Western blot

30 % AA-solution (29:1)

29 %	acryl amide
1 %	N,N'–methylene bisacrylamide

4x separation gel buffer

1.5 M	Tris-HCl pH 8.8
0.4 %	SDS

4x stacking gel buffer

0.5 M	Tris-HCl pH 6.8
0.4 %	SDS

2x protein sample buffer

0.12 mM	Tris-HCl pH 6.8
4 %	SDS
17.4 %	glycerol
2 %	β-mercapto-ethanol
0.02 %	bromphenol blue

10x SDS-PAGE running buffer

0.25 M	Tris base
1.92 M	glycine
1 %	SDS

Coomassie staining solution

25 %	methanol
10 %	acetic acid
0.15 %	Coomassie Brilliant Blue R-250

Coomassie destaining solution

25 %	methanol
10 %	acetic acid

Ponceau S solution

0.5 %	Ponceau S
3 %	TCA

Transblot buffer

48 mM	Tris base
39 mM	glycine
0.037 %	SDS
20 %	methanol

10x TBS(-Tween)

250 mM	Tris-HCl
1.37 M	NaCl
27 mM	KCl
(0.05 %	Tween 20)
pH 8	

5 mM sodium phosphate buffer

A) 0.2 M $NaH_2PO_4.2H_2O$

B) 0.2 M Na_2HPO_4

Mix 2.55 ml of A and 2.45 ml of B.

Add ddH_2O to 200 ml.

2.1.9.5.2. Protein purification under native conditions

Lysis buffer

50 mM	NaH_2PO_4
300 mM	NaCl
10 mM	imidazole
adjust pH to 8.0 using NaOH	

Wash buffer

50 mM	NaH_2PO_4
300 mM	NaCl
20 mM	imidazole
adjust pH to 8.0 using NaOH.	

Elution buffer

50 mM	NaH_2PO_4
300 mM	NaCl
250 mM	imidazole
adjust pH to 8.0 using NaOH.	

Column buffer*

20 mM	Tris-HCl pH 7.5
200 mM	NaCl
1 mM	EDTA
1 mM	sodium azide
1 mM	DTT

* column buffer was used for native purification of maltose tagged proteins

2.1.9.5.3. Protein purification under denaturing conditions

Buffer B (lysis buffer)

100 mM	NaH_2PO_4
10 mM	Tris-HCl
8 M	urea
pH 8.0 (NaOH)	

Buffer C (washing buffer)

100 mM	NaH_2PO_4
10 mM	Tris-HCl
8 M	urea
pH 6.3	(HCl)

Buffer E (elution buffer)

100 mM	NaH_2PO_4
10 mM	Tris-HCl
8 M	urea
pH 4.5 (HCl)	

2.1.9.5.4. Inclusion body purification from *E. coli*

Buffer A (+ sucrose)

50 mM	Tris-HCl pH 7.5
50 mM	EDTA
5 % (v/v)	Triton X-100
(8 % (w/v)	sucrose)

Washing buffer

50 mM	Tris-HCl pH 7.5

2.1.9.5.5. Protein renaturation

a) Dialysis

5x GBB

50 mM	Tris-HCl pH 8
25 mM	EDTA
30 %	glycerol
5 mM	DTT (add just before use)

b) Size exclusion chromatography

Renaturation buffer

50 mM	NaH_2PO_4
300 mM	NaCl

adjust pH to 8.0 using NaOH.

c) Gel extraction

Coomassie staining solution (w/o fixation)

0.2 %	methanol
0.5 %	TCA
0.1 %	Coomassie Brilliant Blue R-250

2.1.10. Bandshift assays

5x GBB

50 mM	Tris-HCl pH 8
25 mM	EDTA
30 %	glycerol
5 mM	DTT (add just before use)

10x Bandshift running buffer

70 mM	Tris-Hcl pH 8
30 mM	NaOAc
20 mM	EDTA

BSA stock solution

1 mg/ml BSA in ddH_2O

Store at -20° C

2.2. Methods

2.2.1. Preparation and transformation of $CaCl_2$ competent *E. coli* cells

100 ml LB-medium containing the respective antibiotics (refer to section 2.1.3.1.) were inoculated with an overnight culture of the particular *E. coli* strain to an OD_{600} of approx. 0.1. The culture was then incubated at 37° C under constant shaking until and OD_{600} of 0.6 – 0.8 was reached. Subsequently, cells were collected by centrifugation for 10 min at 10 krpm at 4° C. The pellet was resuspended in 40 ml ice cold MOPS I solution and incubated another 10 min on ice. The cell suspension was again centrifuged and the resulting pellet was resuspended in 40 ml ice cold MOPS II solution and incubated for 30 min on ice. After another centrifugation step the pellet was thoroughly resuspended in 2 ml MOPS IIa solution. Aliquots of 100 µl each were directly used for transformation or immediately frozen to -80° C for use at a later time point.

For transformation aliquots of competent cells, if frozen, were thawed on ice and DNA was added in proper amounts*. After incubation on ice for 30 min cells were heat shocked by putting them to 42° C for 2 min. Thereafter the cells were put back on ice for a short time. Then 300 µl of LB medium were added per batch and for regeneration the cells were put on 37° C for 30 min without shaking. Eventually each batch was divided into three equal portions that were plated on selective agar plates.

*for XL-1 Blue strains 1 µl of plasmid DNA prepared with commercially available mini prep kits was sufficient to give enough transformants. When transforming a ligated vector the total 15 µl ligation mix (see section 2.2.4.7.5.) were used. Strains intended for overexpression of recombinant protein (Tuner, Bl21(DE3), Rosetta, C41(DE3), C43(DE3)) had to be transformed with 5 - 7 µl of plasmid DNA prepared with commercially available mini prep kits obtained from Fermentas, PEQLAB or QIAGEN.

2.2.2. Preparation and transformation of competent *Nab. magadii* cells

Pre-cultures that had been grown to late log phase were used for inoculation of NVM+ medium containing 70 µg/ml bacitracin to an OD_{600} of approx. 0.1. For faster growth this was done in baffled flasks. Cultures were then incubated at 37° C, shaking, until they reached an OD_{600} of 0.5 – 0.6. Then the cells were collected by centrifugation with 6 krpm for 15 min at room temperature and the pellet was thoroughly resuspended in high salt buffered spheroblasting solution with glycerol (in half of the volume of the initial culture). Subsequently proteinase K was added to a concentration of approx. 20 µg/ml (or 0.6 mAU/ml) and the batch was incubated at 42° C, while shaking, until spheroblasts had formed. Usually this took approx. 48 h and spheroblast formation was visually confirmed by light microscopy.

Transformation of *Nab. magadii* competent cells was done on the basis on the earlier described PEG-600 method (Charlebois *et al.*, 1987; Cline & Doolittle, 1987). Batches of 1.5 ml of spheroblasts were taken for each single transformation. Cells were collected by centrifugation at 10 krpm for 3 min at room temperature using a microcentrifuge. The pellet was carefully resuspended in 150 µl of buffered high salt spheroblasting solution (w/o glycerol) and 15 µl of a 0.5 M EDTA solution (pH 8) were added and incubated at room temperature for 10 min. Subsequently the

respective DNA was added, usually in an amount of 3 µg total, and incubated for another 5 min at room temperature. Then 150 µl of a 60 % PEG-600 solution (therefore frozen amounts of PEG-600 were heated to 65° C and solubilized in unbuffered high salt spheroblasting solution immediately before use) were added and the batch was incubated for 30 min at room temperature. For washing 1 ml of NVM+ was added and cells were collected by centrifugation at 10 krpm for 5 min at room temperature using a microcentrifuge. The supernatant was removed completely and the pellet was thoroughly resuspended in 1 ml of fresh NVM+. For regeneration the batch was put onto a thermomixer and incubated at 37° C, while shaking, until no more spheroblasts were visible (or alternatively 24 – 48 h). Then 100 µl of the batch were plated per selective agar plate and incubated at 37° C until colonies were visible (usually 10 – 21 days).

For confirmation of transformants single colonies were inoculated in 1 ml or 15 ml of fresh selective NVM+ and incubated at 37° C while shaking until the culture reached stationary phase. Then 100 µl of the culture were taken and cells were harvested by centrifugation at 11.2 krpm for 3 min at room temperature. The pellet was resuspended in 100 µl of autoclaved ddH_2O resulting in lysis of the cells. 1 – 2 µl of lysed cells were taken as template for PCR using primers suitable for verification of the clone.

2.2.3. Phage methods

2.2.3.1. Isolation of ϕCh1 virus particles

A fresh plaque of *Nab. magadii* lysogenic strain L11 was inoculated in NVM+ and grown to an OD_{600} of approx. 0.5 – 0.6 (before onset of lysis). If sealed, this culture can be kept for months at room temperature for repeated use.

From the above culture a larger amount of NVM+ (usually 60 – 120 ml) could be inoculated and grown to an OD_{600} of approx. 0.5 – 0.6. This pre-culture was then used to inoculate large quantities of NVM+ (4.5 l – 3 x 1.5 l). The growth curve of this large culture was monitored by measuring OD_{600} 1 – 2 x per day. It was grown beyond the onset of lysis until OD_{600} had dropped to 0.4 – 0.5. Subsequently the batch was spun at 8 krpm for 20 min at room temperature to remove cells and cell debris. To the supernatant (containing the virus particles) 10 % (w/v) of PEG 6000

was added and stirred over night in order to precipitate the phage particles. On the next day the batch was centrifuged (same conditions as before) and the pellet was resuspended in 50 ml of high salt alkaline solution.

Further purification of virus particles was performed with the aid of a discontinuous CsCl density gradient. The gradient was build up in Beckman ultracentrifuge tubes by carefully layering CsCl solutions of different densities over each other. This was done in the following order (from bottom to top): 2 ml of solution 1.5, 5 ml solution 1.3, 6 ml of phage suspension and approx. 1 ml of solution 1.1. The gradient was centrifuged at 30 krpm for 20 h at room temperature resulting in one lower band containing *Nab. magadii* flagella and one or two* upper bands representing the virus particles. The upper band(s) was/were carefully removed in the smallest volume possible and further purified by the use of a continuous gradient. Therefore the virus particles were mixed with an equal volume of solution 1.3 and centrifuged again at 30 krpm for 20 h at room temperature. The resulting band was removed carefully – again in the smallest volume possible and dialyzed two times overnight against high salt alkaline solution in order to remove the cesium. The quantity of infectious ϕCh1 virus particles was determined via phage titer analysis as described below.

*if two bands occurred they lay very closely together, probably representing two fractions of differently methylated viral DNA.

2.2.3.2. Determination of phage titers by soft plating technique

Respective *Nab. magadii* L13 strains were grown to late log or stationary phase. Then 5 ml of NVM+ soft agar were mixed with 400 µl of respective L13 culture as well as with 100 µl of phage dilution (purified virus particles in dilutions of 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3} and 10^{-2} in NVM+). The mixture was poured onto selective or non selective rich medium plates, respectively. Each batch was done in duplicates and plates were incubated at 37° C until plaques were visible (10 – 21 days).

2.2.3.3. Isolation of ϕCh1 virus DNA

300 µl of purified virus particles (see section 2.2.3.1.) were mixed with 900 µl of autoclaved ddH_2O. Subsequently two Phenol/Chloroform extractions were performed using 600 µl Phenol/Chloroform (1:1). The aqueous phase was mixed

with 2.5 x vol. of 96 % Ethanol (-20° C) and put to -80° C for 15 min. Precipitated DNA was collected by centrifugation at 16.4 krpm for 30 min at 4° C using a refrigerated microcentrifuge. The resulting pellet was washed twice with 1 ml of 70 % Ethanol and then the pellet was air dried and solubilized in 25 µl of autoclaved ddH_2O. Successful DNA isolation was confirmed by agarose gel electrophoresis.

2.2.4. DNA Methods

2.2.4.1. Agarose gel electrophoresis

The respective amount of agarose* was mixed with 1x TAE buffer and heated in a microwave until it was completely melted. Then the molten agarose was cooled and cast into the provided gel tray. Combs were put into the agarose and the system was cooled until the agarose had become solid. Subsequently the combs were removed and the gel tray was put into the suitable electrophoresis apparatus that had been filled with 1x TAE buffer. Then 1x TAE was added until the gel was slightly overlaid with buffer. Afterwards the combs were removed and the corresponding slots were ready to be loaded. DNA samples were prepared by mixing 3 or 5 µl of respective DNA with 5 µl of 5x DNA sample buffer. Thereafter the total volume was loaded onto the gel. Separation of DNA started when power was applied (10 V/cm) and migration of DNA fragments could be estimated due to the respective dyes that had been added to the sample buffer. After gels had been sufficiently run they were put into an ethidium bromide bath (5 mg/l) and stained for 15 – 20 min. Subsequently they were shortly rinsed in ddH_2O and DNA bands were visible under UV light where a picture could be taken.

*dependant on the application gels containing 0.8 %, 1.1 % or 1.5 % (for resolution of smaller fragments) agarose were cast.

2.2.4.2. Polyacrylamide gel electrophoresis

For resolution of DNA fragments smaller than 500 bp electrophoresis was carried out using polyacrylamide (PAA) gels. Therefore BIORAD´s Mini – Protean® 3 system was used. Gels (6 % PAA in 1x TBE) were cast according to the manufacturers protocol. After the gel had polymerized the gel cassette was put into

the provided apparatus, which was subsequently filled with 1x TBE buffer. DNA samples were prepared and loaded as described in section 2.2.4.1. Gels were run by applying a power of 15 mA per gel. Migration of DNA fragments could be estimated through the use of dyes that had been added to the DNA sample buffer. Finished gels were left on one of the two sandwich glass plates and carefully put into an ethidium bromide bath (5 mg/l) for 5 – 10 min. After staining the gels were shortly rinsed in ddH_2O and DNA bands were visualized and photographed using UV light.

2.2.4.3. Polymerase chain reaction (PCR)

The polymerase chain reaction allows amplification of a DNA region of interest. In the context of this work PCR was used to amplify certain DNA fragments for (i) cloning or bandshifiting (preparative PCR) or (ii) to confirm positive clones (analytical PCR). Most of the PCRs, especially those for generation of DNA fragments that were used for cloning, were done using, *Pwo* Polymerase (Peqlab), as it has a proof reading function. In cases where *Pwo* did not give sufficient yield *Pfu* Polymerase (Promega) or *Taq* polymerase (Promega, NEB) was used. For analytical purposes *Taq* polymerase was used most frequently. Polymerases were used in amounts described in section 2.2.4.3.3.; elongation temperature and -time were set up as recommended by the provider.

2.2.4.3.1. Primers

Forward and reverse primers were obtained from VBC genomics in a lyophilized state. Upon delivery ddH_2O was added to give a stock concentration of 1 µg/µl and primers were kept at -20° C. Primer stocks were diluted 1:10 right before use. Melting temperature of primers was calculated using the Gene Runner Version 3.01 from Hastings Software, Inc. (available at http://www.generunner.net/). In the PCR program the annealing temperature initially was set to be 4° C below the lower melting temperature of both primers but had to be adjusted in some cases in order to increase yield.

2.2.4.3.2. Template DNA

As template DNA either (i) plasmids purified with a miniprep kit, (ii) purified ϕCh1 DNA or (iii) crude DNA extracts from *E. coli* or *Nab. magadii* were used. Crude extracts from *E. coli* were created by taking 5 µl of an overnight culture,

adding 100 µl of ddH_2O and heating to 95° C for 5 – 10 min. For *Nab. magadii* 100 µl of culture were taken and cells were collected by cenntrifugation at 11.2 krpm for 3 min at room temperature. Afterwards the pellet was resuspended in 100 µl of ddH_2O. Alternatively a single colony from an agar plate could be taken and put directly into 50 – 100 µl of ddH_20.

While crude extracts were taken undiluted (1 – 2 µl) per PCR, plasmids (purified with the GeneJET™ Plasmid Miniprep Kit from Fermentas) or ϕCh1 DNA were diluted 1:30 before use in PCR (1 – 2 µl).

2.2.4.3.3. PCR batches and -programs

Below the batch sizes for preparative PCRs are listed. For analytical purposes the batch sizes were reduced to ½, or even to ¼ in some cases.

Batch for *Pwo*

67 µl	ddH_20
10 µl	10x Pwo buffer
10 µl	2 mM dNTPs
5 µl	Forward primer (500 ng)
5 µl	Reverse primer (500 ng)
1 µl*	template DNA
2 µl	Pwo Polymerase

Program for *Pwo*

	94° C	5 min
33 cycles	94° C	1 min
	Ann.Temp.	1 min
	68° C	1 min/kbp
	68° C	5 min
	4° C	infinite

Batch for *Taq* (Go*Taq*® or Crimson*Taq*™) of *Pfu*

20 / 10 µl	5x *Taq* / 10x *Pfu* buffer
58 / 68 µl	ddH_2O
10 µl	2 mM dNTPs
5 µl	Forward primer (500 ng)
5 µl	Reverse primer (500 ng)
1 µl*	template DNA
1 µl	of the respective polymerase

Program for *Taq* and *Pfu*

	94° C	5 min
33 cycles	94° C	1 min
	Ann.Temp.	1 min
	72° C	1 min/kbp
	72° C	5 min
	4° C	infinite

*if more than 1 µl of DNA was used the volume of ddH2O was adjusted to maintain the total volume of the batch

2.2.4.3.4. Confirmation and purification of PCR products

Successful PCR was confirmed by gel electrophoresis (agarose or polyacrylamide gels). For preparative purposes, and when no interfering side products could be detected on a gel, the PCR product was purified using the QIAquick PCR Purification kit from QIAGEN. This kit is based on the principle that the PCR product is bound to a silica column, washed with ethanol and eluted afterwards. Thus unwanted reagents like the PCR buffer, or excess nucleotides and primers were removed. The PCR product was usually eluted with ddH_2O in half of the initial batch volume, i.e. 50 µl. However, in cases where PCR yield was substantially low elution volume was adjusted to raise DNA concentration to reasonable values. The success of the purification was again confirmed via gel electrophoresis.

2.2.4.4. Plasmid isolation from *E. coli*

For plasmid preparation from *E. coli* was done using the GeneJET™ Plasmid Miniprep Kit from Fermentas and following the provided protocol. However, instead of 1.5 ml 3 ml of an overnight culture of the respective *E. coli* strain (carrying the plasmid of interest) were taken for plasmid isolation. Elution of plasmid DNA from the silica column was done with 50 µl of ddH_2O. Samples of the isolated plasmids were loaded onto an agarose gel in order to confirm the plasmid preparation.

2.2.4.5. Gel extraction of DNA fragments from agarose gels

In cases where PCR gave more than one product or where the DNA fragment of interest had to be retained from a digested plasmid the respective fragment had to be cut out from a gel. The complete DNA sample was mixed with 5x DNA sample buffer and the total volume of the resulting mix was loaded onto a 0.8 or 1.5 % agarose gel (according to the size of the DNA of interest) in portions of 20 or 25 µl per slot. For example, if the DNA sample had a volume of 100 µl it was mixed with 20 µl of 5x DNA sample buffer and loaded onto a gel using 6 slots each loaded with 20 µl of sample. After the gel had sufficiently run it was put into a freshly prepared ethidium bromide bath (5 mg/l) and stained as short as possible (usually 3 min). Afterwards the gel was rinsed shortly with dH_2O und bands were visualized using UV light. The band of interest was cut out of the gel with a scalpel, making the

UV exposure as short as possible. Gel slices were transferred into Eppendorf tubes. Extraction of the DNA fragments from the gel slices was done by using the QIAquick Gel Extraction kit from QIAGEN and following the provided protocol. The result and purity of the extracted DNA fragment were controlled by gel electrophoresis.

2.2.4.6. Gel extraction of DNA fragments from polyacrylamide gels

For DNA fragments around 500 bp in length or shorter extraction from PAA gels had to be performed. As described in section 2.2.4.5. the DNA samples were mixed with 5x DNA sample buffer and loaded onto PAA gels. However, since the slot volume was smaller a maximum of 10 µl per slot could be loaded. After the gel had been run it was put into a freshly prepared ethidium bromide bath (5 mg/l), stained for approx. 1 – 2 min and shortly rinsed in dH_2O. Bands of interest were carefully cut out of the gel by using a scalpel and put into Eppendorf tubes. Subsequently the gel slices within the tubes where thoroughly crushed using a blue or yellow pipette tip. Then approx. 300 – 400 µl of elutionbuffer (alternatively 2 x the weight of the gel slice) were added and the tubes were incubated on a thermomixer at 37° C overnight, shaking. On the next day the PAA had to be removed completely. Therefore the tubes were spun at 13.2 krpm for 5 min at room temperature and the supernatant was transferred into a new tube. This step was usually done 2 – 3 times combining the supernatants. As a last step the united supernatants were centrifuged again and the resulting supernatant was put into a fresh tube. The DNA was then purified from the supernatant by using the QIAquick PCR Purification kit from QIAGEN. DNA was eluted from the silica columns with 50 µl of ddH_2O. The result and purity of the extracted DNA fragment were controlled by gel electrophoresis.

2.2.4.7. Cloning in *E. coli* strains

2.2.4.7.1. Restriction of DNA

Digestion of the respective DNA fragments was done using enzymes and corresponding buffers provided by Fermentas in the recommended concentrations. Full digestion of DNA was achieved by an incubation time of at least 3 h or even overnight (if no star activity) at the recommended temperature. For analytical purposes, i.e. testing for a specific DNA fragment within a plasmid, restriction time was lowered to 1 h.

2.2.4.7.2. Purification of DNA

After digestion of DNA buffers and nucleotides or tiny DNA fragments had to be removed in order to obtain the pure DNA fragment of interest. This was done with the aid of the QIAquick Purification kit from QIAGEN following the provided protocols. Usually, DNA was eluted from the silica column with ddH_2O, in half of the initial volume. In cases where the digested DNA fragment of interest was obtained by extraction from a gel no further purification was necessary (since the result of the extraction is an already purified DNA fragment).

2.2.4.7.4. Blunting DNA with the *Klenow* fragment

For some cloning strategies blunt ends were required. In order to produce blunt ends from sticky ends created by digestion with certain restriction enzymes the *Klenow* fragment of *E. coli* polymerase I was used as it fills up 5´ overhangs if sufficient dNTPs are provided. Batches contained 100 – 200 µM dNTPs (each) and were incubated at 37° C for 30 min followed by inactivation of the enzyme at 75° C for 15 min.

2.2.4.7.5. Ligation of respective DNA fragments

DNA fragments with matching ends (e.g. DNA sequence of interest and respective vector) were ligated using T4 Ligase and the corresponding buffer both provided by Fermentas. Ligation was done either at room temperature for 3 h or at 16° C overnight.

Batches were composed as follows:

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 could be used for transformation of *E. coli* (see section 2.2.1.).

2.2.4.7.6. Pre-screening of candidates by quick plasmid preparations

After transformation colonies were screened for possible positive candidates. 20 – 40 test tubes each containing 5 ml of LB medium and respective antibiotics were inoculated with one colony each and incubated overnight at 37° C while shaking. Subsequently 300 µl of each culture were transferred to an Eppendorf tube and cells were harvested by centrifugation at 13.2 krpm for 3 min at room temperature using a microcentrifuge. The pellets were each resuspended in 30 µl of 5x DNA sample buffer. Upon addition of 14 µl of Phenol/Chloroform (1:1) the tubes were vigorously vortexed for 30 sec and afterwards centrifuged at 13.2 krpm for 7 min at room temperature. 12 µl of the aqueous phase (containing the nucleic acids) were loaded onto a 0.8 % agarose gel. After the gel was run and stained it showed plasmid DNA, as well as chromosomal DNA and larger RNA fractions. Although the size of nonlinear DNA, i.e. plasmids, cannot be estimated putting many plasmids side by side on such gels reveals size differences, assuming that plasmids carrying the sequence of interest run higher than others. This way for each different plasmid species the corresponding culture was picked for deeper testing and verification of the clones.

2.2.4.7.7. Verification of positive clones

From every culture that has been selected by quick plasmid preparation (see above) plasmid DNA was isolated as described in section 2.2.4.4. The obtained plasmids were tested for the respective DNA insert of interest either by restriction analysis (see section 2.1.7.1.) or by PCR (see also section 2.2.4.3.).

2.2.4.7.8. Storage of verified clones

Plasmid DNA of verified clones was kept at -20° C. In addition to that 20 – 25 ml of selective LB were inoculated with the original culture and grown overnight at 37° C. From this culture 1 ml was transferred into a cryotube, mixed with 800 µl of sterile 50 % gylcerol and immediately frozen to -80° C.

2.2.4.7.9. Cloning strategies

pRo-7:

A preparative PCR was done from total ϕCh1 DNA using TR-31 and Ro-7 as primers. The obtained 1469 bp fragment (nu. 34689 – 36157) was digested with *Sma*I and *Xba*I. The vector pKSII+ - TR4/Ro6 - TR3/Ro7 was digested with these enzymes as well and the larger fragment of approx. 4800 bp, being the remaining vector, was extracted from a 0.8 % agarose gel. Both, the PCR fragment, as well as the vector were ligated and transformed into *E. coli*. Thereafter the resulting vector was isolated and digested with *Not*I/*Hind*III. The fragment of interest (approx. 3700 bp) was extracted from an agarose gel and cloned into plasmid pNov-1, which had previously been digested with *Not*I/*Hind*III.

pRo-8:

A preparative PCR was performed with primers TR-1 and Δ-53-1 using total ϕCh1 DNA as template. The resulting 1089 bp fragment was digested with *Not*I/*EcoR*I and cloned into the plasmid pKSII+ giving rise to vector pΔ53-1. The latter was digested with *Hind*III/*EcoR*I and ligated with the 2717 bp fragment obtained from PCR by using primers TR-4 and Δ53-2 that has been previously digested with the same enzymes. Upon transformation in *E. coli* the new vector p Δ53-2 could be isolated. Subsequently it was digested with *Not*I/*Hind*III and the corresponding fragment of approx. 3600 bp was isolated and cloned into vector pNov-1.

pRo-9:

A preparative PCR was performed with primers TR-1 and TR-4 using existing plasmid pRo-5 as template. The resulting fragment of approx. 3600 bp was digested with *Not*I/*Hind*III and cloned into vector pNov-1 which had previously been digested with *Not*I/*Hind*III as well.

pRo-10:

A preparative PCR was performed with primers TR-1 and TR-4 using existing plasmid pRo-7 as template. The resulting fragment of approx. 3600 bp was digested with *Not*I/*Hind*III and cloned into vector pNov-1 which had previously been digested with *Not*I/*Hind*III as well.

pRo-11:

A preparative PCR was performed with primers TR-1 and Δ-54-1 using total ϕCh1 DNA as template. The resulting 2254 bp fragment was digested with *Not*I/*EcoR*I and cloned into the plasmid pKSII+ giving rise to vector pΔ54-1. The latter was digested with *Hind*III/*EcoR*I and ligated with the 1554 bp fragment obtained from PCR by using primers TR-4 and Δ54-2 that has been previously digested with the same enzymes. Upon transformation in *E. coli* the new vector p Δ54-2 could be isolated. Subsequently it was digested with *Not*I/*Hind*III and the corresponding fragment of approx. 3600 bp was isolated and cloned into vector pNov-1.

pRo-5-ORF49Δ1-5:

Preparative PCRs were performed using total ϕCh1 DNA as template and 49-Kpn as a 5´ primer. In order to generate fragments of different lengths different 3´ primers were used (49D1, 49D2, 49D3, 49D4 and 49D5). The obtained fragments were digested with *Kpn*I and *Hind*III and cloned into plasmid pRo-5.

pRSETA-ORF49:

Though ORF49 had previously been cloned into pRSETA (Meissner, 2008) cloning was re-done for this thesis. A preparative PCR was performed with primers 12-7-3 and 12-7-5 using total ϕCh1 DNA as a template. The resulting 354 bp fragment was digested with *BamH*I/*Hind*III and cloned into the vector pRSETA.

pMal-c2X-ORF49:

A preparative PCR was performed with primers 49-blunt and 49-*Hind*III using total ϕCh1 DNA as a template. The resulting fragment was cloned in frame into a pMal-c2X vector that had been digested with *Hind*III and *Xmn*I.

pRSETA-ORF49Δ1 – Δ5:

Preparative PCRs were performed with 12-7-5 as a 5´ primer and different 3´ primers (49D1, 49D2, 39D3, 49D4 and 49D5) using total ϕCh1 DNA as a template. The resulting fragments were digested with *BamH*I/*Hind*III and cloned into the vector pRSETA.

pRSETA-ORF49ΔN:

A preparative PCR was performed with primers 12-7-3 and 12-7-5C using total ϕCh1 DNA as a template. The resulting fragment was digested with *BamH*I/*Hind*III and cloned into the vector pRSETA.

2.2.5. Protein methods

2.2.5.1. Overexpression of recombinant protein and screening of cultures for optimal time point for cell harvesting

For expression of heterologous recombinant protein in *E. coli*, the respective strain was transformed with the desired expression vector (see section 2.1.4.), plated on selective LB agar and incubated over night at 28° C or 37° C, respectively*. Several of the resulting colonies were used to inoculate 25 ml of LB selective media each. After overnight incubation at 28° C or 37° C 1 ml of each culture was stored as glycerol stock (as described in section 2.2.4.7.8.) for possible later use while the rest of the culture was used to inoculate 100 ml of fresh selective medium to an OD_{600} of 0.1. These cultures where grown at different temperatures and growth curves were monitored by measuring the OD_{600}. At an OD_{600} of 0.3 – 0.5 the overexpression of the desired protein was induced by adding a 1 M IPTG stock solution to a final concentration of 0.5 mM IPTG within the growth medium. To prepare crude extracts samples were taken immediately before induction and thereafter every hour for 4 – 5 h. In addition to that one sample was taken the following morning.

In order to prepare raw extracts, at a certain time point with a known OD_{600} 1.5 ml of the respective culture were transferred into an Eppendorf tube and centrifuged at 13.2 krpm for 3 min at room temperature. The pellet was resuspended in 75 x OD_{600} µl of 5 mM sodium phosphate buffer and mixed with the same volume of 2x protein sample buffer. Raw extracts were then boiled at 95° C for 10 min. Obtained samples were analyzed by SDS-PAGE in order to find the optimal time point, i.e. time point of maximal expression, for harvesting the cells.

*in order to reduce the formation of inclusion bodies some *E. coli* strains were grown at 28° C instead of 37° C.

2.2.5.2. Protein purification under native conditions

2.2.5.2.1. Purification of His-tagged proteins

1 l of selective LB medium was inoculated to an OD_{600} of 0.1 with an overnight culture of the respective *E. coli* strain carrying the desired expression vector. The culture was grown at 28° C or 37° C to an OD_{600} of 0.3 – 0.5, induced with 0.5 mM

IPTG and further grown until best time point for cell harvesting. Subsequently the cells were collected by centrifugation at 6 krpm for 15 min at 4° C and the pellet was resuspended in 50 ml of lysis buffer. As cycles of freezing and thawing enhance lysis the suspension was frozen to -20° C overnight. The following day the suspension was thawed and a spatula tip full of egg white lysozyme was added letting the mixture rotate at 4° C for 3 – 4 h. Afterwards the suspension was sonicated under constant cooling until cells were sufficiently lysed (normally it took 3 – 5 sonication rounds of 3 min each, letting the mixture cool on ice between each round). Cell lysis was observed and confirmed under the microscope. Cell debris and insoluble proteins were removed by centrifugation at 8 krpm for 20 min at 4° C. 400 – 500 µl of Nickel Agarose (Ni-NTA) slurry (obtained from QIAgen) were added to the supernatant and stirred overnight at 4° C in order to bind the His-tagged protein to the Ni^{2+} ions. The following day the suspension was loaded onto a designated column provided by the QIAexpressionist™ Kit from QIAgen. The flow-through was collected leaving the Ni-NTA with the bound protein as a thin layer on the column. This was washed twice with 4 ml of washing buffer subsequently followed by elution with 6 x 0.5 ml of elution buffer (elution of recombinant protein from Ni-NTA is achieved by raising the Imidazol concentration). Samples of all fractions were loaded onto an SDS-PAGE gel for analysis of the protein purification.

2.2.5.2.2. Purification of Maltose-tagged proteins

100 ml of selective LB medium (containing 0.2 % glucose) were inoculated to OD_{600} of 0.1 with an overnight culture of *E. coli* BL21 carrying the expression vector pMal-c2X-ORF49. Cells were grown to an OD_{600} of approx. 0.6, induced with 0.3 mM IPTG and further grown for 3 h at 37° C. Afterwards cells were collected by centrifugation at 7 krpm at 4° C for 15 min and the resulting pellet was resuspended in 30 ml of column buffer. The suspension was frozen to -20° C for overnight. Thereafter it was thawed and sonicated 2 x for 3 min each time. Remaining cells and cell debris were removed by a centrifugation step (10 krpm, 4° C, 15 min). The supernatant was transferred into an amylase resin obtained from New England BioLabs. After collection of the flow-through the resin was

washed 3 x with 1 ml of column buffer. Elution of the protein was achieved with 4 x 150 µl of column buffer containing 10 mM maltose. Success of purification was checked by SDS-PAGE analysis of all fractions.

2.2.5.3. Purification of His-tagged proteins under denaturing conditions

Cells for purification under denaturing conditions were prepared and harvested as described in section 2.2.5.2.1. However, the cell pellet was directly frozen to -20° C overnight. After thawing the pellet was resuspended in 50 ml of buffer B* and stirred at room temperature for at least 3 h or overnight. Subsequently the suspension was sonicated as described in section 2.2.5.2.1. After centrifugation at 8 krpm for 20 min at 4° C the supernatant was retained adding 400 – 500 µl of Ni-NTA slurry and stirring the lysate overnight at room temperature. Subsequently the suspension was loaded onto a designated column provided by the QIAexpressionist™ Kit from QIAgen. The flow-through was collected and the remaining Ni-NTA with the bound protein was washed twice, each time with 4 ml of buffer C*. Elution of the recombinant protein from the Ni-NTA slurry is achieved by lowering the pH. This was done by adding 6 x 0.5 ml of buffer E*. Samples of all fractions were loaded onto an SDS-PAGE gel for analysis of the protein purification.

*as buffer B, C and E contain urea their pH was checked and adjusted each time before use.

2.2.5.4. Purification of inclusion bodies from *E. coli*

Inclusion body purification was performed as previously described (Trachuk *et al.*, 2005): 1 l of fresh LB containing respective antibiotics was inoculated to an OD_{600} of 0.1 with an overnight culture of the respective *E. coli* strain. Subsequently the culture was grown to an OD_{600} of 0.3 – 0.5 at 28° C and induced with 0.5 mM IPTG and further grown at 28° C for 3 h. Cells were harvested by centrifugation at 6 krpm for 15 min at 4° C. The resulting cell pellet was resuspended in buffer A containing sucrose. Thereafter the suspension was sonicated as described in section 2.2.5.2.1. followed by centrifugation at 2500 g for 40 min. The pellet (containing inclusion bodies) was washed once with 40 ml of buffer A (without sucrose), twice

with 50 mM (pH 7.5) and one time with 40 ml of ddH_2O. Each washing step was interrupted by centrifugation at 10.000 g for 30 min. The resulting pellet was frozen to -20° C.

2.2.5.5. Protein renaturation

2.2.5.5.2. by dialysis

Purified proteins were transferred into dialysis tubes and dialyzed once against 1x GBB containing 4 M urea overnight. Subsequently a second dialysis step against 1x GBB once more overnight was carried out. The protein solution was then transferred into Eppendorf tubes and the solution was clarified of remaining insoluble protein by centrifugation at 13.2 krpm for 5 min. Samples of the protein solution were analyzed by SDS-PAGE as described in section 2.2.5.6.

2.2.5.5.3. by solubilization of inclusion bodies and rapid dilution

Purified inclusion bodies (section 2.2.5.4.) were solubilized in buffer B (section 2.1.9.5.3.) containing 8 M urea to a final concentration of 5 mg/ml. Another aliquot of inclusion bodies was solubilized in 6 M guanidine hydrochloride to a final concentration of 1 mg/ml. Both protein solutions were afterwards diluted 1:10, 1:20 and 1:100 into native elution buffer (section 2.1.9.5.2.) and the result was observed on an SDS-Page gel.

2.2.5.5.4. by size exclusion chromatography

Approx. 30 ml of sephacryl S200 HR (obtained from Pharmacia Fine Chemicals) were equilibrated in renaturation buffer and filled into a BioRad 25 ml column to a bed size of approx. 12 cm. 0.5 – 1 ml of protein solution obtained as described in section 2.2.5.3 were loaded onto the column. They were eluted under a constant flow of renaturation buffer and fractions of 0.5 ml were collected. Fractions were sampled for presence of the protein of interest by SDS-PAGE.

2.2.5.5.5. Gel extraction from polyacrylamide gels

Proteins obtained as described in section 2.2.5.3 were mixed with an equal volume of 2x protein sample buffer and loaded onto an SDS-PAGE gel. After separation the

respective bands were cut out of the gel and purified according to the protocol using the ElutaTube™ Protein Extraction kit from Fermentas. The success of the extraction procedure was checked by SDS-PAGE (see below).

2.2.5.6. SDS-PAGE

SDS-PAGE (sodiumdodecylsulfate polyacrylamide gel electrophoresis) allows for separation of proteins according to their size. Protein samples are denatured by heat in the presence of ß-mercaptoethanol (or di-thiotreitol, DTT) and SDS. As an anionic detergent SDS covers the native charges of the respective proteins rendering their running behavior in an electric field to be independent from their net charge. Proteins treated in such way migrate through a PAA gel with respect to their molecular weight. However, this is only completely true for (mesophilic) proteins with an average distribution of charges. As mentioned in section 1.1.4.2.1.1. haloalkaliphilic proteins contain a higher than average portion of acidic residues. These proteins migrate slower through the gel than a mesophilic protein of equal size.

2.2.5.6.1. Preparation of a discontinuous polyacrylamide gel

For casting the gels a vertical system from BioRad (Mini Protean 3) was used as instructed from the manufacturer. A discontinuous gel consists of a low PAA percentage stacking gel above the separation gel. The respective mixtures are listed below.

Reagent	***12 % Separation gel***	***4 % Stacking gel***
ddH_2O	1750 µl	1233 µl
Separation gel buffer	1250 µl	
Stacking gel buffer		500 µl
30 % PAA solution	2000 µl	267 µl
10 % APS	60 µl	20 µl
TEMED	10 µl	5 µl

First the separation gel was mixed without APS and TEMED and the mixture was placed on ice for cooling. Upon addition of APS and TEMED the gel started to polymerize. Then it was rapidly poured into the casting apparatus and overlaid

with a thin layer of 2-propanol (approx. 1 mm) in order to obtain a leveled surface. After ca. 5 – 15 min when the gel had polymerized the 2-propanol was poured away. Subsequently the stacking gel was mixed and cast overlaying the separation gel. Before polymerization a comb (10 or 20 wells) was placed into the gel. The finished gel was transferred into the provided tank that was filled with 1x SDS running buffer according to the manufacturer´s instructions.

2.2.5.6.2. Sample preparation and running conditions

Protein samples were mixed 1:1 with 2x protein sample buffer and heated to 95° C for 10 min. Then they were either frozen for later use or cooled down on ice and applied to a prepared SDS-PAGE gel. The gel was run by applying amperage of 15 mA/gel (two gels could be run simultaneously per tank). Migration of proteins could be estimated due to bromphenol-blue present in the sample buffer. Protein separation was stopped either when the bromphenol blue had left the gel (for bigger proteins > 20 kDa) or when the dye was about 1 – 1.5 cm away from the end of the gel (for smaller proteins < 20 kDa).

2.2.5.6.3. Staining of separated proteins

After the gel was run it was carefully removed from the apparatus and put into a bath containing Coomassie staining solution. After 15 – 30 min it was transferred into destaining solution for another 15 – 20 min or into H_2O for destaining over night.

2.2.5.7. Western blot

The Western blotting technique allows for detection of a specific protein by antibodies directed against this protein of interest. In many cases these antibodies are then detected with a second antibody that is specific for the first one. This second antibody can be labeled in different ways, e. g. radioactively or enzymatically and can be visualized according to the labeling method. However, in most of the Western blots performed in this work only a single antibody was used. Directed against a 6x Histidine tag (which had been fused to the protein of interest) it is also conjugated with the horseradish peroxidase (HRP). Thus proteins that are recognized by the antibody can be visualized without the necessity of a second antibody.

2.2.5.7.1. The blotting procedure

Proteins of interest were separated in an SDS-PAGE gel (see section 2.2.5.6.). Six layers of Whatman paper and one layer of Protran nitrocellulose membrane (obtained from Whatman) were cut into the size of the gel. Subsequently all layers were soaked in transblot buffer and the blot was set up the following way (bottom to top): three layers of Whatman paper, one layer of nitrocellulose membrane, the gel containing the protein(s) of interest, 3 layers of Whatman paper. The set up was transferred into a semi-dry blotting apparatus and power was applied at 20 V for 20 min. To confirm successful blotting and, in cases where no pre-stained protein marker was used, in order to mark the protein ladder, the membrane was stained with Ponceau S solution, which provides a method for unspecific, rapid, reversible protein staining. Thereafter the membrane was completely destained with H_2O and for blocking transferred into a bath containing 5 % milk powder in 1x TBS-T. Blocking was done at 4° C overnight while gently shaking.

2.2.5.7.2. Detection of the protein of interest

After blocking the membrane was put into 1x TBS-T for approx. 10 min. Thereafter the TBS-T was poured away and the membrane was incubated for 1 h at room temperature with 15 ml of the α-His HRP conjugated antibody solution (see also section 2.1.8.). Subsequently the antibody solution was removed (but could be kept for reuse) and the membrane was washed 3 x 10 min with 1x TBS-T. Antibodies bound to the proteins on the membrane were detected by the using the ECL kit from PIERCE according to the manual. Upon that the membrane was put into an exposure cassette. In the darkroom an X-ray hyper film (Amersham Biosciences) was placed over the membrane and exposed (the optimal exposure time was determined by trial and error). After development of the film, bands became visible.

2.2.6. Bandshift assays

Concentration of DNA to be shifted was determined using a NanoDrop ND-1000 from Peqlab. Protein concentration was determined by comparison with designated BSA standards of known concentration in an SDS-PAGE gel.

Molarity of DNA or protein concentration, respectively was calculated using the following formulas:

$$pmol\ DNA = \mu g\ DNA \times \frac{pmol}{660\ pg} \times \frac{10^6\ pg}{1\ \mu g} \times \frac{1}{N}$$

(N being the number of nucleotides)

$$pmol\ protein = \frac{\mu g\ protein}{protein\ size\ (kDa)} \times \frac{kg}{10^9\ \mu g} \times \frac{10^{12}\ pmole}{mol}$$

Bandshift reactions were set up in an Eppendorf tube as follows:

x µl protein solution
x µl DNA
x µl* 5x GBB (DTT freshly added)
2 µl BSA (1 mg/ml)
ad 20 µl ddH_2O

*Within the batch the dilution of GBB is 1x. The volume of 5x GBB to be added depends on the volume of protein added, since the protein is dissolved in 1x GBB buffer.

Reactions were incubated at 37° C for 1 h. Thereafter 5 µl of 5x Orange G loading dye were added and the total volume of the reaction (25 µl) was loaded onto an agarose gel (agarose molten in band shift running buffer). The gel was run in bandshift running buffer while applying relatively low voltage of approx. 5 V/cm. After running the gel was stained with ethidium bromide as previously described and bands were visualized under UV light.

2.2.7. Structure prediction of gp49

For computational structure prediction of gp49 the amino acid sequence (nu. 34480 – 34833 of the ϕCh1 genome) was sent to the PHYRE server for analysis (Kelley & Sternberg, 2009).

Disorder prediction of gp49 was carried out by sending the respective sequences to the web based application DISOPRED2 provided by the Bioinformatics Unit at the Department of Computer Science at the University College, London (Ward *et al.*, 2004).

3. Results & Discussion

3.1. Further analysis of the ϕCh1 replication region

The analysis of the ϕCh1 replication region culminated in the development of the pRo-5 shuttle vector system (see also section 1.2.2.5.). However, in order to obtain a more complete picture of the ϕCh1 replication further research had to be done, as some questions remained to be answered. In addition to that, there were some assumptions that had to be proven experimentally. This was done in the present work.

3.1.1. Introduction of +1 frameshift mutations into ORF53 and ORF54

As mentioned previously sequence analysis revealed high similarities of both ORF53 and ORF54 with the *repH* gene of *H. marismortui.* If cloned onto a pKSII+ vector together with a novobiocin resistance marker it is possible to transform this plasmid into *Nab. magadii.* Further experiments with respect to the ϕCh1 replication region had shown, that if the start-codons of the 3´ and 5´ flanking ORFs (ORF49 and ORF55, respectively) are deleted, the transformation efficiency of the resulting vector increases dramatically (see also section 1.2.2.5.).

However, as all shuttle vectors of the pRo – series that had been constructed so far (pRo-1 to pRo-6) always contained the fully functional *repH* homolog comprising both ORF53 *and* ORF54, the assumption that both ORFs are really necessary for replication had to be confirmed experimentally. Therefore frameshift mutations were introduced into ORF53 and ORF54 by PCR and the respective fragments were finally cloned onto the pNov-1 plasmid thus giving rise to two new shuttle vectors, pRo-8 and pRo-11, respectively. Both of which are identical to the functional pRo-5 vector except from carrying a +1 frameshift mutation either in the ORF53 (pRo-8) or in the ORF54 (pRo-11) in order to render the respective ORF inactive.

3.1.2. Deletion of up- and downstream palindromic sequences

The *repH* homolog of ϕCh1, comprising ORFs 53 and 54, is flanked by palindromic, AT-rich sequences. Previously it has been shown by Ng and DasSarma, that at least the 5´ AT-rich sequence is required for the minimal origin of replication in the *Hbt. salinarium* plasmid pNRC-100 (Ng & DasSarma, 1993). Thus the question arose whether this would also be true for the ϕCh1 origin of replication. Additionally, as there is also a 3´ AT-rich sequence it was only reasonable to take a closer look to it as well. In order to further elucidate the role of these two palindromic sequences three different constructs, i.e. shuttle vectors were created. One of which, pRo-7, does no longer contain the 5´ palindromic sequence whereas pRo-9 is deprived of the 3´ AT-rich region. In addition to that pRo-11 lacking both palindromic sequences was also constructed. All deletions were created by PCR.

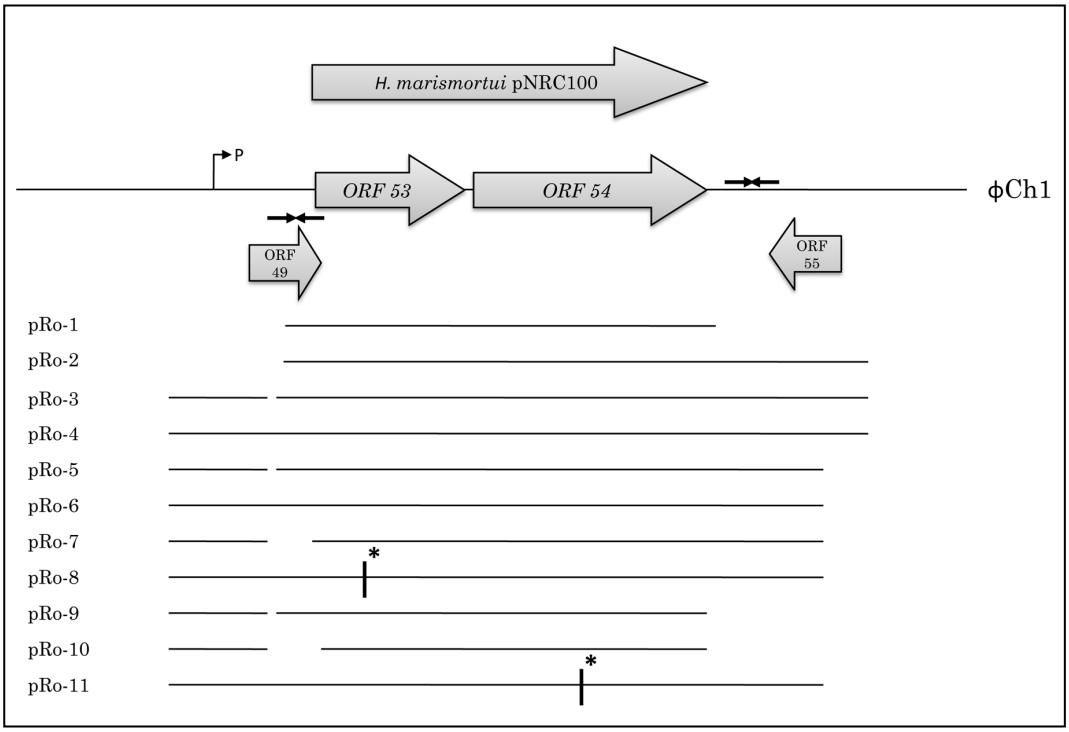

Figure 6| **The pRo cloning series.** The picture shows a schematic representation of the ϕCh1 origin of replication together with the corresponding homologous region of *H. marismortui* plasmid pNRC100 *repH*. The two palindromic sequences flanking ORF53/ORF54 are indicated by opposing arrows. To study this region in ϕCh1 different fragments of it were cloned into a pKSII+ vector and transformation efficiency of *Nab. magadii* transformed with the respective resulting vectors was analyzed. This led to the development of the pRo-5 plasmid by Iro *et al.*, a fully operational *Nab. magadii* shuttle vector. It lacks the starting codons of two ORFs (49 and 55) involved in gene regulation of ϕCh1 (Iro *et al.*, in prep.). In this thesis the derivates pRo-7 – pRo-11 were developed in order to further characterize the origin of replication. pRo-7 and pRo-9 are lacking the 5´ and 3´ palindromic sequences, respectively, while pRo-10 is lacking both at the same time. pRo-8 carries a +1 frameshift mutation (indicated by asterisks) in ORF53 whereas the same is true for ORF54 in pRo-11.

3.1.3. Transformation of pRo-5 derivates into *Nab. magadii*

All of the constructed vectors (pRo-7 to pRo-11) were transformed into *Nab. magadii* L13. As a positive control, pRo-5 was used. Thus (i) it was confirmed, that *Nab. magadii* cells were competent at the timepoint of transformation and (ii) the "normal" transformation efficiency could be determined and correlated with that of the other constructs. As a negative control *Nab. magadii* L13 competent cells not transformed with a vector (but with an equivalent volume of spheroblast solution) were taken. In order to achieve comprehensive comparability between all experiments the used amount of vector DNA was always 3 µg in total. After incubation colony forming units (cfu/ml) were determined and correlated to each other. In order to obtain statistically relevant data each experiment was done in triplicates for each batch. Additionally each batch was plated on three agar plates and the total experiment was repeated at least three times (only statistically relevant data sets were used for calculation of cfu/ml).

3.1.4. Discussion

Concerning the transformation efficacy the inactivation of ORF53 and ORF54, respectively, shows an unambiguous result. Cultures transformed with shuttle vectors carrying one of the two frameshift mutations (pRo-8 and pRo11) are not able to grow on novobiocin suggesting that the vector can no longer be replicated within *Nab. magadii*. These facts lead to the conclusion, that both ORF53 as well as ORF54 are vital for plasmid replication and in the broader sense crucial for ϕCh1 propagation. Furthermore this is in accordance with previous findings of Ng *et al.*. identifying *repH* of plasmid pNRC100 as crucial element for a minimal origin of replication in *Hbt. salinarium* (Ng & DasSarma, 1993). However, these findings are not unexpected given that the virus ϕH infecting *Hbt. salinarium* has been shown to exhibit a similar gene arrangement of its replication origin as ϕCh1 (Klein *et al.*, 2002) and is the closest relative to ϕCh1, sharing up to 97 % sequence identities (Klein *et al.*, 2002) over large parts of their genomes. As a conclusion one can say, that in the case of ϕCh1 infecting *Nab. magadii* two open reading frames mediate RepH activity. Yet this statement immediately leads to the question why there are two separate ORFs involved. As both ORFs eventually give rise to two distinct proteins, it seems possible, that they act as a heterodimer. On the other

hand both proteins could perform distinct actions. Effort has been taken to identify a possible binding site of either gp53 or gp54 on the ϕCh1 genome. Here the two flanking palindromic sequences upstream of ORF53 and downstream of ORF54 seem to be the most likely sites. Such interactions have been shown to occur in other systems, where binding of certain proteins to the DNA results in a melted replication complex, additionally supported by AT-rich regions (Schnos *et al.*, 1988; Kornberg & Baker, 1992; Ng & DasSarma, 1993). However, all attempts confirming this for ϕCh1 gp53/gp54 have so far not been crowned with success which is probably less due to the fact that there is no interaction rather than to the difficulties that accompany the work with halophilic proteins.

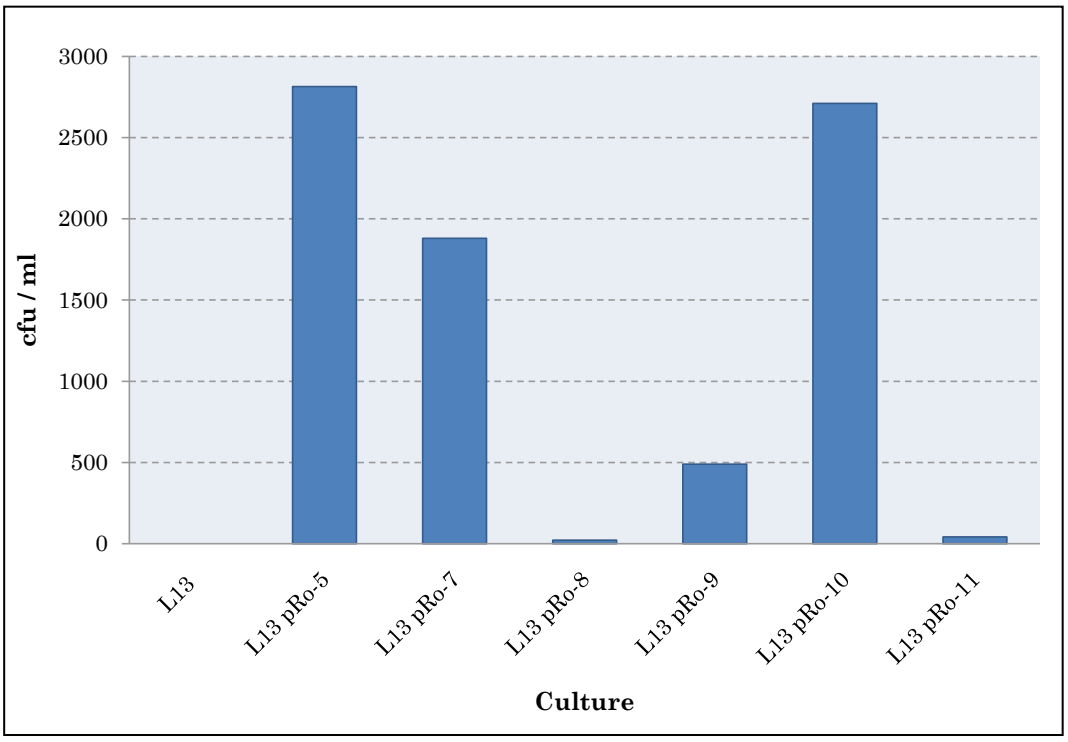

Figure 7| **Transformation efficiency of pRo derivates.** *Nab. magadii* L13 was transformed with different constructs of the pRo cloning series, plated on selective agar and colony forming units (cfu) per ml were determined. All pRo -plasmids contain fragments of the ϕCh1 origin of replication (see also Figure 6). The pRo-5 vector is used as the standard shuttle vector in the lab of A. Witte and serves as a control, along with untransformed L13 cells. When a +1 frameshift mutation is introduced into either ORF53 (pRo-8) or ORF54 (pRo-11) nearly no viable colonies are observed. This confirms the assumption that in ϕCh1 two ORFs rather than one mediate RepH activity. If the 5´ palindromic, AT-rich sequence flanking ORF53 is removed, the transformation efficacy decreases (pRo-7) by ~33 %. In the case of pRo-9, where the 3´ palindromic sequence flanking ORF54 has been removed the transformation efficacy is reduced by ~82 %, suggesting that the 3´ AT-rich sequences are more important. Surprisingly, if both the 5´ as well as the 3´ sequences are removed, as is the case for pRo-10 the transformation efficacy is nearly restored.

The analysis of the AT-rich palindromic sequences themselves delivered somewhat surprising results. If deprived of the 5´ flanking region of ORF53 the shuttle vector shows a significant reduction of the transformation efficiency. However, the degree of this decrease is less dramatic than expected since Ng and DasSarma described a total loss of replication ability when they deleted the 5´ AT-rich sequence of *repH* (Ng & DasSarma, 1993). Nonetheless, the deletion of the 3´ palindromic sequence of ORF54 gave results comparable to that of Ng and DasSarma (Ng & DasSarma, 1993). Although no total abolishment of replication ability could be observed the transformation efficacy is dramatically reduced. These findings suggest the 3´ AT-rich sequence rather than the 5´ sequence plays a key role in ϕCh1 replication. These results however are most surprisingly and completely unexpectedly contradicted as the deletion of both palindromic sequences nearly restores the wild type. For pRo-10 one would expect at least a transformation efficiency in the range of pRo-9 rather than pRo-5. This result questions the importance of both AT-rich regions in ϕCh1 as a whole but would on the other hand provide an admittedly sparse explanation for the fact, that so far no interaction of gp53 or gp54, respectively, with any of these palindromic sequences could be found. Thus the relevancy of these regions could have been simply overestimated. Yet the deletion of only one of these sequences has an undeniable effect on the transformation efficiency so that it seems absurd to assume that they are of no significant importance. In addition to that in the pRo-5 vector the ϕCh1 *ori* has been reduced to a minimum that would give an operational origin of replication. In other words there is not much left in all the shuttle vectors that gp53 and gp54 could interact with.

3.2. Characterization of ORF49

3.2.1. Gp49 acts as ϕCh1 repressor

In previous studies the gene product of ORF49 of ϕCh1 has been suggested to be involved in the switch between the lysogenic and the lytic life cycle of the virus (Iro *et al.*, 2007). The present work contributed in the further elucidation of the gp49 activity. Prior to this work the ORF49 of ϕCh1 including its promoter has been cloned onto the pRo-5 shuttle vector (Meissner, 2008). Furthermore in addition to the pRo-5 shuttle vector the plasmids pRo-4 and pRo-6 were also available (Ladurner, 2008). Both plasmids carry a fully operational copy of the ORF49. However in comparison to pRo-4 the pRo-6 vector has the starting AUG of ORF55 deleted. The latter modification is vital for a high transformation efficiency of the pRo vectors in *Nab. magadii* (refer also to section 1.2.2.5.). In order to investigate the effect of gp49 on the infectivity of ϕCh1 a phage titer analysis was performed. *Nab. magadii* L13 cells transformed with the vectors pRo-4, pRo-5, pRo-6 and pRo-5-ORF49, respectively were infected with previously isolated virus particles and the titer was determined by the soft plating technique. An untransformed *Nab. magadii* L13 strain served as a control and for determination of the titer of the stock. The experiment was done in duplicates and after approx. 2 – 3 weeks the plaques were counted. The results are depicted in Figure 8 and show a dramatic effect of ORF49 on the infectivity of ϕCh1. In comparison to the normal titer for cells carrying pRo-5-ORF49 plaque formation is reduced by a factor of approx. 1.1×10^4. Cells that had been transformed with pRo-5 show no significant deviations from the normal titer since in the pRo-5 vector ORF49 is deleted for its start codon. On the other hand, pRo-6, carrying a fully operational ORF49, shows a decrease in phage infectivity which is as dramatic as it is for pRo-5-ORF49. However, cells that were transformed with pRo-4 behaved like normal *Nab. magadii* L13 cells. This was a little bit surprising as pRo-4 also contains a fully functional ORF49. Consequently ORF55, which is intact on pRo-4 but not on pRo-6, must be responsible for the restoration of the normal titer. One possible explanation for this is that because pRo-4 carries the fully operational ORF55 the ability for replication of the vector itself is reduced and thus gp49 is simply diluted over the time. Indeed, former studies have shown that the transformation

efficiency of pRo-4 is much lower than that of pRo-5 (Ladurner, 2008). Nevertheless, this effect might be too weak to solely account for the restoration of the normal titer. As a matter of fact ORF55 has been shown to share sequence similarities to known regulators (Klein *et al.*, 2002; Iro *et al.*, in prep.) so that it is likely to play a more active role in the regulation of the phage life cycle. This assumption is supported by the present work. However in order to fully elucidate the function of ORF55 further studies on this subject will have to be carried out.

Nonetheless, the above described experiment made it obvious that gp49 has a strong repressor activity on its own. This is also in accordance to the findings of Iro *et al.* (Iro *et al.*, 2007) where a regulatory effect of the known repressor ORF48 (*rep*) on ORF49 was shown. If expressed extra-chromosomally from a plasmid gp49 severely reduces lysis of ϕCh1 infected *Nab. magadii* L13 cells.

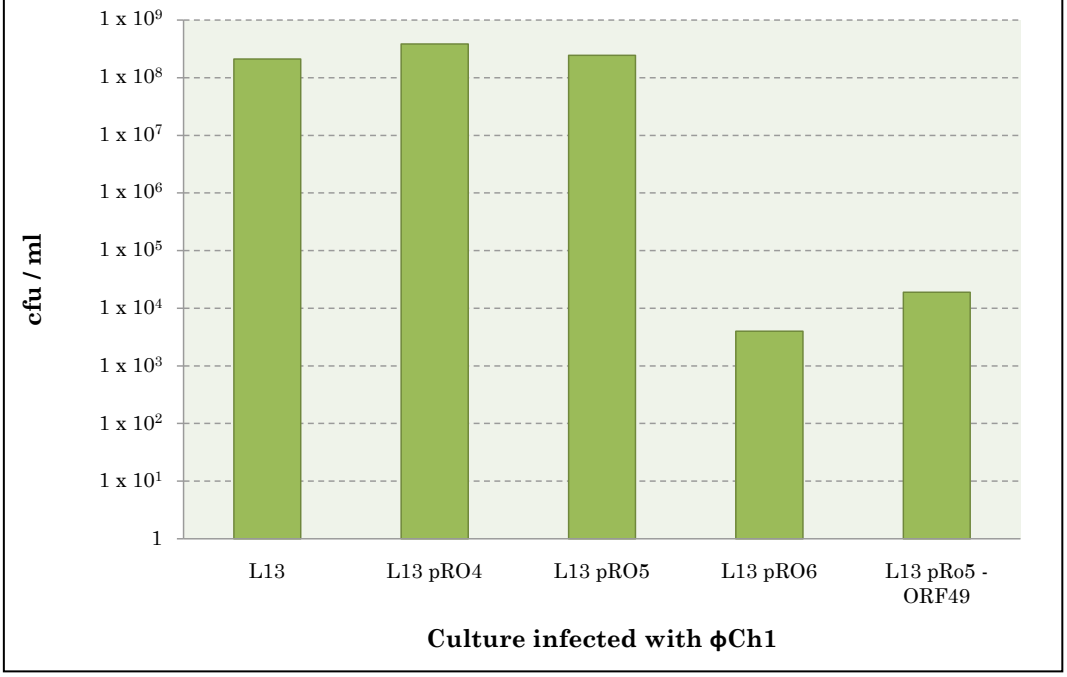

Figure 8| **Repressor Activity of gp49.** Phage titer analysis with different strains of L13 that have been transformed with certain vectors of the pRo series (see also Fig. 6). The normal titer is represented by L13 infected with ϕCh1. The pRo-5 shuttle vector, where gp49 has been deleted for its starting AUG, shows no differences to the normal titer. On the other hand pRo-6, carrying the same sequence as pRo-5 but in addition also an intact gp49, shows a dramatic reduction of the titer. If cloned into a pRo-5 vector ORF49 dramatically reduces the pfu/ml, suggesting that gp49 is able to solely account for repression of the phage. However, pRo-4 shows a surprising result, as it carries a fully operational ORF49 but seems not to repress phage activity at all. One explanation might be the additional ORF55 in comparison to pRo-5. ORF55 is thought to be involved in ϕCh1 regulation.

3.2.2. Putative 3D structure of gp49 – an attempt in silico

The "protein homology/analogy recognition engine" (PHYRE) provides a powerful tool for protein structure prediction (Kelley & Sternberg, 2009). Besides the common iterative PSI-Blast search for homologies one of its features comprises the comparison of the amino acid sequence of the protein of interest to known protein structures available in databases. Thus the 3D structure prediction gains on accuracy. This was also done with the sequence of ORF49 (nu. 34480 – 34833 of the ϕCh1 genome). The output of the PHYRE query gave a couple of different models for the predicted 3D structure of gp49. Disappointingly, yet not unexpectedly, the estimated precision of the calculated models is very low. This is probably due to the fact that present databases contain relatively few archaeal protein structures in general. Thus the highest score with an estimated precision of 5 % showed some homologies to a eukaryotic enzyme – the monoclinic mimivirus capping enzyme triphosphatase (Benarroch *et al.*, 2008). Nonetheless, it should be mentioned that among all calculated 3D structures one prediction is reminiscent to a protein named *Tat*, which is involved in transcription regulation (Sticht *et al.*, 1994). The results of the secondary structure prediction of gp49 by PHYRE are depicted in Figure 9 together with some 3D structure predictions. However, it is obvious that these predictions are very unlikely to match the reality. As a matter of fact there are apparently too few (archaeal) protein structures available in the current databases to ensure a reliable structure prediction for gp49.

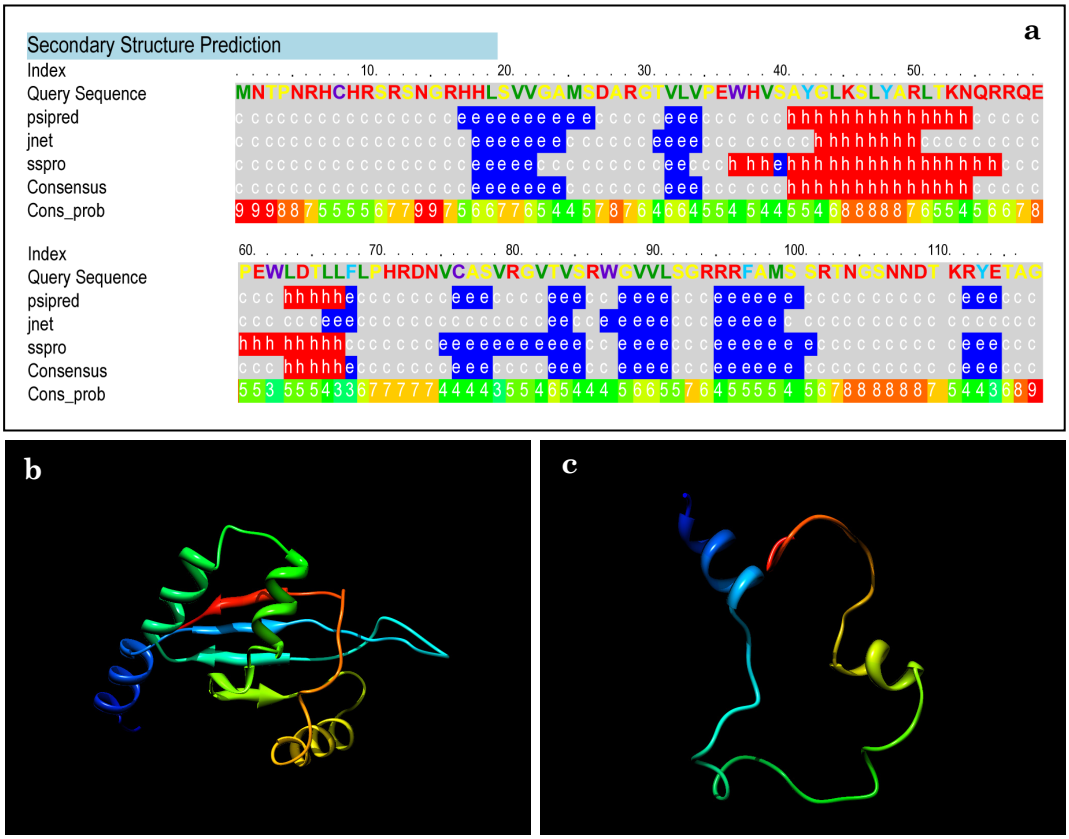

Figure 9| **Structure predictions of gp49. a**| Secondary structure predicted by PHYRE. The "Cons_prob" row illustrates the confidence of the consensus prediction, that is used for tertiary structure prediction (0=low confidence; 9=high confidence). The letter "e" indicates amino acids predicted to fold into ß-sheets whereas "h" stands for amino acids that fold into a helix. The overall confidence of structure prediction is relatively low since too few archaeal protein structures are currently available in public databases. **b**| The tertiary structure prediction with the highest estimated precision of 5 % is shown here. It is reminiscent of the monoclinic mimivirus capping enzyme triphosphatase (Benarroch *et al.*, 2008) with a sequence identity of 19 %. **c**| At least one of the structures predicted by PHYRE shows similarities to a protein named *Tat* which is involved in transcriptional regulation (Sticht *et al.*, 1994).

(Structure predictions were done with PHYRE (Kelley & Sternberg, 2009) and molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) (Pettersen *et al.*, 2004).

In addition to the 3D structure a prediction for disordered residues was also carried out. This was done with the DISOPRED2 server according to Ward *et al.* (Ward *et al.*, 2004). Disordered regions do not per se have a native conformation into which they fold rather than being intrinsically unstructured. This disorder can be local or global and for the former it has been shown that such regions can be biologically active (Wright & Dyson, 1999). The results are depicted in Figure 10 and show predicted disordered regions at the N-terminal, as well as at the C-terminal end of gp49.

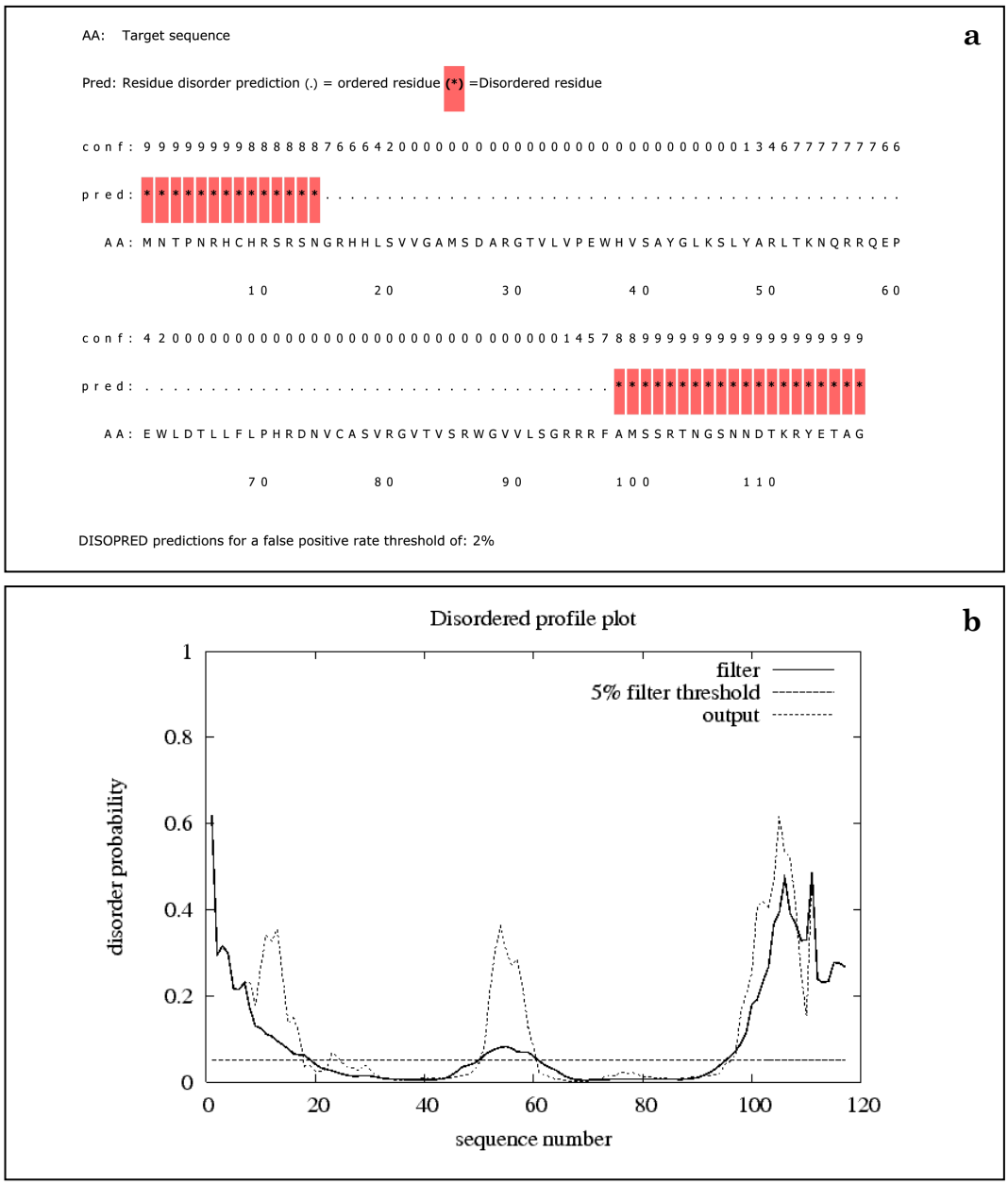

Figure 10| **Intrinsic disorder prediction for gp49. a**| Prediction of disordered residues within the primary sequence of gp49. Numbers in the upper line indicate the confidence of the computer prediction (9 = high confidence). The N-terminal as well as the C-terminal end of gp49 probably contain disordered regions. Such regions are likely to exist in an unfolded state within the cell. **b**| Schematic representation of the disorder probability of amino acid residues within gp49.

(Predictions were done with the DISOPRED2 server according to Ward *et al.* (Ward *et al.*, 2004))

3.2.3. Overexpression of ORF49 in *E. coli* sp.

In order to purify gp49 in sufficient amounts it had to be overexpressed heterologous in *E. coli*. Therefore ORF49 was cloned onto the pRSETA plasmid, an expression vector commercially available from Invitrogen that carries the T7 promoter to ensure strong expression of the protein of interest. This was done for gp49 in order to make the purification of the protein as well as its identification more convenient.

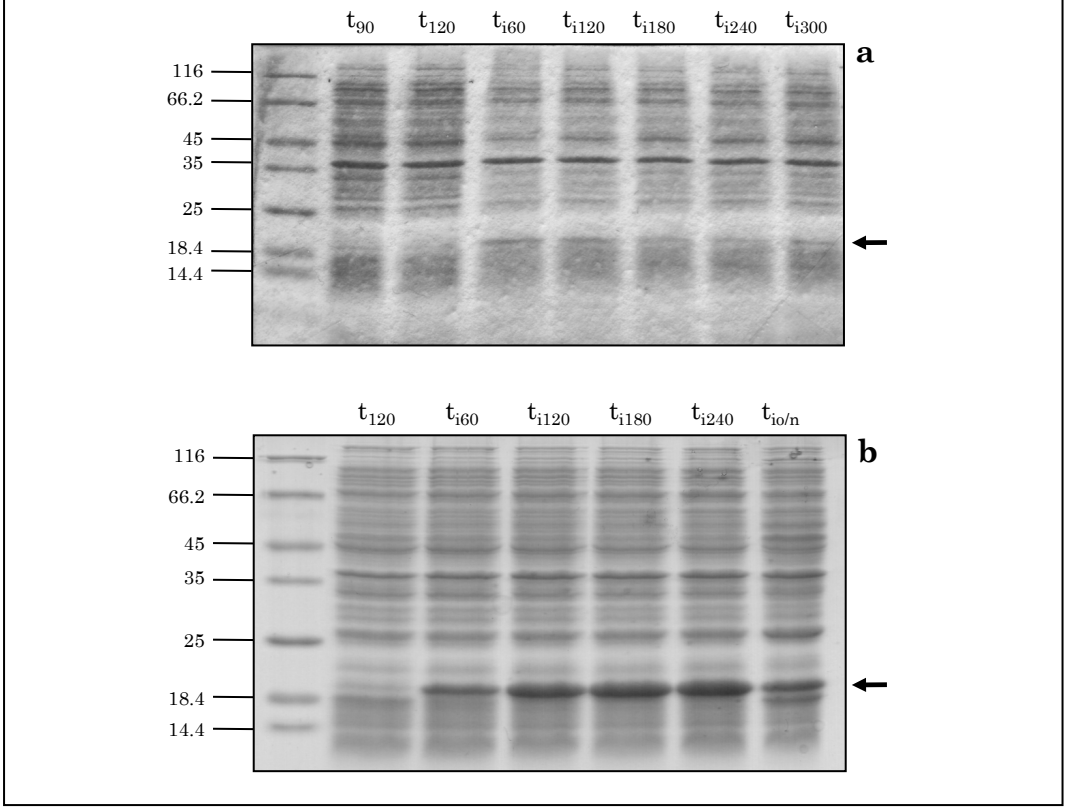

Figure 11| **Overexpression of gp49.** Crude extracts taken from two different *E. coli* strains that had been transformed with pRSETA-ORF49. Extracts where taken at different time points before and after induction with 0.5 mM IPTG (t_{min}= minutes after inoculation (before induction); t_{imin}= minutes after induction). Arrows indicate the expected height for gp49. **a**|*E. coli* Rosetta shows, if any, a minimal expression of gp49 60 min after induction. The expression level stays nearly the same over the observed amount of time**. b**| In contrast to that *E. coli* Bl21-pLysE shows an increasing expression of gp49 with a maximum 4 h after induction. However, upon an expression over night the protein is degraded.

The vector was then transformed into different *E. coli* strains (Tuner, Rosetta, BL21, C41 and C43) to find the most suitable strain for expression. Surprisingly, sufficient expression of gp49 could only be observed in *E. coli* BL21 pLysE but not in Rosetta, though both strains are extremely closely related. The best time point for harvesting expressing cells was determined to be 3 – 5 h after induction with

IPTG. If expressed overnight gp49 is no longer detectable in a PAA gel probably due to degradation. Expression of gp49 was confirmed by Western blot analysis using α-His antibodies. The results of the overexpression are shown in Figure 11.

In order to find a way to purify gp49 natively ORF49 was cloned additionally into the pMal-c2X vector which fuses a maltose tag to the protein of interest (see also section 3.2.4.). This vector was cloned into *E. coli* BL21 and overexpression of the fusion protein was monitored by SDS-Page. Again the best time point for harvesting cells was 3 – 5 h after induction with IPTG. Yet again as described above upon induction overnight the protein was no longer detectable in a PAA gel.

3.2.4. Purification of gp49

In order to use gp49 for bandshift assays it had to be purified from *E. coli* first. For such an assay to work it is further necessary that the purified protein retains its native conformation and folding. Therefore purification under native conditions is desirable. A native purification was attempted in the first place with gp49 that had been cloned into pRSETA (see also section 2.2.5.2.) but could never be achieved and did not show any detectable results although overexpression of the protein gave massive amounts of protein seen in an SDS-Page gel when *E. coli* BL21 crude extracts were analyzed (see Figure 11b and 12a). This suggests that gp49 when expressed from a T7 promoter and carrying an N-terminal 6x His tag is either insoluble or aggregates in insoluble inclusion bodies within the host cell. This was confirmed when a direct purification of the inclusion bodies that was performed resulted in massive amounts of purified gp49. In order to reduce inclusion body formation different growth temperatures were tested as well. As lower temperatures result in slower growth of *E. coli* and should hence reduce formation of inclusion bodies it was attempted to grow the cultures at 28° C as well as at 16° C. These attempts were unsuccessful though, as no soluble protein could be detected. In a last attempt to isolate gp49 natively it was purified from *E. coli* Bl21 that has been transformed with the pMal-c2X-ORF49 vector carrying a maltose tag. Using this plasmid purification of the fusion protein under the previously described native conditions was possible. However, the obtained protein fractions seemed to be extremely impure, unsuitable for bandshifting experiments.

For all of the above described reasons native purification of gp49 was abandoned and exchanged by purification under denaturing conditions with 8 M urea using the vector pRSETA.

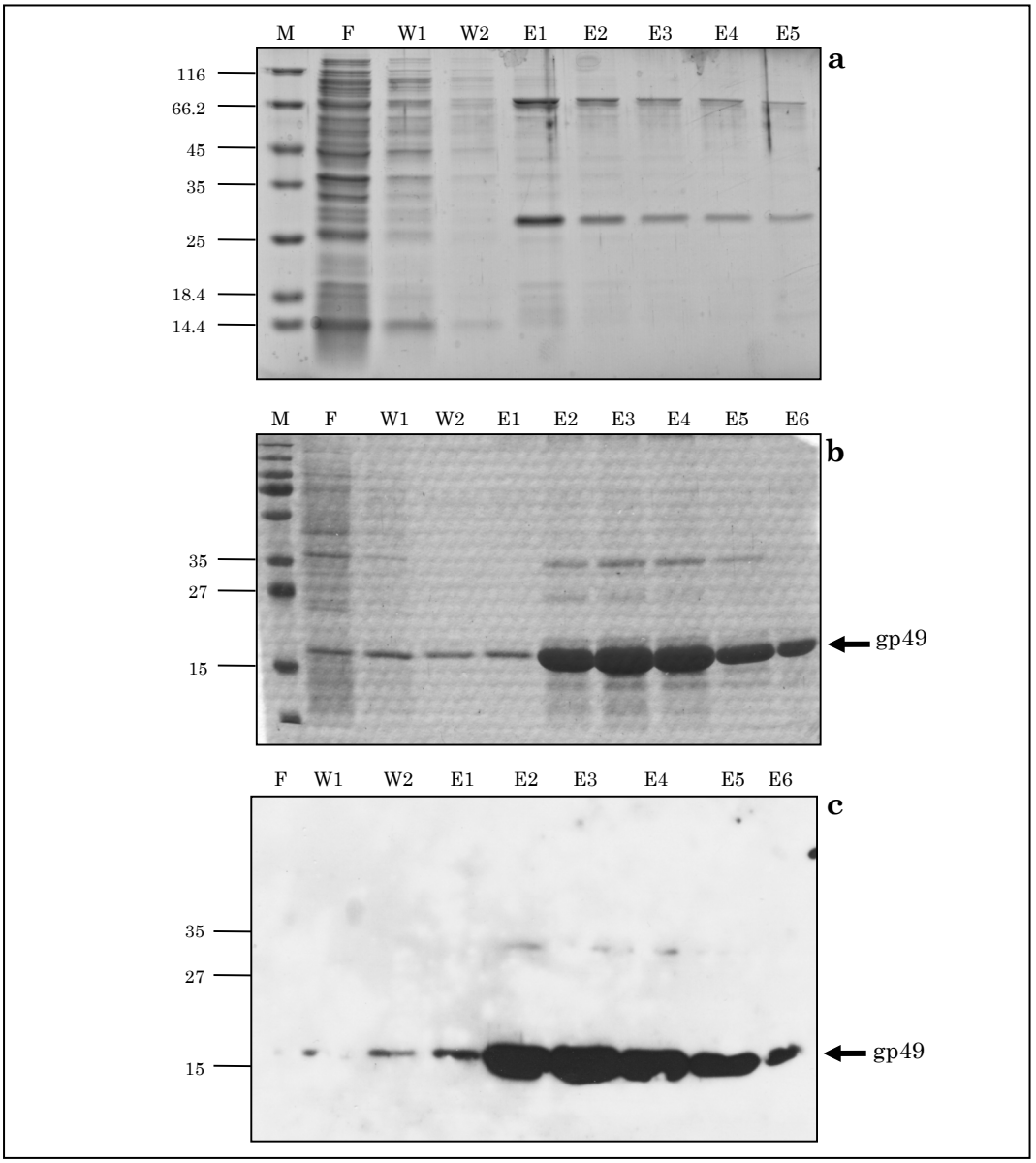

Figure 12| **Purification of gp49.** Capital letters indicate the different fractions of affinity chromatography (F=Flowthrough, W=Wash, E=Elution) **a**| Native purification using Ni-NTA agarose to isolate His-tagged gp49 failed. No bands on the correct height of around 18 kDa are visible in a coomassie stained PAA gel. **b**| Denaturing purification of His-tagged gp49 using 8 M urea reveals massive amounts of gp49 visible in the PAA gel. The fact that denaturing purification worked that well while native attempts remained unsuccessful suggests that gp49 aggregates into insoluble inclusion bodies within the cells. **c**| A Western blot using α-His antibodies confirms that the purified protein is indeed gp49.

In contrast to the native purification this worked extraordinary well as Figure 12b depicts. Yet it raised another problem as denatured proteins are of no use in a bandshift assay. Therefore several methods for renaturing the protein were tested. (i) Dialysis provides an easy and convenient way to get rid of denaturing agents but

although at least two dialysis steps were performed (see also section 2.2.5.5.2.) gp49 always precipitated when the concentration of urea was reduced to less than 4 M. Thus it seemed that dialysis was unsuitable for removing the denaturing agents. (ii) A promising method for renaturing the protein was to cut it out from a PAA gel. This additionally provided another purification step as exactly the band of interest can be used. This was done with a kit obtained from Fermentas and the protocol included an electrophoresis step combined with a dialysis step followed by precipitation of the protein with trichloroacetic acid (TCA). However, though this method was repeated several times isolation of gp49 worked only once. For this reason and for the fact that an additional denaturing step (TCA precipitation) is involved the method of eluting gp49 from a PAA gel was abandoned as well. (iii) In order to test another simple method for renaturation – rapid dilution – the inclusion bodies containing gp49 were directly purified from *E. coli*. Thereafter one batch was solubilized using 8 M urea and another one using 6 M guanidine hydrochloride (GuHCl). Thus protein solutions with a concentration of 5 mg/ml and 1 mg/ml respectively were created. These were simply diluted into a relatively large volume of buffer containing no urea or GuHCl. However, this did not work as the volume of native buffer in which the protein no longer precipitated was too large to give a reasonable concentration of gp49 suitable for bandshifting.(iv) Finally, renaturation using size exclusion chromatography, i.e. gel filtration, gave a satisfactory yield of gp49 solubilized in a buffer containing no denaturing agents.

3.2.5. Bandshifit assays (EMSA)

In order to identify the binding region of gp49 within the ϕCh1 DNA a series of electrophoretic mobility shift assays (EMSAs), i.e. bandshift assays, were performed. In general this method is based on the altered migration behavior of DNA bound to protein in comparsion to free DNA in an agarose gel.

3.2.5.1. Bandshifting of total ϕCh1 DNA and fragments

In a first attempt viral DNA from ϕCh1 was isolated and digested with *Bgl*II giving 24 fragments including the non digested, methylated fraction. Thereafter the digested DNA was incubated together with rising concentrations of purified and renatured gp49. Subsequently the batches were loaded onto an agarose gel. The

results are shown in Figure 13. Most obvious is that upon addition of very high concentrations of gp49 the migration of the larger fragments is drastically slowed down leading to distortions of the respective lanes. Nonetheless at least two shifts are visible on the gel. One affects a fragment that ranges around 3.6 kb (the $F_{1,2}$ fragment) and another one occurs in a fragment that is out of the range of the marker at approx. 11 kb (the A fragment). Unfortunately, although the experiment was repeated several times, these results could not be reproduced.

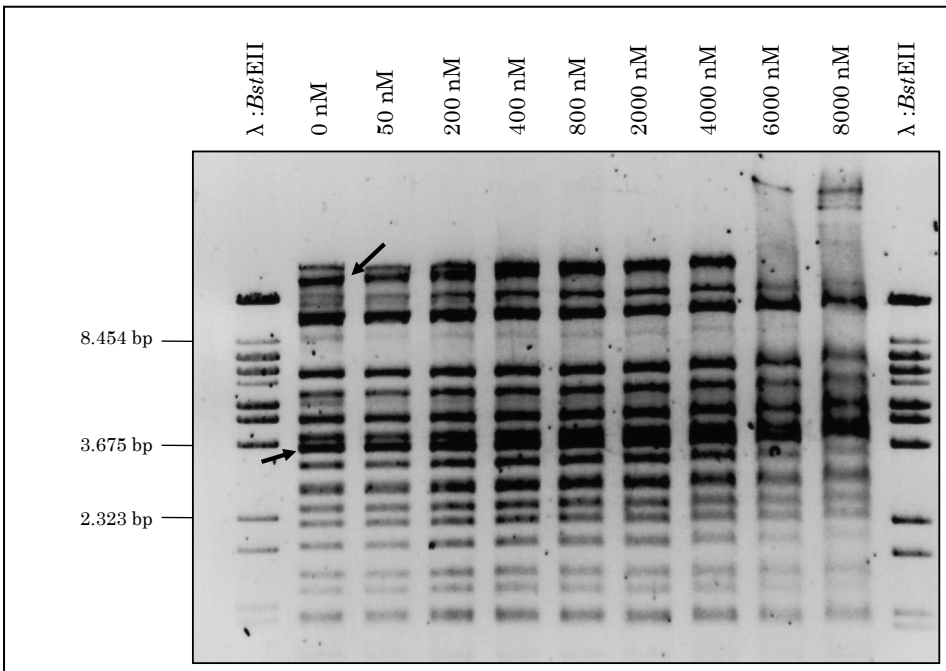

Figure 13| **Electrophoretic mobility shift assay (EMSA) using total ϕCh1 DNA.** The DNA was isolated from the virus and digested with *Bgl*II giving 24 fragments. Thereafter DNA was incubated with different concentrations of purified gp49 that had been renatured by gel filtration. A λ:*Bst*EII digest serves as a marker. The highest concentrations of gp49 result in massive distortions of the upper part of the respective lanes. At least two different shifts are visible. One of these shifts ranges around 11 kb, the other at around 3.6 kb (indicated by arrows).

Subsequent to this experiment other ϕCh1 DNA fragments were directly tested. These had been previously cloned into several vectors in the lab of A. Witte by R. Klein. The respective fragments were cut out of their vectors and purified. In addition to that two fragments were generated by PCR, the F1 fragment as well as the p17-5/21-Hind fragment (according to the primers used for its generation). All these fragments were incubated together with renatured gp49 to see whether they would shift or not. The result is depicted in Figure 14a, b. Disappointingly, there is no convincing shift visible on the gel. Yet, the I/9 fragment as well as the F1 digest and the p17-5/21-Hind digest show, though little, shifts. In contrast to that the

control, which contains linear pKSII+ DNA does not show any shift. In addition to that the BSA control does also show no shifts of any of the fragments.

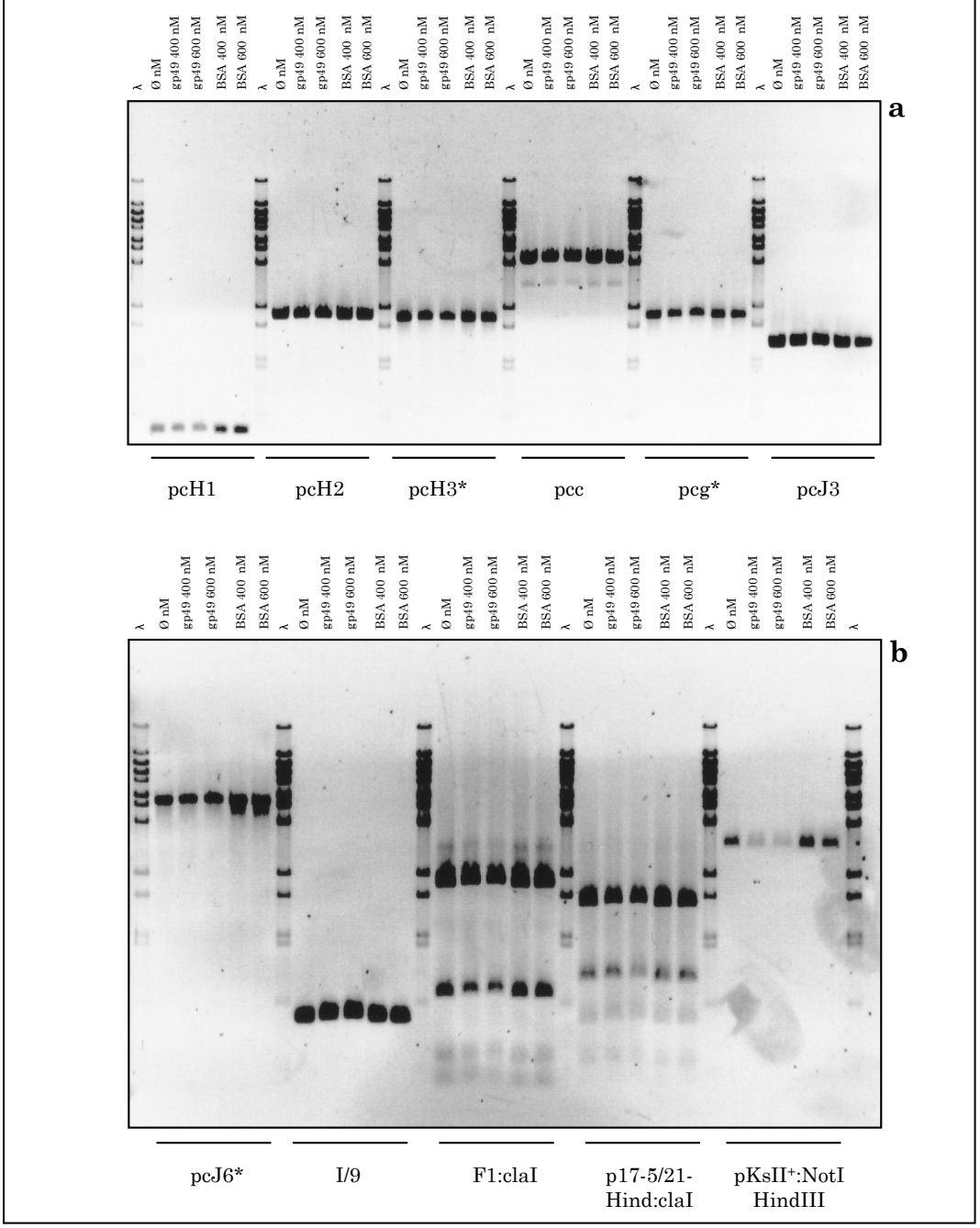

Figure 14| **Bandshift assay with selected ϕCh1 DNA fragments.** The ϕCh1 fragments have been previously cloned by R. Klein (Klein, 1999). Plasmids were isolated from *E. coli* XL1 – Blue (asterisks mark fragments isolated from *E. coli* GM48F´) and fragments were cut out of their respective vectors. EMSAs were performed using 400/600 nM of gp49 that have been purified and renatured by gel filtration. The same concentrations of BSA were used as a control. A λ:*Bst*EII digest serves as a marker. **a**| No convincing shifts are visible with these fragments. **b**| Shifts can be observed with the fragments I/9, F1 and p17-5/21-Hind. Here both concentrations of gp49 seem to result in a small shift of the respective fragments, while the BSA controls do not shift. Linearized pKSII+ DNA does also show no shift at all.

3.2.5.2. Bandshifting selected ϕCh1 ORFs and DNA fragments

As the first bandshift assays gave only moderate results strategies were changed. A number of selected ϕCh1 ORFs suspected to be involved in the regulation of the phages life cycle were directly tested. The second change in the experimental procedure was to no longer try to fully renature gp49. The protein was dialyzed only to 4 M urea after it has been purified under denaturing conditions (using 8 M urea). Under these conditions it remained soluble. This alteration made the whole assay simpler by reducing the steps of renaturation.

The first ORFs to be tested were ORF43 and ORF44, since these ORFs were shown to be involved in the gene regulation of ϕCh1 as enhancing elements for the ORF48 – ORF49 intergenic region that contains at least two promoter elements (Iro *et al.*, 2007). Surprisingly, the very first experiment of this series, in which the PCR generated ORF43 DNA was used already showed shifts and is depicted in Figure 15a. However, unfortunately the control DNA, which is the CRISPR sequence from *Strp. pyogenes* also showed a comparable shift. On the other hand shifting ORF43 with BSA solubilized in 4 M did not work under the same conditions (Figure 15b). In another experiment the influence of methylation on binding was studied when the PCR generated ORF43 was compared to the ORF43/44* fragment that had been cloned into a pKSII+ vector prior to this work (Iro, 2006). This way it could be shown, that methylation does not affect binding of g49 (Figure 15c,d).

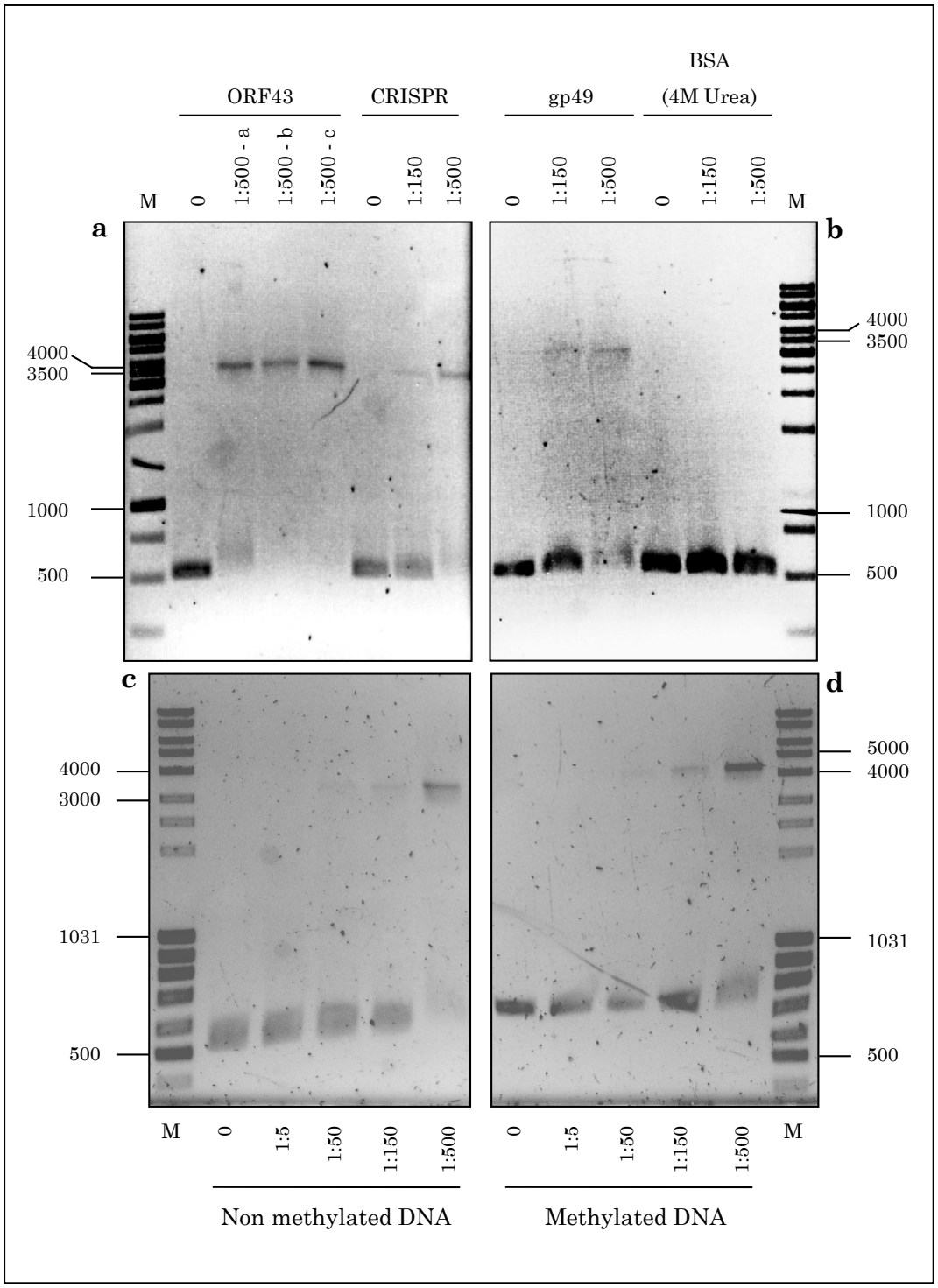

Figure 15| **EMSA – testing gp49 on ORF43.** Gp49 was purified under denaturing conditions, dialyzed to 4 M urea and used to shift the DNA of ϕCh1 ORF43. Numbers indicate the molar ratio of DNA to protein. **a**|Three different isolates of gp49 (a,b,c) where initially tested on ORF43 DNA that had been generated by PCR. The CRISPR sequence of *Strp. pyogenes* was used as a control. Both sequences have a comparable length of approx. 500 bp and seem to be shifted in the presence of gp49. Shifts appear at a height of approx. 3.7 kbp. **b**| Control proving that gp49 is responsible for the shifts. Equal concentrations of BSA dissolved in 4 M urea do not induce a shift in ORF43 DNA. **c, d**| Methylation of DNA does not affect the shifts. Non methylated DNA of ORF43 was generated by PCR while methylated ORF43 was obtained from from *E. coli* XL1-Blue where it had previously been cloned into a pKSII+ vector (Iro, 2006). However the DNA isolated from the plasmid is a little longer resulting in higher shifts at approx. 4.5 kbp.

To confirm, that the above observed shifts are not unspecific the binding of gp49 to the intergenic region of ORF48 and ORF49 was tested. Therefore two fragments of different lengths were created by PCR: (i) the prom-1/prom-2 fragment with a length of 579 bp and (ii) the 1082 bp prom-2/prom-3 fragment. Unexpectedly both fragments did also show a shift as depicted in Figure 16. In a last attempt gp49 was tested against the region of the *Nab. magadii* methyltransferase, where it definitely should not bind to. Again the fragment was generated by PCR using the primers N6-1b/MT-RT-I1 but there also a shift could be observed. These observations lead to the conclusion that binding of gp49 to DNA is of an unspecific nature under the applied conditions.

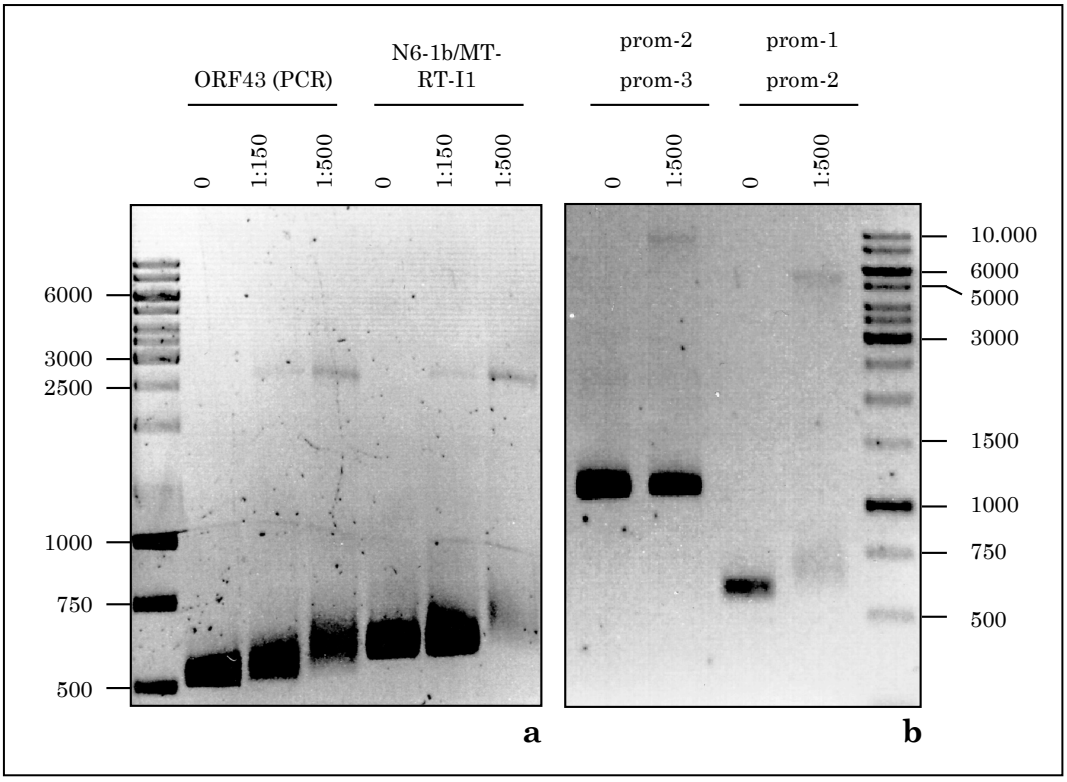

Figure 16| **EMSAS with unspecific DNA.** Gp49 was purified under denaturing conditions and dialyzed to 4 M urea. DNA used for these assays was generated by PCR. Numbers indicate the molar ratio of DNA to protein. **a**| Unexpectedly besides the ORF43 DNA gp49 is also able to shift a 580 bp long part of the ϕCh1 methyltransferase region which has been obtained by PCR. This fragment is shifted to a height of approx. 2.8 kbp. **b**| Gp49 is also able to induce a gelshift with DNA from the intergenic region of ORF48 and ORF49. Two fragments of different sizes were generated, one of which with a length of 1082 bp is shifted to approx. 10 kbp. The other fragment with 579 bp in length is shifted to approx. 5.8 kbp.

3.2.5.3. Comparison of gp49 wt and 3´ deletion mutants

In order to answer the question where the gp49 DNA binding domain is located C-terminally truncated versions of the protein were created. Subsequently the

truncated gp49 derivatives together with the wild type protein were tested for bandshift activity on ϕCh1 ORF43 DNA. Surprisingly all of the deletion mutants seem to shift equally well. Even the shortest version, gp49Δ5 which consists of only 57 of the native 118 codons is able to induce a gelshift equal to the wild type protein (Figure 17b, c).

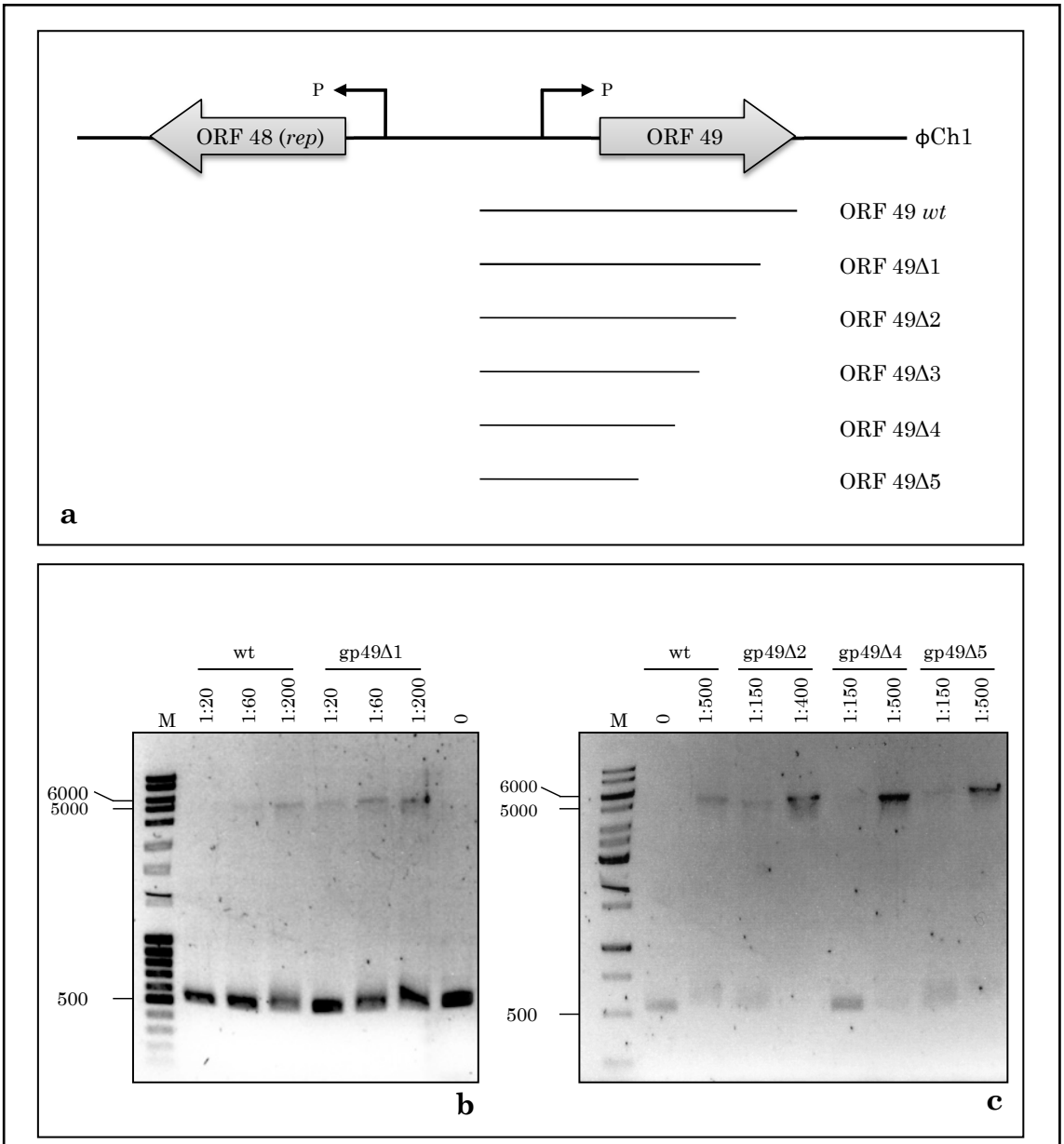

Figure 17| **Comparison of gp49 wt and truncated versions on ORF43 DNA. a|** Schematic representation of the ORF48 – ORF49 region on the ϕCh1 genome showing the different constructs used in this assay. Each mutation being shorter than the former one by approx. 30 – 40 nucleotides. The fragments were each cloned into a pRSETA expression vector and the respective gp49 derivatives were purified under denaturing conditions using 8 M urea. Thereafter proteins were dialyzed to 4 M urea. Subsequently the proteins were used to shift ϕCh1 ORF43 DNA obtained by PCR. **b|** No difference in the shifts is visible between gp49 wt and gp49Δ1. **c|** The same is true for deletions Δ2, Δ4 and Δ5. Unfortunately gp49Δ3 could not be purified in sufficient amounts for testing. However, evidence is strong that Δ3 also would give an equal shift as gp49 wt.

3.2.5.4. Comparison gp49 wt and N-terminal deletion

Since all of the C-terminally truncated versions of gp49 seemed to be able to shift DNA an N-terminal deletion mutant of gp49, gp49ΔN, was constructed in order to compare its binding affinity to that of wild type gp49. Gp49ΔN was deleted for the first 52 codons in comparison to the wild type. Both proteins were purified under denaturing conditions as before and then tested on the ORF43 DNA generated by PCR that has been shown to give a clearly detectable shift. This bandshift assay is depicted in Figure 18. While gp49 shifts the DNA as observed previously the mutant form does not seem to induce such a shift. Taken together with the results from the experiments with the C-terminal truncations (see above) this suggests that the DNA binding domain of gp49 is located N-terminally rather than C-terminally.

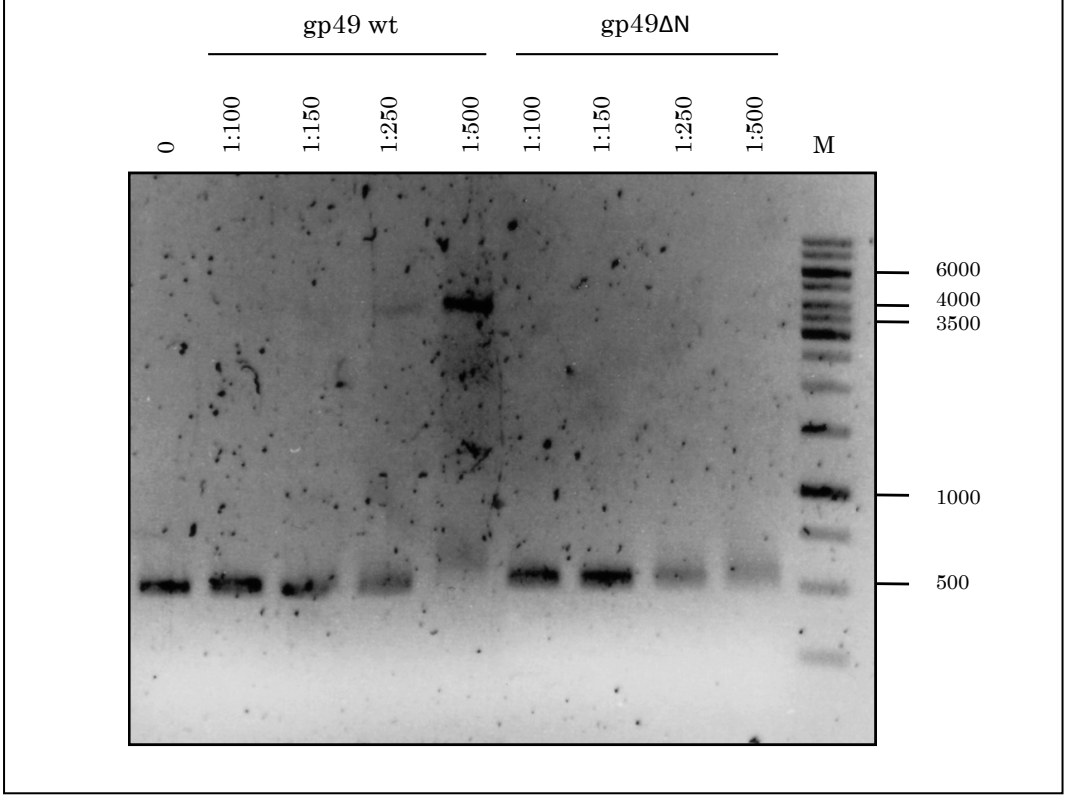

Figure 18| **Comparison of gp49 and gp49ΔN.** Both proteins were purified under denaturing conditions and thereafter dialyzed to 4 M urea. Subsequently they were used for bandshift assays with ϕCh1 ORF43 DNA that had been generated by PCR. The numbers indicate the molar ration between DNA and protein. Gp49 wt does induce a gelshift as before which becomes clearly visible at a molar ratio of 1:250. On the other hand no reaction can be observed when gp49ΔN is used suggesting that the binding domain of gp49 is located at the N-terminal end of the protein.

3.2.6. Construction and analysis of 5 different deletion mutants of ORF49

To further elucidate the characteristics of gp49 deletion mutagenesis was used. Using the sequence of ORF49 five derivates (ORF49Δ1 – ORF49Δ5) with truncated 3´ ends were generated by PCR. Each fragment subsequently being shorter than the former one by approx. 30 – 40 nucleotides (for exact values refer to section 2.1.4.).

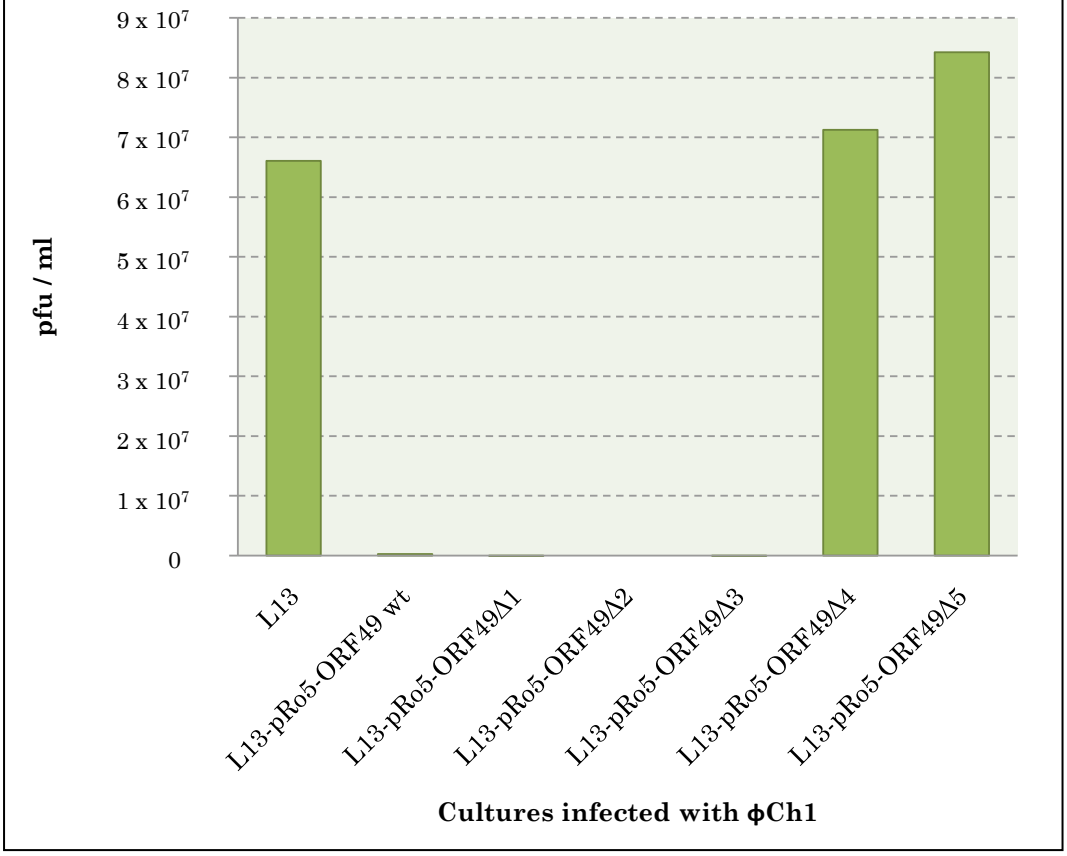

Figure 19| **Repressor activity of gp49 C-terminal deletion mutants.** 5 different deletion mutants of ORF49 were generated, each being subsequently truncated by approx. 30 – 40 nucleotides. This mutants ORF49Δ1 – ORF49Δ5 were cloned into the pRo-5 shuttle vector and transformed into *Nab. magadii* L13. Strains obtained this way were infected with ϕCh1 virus particles and titers were determined. The data show that while mutants Δ1 – Δ3 are equally potent to repress plaque formation as the wild type ORF49Δ4 and ORF49Δ5 seem to have lost their repressor activity.

The thus obtained fragments were then cloned into the pRo-5 shuttle vector and subsequently transformed into *Nab. magadii*. Thereafter single colonies were taken and used for inoculation of selective medium. These liquid cultures were later infected with ϕCh1 particles and plaques were counted. The result is unambiguous as Figure 19 clearly depicts. While mutants ORF49Δ1 – Δ3 seem to

be fully functional concerning their repressor activity ORF49Δ4 and ORF49Δ5 have lost their ability to repress plaque formation. Hence the crucial region for gp49 activity is located in the codons 1 – 81 from altogether 117 codons.

3.2.7. Discussion

The data gathered by this work clearly underline the assumption that gp49 is indeed a repressor of ϕCh1. It also gives some hints on the nature of this repressing function. Although 3D structure prediction only marginally supports the suggestion that gp49 is a DNA binding protein (but at the same time does not exclude it), the performed EMSAs show a different picture.

The purification of gp49 in a way that it remains its native structure from the beginning to the end was shown to be extraordinary difficult and could not be achieved with the disposable means. Heterologous expression of gp49 in *E. coli sp*. led to the formation of insoluble inclusion bodies and no soluble form of gp49 could be detected using methods for native protein isolation. Inclusion body formation seemed to be independent from the used *E. coli* strain as well as the growth temperature. On the other side the denaturing purification of the inclusion bodies containing gp49 does not only have disadvantages. In many cases the expression level of the heterologous protein can reach up more than 30 % of the total cellular protein (Singh & Panda, 2005). Additionally inclusion bodies can be isolated very simply and they contain the accumulated protein of interest in a relative pure form (Singh & Panda, 2005). However, the renaturation of a protein of unknown function holds certain biases as a successful restoration of the protein´s active form cannot be confirmed by an assay testing its function. This problem is reflected by the attempt to perform an EMSA with total ϕCh1 DNA using gp49 that has been renatured by gel filtration. Though several bands seem to shift (see Figure 13) the experiment itself could only be performed once. Repeated efforts to reproduce the initial results failed. In a subsequent experiment, when selected fragments of ϕCh1 DNA were tested using gp49 that was purified the same way the results were also not convincing (see Figure 14) giving, if any, very tiny shifts. It could be stated that as gp49 is with 13.3 kDa a relatively small protein the resulting shifts are also small. However, even the small shifts are not very convincing as they seem to be "all or nothing" shifts, meaning that either the whole DNA fraction is shifted or nothing at all. Taken altogether it seems very unlikely that the data shown in

Figure 14 really represent valid band shifts. This brings up the question what exactly was observed in these two experiments. When renaturating a protein by size exclusion chromatography the denaturing agent is slowly and thoroughly removed thus allowing the protein to regain its native conformation. However, instead of returning into its active form refolding could as well result in any other stable conformation. Thus it seems obvious that smaller proteins are more likely to restore their native states than bigger ones simply because fewer amino acids are participated in the refolding process. Furthermore it is also apparent that after removal of the denaturing agent all possible conformations of the now soluble protein are simultaneously available in different concentrations, favoring the most stable one. Therefore though gp49 is a relatively small protein it cannot be tested nor assumed that after size exclusion chromatography gp49 has regained its native and active conformation. Consequently it is also possible that the shifts observed using total ϕCh1 DNA are due to an impurity, since it seems to be likely that proteins smaller than gp49 refold better into their native conformation.

The experiments, however, that were performed with gp49 solubilized in 4 M urea gave somewhat surprising results as it seems that even under these conditions gp49 is able to interact with DNA. This leads to the immediate question how proteins are able to be active in a denatured state? Maybe the answer lies in a paradigm change that lately emerges within protein biochemistry. Recent studies have shown that in some cases an intrinsic three dimensional structure of a protein is not necessary for its function. Such proteins are called intrinsically disordered proteins (IDPs) (Dunker *et al.*, 2001) and have been shown to be involved in regulation, signaling and control pathways (Radivojac *et al.*, 2007) including DNA binding – for example GCN4, a yeast transcription factor. Here the basic region adjacent to the leucine zipper is unstructured and folds into a helical structure only upon binding to its cognate site (Weiss *et al.*, 1990). In a study of Ward *et al.* the frequency of native (i.e. intrinsic) disorder in a number of genomes, including *Archaea*, has been estimated. Disordered segments of at least 30 residues were predicted for 2 % of the archaeal proteins. Within the *Archaea* the highest abundance of IDPs (5 %) occurs in *Halobacterium sp.*, a close relative of *Nab. magadii* (Ward *et al.*, 2004). A computer prediction for gp49 showed a number of disordered residues at the N-terminal, as well as at the C-terminal end. Thus it seems likely that this protein does not necessarily need a tertiary structure to bind DNA in vitro. However, the results of several EMSAs performed in this work point

into the direction of a rather unspecific binding under the applied conditions. Thus a situation similar to that of GCN4 seems conceivable. This would suggest that one of the two disordered regions of gp49 is responsible for the DNA binding while it is possible that recognition of the cognate site is carried out by a tertiary structure which is lost upon treatment with 4 M urea.

The present work also took effort to identify the regions responsible for gp49 activity. By performing a phage titer analysis comparing five different 3´ deletion mutants of gp49 to the wild type protein it was found that the repressor activity of gp49 is situated within the first 81 of altogether 117 codons. On the other hand EMSAs performed with virtually the same mutants showed that even the smallest gp49 derivative, gp49Δ5 consisting only of the first 57 codons, is able to shift DNA equally well as the wild type does. These facts suggest that gp49 contains at least two domains, one of which being solely responsible for the binding of DNA and the other mediating repressor activity. Consequently the DNA binding domain (DBD) is located at the N-terminal end of the protein with the initial 57 codons. Studies with an N-terminal deletion mutant of gp49 deprived of the first 52 codons support this theory since no bandshifts were observed when an EMSA was performed using such a protein.

3.2.8. Closing words

The present work represents the initial steps of the characterization of a protein that has been herein verified to be a repressor of the halophage ϕCh1. To the knowledge of the author this is also the first time that an EMSA using a haloalkaliphilic protein could ever be successfully performed. However, the characterization of gp49 is by far not completed as for example the nature of the repressor activity of gp49 is still unknown, a mechanism further studies will have to reveal.

The third domain of life certainly holds a tremendous amount of yet undiscovered knowledge that may proof indispensable for the future. Moreover, though the work with *Archaea* sometimes seems to be extraordinary difficult and cumbersome, especially when it comes to extremophiles, it should not be forgotten that research on *Archaea* also contributes an undeniable part to the revelation of evolutionary relationships between all living organisms.

4. References

Allers T, Mevarech M. 2005. Archaeal genetics - the third way. Nature Reviews Genetics. 6(1):58-73.

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 ER, Bell SD. 2006. DNA Replication in the Archaea. Microbiology And Molecular Biology Reviews. 70(4):876-87.

Bartlett MS. 2005. Determinants of transcription initiation by archaeal RNA polymerase. Current Opinion in Microbiology. 8:677-84.

Bath C, Cukalaca T, Portera K, Dyall-Smith ML. 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus. Virology. 350(1):228-39.

Bath C, Dyall-Smith ML. 1998. His1, an Archaeal Virus of the *Fuselloviridae* Family That Infects *Haloarcula Hispanica*. Journal of Virology. 72(11):9392-5.

Bell SD, Jackson SP. 1998. Transcription and Translation in Archaea: a mosaic of eukaryal and bacterial features. Trends in Microbiology. 6(6):222-8.

Benarroch D, Smith P, Shuman S. 2008. Characterization of a Trifunctional Mimivirus mRNA Capping Enzyme and Crystal Structure of the RNA Triphosphatase Domain. Structure. 16(4):501-12.

Brown JR, Doolittle WF. 1997. *Archaea* and the Prokaryote-to-Eukaryote Transition. Microbiology And Molecular Biology Reviews. 61(4):456-502.

Chaban B, Ng SY, Jarrell KF. 2006. Archaeal habitats — from the extreme to the ordinary. Canadian Journal of Microbiology. 52(2):73-116.

Charlebois RL, Lam WL, Cline SW, Doolittle WF. 1987. Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaebacterium. Proc. Natl. Acad. Sci. USA. 84(23):8530-4.

Chatton E. 1938. Titre et travaux scientifique (1906–1937). Sete: E. Sottano (in French).

Cline SW, Doolittle WF. 1987. Efficient Transfection of the Archaebacterium *Halobacterium halobium*. Journal of Bacteriology. 169(3):1341-4.

Copeland HF. 1938. The kingdoms of organisms. Q.Rev. Biol. 13:383-420.

Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, *et al*. 2001. Intrinsically disordered protein. J. Mol. Graph. Model. 19(1):26-59.

Dyall-Smith ML. 2008. The Halohandbook. Protocols for haloarchaeal genetics. Version 7.
http://www.haloarchaea.com/resources/halohandbook/Halohandbook_2008_v7.pdf.

Dyall-Smith M, Tang SL, Bath C. 2003. Haloarchaeal viruses: how diverse are they? Research in Microbiology. 154(4):309-13.

Edgell DR, Doolittle WF. 1997. Archaea and the Origin(s) of DNA Replication Proteins. Cell. 89(7):995-98.

Eichler J. 2001. Biotechnological uses of archaeal extremozymes. Biotechnology Advances. 19(4):261-78.

Fendrihan S, Legat A, Pfaffenhuemer M, Gruber C, Weidler G, Gerbl F, Stan-Lotter H. 2006. Extremely halophilic archaea and the issue of long-term microbial survival. Reviews in Environmental Science and Biotechnology. 5(2-3):203-18.

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

Garrity GM, Boone DR, Castenholz RW, editors. 2001. Bergey's Manual of Systematic Bacteriology Volume One: The Archaea and the Deeply Branching and Phototrophic Bacteria. 2nd ed. Vol 1. New York: Springer.

Gomes J, Steiner W. 2004. The Biocatalytic Potential of Extremophiles and Extremozymes. Food Technol. Biotechnol. 42(4):223-35.

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

Guo PX, Bailey S, Bodley JW, Anderson D. 1987b. Characterization of the small RNA of the bacteriophage ϕ29 DNA packaging machine. Nucleic Acids Res. 15(17):7081-90.

Guo PX, Erickson S, Anderson D. 1987a. A small viral RNA is required for in vitro packaging of bacteriophage ϕ29 DNA. Science. 236(4802):690-4.

Haeckel E. 1866. Generelle Morphologie der Organismen. Reimer, Berlin.

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, Dyall-Smith ML. 1990. A Plasmid Vector with a Selectable Marker for Halophilic Archaebacteria. Journal of Bacteriology. 172(2):756-61.

Holmes ML, Dyall-Smith ML. 1991. Mutations in DNA Gyrase Result in Novobiocin Resistance in Halophilic Archaebacteria. Journal of Bacteriology. 173(2):642-48.

Horikoshi K. 1999. Alkaliphiles: Some Applications of Their Products for Biotechnology. Microbiology and Molecular Biology Reviews. 63(4):735-50.

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.

Iida S. 1984. Bacteriophage P1 carries two related sets of genes determining its host range in the invertible c segment. Virology. 421-34.

Iro M. 2006. Characterisation of the regulatory region of the virus ϕCh1 infecting Natrialba magadii. [PhD Thesis]. Department für Mikrobiologie und Immunbiologie der Universität Wien.

Iro M, Klein R, Gálos B, Baranyi U, Rössler N, Witte A. 2007. The lysogenic region of virus ϕCh1: identification of a repressor-operator system and determination of its activity in halophilic *Archaea*. Extremophiles. 11(2):383-96.

Iro M, Ladurner A, Meissner C, Derntl C, Reiter M, Haider F, Rössler N, Klein R, Baranyi U, Scholz H, Witte A. in prep. Transformation of *Natrialba magadii*: use of ϕCh1 elements for a shuttle vector system for haloalkaliphilic *Archaea*.

Kamekura M, Dyall-Smith ML, 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 *Natronomonas pharaonis* comb. nov., Respectively. International Journal of Systematic Bacteriology. 47(3):853-7.

Kandler O, König H. 1998. Cell wall polymers in Archaea (Archaebacteria). Cell. Mol. Life Sci. 54(4):305-8.

Karner MB, DeLong EF, Karl DM. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. Nature. 409 (6819):507-10.

Keeling PJ, Doolittle WF. 1995. Archaea: Narrowing the gap between prokaryotes and eukaryotes. Proc. Natl. Acad. Sci. USA. 92:5761-4.

Kelley LA, Sternberg MJE. 2009. Protein structure prediction on the Web: a case study using the Phyre server. Nature Protocols. 4(3):363-71.

Klein R. 1999. ϕCh1-spezifische Transkriptionsstudien in *Natrialba magadii*. [PhD Thesis]. Institut für Mikrobiologie und Genetik der Universität Wien.

Klein R, Baranyi U, Rössler N, Greineder B, Scholz H, Witte A. 2002. *Natrialba magadii* virus ϕCh1: 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 ϕCh1: Evidence for Processing in *Natrialba magadii* during Virus Maturation. Virology. 276(2):376-87.

Kornberg A, Baker TA. 1992. DNA Replication. 2nd ed. New York: W. H. Freeman & Co Ltd.

Kushner DJ. Life in high salt and solute concentrations. 1978. In: Kushner DJ, editor. Microbial Life in Extreme Environments. London: Academic Press. p 317-368.

Ladurner A. 2008. Characterization of the inversion reaction of ϕCh1 and the establishment of a transformation system for *Natrialba magadii*. [Diploma Thesis]. Universität Wien.

Lam WL, Doolittle WF. 1992. Mevinolin-resistant Mutations Identify a Promoter and the Gene for a Eukaryote-like 3-Hydroxy-3-methylglutaryl-coenzyme A Reductase in the Archaebacterium *Haloferax volcanii**. The Journal of Biological Chemistry. 267(9):5829-34.

Lange U, Hausner W. 2004. Transcriptional fidelity and proofreading in Archaea and implications for the mechanism of TFS-induced RNA cleavage. Molecular Microbiology. 52(4):1133-43.

Lanyi JK. 1974. Salt-Dependent Properties of Proteins from Extremely Halophilic Bacteria. Bacteriological Reviews. 38(3):272-90.

Lodwick D, Ross HNM, Harris JE, Almond JW, Grant WD. 1986. *dam* Methylation in the Archaebacteria. Journal of General Microbiology. 132(11):3055-9.

Lundgren M, Andersson A, Chen L, Nilsson P, Bernander R. 2004. Three replication origins in *Sulfolobus* species: Synchronous initiation of chromosome replication and asynchronous termination. Proc. Natl. Acad. Sci. USA. 101(18):7046-51.

Madern D, Ebel C, Zaccai G. 2000. Halophilic adaption of enzymes. Extremophiles. 4(2):91-8.

Margesin R, Schinner F. 2001. Potential of halotolerant and halophilic microorganisms for biotechnology. Extremophiles. 5(2):73-83.

Meissner CS. 2008. Characterisation of Int2 and determination the binding sites of proteins gp43 and gp44. [Diploma Thesis]. Universität Wien.

Mescher MF, Strominger JL. 1976. Structural (shape-maintaining) role of the cell surface glycoprotein of *Halobacterium salinarium*. Proc. Natl. Acad. Sci. USA. 73(8):2687-91.

Miroux B, Walker JE. 1996. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. J. Mol. Biol. 260(3):289-98.

Moens S, Vanderleyden J. 1997. Glycoproteins in prokaryotes. Archives of Microbiology. 168(3):169-175.

Ng WL, DasSarma S. 1993. Minimal Replication Origin of the 200-kilobase *Halobacterium* Plasmid pNRC100. Journal of Bacteriology. 175(15):4584-96.

Nuttall SD, Dyall-Smith ML. 1993. HF1 and HF2: Novel Bacteriophages of Halophilic Archaea. Virology. 197(2):678-84.

Oesterhelt D, Bräuchle C, Hampp N. 1991. Bacteriorhodopsin: a biological material for information processing. Quarterly Reviews of Biophysics. 24(4):425-78.

Olsen GJ. 1994. Archaea, Archaea, everywhere. Nature. 371:657-8.

Oren A. 1999. Bioenergetic Aspects of Halophilism. Microbiology And Molecular Biology Reviews. 63(2):334-48.

Oren A. 2002. Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. Journal of Industrial Microbiology & Biotechnology. 28(1):56-63.

Oren A. 2008. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Systems. 4(2).

Oren A, Bratbak G, Heldal M. 1997. Occurrence of virus-like particles in the Dead Sea. Extremophiles. 1(3):143-9.

Palmer BR, Marinus MG. 1994. The *dam* and *dcm* strains of *Escherichia coli* — a review. Gene. 143(1):1-12.

Perry KL, Elledge SJ, Mitchell BB, Marsh L, Walker GC. 1985. *umuDC* and *mucAB* operons whose products are required for UV light- and chemical-induced mutagenesis: UmuD, MucA, and LexA proteins share homology. Proc. Natl. Acad. Sci. USA. 82(13):4331-5.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera – A visualization system for exploratory research and analysis. J. Comput. Chem. 25(13):1605-12.

Prangishvili D, Forterre P, Garrett RA. 2006a. Viruses of the Archaea: a unifying view. Nature Reviews Microbiology. 4(11):837-48.

Prangishvili D, Garrett RA, Koonin EV. 2006b. Evolutionary genomics of archaeal viruses: Unique viral genomes in the third domain of life. Virus Research. 117(1):52-67.

Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. 2007. Intrinsic Disorder and Functional Proteomics. Biophysical Journal. 92(5):1439-56.

Raymond-Denise A, Guillen N. 1991. Identification of *dinR*, a DNA Damage-Inducible Regulator Gene of *Bacillus subtilis*. Journal of Bacteriology. 173(22):7084-91.

Rodriguez-Valera F, Lillo JG. 1992. Halobacteria as producers of polyhydroxyalkanoates. FEMS Microbiology Reviews. 103(2-4):181-6.

Rössler N, Klein R, Scholz H, Witte A. 2004. Inversion within the haloalkaliphilic virus ϕCh1 DNA results in differential expression of structural proteins. Molecular Microbiology. 52(2):413-26.

Rothschild LJ, Mancinelli RL. 2001. Life in extreme environments. Nature. 409(6823):1092-101.

Sandmeier H. 1994. Acquisition and rearrangement of sequence motifs in the evolution of bacteriophage tail. Molecular Biology. 12(3):343-50.

Schnabel H, Zillig W, Pfäffle M, Schnabel R, Michel H, Delius H. 1982. *Halobacterium halobium* phage ϕH. EMBO Journal. 1(1):87-92.

Schnos M, Zahn K, Inman RB, Blattner FR. 1988. Initiation protein induced helix destabilization at the lambda origin: A prepriming step in DNA replication. Cell. 52(3):385-95.

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

Singh SM, Panda AK. 2005. Solubilization and Refolding of Bacterial Inclusion Body Proteins. Journal of Bioscience and Bioengeneering. 99(4):303-10.

Sticht H, Willbold D, Ejchart A, Rosin-Arbesfeld R, Yaniv A, Gazit A, Rösch P. 1994. Trifluoroethanol stabilizes a helix-turn-helix motif in equine infectious-anemia-virus *trans*-activator protein. European Journal of Biochemistry. 225(3):855-61.

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

Tindall BJ, Ross HNM, Grant WD. 1984. *Natronobacterium* gen. nov. and *Natronococcus* gen. nov., two new genera of haloalkaliphilic archaebacteria. Systematic and Applied Microbiology. 41-57.

Torsvik T, Dundas ID. 1974. Bacteriophage of *Halobacterium salinarium*. Nature. 248(450):680-1.

Trachuk L, Letarov A, Kudelina IA, Yusupova MP, Chestukhina GG. 2005. In vitro refolding of carboxypeptidase T precursor from *Thermoactinomyces vulgaris* obtained in *Escherichia coli* as cytoplasmic inclusion bodies. Protein Expression and Purification. 40(1):51-9.

Valentine DL. 2007. Adaptations to energy stress dictate the ecology and evolution of the Archaea. Nature Reviews Microbiology. 5(4):316-23.

van de Vossenberg JLCM, Driessen AJM, Grant D, Konings WN. 1999. Lipid membranes from halophilic and alkali-halophilic Archaea have a low H+ and Na+ permeability at high salt concentration. Extremophiles. 3(4):253-7.

Ventosa A, Nieto JJ. 1995. Biotechnological applications and potentialities of halophilic microorganisms. World Journal of Microbiology and Biotechnology. 11(1):85-94.

Ventosa A, Oren A. 1996. *Halobacterium salinarum* nom. corrig., a Name To Replace *Halobacterium salinarium* (Elazari-Volcani) and To Include *Halobacterium halobium* and *Halobacterium cutirubrum*. International Journal of Systematic Bacteriology. 46(1):347.

Wais AC, Kon M, MacDonald RE, Stollar BD. 1975. Salt-dependent bacteriophage infecting *Halobacterium cutirubrum* and *H. halobium*. Nature. 256(5515):314-5.

Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. 2004. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. Journal of Molecular Biology. 337(3):635-45.

Weiss MA, Ellenberger T, Wobbe CR, Lee JP, Harrison SC, Struhl K. 1990. Folding transition in the DNA-binding domain of GCN4 on specific binding to DNA. Nature. 347(6293):575-8.

White MF, Bell SD. 2002. Holding it together: chromatin in the Archaea. Trends in Genetics. 18(12):621-6.

Whittaker RH. 1959. ON THE BROAD CLASSIFICATION OF ORGANISMS. The Quarterly Review of Biology. 34(3):210-26.

Whittaker RH, Margulis L. 1978. Protist classification and the kingdoms of organisms. Biosystems. 10(1-2):3-18.

Witte A, Baranyi U, Klein R, Sulzner M, Luo C, Wanner G, Krüger DH, Lubitz W. 1997. Characterization of *Natronobacterium magadii* phage ϕCh1, 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 organimsms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA. 87(12):4576-9.

Wright PE, Dyson HJ. 1999. Intrinsically Unstructured Proteins: Re-assessing the Protein Structure-Function Paradigm. J. Mol. Biol. 293(2):321-31.

Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene. 33(1):103-19.

5. Appendix

Index of Figures

Figure 1| The three domains of life. 13
Figure 2| Morphology of *Nab. magadii*. 26
Figure 3| Morphology of ϕCh1 particles 32
Figure 4| Genetic organization of virus ϕCh1. 34
Figure 5| Gene regulation in ϕCh1 by ORF48 (*rep*) and ORF49. 38
Figure 6| The pRo cloning series. 78
Figure 7| Transformation efficiency of pRo derivates. 80
Figure 8| Repressor Activity of gp49. 83
Figure 9| Structure predictions of gp49. 85
Figure 10| Intrinsic disorder prediction for gp49. 86
Figure 11| Overexpression of gp49 87
Figure 12| Purification of gp49 89
Figure 13| Electrophoretic mobility shift assay (EMSA) using total ϕCh1 DNA. .. 91
Figure 14| Bandshift assay with selected ϕCh1 DNA fragments 92
Figure 15| EMSA – testing gp49 on ORF43. 94
Figure 16| EMSAS with unspecific DNA. 95
Figure 17| Comparison of gp49 wt and truncated versions on ORF43 DNA. 96
Figure 18| Comparison of gp49 and gp49ΔN 97
Figure 19| Repressor activity of gp49 C-terminal deletion mutants. 98

Acknowledgements

My gratitude goes to all people who supported me and enabled me to work on this thesis.

In the first place I would like to thank Prof. Dr. Angela Witte. She made it possible for me to work in her group and do research on one of the most interesting yet far too less considered fields of research – the *Archaea*. Moreover she was always there for me and my questions, never getting tired to answer them even when they were asked twice or even three times. She created a warm and inviting atmosphere in her laboratory and I consider it as a unique gift to have been able to work there. I am deeply grateful for the great time I had in her lab.

Furthermore I want to express my gratitude to my co-workers, who over the time became more than just fellow colleagues rather than true friends. At first I want to thank Angela Ladurner and Christina Meissner who introduced me to the Angela Witte lab teaching me all the "*important stuff*". Moreover I particularly want to thank the "core" of the "Halo-lab" at that time – Christian Derntl, Flora Haider and Regina Selb, as we were having a great time and much fun together. Furthermore my thanks go to Petra Till, who was a great help for my work, as well as to Daniel Kiesenhofer and Tatjana Svoboda.

I am also thankful to Univ.-Prof. Dr. Udo Bläsi and Mag. Dr. Isabella Moll for their support and moreover for all the support I got from all the people working at the other labs of the department of Microbiology, Immunobiology and Genetics: Eliza Deltcheva, Stephanie Füreder, Hermann Hämmerle, David Hasenöhrl, Sanda Pasc, Silvia Spiess, Branislav Vecerek and Oliver Vesper. Special thanks go to Konstantin Byrgazov and Krzysztof Chylinski for their constructive criticism as well as for giving me important hints and new ideas from time to time.

Last but not least I want to thank my family for supporting me through all the years. I am very grateful that my parents Marion and Manfred made the studies of microbiology possible for me. Furthermore I want to thank my brother Martin who would always listen to me giving me emotional support.

Abstract

The halophage ϕCh1 infects *Natrialba magadii*, a haloalkaliphilic archaeon facing extreme living conditions – high salinity and high pH values. The present work contributed to the elucidation of the ϕCh1 origin of replication as well as to the characterization of the ϕCh1 repressor gp49.

The ϕCh1 origin of replication consists of two open reading frames (ORF53 and ORF54) similar to RepH, a protein that has been shown to be required for replication of plasmid pNRC100 of *Haloarcula marismortui*. ORF53 is flanked by an AT-rich palindromic sequence at the 5´ end while the same is true for ORF54 at the 3´ end. Both palindromic sequences are suggested to be involved in the mechanism of replication. By introducing a +1 frameshift mutation into ORF53 as well as ORF54 this work could provide evidence that both ORFs are absolutely required for replication suggesting that in ϕCh1 ORF53 and ORF54 together mediate RepH activity. Furthermore by deleting the 5´ as well as the 3´ AT-rich palindromic sequences this work could show that the 3´ sequence seems to be more important for replication if only one of these two sequences is deleted. However, deleting both sequences has no effect on replication suggesting a so far unknown mechanism of replication.

The second part of this work describes the characterization of gp49 the gene product ϕCh1 ORF49. In an initial phage titer experiment where ORF49 was cloned into a *Nab. magadii* shuttle vector, gp49 could be shown to be a powerful ϕCh1 repressor. By deletion mutagenesis the active domain of gp49 activity was mapped to comprise the first 81 of altogether 117 codons. Furthermore gp49 was cloned into an *E. coli* expression vector from which it could be successfully purified. Moreover using gp49 the first EMSA with a haloalkaliphilic protein could be performed identifying gp49 as DNA binding protein. Further binding studies with truncated forms of gp49 revealed that C-terminal truncations have no effect on DNA binding whereas repressor activity is lost. On the other hand an N-terminal deletion of the first 52 codons leads to the abolishment of the DNA binding activity. These data suggest that gp49 is comprised at least of two domains, one of which being responsible for binding the other mediating the repressor activity.

Zusammenfassung

Der Halophage ϕCh1 infiziert das Archaeon *Natrialba magadii*, welches in extremen Habitaten vorkommt, wo sowohl ein hoher Salzgehalt als auch ein hoher pH Wert vorherrschen. Die vorliegende Diplomarbeit beteiligt sich an der näheren Aufklärung des Replikationsursprungs von ϕCh1 sowie an der Charakterisierung des ϕCh1 Repressors gp49.

Der Replikationsursprung von ϕCh1 besteht aus zwei offenen Leserahmen (ORFs) statt wie gewöhnlich nur aus einem. Beide Leserahmen (ORF53 und ORF54) ähneln RepH, einem Faktor der für die Replikation des *Haloarcula marismortui* Plasmids pNRC100 eine große Rolle spielt. Diese beiden ORFs sind für die Replikation eines zuvor entwickelten Shuttle Vektors in *Nab. magadii* notwendig. Die 5´ Region von ORF53 sowie die 3´ Region von ORF54 wird von AT-reichen palindromischen Sequenzen flankiert. Beide palindromischen Sequenzen sind höchstwahrscheinlich an der Replikation von ϕCh1 beteiligt. In der vorliegenden Arbeit wurden ORF53 sowie ORF54 durch eine Mutation, welche den Leserahmen der Gene um jeweils +1 verschiebt, zerstört. So konnte zweifelsfrei gezeigt werden, dass beide ORFs für die Replikation des *Nab. magadii* Shuttle Vektors absolut notwendig sind. Dadurch steht es nunmehr außer Frage, dass in ϕCh1 beide Leserahmen für die RepH Aktivität verantwortlich sind. Durch die Deletion der 5´ sowie der 3´ AT-reichen Sequenzen konnte in dieser Arbeit gezeigt werden, dass die 3´ Sequenz eine größere Rolle spielt, wenn nur jeweils eine der beiden Sequenzen deletiert ist. Gleichzeitige Deletion beider Sequenzen stellt allerdings das normale Replikationsverhalten wieder her, was auf einen bisher unbekannten Replikationsmechanismus hindeutet.

Der zweite Teil dieser Arbeit befasst sich mit der Charakterisierung von gp49, dem Genprodukt des ORF49. In einem anfänglichen Phagentiterexperiment, bei dem ORF49 auf den *Nab. magadii* Shuttle Vektor kloniert wurde, konnte eine eindeutige Repressoraktivität von gp49 festgestellt werden. Durch Kartierung mittels Deletionsmutagenese konnte die aktive Domäne des Proteins auf die ersten 81 von insgesamt 117 Codons festgelegt werden. Darüberhinaus wurde gp49 in einen *E. coli* Expressionsvektor kloniert, von dort aus exprimiert und erfolgreich

aufgereinigt. Mit dem aufgereingten Produkt konnte der erste EMSA mit einem haloalkalophilen Protein durchgeführt werden, was den Nachweis der DNA-Bindungsaktivität von gp49 lieferte. Weitere Bindungsstudien mit verkürzten Versionen von gp49 ergaben, dass C-terminale Deletionen keinen Effekt auf die DNA – Bindungsaktivität haben, während die Repressoraktivität zerstört wird. Werden hingegen die ersten 52 N-terminalen Codons deletiert, geht die Bindungsaktivität verloren. Damit konnte gezeigt werden, dass gp49 aus mindestens zwei funktionellen Domänen besteht, einer Domäne, die für die DNA – Bindung verantwortlich ist und einer Domäne, welche die Aktivität des Proteins vermittelt.

Curriculum Vitae

Personal details

Name:	Michael Reiter
Date and place of birth:	April, 7th, 1981 in Coburg, Germany
Nationality:	Austria

Education

"Volksschule Wallsee – Sindelburg", Austria
1988 - 1991

"Hauptschule Wallsee – Sindelburg", Austria
1991/92 - 1995

"BORG Perg", Austria (track with emphasis on natural science)
1995/96 -1999

Studies of „Genetics and Microbiology"at the University of Vienna, Austria
October 2000 - May 2009 (interrupted from July 2005 to end of June 2006 due to compulsory community service)

Diploma thesis: Laboratory of Prof. Dr. Angela Witte; Department of Microbiology, Immunobiology and Genetics; University of Vienna, Austria
May 2008 – May 2010

Working experience

University of Vienna, Austria:

Tutor for student practical trainings: "Übungen in Mikrobiologie", "Übungen im Pflichtfach Mikrobiologie", "Übungen I B – Mikrobiologie und Genetik"
Winter term 2008/2009; summer term 2009, winter term 2009/2010, summer term 2010

FH Campus Wien (University of Applied Sciences, Vienna, Austria):

Tutor for student practical trainings "Molekularbiologische Übungen I und II"
Winter term 2008/2009; summer term 2009, winter term 2009/2010, summer term 2010