universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

„Development of a plasmid shuttle vector system for CRISPR studies in the archaeon *Sulfolobus solfataricus*“

verfasst von

Isabelle Anna Zink

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2013

Studienkennzahl lt. Studienblatt:	A 441
Studienrichtung lt. Studienblatt:	Diplomstudium Genetik - Mikrobiologie
Betreut von:	Univ.-Prof. Dipl.-Biol. Dr. Christa Schleper

für Mario

Acknowledgements

The imagination of life being determined and controlled by tiny molecular entities arising from a progression of only four different chemical compounds has fascinated me ever since I heard about genetics. From the moment I was holding a pipette in my hands I enjoyed working in the lab and I never regretted throwing my initial plans of studying zoology overboard. The conviction of having taken the right route in the big field of biology even got stronger when starting my diploma thesis in the department “Genetic in Ecology”. Therefore, my biggest debt of gratitude goes to Christa Schleper for allowing me to work on the fascinating subject of hyperthermophile Archaea which often resulted in burnt fingers but far more than that in burning excitement. Thank you, Christa for your enthusiasm which is “infectious” and the professional leading through my diploma thesis.

Working in this lab has greatly improved my technical and theoretical knowledge which is also the result of Ziga`s excellent tutelage. Thank you very much Ziga for your teaching, your support, your ideas, your patience, your help and last but not least for the cigarette breaks when I was nervously waiting for the Southern - or Northern blot development.

Another big “thank you” goes to my diploma-companion Bella with whom I shared a lot of thoughts, coffees and beers. It was always advising and a pleasure to discuss with you, Bella and I will never forget to “just pipette” in case of emergency ;). Furthermore, I want to thank Andrea Manica for his help, advices and discussions during my practicum. Beside the CRISPR group, a big “thank you” also goes to all other members of the department who gave me such a warm welcome when I started my work here and became such good fellows. Thank you, Ina (also for the indoor pedaling) Anna, Chrisi, Steffi, Andrea, Melina, Michaela, Romana, Ricardo, Nika, Nico, Michael, Clarissa, Pierre, Anja, Amir, Tim and Silvia.

Von ganzem Herzen möchte ich meiner Familie, Philipp und meinen Freunden danken, die das wichtigste auf der Welt für mich sind! Danke an meinen Bruder Marko für alles, besonders für das, dass du mir immer so viel Mut machst und keine „Sekunde an mir zweifelst“, wie du immer sagst! Danke Mama für einfach alles und für dein Interesse an meinem Studium und deinem Beistand während desselbigen! Danke an mein Lesterschwein, dass du mir immer mit Rat und Tat zur Seite stehst - SISTERS! Danke an dich, Claudl meine Biologie (und sowieso) Verbündete, für deinen Beistand vor allem am Anfang meines Studiums, ich sag nur „Miss Chemie“. Danke an meinen Bruder Mario der ein unbeschreiblich wichtiger Bestandteil meines Lebens war und noch ist und an meinen Vater, der mir den „Glauben an das Gute“ vermittelt hat. Danke an EJ, den

Schützen, für alles (unter anderem meinen ersten Hund) und dass du und Sabine mir geholfen habt, Fuß in Wien zu fassen. Danke Gerd (oder seit neuestem GINDER) neben vielem anderen dass du mein personal fashion master bist!

Danke Philipp, besonders für deinen Beistand in letzter Zeit solltest du einen Orden erhalten. Ich kann mich einfach immer bei dir „hängen lassen“ und dafür - und wegen noch viiiieeelem mehr – liebe ich dich und danke ich dir.

Danke an Kitte (net zuletzt für üsre coole WG), Macatha, Mr.Spock, Rollmops, Geli, Nörschgi, Ännschgi, Kathibü – i kann gar net säga, für was i eu alls danka will, es git viel z viel - drum dank i eu einfach für eure bloße Existenz-ihr sen die besta! Danke Marion, dass aus einer Zwecks-Referats-Gemeinschaft so eine tolle Freundschaft entstanden ist – das Beste, was die „Biologie“ wohl hervor gebracht hat ;). Danke Andi, für die vielen „tea times“ (die sich manchmal zu „beer times“ erweitert haben) und unsere gemeinsamen „Kämpfe“ durch die letzten Etappen des Studiums. Danke Brigitte für die vielen aufbauenden Gespräche in der letzten Phase unseres Studiums – MIR SCHAFFAN DES!!

Table of contents

1 Abstract

The hyperthermophilic archaeon *Sulfolobus solfataricus* has become an important model organism in different fields not least because of its amenability for genetic manipulation. The acidophilic Crenarchaeote was the first hyperthermophilic archaeon equipped with a functional and widely applicable genetic system on the basis of the lysogenic fusellovirus *Sulfolobus*-shibatae virus 1 (SSV1). SSV1 based vectors spread throughout the culture which is advantageous in that successful transformation does not rely on stringent selection. However, the infectious nature of the virus system interferes with overexpression and knockout studies due to highly variable copy numbers of free virus particles and also to variation of the SSV1 genome in its episomal form. Previous attempts to establish a non-spreading transformation system which would overcome this unfavorable effect of the virus, however, turned out to be difficult.

The goal of this study was to design and establish a plasmid based shuttle vector that is stably maintained in low copy numbers in uracil auxotroph *S. solfataricus* P1 mutants and to explore its use for studying the virus defense system in *Sulfolobus*.

The *Sulfolobus-E.coli* shuttle vector pCmalLacS which is based on the *S. islandicus* pRN1 plasmid was modified by introducing a Gateway® cassette with an arabinose inducible promoter upstream of it. The resulting Gateway® destination vector pCAra-GW was used for site-specific recombination of different inserts producing the so called, pIZ expression vectors able to overexpress the transcripts of the inserted sequences upon arabinose induction. Five different variants of pIZ vectors were successfully transformed within three transformation approaches following standard SSV1 electroporation protocols and growth of transformants in N-Z-amine/sucrose selective media. We were able to retransform total DNA of transformants into *E.coli* and recovered the intact plasmid from all analyzed colonies. Additionally, Southern hybridization verified the presence of the pIZ

vectors in the culture and elucidated the plasmids to reside as episomes in the cytoplasm. Plasmid copies were determined by qPCR using a chromosomal gene as reference. Numbers of chromosomes and plasmid counts were almost equal (around $3x10^6$ per µg total DNA) in pIZ-WOP transformants implying a plasmid copy number of between one and two per cell.

Furthermore, the plasmids were used to investigate *in vivo* the activity of the CRISPR system (*C*lustered *R*egularly *I*nterspaced *S*hort *P*alindromic *R*epeats), a recently identified RNA - based defense barrier protecting host cells from extracellular invaders. Six CRISPR loci located in the *S. solfataricus* genome each containing reservoirs of short sequence fragments, called spacers which stem from viruses or plasmids of former infections and are separated by repeats. Upon transcription of those loci small crRNAs (CRISPR RNA) comprising a spacer and parts of the adjacent repeats are incorporated into protein complexes which can cleave either DNA or RNA sequences of invaders (protospacers) when matching the crRNA.

Recently, protospacer constructs matching the spacer 63 of *S. solfataricus* CRISPR locus D were modified in our laboratory in that they circumvented CRISPR - mediated DNA interference by mismatches allowing GU pairing (D63-7U) and even more efficiently, by complementarity to the autoimmunity 5' handle of the crRNA (D63-HA). Transfection studies of *S. solfataricus* M18 with SSV1-based virus vectors harboring these protospacers therefore led to preservation of the vector within the cell but to degradation of the mRNA of the protospacers. Astonishingly, we did not see these effects when transforming *S. solfataricus* M18 with pIZ plasmids harboring the same protospacers. Contrarily, our results indicate DNA interference (although to different degrees) of the protospacers which led to the loss of the plasmid in the culture. Only low numbers of the pIZ-HA construct (maximum $3.1x10^5$ copies/µg DNA) and less than 400 copies for pIZ-7U were detected in total DNA preparations of transformed cells, whereas negative controls without a matching protospacer were stably maintained in the culture with copies of up to $3.18x10^6$ per µg DNA. Surprisingly, growth of transformants and equal chromosomal copies (measured by qPCR) were obtained in all samples, indicating that at least some

cells transformed with matching protospacers have restored their defective *pyrEF* most likely through homologous recombination possibly as a rescue effect upon CRISPR DNA interference. Altogether, our results indicate a different reaction of the CRISPR system upon transformation with virus- and plasmid shuttle vectors harboring the same protospacer.

2 Introduction

2.1 The Archaea – small witnesses of ancient history

"These "bacteria" appear to be no more related to typical bacteria than they are to eukaryotic cytoplasm"

Woese C.R. and Fox G.E.,1977

Life on earth is phylogenetically separated into three domains: The Eukaryotes representing every organism whose genetic material is enclosed by a membrane bound nucleus, Bacteria, and Archaea (Carl R Woese, Kandlert, & Wheelis, 1990). The latter two, summarized as Prokaryotes share the cytological feature of not possessing defined encapsulated nuclei but heavily diverge in their molecular features. This three domain classification did not exist before 1990, when Carl R. Woese`s seminal phylogenetic work emphasizing the separation of prokaryotes into two very distant groups, was postulated. Woese was the first to compare the small-subunit of the ribosomal RNA of organisms on the nucleic acid sequence level (C R Woese & Fox, 1977). Thereby, he elucidated the misclassification of methanogenic "bacteria" being absolutely unique compared to all others. These findings together with additional singularity within the known molecular markers of those methanogens, gave rise to the third domain of life: the Archaea. Not only the definition of another domain of life revolutionized science then but also the assumption of those prokaryotic cells being more closely related to Eukaryotes than to Bacteria (Carl R Woese et al., 1990; Zillig et al., 1992). Thus, Woeses phylogenetic tree of life when rooted to a common ancestor splits into two main branches: the bacterial branch and the Eucarya/Archaea branch (Carl R Woese et al., 1990). This relationship was underpinned by the findings that Archaea, when compared to the other two domains, show striking similarities in their core proteins (e.g. processing machineries for

replication/transcription/translation) to Eukaryotes (Forterre, Brochier, & Philippe, 2002). Since such information processing genes are supposed to be less affected by horizontal gene transfer (HGT) than others (Jain, Rivera, & Lake, 1999) this similarity between Archaea and Eukarya is suggested to result either from being sister groups or less likely from a loss or change of these genes in the bacterial clade when considering an independent divergence of all three domains from the common ancestor (meaning three lineages) (Forterre et al., 2002). Despite the similarities to Eukarya, Archaea have overlapping genes with bacteria in particular concerning metabolism and physiology. However, the question of the relationship between and within the domains is still discussed and the right answer, if ever revealed, is far from being found.

Archaea are not only a sum of "Eukaryotic" and "Bacterial" gene sets, but half of their genes are of unknown function and encode specific features making them unique (Allers & Mevarech, 2005). For example, the cell wall of Archaea lacks the peptidoglycane layer and their membranes show an ether linkage instead of an ester between the fatty acid and the glycerol (Kate, 1993). Moreover, the probably most prominent characteristic of the majority of archaeal members is their ability to thrive in harsh environments, such as highly saline, anaerobic and/or hot habitats. Since such extreme environments are expected to have ruled early earth, microorganisms adapted to those are often suggested to constitute the oldest lineages on this planet (Carl R Woese et al., 1990). This fact was eponymous for the Archaea meaning "ancient" in Greek (Carl R Woese et al., 1990).

Starting with 4 taxa in 1977, new findings and the ability to also classify uncultivable organisms through 16S rRNA sequencing revealed over a thousand of species http://www.ncbi.nlm.nih.gov/) whereof the majority of cultivated organisms is categorized into two predominant phyla: The *Euryarchaeota* and the *Crenarchaeota* (Carl R Woese et al., 1990). Only recently a third phylum was defined as *Thaumarchaeota* harboring exclusively members able to oxidize ammonia (Brochier-Armanet, Boussau, Gribaldo, & Forterre, 2008). *Nanoarchaeum equitans* is a parasitic archaeon which was initially suggested to represent another phylum but recently was shown to belong more likely to the *Euryarchaeota* (Brochier, Gribaldo, Zivanovic, Confalonieri, & Forterre, 2005;

Huber et al., 2002). Additionally, three more phyla have been proposed on the basis of only genomic data, the *Korarchaeota, Aigarchaeota* and *Geoarchaeota*, respectively - although no cultivated representative is yet available (Barns, Delwiche, Palmer, & Pace, 1996; Elkins et al., 2008; Kozubal et al., 2013; Nunoura et al., 2011).

The *Euryarchaota* comprise eight geographically widespread taxonomic classes consisting of the first detected methanogens as well as halophiles and thermophiles (Offre et al., 2012). *Crenarchaeota* are represented by only one class, the *Thermoprotei* being exclusively hyperthermophiles inhabiting hot environments such as hot springs and hydrothermal vents and therefore showing an optimal growth at high temperatures reaching up to 113°C and a pH of sometimes lower than three. These two phyla share the biggest fraction of cultivated species and therefore are prevalent in molecular biology.

2.2 ***Transformation systems under "extreme" conditions-*** *Archaea as genetic models*

Archaea comprise a lot of unique properties making them indispensable models in different research fields. The fact that they show high similarity to eukaryotic information processing machineries elected them as a streamlined tool for the study of replication, transcription or DNA repair (Grabowski & Kelman, 2003; Huet, Schnabel, Sentenac, & Zillig, 1983; Langer, Hain, Thuriaux, & Zillig, 1995). Furthermore, biochemists profit from their "extreme" lifestyles by using thermostable enzymes for crystallization (e.g. Reeks, Naismith, & White, 2013). Many enzymes are harnessed in molecular biology such as the Pfu Polymerase which is generally used in PCR assays (Lundberg et al., 1991). In addition, physiological features such as methane production by methanogens or the ability of *Thaumarchaeota* to oxidize ammonia demonstrate that Archaea play important roles in geochemical cycles. Contrary to biochemistry, genetic studies examining genes, their function and their impact on the cell upon nucleic alterations (e.g. phenotype changes) can only be performed in a living cell. In general, such an *in vivo* system or so-called genetic model organism generally complies with at least three requirements**:** i) easy

growth on media ii) short generation time and iii) amenability to experimental manipulation premising selective transfection with recombinant DNA. Despite cultivation difficulties of extremophiles, some archaeal model organisms exist and transformation strategies comprising transfection methods as well as selective vector systems have been applied for them.

2.2.1 Growth and transformation methods

The first demand "growth on media" already decimates the choice of Archaea as genetic models since cultivation of organisms thriving in harsh environments or anaerobically has a challenge. The sole mesophilic Archaea growing under aerobic conditions (and not being halophile) came up with the new phyla *Thaumarchaeota* (Brochier-Armanet et al., 2008). For example, *Nitrosopumilus maritimus* (Konneke et al., 2005) and *Nitrososphaera viennensis* (Tourna et al., 2011) were isolated from sea and soil, respectively. These organisms are of extensive physiological interest due to their ability of oxidizing ammonia and consequently their crucial role in the nitrogen cycle previously thought to be dominated by bacteria (Stahl & de la Torre, 2012). However, so far no genetic studies have been applied for *Thaumarchaeota,* because they do not grow to high cell densities.

Halophiles

Halophiles are aerobic mesophiles growing in hypersaline environment and are mostly represented as genetic models by *Halobacterium* spp and *Haloferax volcanii*. They display the most straightforward Archaea used in the laboratory due to the aerobic and mesophilic growth. Since they adjust equimolar salt concentrations with their environment through salt-in or salt-out strategies, they grow on defined media covering their salt needs (Leigh, Albers, Atomi, & Allers, 2011,). The first Archaeon ever transformed was *Halobacterium halobium* via a PEG (polyethylenglycole) mediated transformation method whereby spheroplasts gained by reversible removal of the outer paracrystalline glycoprotein surface layer (S layer) were transfected with naked phage DNA (Cline & Doolittle, 1987). This strategy has become the method of choice for halophile transfection owing to the high transformation rate and its broad application.

Methanogens

Mesophilic methanogens are anaerobes and therefore more complex to handle. However, methods based on the initial descriptions of Hungate and Bryant of prereduced media preparation and anoxic handling of cultures (Bryant, 1972) with the sophistication of pressurized cultivation (Balch, Fox, Magrum, Woese, & Wolfe, 1979) led to successful growth of methanogens in the laboratory (Wolfe, 2011). The final breakthrough making genetic studies of anaerobes unconfined possible came with the use of an anaerobic chamber. Most prominent genetic models became *Methanococcus maripaludis*, *Methanococcus voltae* or *Methanosarcina*. For the latter, the eukaryotic approach of lipofection using a cationic artificial lipid (DOTMA, Roche) as shuttle for DNA (Felgner et al., 1987) was successfully applied as transformation system by Metcalf and collaborators (Metcalf, Zhang, Apolinario, Sowers, & Wolfe, 1997). The highest transformation efficiency ($2x10^8$ per µg vector DNA) was yielded by transfecting spheroplast having the membrane as outermost cell surface likely to promote fusion of liposomes (Metcalf et al., 1997). Together with PEG method mentioned above which was also applied for *M. maripaludis* (Tumbula et al., 1994) these two transformation techniques are generally used for methanogens. Furthermore, Bertani and colleagues found *M. voltae* to be naturally competent whereupon transformation efficiencies were increased when using protoplasts instead of whole cells (Patel, Nash, Agnew, & Sprott, 1994). Highest transformation rates were yielded by transfecting protoplasts via electroporation (Patel et al., 1994) (see below). However, this method and heat shock methods turned out not to be efficient for routine work (Allers & Mevarech, 2005).

Hyperthermophiles

Hyperthermophilic *Euryarchaeotes* are also obligate aerobes and are mainly represented by *Thermococcales* with their genera *Pyrococcus* and *Thermococcus* growing at neutral conditions and temperatures above 80°C. Only recently a vector system for *P. furiousus* was established. It can be transfected via heat shock and subsequent $CaCl_2$ treatment (Waege, Schmid, Thumann, Thomm, & Hausner, 2010). Before this, heat shock transformation was used for transformation of *P. furiousus* (Aagaard et al., 1996) and

Thermococcus kodakariensis not being very efficient but elucidating natural competence of the latter (Sato, Fukui, Atomi, & Imanaka, 2003). Transformation of *P. abyssi* applying PEG mediated transformation yielded 10^3 transfectants per µg DNA (Lucas et al., 2002). Within the *Crenarchaeota*, the *Sulfolobales* represented by *Sulfolobus solfataricus, S. acidocaldarius, S. tokodaii* and *S. islandicus* are frequently used for genetic manipulation. They grow aerobically on organic substrates at a pH between 2-3 and high temperatures with doubling times varying from 3 to 8 hours. The first described member of this family was *S. acidocaldarius* isolated from a hot spring in Yellowstone National Park by Thomas Brock who defined the growth medium containing tryptone as carbon source (Brock et al., 1972). For solid media, replacement of conventional agar by a thermostable polysaccharide was inescapable for hyperthermophiles. Gellan gum, naturally produced by *Pseudomonas* species and mainly represented by GELRITE in the laboratories is therefore used for plating of hyperthermophiles (Lin & Casida, 1984). Transformation of *Sulfolobales* is achieved by electroporation where voltage is applied to cells promoting the permeability of the membrane. The specific mechanism of DNA uptake upon the electric shock is not yet clearified, but most likely a transient hydrophilic pore formation enables cationic molecules to pass. The first electroporation protocol for an archaeon was published by Christa Schleper and colleagues in 1992. *S. solfataricus* P1 was transformed with lysogenic SSV1 virus DNA yielding 10^6 transfectants per µg DNA scored by plaque assay (Schleper, Kubo, & Zillig, 1992). That study therefore opened the doors for genetic work on *Sulfolobus* not only by finding an efficient transformation method but also by introducing the first viable genetic system and appropriate quantification for an aerobe hyperthermophilic archaeon (Schleper et al., 1992). This method is restricted to *Sulfolobales* since it cannot be applied for halophiles due to ionic charges of the salt and also failed for most other *Euryarchaeota* (Allers & Mevarech, 2005).

2.2.2 Prerequisites for a feasible vector system

A vector system is a genetic element such as a virus, a plasmid or a transposon used as vehicle to introduce foreign genetic material into a cell. For Archaea, mainly plasmid and viral based genetic tools have been designed and applied for gene overexpressions, knock outs or reporter gene assays (Allers & Mevarech, 2005). The repertoire of genetic

toolboxes for *Euryarchaeota* is versatile mainly because of the presence of accurate selection systems whereby selection leakiness limits the choice of genetic elements for *Crenarchaeota* (see below).

Shuttle vector characteristics

Most genetic vehicles are designed as shuttle vectors. Such elements can propagate within two species due to the presence of two appropriate origins of replication (ori) and two selection markers. In most cases, one replicon originates from *E.coli* for propagation and cloning reasons while the other is functional in the organism to be manipulated. Therefore, genetic elements originally found in an Archaeon are modified to shuttle vectors used for transformation of the same Archaeon being cured from the plasmid beforehand, or for near relatives that don`t harbor the element. Despite the variety of naturally occurring extrachromosomal elements within the Archaea, their usability for shuttle vector construction is constricted to a few owing to the lack of sequence analysis and the only rudimentary understanding of replication mechanisms (Garrett and Klenk, 2007). Therefore, some shuttle vectors have been constructed by fusing the entirety of a sequenced archaeal element more or less randomly with an *E.coli* plasmid part. After selection of non-functionals, novel shuttle vectors have come to light using this approach. For example, pC2A based vectors fused with the R6K *E.coli* replicon were successfully designed and used to transform *Methanosarcina* (Metcalf et al., 1997). Furthermore, extensive empirical studies conducted via random or targeted gene disruption designating plasmid/viral ORFs important for proper replication of the genetic element were performed in case of the SSV1 virus or the pRN1 plasmid leading to the construction of shuttle vectors by omitting or changing indispensable genomic parts (Berkner, Grogan, Albers, & Lipps, 2007; Stedman, Schleper, Rumpf, & Zillig, 1999). For halophile shuttle vectors, similar approaches were used to identify the replication region of the plasmid pHK2 isolated from *H. volcanii* bringing up shuttles such as pMDS20 and pMLH3 (Holmes, Pfeifer, & Dyall-Smith, 1994). So far, only for two archaeal plasmid families, namely pGT5 of *P. abyssi* and pRN1 originally found in *S. islandicus*, all genes encoding enzymes involved in replication have been elucidated (G Erauso et al., 1996; Zillig et al., 1993,

Garrett and Klenk, 2007). pGT5 replicates via rolling circle mechanism and was used to design the several shuttle vectors for *Pyrococcus* (G Erauso et al., 1996; Lucas et al., 2002). Beside the previous mentioned SSV1 based shuttle vectors, pRN1 and the co-existing pRN2 are cryptic plasmids commonly used for shuttle vector construction for *Sulfolobales* (Berkner et al., 2007; Berkner, Wlodkowski, Albers, & Lipps, 2010a; Deng, Zhu, Chen, Liang, & She, 2009). Both are members of the pRN family and were isolated from *S. islandicus* by Zillig and co-workers (Zillig et al., 1993). The advantages of these plasmids with respect to their use as shuttle vectors are their relatively small size (5.3kb for pRN1 and 6.7 kb for pRN2), their applicability within different *Sulfolobales* hosts and the description of the conserved ORFs, namely ORF904, ORF80 and ORF56 (Berkner & Lipps, 2007; Lipps, 2009). The latter is suggested to control the copy number of the plasmid since it was shown to act as a repressor, whereas ORF904 is referred as "replication protein" showing primase, polymerase and sequence specific helicase activity. ORF80 also shows specific DNA binding activity but seems to be dispensable for proper propagation of the plasmid (Berkner & Lipps, 2007; Lipps, 2009). Such detailed characterization led to the construction of several pRN1 based plasmids mainly fused to the pBlueskript *E.coli* vector (Berkner et al., 2007, 2010a) harboring the ColE1 ori and an ampicillin resistance gene (β-lactamase). One of these pRN1 based shuttle vectors is pCmalLacS additionally comprising the *pyrEF* operon of *S. solfataricus* P2 for selection in uracil auxotrophic *Sulfolobus* mutants (see below) and the reporter gene *lacS* of *S. solfataricus* P2 encoding β-galactosidase (Fig.1B). The vector was originally designed for promoter studies in *S. acidocaldarius* and therefore carries a maltose inducible promoter upstream to the reporter (Berkner et al., 2010a). pCmalLacS is depicted below to give a picture of a typical shuttle vector in Archaea (Fig.1B).

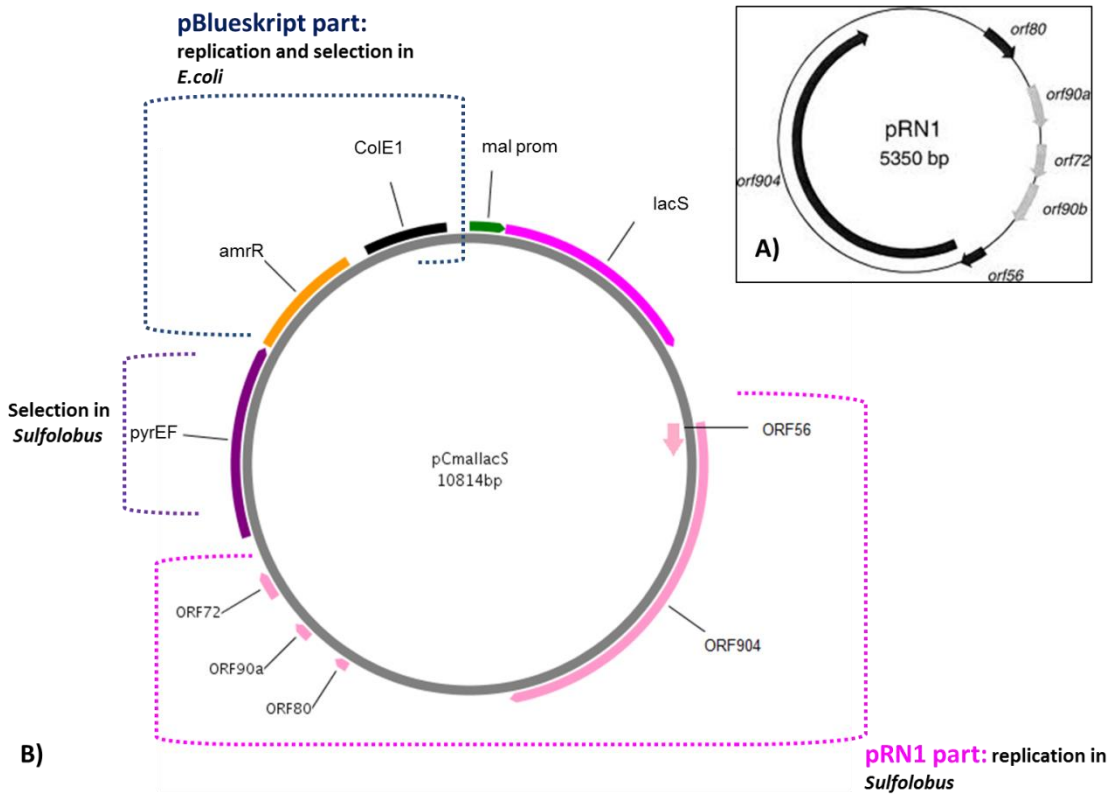


Fig. 1: Map of pRN1 plasmid and pCmalLacS as example for a shuttle vector. A) Map of 5.3kb sized cryptic plasmid pRN1 (Berkner & Lipps, 2007).Three core ORFs are indicated: ORF904 is a replication protein with helicase, primase and polymerase activity; ORF80 is a DNA binding protein of unknown function as well as ORF56, which is supposed to be involved in copy number control (Lipps, 2009). ORF 72,90a and 90b are of unknown function and dispensable for proper replication. **B)** Map of pCmalLacS, a shuttle vector constructed for promoter studies in *S. acidocaldarius* (Berkner et al., 2010a). Parts essential for a conventional shuttle vector are boxed, pink: pRN1 plasmid part comprising the entire plasmid except of ORF90b; essential for replication in *Sulfolobus.* Blue: pBlueskript vector part containing an origin of replication (ColE1) and ampicillin selection marker for *E.coli.* Purple: *pyrEF* selection marker (*S. solfataricus* P2) in uracil auxotrophic *Sulfolobus* mutants. Between pBlueskript and pRN1: maltose inducible promoter upstream to a *lacS* gene encoding β-galactosidase (*S. solfataricus* P2) for promoters studies in *Sulfulobus* (Berkner et al., 2010a). Vector map was created using PlasMapper (http://wishart.biology.ualberta.ca/PlasMapper/).

To simplify standard restriction-ligation cloning, a multiple cloning site composed of a recombinant reporter gene extended by several restriction sites is often introduced into a vector enabling screening for positive uptake of the insert (such as *lacZ* reporter on the pUC vector). A relative new and restriciton – ligation free cloning system has been commercialized by Invitrogen having the advantage of maintaining the orientation of the gene of interest (GOI). This Gateway® cloning procedure comprises two different reactions: the attB-attP and attL-attR recombination reactions based on site specific recombination system of the bacteriophage λ. The GOI is flanked by attB1 and attB2 sites,

respectively, previous to the site specific recombination transfer (attB1 to attP1) into a donor vector harboring attP1 and attP2 sequences. The resulting vector is called "entry clone" and carries the GOI enclosed by attL sites as well as a selection marker. Entry clones are subsequently used for LR *in vitro* recombination with any destination vector possessing a different selection marker and a Gateway® cassette flanked by attR sites. The cassette consists of the *ccdB* suicide gene and a chloramphenicol resistance gene (CmR). Upon LR reaction, the Entry vector interchanges the GOI with the Gateway® cassette and only the expression vector (destination vector+GOI) is able to grow on appropriate selective media (Fig.2). Destination vectors without insert will not propagate, since the *ccdB* suicide gene kills the cells upon expression due to its inhibition of the bacterial gyrase. CcdB induces the topoisomerase II (gyrase) to stably bind the DNA leading to double stand breaks within bacterial chromosomes. Destination vectors therefore can only be propagated in *E.coli* DB 3.1. (Invitrogen™) cells containing a mutation within the subunit A of their topoisomerase II and therefore are not affected by the toxin (Bernard & Couturier, 1992).

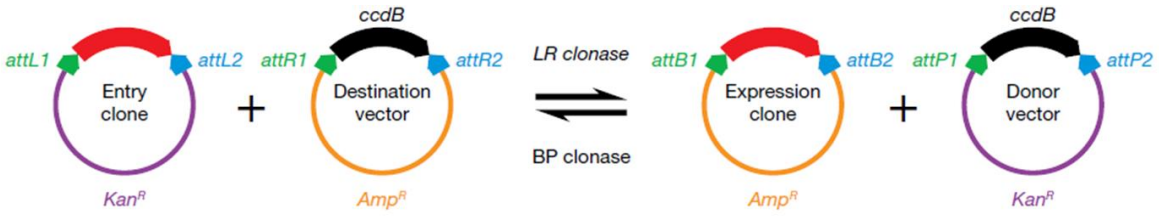


Fig. 2: LR reaction of the Gateway® *in vitro* recombination. An entry clone, established by previous BP reaction harboring the gene of interest (in red) flanked by attL1 and attL2 sites and a kanamycin resistance gene (KanR) is mixed with the destination vector carrying the Gateway® cassette an ampicillin resistance marker (*AmpR*) within *the in vitro* recombination assay. The LR clonase enzyme catalyzes the site specific recombination of attL1 with attR1 and attL2 with attR2 sites between the vectors, respectively leading to the exchange of the gene of interest with the Gateway® cassette made of the *ccdB* suicide gene (in black) and an additional chloramphenicol resistance gene (not cited). Upon transformation of the recombination mixture and plating on ampicillin supplemented media, only cells will grow that harbor the expression clone comprising the gene of interest and the backbone of the destination vector carrying the ampicillin resistance. Destination vectors not having exchanged the Gateway® cassette with the gene of interest are not able to propagate and therefore will not lead to cell growth owing to the toxic effect of the expressed *ccdB* gene. Source: Invitrogen®.

Shuttle vectors that harbor a Gateway® cassette and therefore are destination vectors represent sophisticated tools which can be easily manipulated.

Selection markers

Once a vector is transformed, maintenance of the transformed cells has to be ensured. Therefore, a selection marker gene must be comprised within the vector applying benefits to the transformed cell over the untransformed. Especially in case of low transformation efficiencies due to an inefficient transformation technique, only high selection pressure can increase the yield of transformants. Such a selection marker confers either resistance to a growth inhibiting agent in the media (e.g. antibiotics) or provides a cell function essential for survival disrupted in a mutant cell (e.g. metabolic enzyme). Selected propagation in bacteria is guaranteed by common antibiotics such as ampicillin, kanamycin or chloramphenicol. These kinds of bacterial antibiotics are not effective in Archaea since most attack bacteria specific cell walls or information processing machineries. For *euryarchaeotes,* growth inhibitors are utilized for selection inhibiting general features of life. Puromycin represents such a toxin causing premature chain termination during translation in Prokaryotes as well as in Eukaryotes. It is inhibited by the puromycin transacetylase due to N-acetylation (Jimenez, 1985) making this constellation a common selection system for methanogens (Metcalf et al., 1997; Patel et al., 1994; Possot, Gernhardt, Klein, & Sibold, 1988, Gernhard et al., 1990). Another selection strategy has been proven to be feasible using spontaneous mutated genes resistant to a special drug. The isolated gyrase B gene of a *H. volcanii* mutant is unaffected by the gyrase inhibitor novobiocin and therefore displays a useful resistance gene for other halophiles in novobiocin assays (Bitan-banin, Ortenberg, & Mevarech, 2003)

However, for thermophilic Archaea the choice of such growth inhibitors is limited due to instability of the antibiotic itself and/or the expressed resistant enzyme at high temperatures and low pH (Berkner & Lipps, 2008a). Up to now, only three vector systems based on antibiotics are published for thermophilic Archaea two of which have failed to be reproduced in other laboratories and therefore are not further used (Berkner & Lipps, 2008a). One of these is the alcohol dehydrogenase gene of *S. solfataricus* complementing the sensitivity to butanol or benzyl alcohol used in *P. furiosus* and *S. acidocaldarius* (Aravalli and Garrett, 1997) whereas the other is a thermostabilized hygromycin

phosphotransferase from *E.coli* applied as selection marker for *S. solfataricus* when treated with the translation inhibitor hygromycin B (Cannio, Contursi, Rossi, & Bartolucci, 1998). A relatively young selection system based on the effect of Simvastatin inhibiting 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase involved in lipid biosynthesis of Archaea was found to be suitable for *T. kodakariensis* and *P. furiosus* (Matsumi, Manabe, Fukui, Atomi, & Imanaka, 2007; Waege et al., 2010). Simvastatin treatment results in growth retardation of cells and was shown to be complemented by the overexpression of the affected HMG-CoA gene. Very recently, Zheng and colleagues successfully applied this selection system to the acidophile *S. islandicus* by overexpressing the *hmg* gene from *S. tokodaii.* Besides, they proved Simvastatin to affect other *Sulfolobales* such as *S. solfataricus* and *S. tokodaii* (Zheng et al., 2012). This was underpinned by simultaneous studies where simvastatin concentrations were determined causing growth retardation of *S. solfataricus*, different *S. islandicus* strains, *S. acidocaldarius* and *S. tokodaii.* Within this study, Simvastatin was successfully applied for generating markerless knock out mutants in *S. islandicus* (C. Zhang & Whitaker, 2012a). Due to these results, Simvastatin possibly displays the currently known antibiotic selection system useful in hyperthermophiles. However, availability of antibiotic selection in thermophiles has long been inapplicable and scientists have helped themselves by the use of auxotrophic mutants.

The most prominent metabolic selection system relies on uracil auxotrophy representing the main selection technique used for *Sulfolobus* strains (Berkner & Lipps, 2008a). For this, mutants are established defective in their uracil synthesis upon transcriptional or functional breakdown of their *pyrE* and/or *pyrF* gene products, namely orotate phosphoribosyl transferase and orotidin-5'-monophosphate decarboxylase, respectively. These two enzymes catalyze important steps of the uridine monophosphate *de novo* synthesis pathway in *Sulfolobus* (Grogan & Gunsalus, 1993). The auxotrophy is complemented by functional *pyrEF* genes expressed on the transfected vector. Mutants are gained by growing wild type cells in 5-fluoroorotic acid (5'-FOA) being a uracil analogue and converted to its toxic derivate 5-flurouracil by functional *pyrEF* gene products (Thia-Thong et al., 2002). Consequently only cells mutated in their *pyrEF* gene

cluster are able to survive this procedure (Berkner & Lipps, 2008a). The type of *pyrEF* mutation varies among different strains. In *S. solfataricus*, preferentially integration of insertional elements disrupt the *pyrEF* gene cluster causing the auxotrophic phenotype (Martusewitsch et al., 2000). Spontaneous deletion mutants were observed in some *S. acidocaldarius* strains upon 5`FOA selection (Grogan & Gunsalus, 2003) and *S. islandiucs* showed *pyrEF* dysfunction mainly due to random point mutations (Berkner & Lipps, 2008b). Recently, also for *S. tokodaii* a *pyrEF* mutant was established in our laboratory (Zebec, in preparation). Shuttle vectors for *Sulfolobales*, such as pMJ03 or pCmalLacS (Berkner, Wlodkowski, Albers, & Lipps, 2010b; Jonuscheit, Martusewitsch, Stedman, & Schleper, 2003) harbor the *pyrEF* genes from *S. solfataricus* P2 to complement the auxotrophy. Besides hyperthermophiles, this selection method was also used for some halophiles (Bitan-banin et al., 2003; Peck, DasSarma, & Krebs, 2000) but not for methanogens since they are autotrophic (Allers & Mevarech, 2005).

A second auxotrophic selection system represents growth of a *lacS* deficient mutant on lactose as sole energy source. *lacS* of *Sulfolobus* encodes β-galactosidase (Cubellis et al., 1990, Grogan DW, 1991b) and displays a selection - and reporter gene system rolled into one, since functional β-glycosidase is able to hydrolyse the organic compound X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) resulting in an insoluble blue color. The system is quite limited owing to its dependence on strains able to grow on lactose minimal media. Up to now, it has only been applied for *S. solfataricus* and *S. islandicus* for which solely selection in liquid culture was accomplishable instead of plating due to very slow growth rates (Berkner & Lipps, 2008a).

Choice of the vector backbone – avoiding revertants

A disadvantage of selection based on auxotrophic mutants is, that *pyrEF* revertants have been frequently observed upon transfection of mutants carrying a transposon insertion disrupting either gene (Berkner & Lipps, 2008b). As a result, transformed vectors were lost due to leaky selection. It was supposed but never confirmed that mobile elements possibly mobilize upon electroporation stress leaving an intact gene behind. Solely, *pyrEF* deletion mutants seem to be not as affected from reversion. Unfortunately, only for *S.*

acidocaldarius and *S. islandicus pyrE* deletion mutants are available, so far (Deng et al., 2009; Wagner et al., 2009). A way to overcome this unfavorable effect is to construct a self-spreading shuttle vector based on a virus or a conjugative plasmid backbone. These elements are able to persist without stringent selection since they distribute primary by "infection" of the neighbor cell and are not solely dependent on mitosis. Such advantage has been taken for *S. solfataricus* by using SSV1 virus based vectors (S. Albers et al., 2006; Clore & Stedman, 2007; Jonuscheit et al., 2003; Manica, Zebec, Teichmann, & Schleper, 2011; Schleper et al., 1992; Stedman et al., 1999). Also conjugative plasmids were used in transformation assays showing self-spreading contribution in *Sulfolobus* (Erauso, Stedman, van de Werken, Zillig, & van der Oost, 2006, Elferink MG et al, 1996.). However, conjugative plasmid based shuttle vector systems have not been constructed so far. Self-spreading vectors are well suitable for expression assays (Jonuscheit et al., 2003) but contrary to cryptic plasmids are not useful for gene disruption studies which is a big disadvantage. Therefore when creating a shuttle system, the archaeal vector part has to be distinguished dependent on the study to be conducted.

Modification of the plasmid

Restriction modification systems contribute to the defense systems against foreign nucleic acid. In principle, they comprise two enzymes. One being a restriction endonuclease (RE) recognizing certain sequences on DNA and cutting within or outside of these motifs. Dependent on the enzyme, a blunt (both strands are hydrolyzed at the same base) or overhang (the lagging and leading strand are cut asymmetrically) incision is produced. The host DNA is protected from the hydrolyzing activity by a methylase adding a methyl group to cytosine or adenin (Sambrook et al, .). Beside many bacteria, some Archaea were investigated to comprise such a system and are able to degrade foreign DNA. Upon the first transformation of the target host, shuttle vectors carry the methylation pattern of an *E.coli* species used for final cloning steps. Therefore recombinant *E.coli* species are available expressing methylases corresponding to a RE in the recipient. As an example, *S. acidocaldarius* harbors the RE R.*SuaI* recognizing GGCC motifs but unable to cut in case of methylation of the inner C of this pattern at the N4

position (Grogan & Hansen, 2003; Prangishvili, Vashakidze, Chelidze, & Gabriadze, 1985). Grogan and colleagues identified enzymes such as the specific N4 cytosine methylase M.*Esa*BC41 and M.*Hea*III methylating C5 shown to be successful in protecting the GGCC pattern from R.*SuaI* degradation. The former, being more protective to the R.*SuaI* enzyme than the latter, is recombinantly expressed in the *E.coli* strain ER 1821 (NewEngland Biolabs). For generating stably maintained shuttle vectors for *S. acidocaldarius,* vectors were channeled through this strain prior to transformation (Grogan & Hansen, 2003; Kurosawa & Grogan, 2005). *S. islandicus* REN2H1 expresses *Sui*I digesting GCwGC which is an isoschizomer of *Tse*I from the bacterial species *Thermus*, but unfortunately is not protected by methylation through the corresponding M.*Tse*I (Söllner, Berkner, & Lipps, 2006). Also, *S. tokodaii* seems to harbor an efficient restriction modification system but different from R.*SuaI* as indicated in recent studies where a four-fold increase of transformation efficiency when transforming plasmids directly recovered from *S. tokodaii* over R.*SuaI* protected plasmids was shown (Zebec, in preparation). However, no isolation of most predicted restriction modification systems has been established yet and it has to be considered that there are more systems suppressing efficient transformation than currently known.

2.2.3 Quantification of transformation

There are two ways to accurately determine transformation efficiencies: either by direct plating after transformation or by performing a plaque assay when using an intact viral vector. When quantifying transformants via colony counting, the selection must be very efficient to avoid background growth. In case of *pyrEF* auxotrophs, background colonies are most likely to rise due to traces of uracil in GELRITE used for the overlay (see previous chapter) as well as in tryptone. It was reported that direct plating assays failed for *S. solfataricus* as well as for *S. islandicus* supposedly due to such contaminations (Berkner & Lipps, 2008b; Jonuscheit et al., 2003). Lately, for *S. islandicus* and *S. tokodaii,* tryptone was successfully replaced by the synthetic N-Z-Amine which contains less uracil residues than tryptone abating background growth (Berkner & Lipps, 2008b; Zebec, in preparation). Direct plating of *S. acidocaldarius pyrEF* deletion mutants was shown to be possible and reached up to 10^4 transformants per µg DNA using a pRN1 based shuttle

vector (Berkner et al., 2007). The transformation efficiencies using virus based plasmids, such as SSV1 are generally scored via plaque assay. Since SSV1 is not a lytic virus, plaques form upon growth retardation of the surrounding cell lawn (Jonuscheit et al., 2003; Manica et al., 2011; Schleper et al., 1992). SSV1 based transformation of *S. solfataricus* yielded a maximum value of 10^6 transfectants per µg virus DNA (Schleper et al., 1992).

2.3 Genetic toolbox for S. solfataricus P1 – need for a cryptic plasmid vector

2.3.1 Genetic characteristics of *S. solfataricus*

In 1980 the *Crenarchaeote S. solfataricus* P2 was isolated from a solfataric field in Piscarelli by Wolfram Zillig and colleagues (Zillig et al, 1980). It is a member of the thermoacidophilic *Sulfolobales* and therefore grows aeorbically at temperatures around 80°C and pH between 2 and 3. *S. solfataricus* is an often used lab strain and thrives on organic substrates like tryptone and sucrose. The strain P1 has been the first of its sort amenable for genetic studies due to an applicable genetic tool box (Schleper et al., 1992) and the total genome of its sibling P2 was sequenced in 2001 (She et al., 2001). It is noteworthy that *Sulfolobales* harbor two genome copies during their G2 stage of the stationary phase in the cell cycle (Bernander et al., 1997). *S. solfataricus* also harbors ABC arabinose sugar transporters leading to efficient uptake of arabinose inducing the *ara*S promoter used in several overexpression and promoter studies (Albers et al., 1999; Gudbergsdottir et al., 2011; Lubelska, Jonuscheit, Schleper, Albers, & Driessen, 2006; Manica et al., 2011). With a size of almost 3Mbp, *S. solfataricus* carries the largest genome of sequenced *Sulfolobales* which is made of 11% by mobile elements (She et al., 2001). Since active transposable elements have the feature to change their location within a genome, they are credited to the high plasticity of the archaeon causing a spontaneous mutation rate of 10^{-4} per cell (Martusewitsch et al., 2000). As a comparison, *S. acidocaldarius* not comprising any active mobile elements shows a mutation rate of 10^{-7} (Jacobs & Grogan, 1997; Martusewitsch et al., 2000). These rates

were measured by quantifying spontaneous mutations within the *pyrEF* genes upon 5´-FOA treatment described above (Martusewitsch et al., 2000). This genetic instability compromises genetic studies by a higher rate of revertants in terms of auxotrophic selection procedures (see section “Selection”).

Nevertheless, important *S. solfataricus* P1 *pyrEF* mutants were isolated in the course of mutagenesis studies, e.g. PH1 (Schleper et al., 1994), PH1-16 and P1-18 (Martusewitsch et al., 2000). The PH1 mutant harbors the 1147 bp sized insertional element ISC1217 within the β-galactosidase gene and was expanded to the double mutant PH1-16 or M16 by the additional element ISC1359 disrupting the *pyrF.* In our laboratory, most experiments are conducted using the auxotrophic P1 mutant P1-18,in short M18 of which the promoter region of the *pyrEF* gene cluster is interrupted by the ISC1359 insertional element as in M16 (see results). As already described, *pyrEF* selection is the predominant application for transformation of *Sulfolobales* making those auxotrophic mutants very important strains for genetics.

2.3.2 Genetics in *S. solfataricus* – SSV1 virus - versus plasmid based shuttle vectors

Genetics in *S. solfataricus* are dominated by shuttle vector constructs of the self-spreading virus SSV1 *(Sulfolobus shibatae* virus 1). This fusellovirus was originally isolated from *S. shibatae* and was the first archaeal virus being sequenced (Palm et al., 1991; Yeats, McWilliam, & Zillig, 1982). Not until the plaque assay of Christa Schleper and colleagues elucidated the viral nature of the element, it has been mistaken for a plasmid (Schleper et al., 1992; Yeats et al., 1982). Upon the successful transformation of *S. solfataricus* using SSV1 DNA within the same study, different shuttle constructs based on the virus have been designed. These vectors either comprise the whole genome of the virus, such as pKMSD48, pMJ03, pMJ05, pMJ0305, pMZ (Albers et al., 2006; Jonuscheit et al., 2003; Manica et al., 2011; Stedman et al., 1999; Manica, Zebec, Steinkellner, & Schleper, 2013; Zebec, submitted) or deletion derivates of it, namely pAJC96 (Clore & Stedman, 2007) and pEXSs (Cannio et al., 1998). Also a virus – plasmid hybrid shuttle vector was established by Aucelli and coworkers functional dependent on the presence of

the helper virus SSV2 (Aucelli, Contursi, Girfoglio, Rossi, & Cannio, 2006). However, the sole shuttle vector applicable in other laboratories is pMJ03 and its followers. As the plain virus DNA itself, also this shuttle is able to site specifically integrate into the arginyl – tRNA gene of the host chromosome, beside existing as episomal form in the same cell (Yeats et al., 1982). SSV1 constructs have been successfully applied for transcription-, overexpression-, promoter- and UV inducible studies. Lately even *in vivo* CRISPR studies based on SSV1 shuttle vectors are in progress which is described below (Manica et al., 2013, 2011; Zebec, submitted).

The advantages of those vector systems are overshadowed by their inapplicability to knock out a target gene. Also, they cannot be determined easily because the virus spreads throughout the culture leading to an overrepresentation of episomal copies as DNA also resides in free viral particles. Furthermore, handling with an infective virus in the laboratory represents a contamination risk for other cultures.

All these disadvantageous characteristics could be set aside by the use of a cryptic plasmid replicating in low copy numbers and dependent on the cell cycle. Some attempts were made in constructing cryptic plasmid transformation systems for *S. solfataricus* mainly based on the above mentioned pRN1 plasmid (Berkner et al., 2007). However, transformation of *S. solfataricus* P1 has shown to be difficult since the plasmids if at all transformed, were lost within incubations as demonstrated by Southern blot and the impossibility of retransformation with total DNA (Berkner et al., 2007, Schleper personal communication). This instability is attributed to the weak auxotrophic selection pressure executed on the recipient fostered by residual uracil in media components such as tryptone. As a result, cells fail to keep the plasmid in the culture which is often accompanied by reversion of their mutated *pyrEF* gene cluster to a functional one. Up to now, only for expression studies, *S. solfataricus* P1 mutants were transformed successfully using a plasmid. In any case, positive transformation using this vector stemming from pHZ2lacS (pRN2 replicon) previous applied for *S .islandicus* (Deng et al., 2009), were solely verified by PCR reaction after a relative short incubation time (Lintner et al., 2011).

The selection dilemma was shown to be efficiently encountered by the complete deletion of the *pyrEF* within the recipient and/or by strengthening selection pressure coming from the media. As shown for *lacS* deletion mutants from *S. solfataricus 98/2* (e.g. PBL2025) transformation and even targeted gene disruptions were established using different plasmids (S.-V. Albers & Driessen, 2008; Schelert et al., 2004; Schelert, Drozda, Dixit, Dillman, & Blum, 2006; Worthington, Hoang, Perez-pomares, & Blum, 2003). Adversely, targeted gene disruption of this mutant solely relies on selection of *lacS* which implicates time consuming incubation before plating. So far, no knock out technique has been established using *pyrEF* as selection marker for any strain of *S. solfataricus.*

The hyperthermophile *S. solfataricus* has gained more and more attention in terms of genetic studies not at least due to its extensive use as an archaeal model organism within recently emerged CRISPR studies (see below). Therefore, a reliable genetic toolbox on the basis of a cryptic plasmid to overcome the limits of SSV1 based vectors and to extend genetic studies in *S. solfataricus* is required.

2.4 ***A glance at the CRISPR system –*** *S. solfataricus and its role as a defender*

2.4.1 An innate immune system in *S. solfataricus*

Only recently, conspicuous arrays of repeats in the genome of *E.coli* (Ishino, Shinagawa, Makino, Amemura, & Nakata, 1987; Nakata, Amemura, & Makino, 1989) was revealed as a repertoire of foreign sequences mainly stemming from extrachromosomal elements which acts as an RNA-guided adaptive immune system (Bolotin, Quinquis, Sorokin, & Ehrlich, 2005; Jansen, Embden, Gaastra, & Schouls, 2002; Francisco J M Mojica, Díez-Villaseñor, García-Martínez, & Soria, 2005; Pourcel, Salvignol, & Vergnaud, 2005; Wiedenheft et al., 2012; Barrangou et al., 2007). These arrays are called CRISPR (*C*lustered *R*egularly *I*nterspaced *S*hort *P*alindromic *R*epeats) and represent conserved patterns arranged in defined loci within the genome of half of the bacteria and the majority of Archaea (Mojica et al.,2002; Jansen et al., 2002). Beside cellular

chromosomes, these clusters were also detected on the conjugative plasmids pNOB8 and pKEF9 of *Sulfolobus*, although associated CRISPR specific proteins were not encoded on these plasmids (Greve, Jensen, Brügger, Zillig, & Garrett, 2004; Lillestøl, Redder, Garrett, & Brügger, 2006; She et al., 1998). The structure of a CRISPR locus is conserved and initiated by an AT-rich and repeat specific leader region of about 200-600 bp (Jansen et al., 2002) adjoining a subsequent progression of 21-50 bp sized direct repeats mostly uniform within the loci and separated by unique 34-44 bp long spacer sequences acquired from viral or plasmid – derived DNA. CRISPR loci can vary in number and length and are accompanied by different Cas genes (CRISPR associated proteins) which, with exception of Cas1 and Cas2, have lately been used to classify the loci into three major types (I,II,III) (Makarova, Aravind, Wolf, & Koonin, 2011). *S. solfataricus* comprises six CRISPR loci named A – F and besides the core proteins Cas1 and Cas2, four type I Cas sets similar to the CASCADE CRISPR complex found in *E.coli*, one type IIIA Csm complex, as well as one Cmr complex, referred to as type IIIB. The function of those protein complexes is extensively studied as they act as machineries of the three main CRISPR processes divided into i) spacer acquisition, ii) expression and maturation of crRNA and iii) crRNA guided interference of invading DNA (van der Oost, Jore, Westra, Lundgren, & Brouns, 2009).

Adaption

Within the acquisition process, new spacers from invading elements, so called protospacers, were shown to be integrated into the CRISPR locus at the side of the leader (Barrangou et al., 2007; Erdmann & Garrett, 2012; Lillestøl et al., 2009, 2006; Pourcel et al., 2005; Yosef, Goren, & Qimron, 2012). The exact adaption mechanism is not yet elucidated, but the endonucleolytic Cas1 and Cas2 proteins were suggested to be involved in that process by recognizing a distinct di-nucleotide sequence, namely protospacer adjacent motif (PAM) (F J M Mojica, Díez-Villaseñor, García-Martínez, & Almendros, 2009; Haft et al., 2005) on the invader DNA followed by the excision of a nearby sequence stretch and subsequent integration of the respective sequence into the CRISPR array (Babu et al., 2011; Barrangou et al., 2007; Francisco J M Mojica et al., 2005; Yosef et al., 2012). The PAM itself differs among CRISPR types and is not part of the

integrated spacer but beside its role as recognition motif, seems to guide the orientation of protospacer integration (Barrangou et al., 2007; Lillestøl et al., 2009; F J M Mojica et al., 2009). In *S. solfataricus* recent studies showed an enormous CCN-PAM specific uptake of protospacers (=spacers prior to incorporation) upon infection with an environmental virus mix. Beside the observation of locus C, D and E to be the only loci active in accumulation, Erdmann and coworkers reported random acquisition of spacers throughout the locus E constituting an exception of common "leader proximal acquisition". Furthermore, most spacers seemed to stem from conjugative plasmids which was surprising since the infection mix was dominated by viruses (Erdmann & Garrett, 2012).

The reservoir of acquired spacers which is heritable to the progeny, functions as a "memory of infections" just as in the adaptive immune system of eukaryotes and upon transcription, efficiently cures the cell from an invader genome matching a spacer sequence.

Processing of pre-crRNA

A promoter region is located within the leader sequence of the CRISPR array being constitutively active but inducible upon stress, as it was shown in *E.coli* (Perez-Rodriguez et al., 2011; Pul et al., 2010). One locus is generally transcribed in sense direction forming the pre-crRNA (Pul et al., 2010) which is subsequently cleaved within the repeats into functional crRNAs (CRISPR RNA) by the Cas6 protein in CRISPR types I and III (Fig.3) (Brouns et al., 2008; Carte, Wang, Li, Terns, & Terns, 2008; Lillestøl et al., 2006; Lintner et al., 2011). *S. solfataricus* harbors four Cas6 genes located next to aCASCADE complexes (archaeal CRISPR associated complex for antiviral defense) indicating a possible linkage to them (Lintner et al., 2011). The CRISPR loci E and F in *S. solfataricus* were shown to be silent, i.e. transcribed if at all in low amounts and therefore inactive (Gudbergsdottir et al., 2011). The mature crRNAs are composed of an 8 nucleotide 5' handle derived from the upstream located repeat, the spacer and a 3' handle of variable size stemming from the 3' repeat of the loci. It constitutes the crucial recognition agent in CRISPR interference

(Brouns et al., 2008; Carte et al., 2008; Hale & Duff, 2010; Lillestøl et al., 2006; Lintner et al., 2011; J. Zhang et al., 2012).

crRNA guided interference

Elimination of already registered invaders represents the third, most studied and even practically applied function of the CRISPR system. Within this process, the mature crRNA associates with either type of interference complex determining the target nature of the nucleic acid and leads the enzyme complexes to the invader by specific base pairing to the target (Garneau et al., 2010). *S. solfataricus* shows three interference mechanisms mediated by the aCASCADE (type I), Csm (type IIIA) and Cmr (type IIIB) enzyme complexes (Fig3). Except of IIIB, all so far known types of CRISPR-Cas systems inactivate the invader by degrading the single strand DNA sequence complementary to the crRNA (Garneau et al., 2010; Manica et al., 2013, 2011; L. A. Marraffini & Sontheimer, 2008). The endonucleolytic agents of Cas complexes are not yet verified but it was demonstrated that Cas3 of the CASCADE complex mediates cleavage of ssDNA in *E.coli* (Brouns et al., 2008; Sinkunas et al., 2011). Beside several *in vivo* studies in bacteria, only recent studies in *S. solfataricus* and *S. islandicus* confirmed a degradation of invaders on the DNA level *in vivo* for Archaea (Gudbergsdottir et al., 2011; Manica et al., 2011). Contrary, the IIIB CRISPR/Cmr system attacks the mRNA instead of DNA of the intruder which was proven *in vitro* for *P. furiosus and S. solfataricus* and recently also *in vivo* for the latter (Hale & Duff, 2010; J. Zhang et al., 2012; Zebec, submitted). It is noteworthy, that the Cmr complex in *S. islandicus* which is dissimilar to that of *S. solfataricus* was previously shown to cut DNA in a transcription-dependent manner (Deng, Garrett, Shah, Peng, & She, 2013). *In vitro* studies isolating the Cmr complex of *S. solfataricus* showed the dominance of incorporated crRNAs form loci A and D. Within this study, also the cutting sites were elucidated being favorably AU nucleotides which was underpinned by mapping the degraded protospacer mRNA by RACE experiments in our laboratories (J. Zhang et al., 2012; Zebec, submitted).

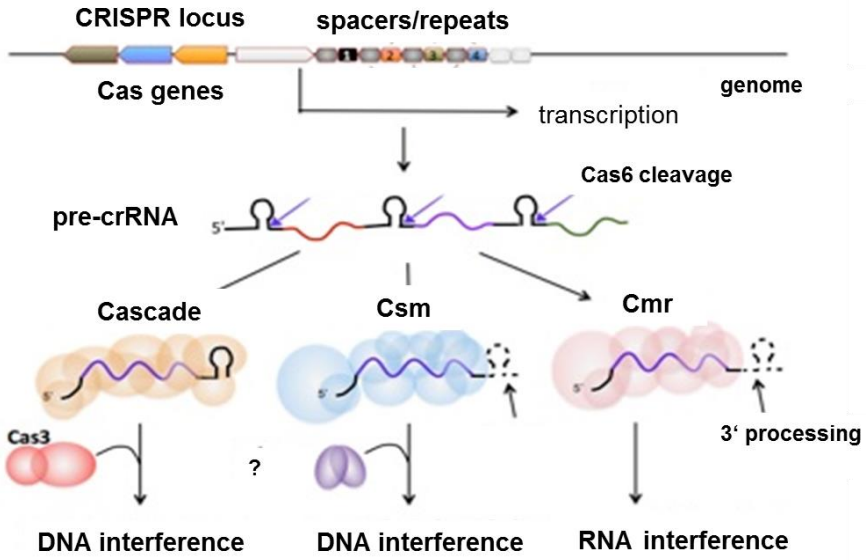


Fig. 3: Schematic overview of the CRISPR pathways in *S. solfataricus*. CRISPR spacer/repeat arrays are transcribed as pre-crRNA and cleaved by Cas6 to mature crRNAs consisting of the 5`repeat (handle), a spacer sequence and remnant 3`repeat mostly forming a hairpin structure. crRNAs can be incorporated into three interference complexes whereby the 3`repeat is further truncated within the Csm and Cmr complex. The crRNA guides the appropriate interference complex to a matching invading sequence and cleavage of the respective is mediated by endonucleolytic enzymes upon specific crRNA-target base pairing, which in case of the Cascade complex was elucidated to be Cas3. The interference complexes determine the nature of the nucleic acid target, since Cascade as well as Csm complexes cleave DNA whereby the Cmr mediates mRNA degradation. Source: modified from www.st-andews.ac.uk.

2.4.2 In vivo interference studies in *S. solfataricus*

Besides seminal *in vivo* studies of CRISPR interference in *Streptococcus thermophilus* (Barrangou et al., 2007) and - *epidermidis* (L. a Marraffini & Sontheimer, 2010a; L. A. Marraffini & Sontheimer, 2008) *as well as E.coli*, there are less than a handful of studies in Archaea and all of them have been conducted with *Sulfolobus* so far.

The first two *in vivo* studies demonstrated CRISPR immunity upon protospacer – spacer binding and the impact of mismatches of the target on successful interference (Gudbergsdottir et al., 2011; Manica et al., 2011). The degree of interference was determined by evaluating the transformation efficiency of the recipient cells either by plaque assay or by direct plating, depending on the nature of the shuttle vector. pMJ - SSV1 based shuttle vectors were used within the study of Manica and coworkers, with protospacer variants of ORF406 – a 37nt sequence derived from the conjugative plasmid pNOB8 matching spacer in the CRISPR locus A of *S. solfataricus* P2 but not in the near relative P1. Infection of P2 cells with pMJ406-0M (perfect matching protospacer) did not

yield any plaques whereas P1 infection showed very high transformation efficiency expressed in turbid plaque formation. Contrary to *in vitro* studies of *P.* showing mRNA to be the target of crRNA (Hale & Duff, 2010), within this study CRISPR interference was shown to be transcription independent i.e. acting on the DNA level since no plaques were established upon infection of pMJ406-0M lacking a promoter.

Similar studies examining protospacers of the ATV virus 100% matching the corresponding spacers in the CRISPR of a *pyrEF* mutant strain of *S. solfataricus* P2 (transposon insertion) and *S. islandicus* (*pyrEF* deletion mutant) were cloned into a plasmid vector pEXA2 (Gudbergsdottir et al., 2011). It should be mentioned here, that only indirect quantification of transformants was possible in these experiments, since direct plating was not feasible due to background colonies owing to residual uracil traces in the media. Colonies were picked and raised in an additional round of selective liquid inoculation. Results show that under selective pressure, only few colonies which showed large deletions ranging over whole CRISPR loci or at least in the corresponding spacer were able to grow. Within this study also mutations in the PAM, previously argued to play an important role in interference (Lillestøl et al., 2009) were introduced leading to an increase of transformation efficiency. The authors therefore claim an intact PAM to be crucial for interference even within a perfect spacer – protospacer match (Gudbergsdottir et al., 2011). Contrary to these findings, no PAM was needed for DNA or RNA degradation in neither of the study within *S. solfataricus* conducted in our laboratory (Manica et al., 2013, 2011; Zebec, sumitted). In contrast, another protospacer adjacent sequence playing crucial role in autoimmunity in other systems, which is also situated upstream of the protospacer was shown to block interference when it was complementary to the crRNA (Manica et al., 2013 and see below). The authors argued that a need for a PAM could be dependent on the nature of the DNA interference complex, since aCASCASE and Csm are proven to be involved in DNA degradation, it could be possible that type I systems are dependent on a PAM whilst the IIIA complex (Csm) is not. This is further underpinned by the initial findings of Marraffini and Sontheimer who showed DNA interference in absence of a PAM in *Staphylococcus epidermidis* (L. A Marraffini & Sontheimer, 2010a).

In order to examine the number of mismatches between the crRNA and the protospacer target tolerated by the CRISPR system in conferring immunity, point mutations were introduced into the protospacer and transformation efficiencies were compared to perfect matches (Gudbergsdottir et al., 2011; Manica et al., 2011). It was shown that contrary to the findings in *S. thermophilus* indicating one mismatch being sufficient to abolish interference (Deveau et al., 2008), more mismatches are tolerated in *S. solfataricus* (Gudbergsdottir et al., 2011; Manica et al., 2013, 2011). Moreover, recent studies indicate the position of the point mutation to be relevant as protospacer mismatches at the 5'end preceding the repeat handle of the crRNA (maximum of three mismatches) were less tolerated whereas considerable 15 mismatches at the 3' end still yielded 40% interference (Manica et al., 2013). These findings underline the relevance of a matching 5` end of the crRNA which is reminiscent of the need for a matching "seed" sequence representing the first 8 nucleotides of the 5' end of the *E.coli* and *Pseudomonas aeruginosa* crRNA indispensable for interference (Semenova et al., 2011; Wiedenheft et al., 2011).

Three years ago one of the big CRISPR mysteries, namely how autoimmunity is avoided during a CRISPR interference reaction was elucidated via *in vivo* studies in *S. epidermidis* (L. A. Marraffini & Sontheimer, 2010b). Since the protospacer target contains the same sequence as the spacer in the chromosome, the CRISPR system must be able to discriminate between the own chromosome and the non-self-invader DNA to protect its own genome. This is achieved by base pairing of at least four nucleotides of the 8 nt 5' handle of the crRNA derived from the upstream repeat of the spacer to the respective sequence within the genome. Since invader sequences do not display the repeat sequences, the 8nt handle does not bind licensing degradation (L. A. Marraffini & Sontheimer, 2010b). In *S. solfataricus* this autoimmunity sequence was recently referred to as protospacer adjacent sequence (PAS). It was shown that at least three matches with the 8 nucleotide 5' handle of the crRNA are needed to lead to protection (Manica et al., 2013). This seminal function of the CRISPR system has been recently exploited for CRISPR mediated RNA interference studies (Zebec, submitted). Since *S. solfataricus* possesses a type IIIB Cmr system shown to be able to degrade mRNA of the invader instead of DNA

(Hale & Duff, 2010; J. Zhang et al., 2012), a PAS sequence was cloned upstream of a protospacer matching the spacer 63 of locus D, whose corresponding crRNA was shown to be most abundant in the Cmr complex (J. Zhang et al., 2012). These nucleotides provided that DNA attack would be circumvented. Furthermore, the spacer-matching sequence of the protospacer was modified by exchanging seven C bases with T bases leading to T-G mismatch in the DNA:crRNA hybrid and therefore negatively affecting DNA interference. On the mRNA level U-G can form a relatively stable binding (wobble base pairing), so that crRNA would be able to fully bind the mRNA of the protospacer whereas on the DNA level, the mutations would have a bigger impact. Two GU-pairing protospacer constructs were tested: D63-HA additionally binding 8 nucleotides of the PAS (HA stands for handle) was almost 100% protected from DNA interference and showed 43% of RNA interference, and the construct 7U without a handle match, showing 68% circumvention of DNA degradation and therefore only 26% RNA interference (Fig.4 A,B) (Zebec, submitted). This study demonstrated for the first time CRISPR mediated RNA interference in a prokaryote *in vivo*. Furthermore, an artificial mini-CRISPR locus targeting the β-galactosidase of the chromosome was established within this study which showed up to 50% directed mRNA degradation of the gene revealing the ability to exploit the sophisticated CRISPR mechanisms to silence genes in *S. solfataricus* (Zebec, submitted).

A)

B)

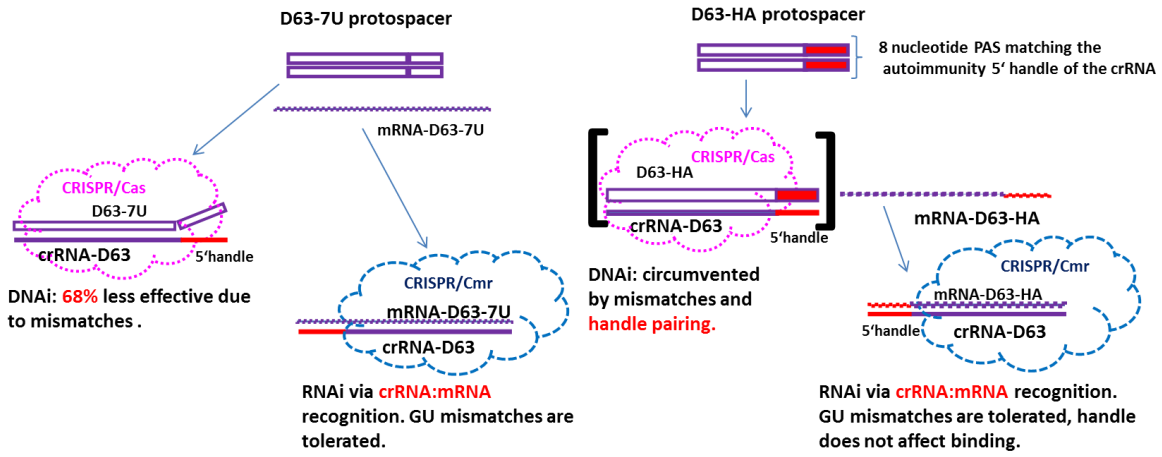


Fig. 4 Protospacer sequences of D63-7U and D63-HA construct used in this study and an overview about the putative interactions between the D63 protospacer and the CRISPR/Cas and CRISPR/Cmr systems, respectively examined in further studies using SSV1 based plasmids. A) Depiction of the 37nt protospacer D63 dsDNA sequences with different PAS (protospacer adjacent motifs) at the 3' end; D63-HA (beneath) carries a 8 nucleotide PAS (dashed blue box) which perfectly matches the 5` handle of the D63 crRNA; D63-7U (above) carries a not-matching PAS (light grey). Red and boxed nucleotides: mismatches (cytosine to thymin substitutions) to the chromosomal spacer. **B)** Schematic overview about the putative crRNA – protospacer interactions of the protospacers 7U and HA based on recent infection studies of M18 using the SSV1 virus based shuttle vectors (Zebec, submitted). CRISPR/Cas complex: suggested to carry out DNA interference; CRISPR/Cmr complex: putative RNA interference. D63-HA: PAS matches crRNA 5' handle leading to efficient protection from DNA interference. D63-7U: no matching PAS, DNA interference is circumvented to 68%.

2.5 Aim of the study

The genetic toolbox of the Crenachaeon *S. solfataricus* P1 has so far been dominated by SSV1 virus based vector systems representing versatile genetic elements. The high transformation efficiencies easily determinable by plaque assay and the stable maintenance in high copy numbers within the cells is attributable to the infectious nature of the virus. It is precisely this spreading nature that in turn renders the system unusable for doing delicate genetic manipulations such as targeted gene disruptions dependent on reliable selection. In addition, the copy number of viral DNA can be highly variable which is unfavorable in expression studies.

Therefore, this study concentrated on the design and application of a novel genetic system for the archaeon *S. solfataricus* P1 based on a cryptic plasmid. Previous studies have shown plasmid transformation of *S. solfataricus* P1 to be difficult attributed to the leakiness of the *pyrEF* selection in disruption mutants (i.e. insertional element within the *pyrEF*). Residual uracil traces in tryptone media were reported to cause background growth of cells and wild type revertants arose for undefined reasons leading to the loss of the plasmid. We aimed to find beneficial electroporation settings and media composition, where we used synthetic N-Z-Amine instead of tryptone for the transformation and further incubation of the *S. solfataricus pyrEF* mutant M18. For construction of the shuttle vector, the pRN1 based plasmid vector pCmalLacS was chosen as starting point as it is assumed to be able to replicate in *S. solfataricus*. The shuttle vector itself should be easy to handle and therefore we intended to generate a Gateway® destination vector by introducing an arabinose inducible Gateway® cassette amenable for a rapid one-step integration of a target gene via site specific recombination. The goal was to test whether the M18 *S. solfataricus* mutant was transformable with the plasmid vector carrying a desired insert and if yes, in which quantities the plasmid is present in the cells.

To further investigate the application of the plasmid in genetic research, we aimed to introduce the previously examined CRISPR protospacer constructs D63-7U and D63-HA

into the plasmid to elucidate whether the interference results would be congruent to those of previous studies carried out with the SSV1 spreading virus vector. The two protospacers match the spacer 63 of locus D in M18 and had been shown to prohibit CRISPR/Cas mediated DNA interference in different degrees owing to a complementary handle sequence (HA construct) and GU pairing (both constructs) (see above). This abortion of virus degradation led to maintenance of the respective but to CRISPR/Cmr interference acting on the mRNA of the protospacers.

Altogether, the goal of the study was to construct a plasmid based shuttle vector system for the archaeon *S. solfataricus* P1 which is amenable for cloning and carries an arabinose inducible promoter upstream to the integration site. To gain insights whether CRISPR interference in M18 changes upon the nature of the used shuttle vector (plasmid versus virus), protospacer constructs had to be placed into the newly developed plasmid vector and outcomes to be compared to the results obtained with the viral vector (Fig.4).

3 Materials and Methods

3.1 Strains and growth conditions

3.1.1 *Sulfolobus* strains

S. solfataricus P1 (DSM-1616) *pyrEF* mutant strain M18 and *pyrEF*/*lacS* doublemutant strain M16 (Martusewitsch et al., 2000) as well as *S. tokodaii* strain 7 (DSM-16993) *pyrEF* mutant c92 (Zebec, in prep.) were used in this study. For starting cultures (untransformed), 300µl cells from glycerol stocks were used for inoculation comprising 10x concentrated cells previously sampled at an OD of 0.2 (exponential growth phase).

For liquid cultures, cells were incubated in 50 ml Brock`s basal salts medium at pH 2-3 (adjusted with 50% H_2SO_4) (Brock et al., 1972) whereby for raising of auxotrophic mutants prior to electroporation and first inoculation rounds of transformants, tryptone (BD Biosciences) at 0.1% (w/v) and sucrose at 0.2% (w/v) (Serva) were supplemented as nitrogen and carbon source, respectively. Second incubation rounds of transformed cells were carried out in the same media but with exchange of tryptone for N-Z-Amine AS (Sigma) at same concentrations. Transformed *S. tokodaii* c92 mutants were also grown in Brock`s basal salts medium at same pH, whereby N-Z-Amine AS was used from the beginning on and (+) D-Glucose (Roth) was supplemented as sugar source at 0.2%. For raising untransformed uracil auxotrophic mutant strains, additionally uracil was added to the media at 0.0125 mg/ml to complement the *pyrEF* dysfunction.

For arabinose (AppliChem) induction studies (see Northern blot analysis), the sucrose concentration in the media was reduced to 0.04% and for induction of samples, 400µl of 20% arabinose (final concentration 0.16%) was added to the media.

All cultures were grown in long-necked conical flasks with continuous shaking at 150 rpm (Innova 3100 oil bath shaker, New Brunswick by Eppendorf) at 78°C.

3.1.2 Measurement of cell density

Cell density of *Sulfolobus* cell suspensions were determined spectrophotometrically by measuring the absorbance at a wavelength of 600nm (Beckam Coulter, DU® 800 Spectrophotometer).

3.1.3 *E.coli* strains

E.coli TOP 10 (Invitrogen™), *E.coli* DB3.1 (Invitrogen™) and *E.coli* ER 1821 (NewEngland Biolabs) competent cells were used in this study. Between 70 – 25µl of cells were used directly from the glycerol stock for transformation (see Transformation). Transformation reactions were plated LB agar media (Bertani, 1951) supplemented with ampicillin at 50µg/ml, chloramphenicol at 25µg/ml or streptomycin at 100µg/ml, dependent on the selection marker located on the transformed vector. In general, *E.coli* DB 3.1 cells were used only for pCAra-GW cloning and - propagation and therefore were only grown on chloramphenicol supplemented LB-agar. For incubation of colonies picked from the agar plate, 5ml liquid LB media supplemented with the appropriate antibiotic were used.

Plates were incubated over night (o/n) and liquid cultures in an air shaker at 250 rpm and 37°C.

3.2 Primers and probes

Tab.1: Primers used in this study. Primer pairs are listed in one line. Underlined sequences are restriction enzyme recognition motifs. Primers marked with an asterisk (*) were designed especially for this study.

Primer pairs	Sequences (5' -3')	Template and application	amplification size
406Seq_FW 406Seq_RV	ATAAGTTTGATGGGGCAGCA TGACCATGATTACGAATTCGA	406Seq_FW binds *pyrEF* part and 406 Seq_RV primer binds the pUC19 part of pDEST-MJ-Ara vector; used for Gateway amplification prior to cloning, sequencing and check PCR on pCAra-GW and pIZ-vectors	2300bp
GW-internal*	AGACGACGGGCTTCATTCT	binds in the middle of the Gateway cassette; used for sequencing	
ORF904_FW* ORF904_RV*	ACAAGAAGAACGGGGGTG ACCTCTTCAGCAATCGCCT	bind ORF904 located on the pIZ-vector backbone; used for check PCR and Southern blot probe	1030bp
pyrB_FW* pyrE_RV*	TCTCTGGTCAACTCAAGCGA GAGTTCCCTTATCGGCATTC	bind *pyrB* and *pyrE* located on the pIZ-vector backbone and in the genome of *S. solfataricus*; used for PCR assays	326bp and 1658bp (disrupted)
qlacS_FW1 qlacS_RV1	AGCCAACCACTTGGTCCAGTAA GGCTGGAATGAGCTATTAGCGT	bind *lacS* gene of *S. solfataricus;* used for check PCR	300bp
Q1 406_FW pCflanking_FW*	GGTTACATAAGGCTCTGTCGAGG AGTCCTGTCGGGTTTCGC	bind pIZ-protospacer constructs, Q1 406_FW binds the ORF 406 and pCflanking binds the pIZ vector backbone; used for check PCR	10836bp
Q1 406_FW Q1 406_RV	CACTCTACGCTGGGCAGACATCT TTGGGAACCTGGCTAGACCTTCA	bind ORF 406 protospacer insert; covering the D63 protospacer; used for pPCR	200bp
Pro-FW1 Pro-RW2	AATTCCCACAATTGCCAAAG TTTCGGACTTTTCCACCAACT	bind ORF 406 protospacer insert, not covering the D63 protospacer; used for dsDNA Northern blot probe	333bp
D-291-FW D-291-RV	ACTATAGCCTTAACGCAGAAGGGT TAGTTGTGTGCCCGCAAAACTG	bind SSV1 and SSV1 based vector backbones, coved sequence was used as dsDNA probe for Southern blot	329bp

All primers created in this study (*) were designed using the Gene Runner 3.0 software and ordered from the Eurofins MWG Operon company (Ebersberg,Germany).

Tab.2: Probes used in this study. Probes marked with an asterisk (*) were designed especially for this study.

Probes	Primer used for probe amplification	size
ORF 904 pIZ-plasmid Southern blot probe*	ORF904_FW,ORF904_RV	1030bp
SSV1 virus Southern blot probe*	D291-FW, D291-RW	329bp
ORF 406 Northern blot probe	Pro-FW1, Pro-RW1	333bp

3.3 Standard Methods

3.3.1 Transformation of *E.coli*

Transformation of *E.coli* was conducted via heat shock (42°C) assay according to the company`s protocol (Invitrogen™), whereby 25-35 µl of chemical competent cells were mixed with 1-7µl of *in vitro* recombination reaction, ligation suspension or pure plasmid. Transformation suspensions were recovered under continuous shaking (300rpm) in 200µl S.O.C media (Invitrogen™) at 37°C before plating.

3.3.2 Extraction procedures

3.3.2.1 Plasmid extractions

For plasmid extraction, colonies were picked from the plate and incubated in 5ml LB-agar/antibiotic media o/n. The suspension was centrifuged at full speed and 4°C for 15 minutes whereupon the supernatant was removed and the plasmids were recovered

from the cell pellets using the E.Z.N.A® Plasmid Miniprep Kit (Omega Bio-Tek, Doraville, GA, USA) following the Plasmid Mini II Spin protocol.

3.3.2.2 DNA extraction from Sulfolobus

For *Sulfolobus* DNA extraction, cell suspensions of an optical density of 0.2 were harvested whereof 7ml were centrifuged at 4500 rpm and 4°C for 10 min (Universal 320R, *Hettich Zentrifugen)*. Cells were lysed by resuspending the pellets in 500µl TEN buffer followed by the transfer in a fresh 2ml tube and subsequent mixing with 500µl TENST buffer (Stedman et al., 1999). The lysis reaction was vortexed roughly and kept on ice for 30 min with intervals of vortexing in-between. 500µl Phenol:Chloroform:Isoamylalcohol (25:24:1, pH=8) was added, vortexed and centrifuged at 13000 rpm (Eppendorf centrifuge,5415 R) at 4°C for 15 min. The upper phase of the sample was transferred into a fresh Eppendorf tube and 500µl of Chloroform were added, vortexed and centrifuged for 15 min. under same conditions. This step was repeated using the upper phase of the previous purification step. DNA was precipitated by transferring the upper phase into 700µl isopropanol and 70µl of Na-acetate was subsequently added. The reaction was kept at 4°C for at least 30 min and centrifuged again at 13000 and 4°C for 30 min. The supernatant was decanted and 500µl of 80% ethanol (EtOH) were added to the pellet without resuspending it. The mixture was centrifuged at same conditions for 10 min. The supernatant was completely removed and the pellet was dried from ethanol by using a speed vacuumizer (Concentrator Plus, Eppendorf). 100µl of DEPC water was used to resuspend the DNA pellet and 5µl of Solution 1 containing RNase (Omega Bio-Tek, Doraville, GA, USA) was added.

3.3.2.3 RNA extraction from Sulfolobus

For RNA extraction, cells with an OD_{600} of 0.2 were sampled to 7ml and centrifuged at 4500 rpm (8 Universal 320R, *Hettich Zentrifugen)* and 4°C for 15 min. The supernatant was decanted and the pellet was subsequently stored at -80°C. Unthawed pellets were used for RNA extraction which was achieved with the mirVana™ miRNA Isolation Kit (Ambion®). Extracted RNA was checked on a 1.5% agarose gel for integrity.

3.3.2.4 Quantification of nucleic acids

The concentration and purity of extracted plasmid or total DNA as well as RNA was spectrophotometrically determined by NanoDrop (Peqlab) whereby the wavelength absorbance of the nucleic acids at 260nm was measured.

3.3.3 Polymerase chain reaction

3.3.3.1 Standard protocols for PCR

Polymerase chain reaction (Mullis, 1983) was used for amplification of the Gateway® cassette (see below) as well as verification of cloning-/*in vitro* recombination reactions and detection of the plasmid in total DNA extracts. Three different set ups were used dependent on the polymerase type (Table 3). For PCR directly using DNA as template, 1ng of plasmid DNA or 10ng of total DNA were used. For most PCRs, 30 cycles were chosen meaning that the progression of denaturing of dsDNA (step 2), annealing of primers (step 3) and extension of the amplicon (step 4) by the polymerase was repeated 30 times. Annealing temperature was dependent on the primer pair used, but ranged between 55 and 57°C. The elongation step was chosen according to the template length and the rapidity of the polymerase (Tab. 3). PCR amplifications were generally checked on agarose. All PCR reactions were carried out in a TProfessional/Standard thermocycler (Biometra).

Tab.3: Standard PCR protocols used in this study. Annealing temperature (step 3) and extension time (step 4) is dependent on the primer and the template, respectively. (*) alternatively GoTaq 5xBuffer (Promega) was used in combination with this polymerase. Extension time in the PCR program depends on the size of the template and the rapidity of the DNA polymerase, respectively (cited in the column "polymerase type"). 1µl of template were applied when DNA templates were used, whereby for plasmids 1ng and for total DNA 10ng were used. For colony PCR, a colony fraction direct from the plate was used as template. Linked steps in the PCR program are repeated dependent on the chosen cycle number.

Polymerase type	Set up	Program
Phusion Polymerase (Thermo Scientific®) → blunt end generation; 1000bp/30sec Used for high fidelity assays	0.5µl 10µM Primer FW 0.5µl 10µM Primer RV 0.5µl 10mM dNTPs (Thermo Scientific) 5µl 5xGC Phusion Buffer (ThermoScientific) 0.2µl Phusion Polymerase (2U/µl) 1µl template ad 25 µl DEPC water	(1) 98°C – 2 min (2) 98°C – 15 sec (3) 55-57°C – 20 sec (4) 72°C – x (5) 72°C – 5 min (6) 4°C - ∞
GoTaq Polymerase (Promega®) → generating a 3' A-overhang, 1000bp/min Used for control PCRs	0.5µl 10µM Primer FW 0.5µl 10µM Primer RV 0.5µl 10mM dNTPs (Thermo Scientific) 5µl 5xGoTaq Buffer (Promega) 0.15µl Taq Polymerase (5U/µl) 1µl template ad 25 µl DEPC water	(1) 95°C – 2 min (2) 95°C – 15 sec (3) 55-57°C – 20 sec (4) 72°C – x (5) 72°C – 5 min (6) 4°C - ∞
DreamTaq Polymerase (Thermo Scientific®) → generating a 3' A-overhang; 1000bp/min Used for Colony PCRs	1.5µl $MgCl_2$ 0.5µl 10µM Primer FW 0.5µl 10µM Primer RV 0.5µl 10mM dNTPs (Thermo Scientific) 5µl 5xDreamTaq Buffer * (Thermo Scientific®) 0.15µl DreamTaq Polymerase (5U/µl) Colony as template ad 25 µl DEPC water	(1) 95°C – 5 min (2) 95°C – 15 sec (3) 55-57°C – 20 sec (4) 72°C – x (5) 72°C – 5 min (6) 4°C - ∞

3.3.3.1 Colony PCR

Colony PCR was performed directly on colonies to detect *E.coli* transformants harboring the desired plasmid without previous DNA extraction. For this, the DreamTaq polymerase assay was used preferentially (Tab. 3), whereby a small amount of a single colony was picked with a sterile pipette tip and was directly transferred into the appropriate PCR tube. After short incubation, the tip was removed and subsequently put into 5ml LB media supplemented with the appropriate antibiotic and incubated at 37°C in an air shaker o/n. The colony PCR was checked on agarose gel and positive samples were plasmid prepped.

3.3.3.2 Purification of PCR reactions

PCR products which were sent for sequencing or were used for cloning (Gateway® cassette) were prior purified using the NucleoSpin® Gel and PCR Clean-up Macherey-Nagel kit.

3.3.4 Restriction digests

Restriction enzymes, appropriate assays and applications of this study are summarized in table 4. All reactions were incubated at 37°C for 1 hour and heat inactivated at 65°C for 20 min. *Sac*II and *Eag*I assays were exclusively used for construction of pCAra-GW vector. *Eco*RI restriction reactions were performed to check the integrity of vectors and to verify transformation reactions. Total DNA used in Southern blots was treated with *Ase*I (Tab. 4). Restriction digests were generally visualized on agarose.

Tab.4: Restriction enzymes and appropriate protocols used in this study.

Restriction enzymes and recognition sites	Applications	Set – up
*Eco*RI (Thermo Scientific) 5' GAATTC 3'	integrity check and verification of retransformation	20µl total volume: 500ng Plasmid DNA 2µl Buffer EcoRI (fin.conc. 5x) 1µl *Eco*RI (fin.conc.) ad DEPC water
BglII (Thermo Scientific) 5' AGATCT 3'	Integrity check	20µl total volume: 500ng Plasmid DNA 2µl Buffer 0 (fin.conc. 5x) 1µl *Eco*RI (fin.conc.) ad DEPC water
SacII/Cfr421 (Thermo Scientific) 5' CCGCGG 3'	Cloning	50µl total volume: 2µg Plasmid DNA 5µl Buffer B (fin.conc. 5x) 1µl SacII (fin.conc.) ad DEPC water
EagI (New England Biolabs) 5' CGGCCG 3'	Cloning	50µl total volume: 2µg Plasmid DNA 5µl Buffer R (fin.conc. 5x) 1µl *Eag*I (fin.conc.) ad DEPC water
*Ase*I (New England Biolabs) 5' ATTAAT 3'	total DNA digest for Southern blot	50µl total volume: 2-18µg total DNA 5µl NEBuffer 3.1 (fin.conc. 5x) 1µl *Ase*I (fin.conc.) ad DEPC water

3.3.5 Agarose gel electrophoresis

3.3.5.1 Gel electrophoresis

To verify the size of PCR reactions and/or restriction digests, samples were prior mixed with loading dye in 6:1 ratio (6x loading dye, Thermo Scientific) and loaded on an 0.8% - 1.5% agarose gel (appropriate amount of agarose dissolved in 1xTEA buffer) which was supplemented with Ethidium bromide (0.1 µg/ml) and placed into the electrophoresis chamber filled with 0.5xTEA. Conventional gel electrophoresis was performed at 100-

120V whereby size specific separation of nucleic acids is achieved based on the migration of the negatively charged nucleic acids towards the positive pole (Sambrook et al., 2001). To determine the size of the sample, a DNA ladder (GeneRuler 1kb or 1kb+, Thermo Scientific) containing nucleic acid fragments of standard size was loaded as marker. Gels were analyzed using the Gel doc XR (BioRad Laboratories) visualizing the Ethidium bromide stained nucleic acids by UV irradiation. The pictures were analyzed using the Quantity 1-D analysis software (BioRad Laboratories).

3.3.5.2 Purification of nucleic acids from agarose gels

To recover specific nucleic acid fragments separated by gel electrophoresis from the gel, the desired band was cut out under UV irradiation using a sterile razor blade and adequate face protection. The excised fragment was purified according to the protocol of the NucleoSpin® Gel and PCR Clean-up Macherey-Nagel kit.

3.4 Cloning procedure generating the pCAra-GW vector

3.4.1 Starting vector constructs

3.4.1.1 pCmalLacS

The pRN1 based shuttle vector pCmalLacS (Berkner et al., 2010) (Fig. 5A) was used to generate the destination vector pCAra-GW amenable for Gateway® recombination. pCmalLacS consists of a pRN1 plasmid part comprising all important genes of which the products are involved in successful replication, a *pyrEF* selection marker ensuring complementation of uracil auxotrophy in appropriate mutants as well as a pBlueskript vector part harboring ColE1 ori and an ampicillin resistance gene (*bla* gene) enabling propagation in *E.coli*. Additionally, the vector consists of a maltose inducible promoter situated upstream to the *lacS* gene of *S. solfataricus* P2, expressing β-galactosidase. The *lacS* gene is enclosed by a terminator sequence. To generate the pCAra-GW vector the Gateway® cassette harboring the arabinose inducible promoter *ara*S upstream of the

attR1 site, was cloned into the pCmalLacS after the removal of the maltose promoter and the *lacS* gene.

3.4.1.2 pDEST-Ara

The SSV1 virus shuttle vector pDEST-Ara (Zebec, submitted) harbors a full length Gateway® cassette (Invitrogen) bordered by the attR sites and an arabinose inducible promoter *ara*S upstream of the attR1 site (Fig.5B). The Gateway® cassette consists of a *ccdB* suicide gene encoding the CcdB protein and CmR selection marker conferring chloramphenicol resistance to positive transformants. The vector was used as template to amplify the Gateway® cassette together with the arabinose promoter sequence.

A) **B)**

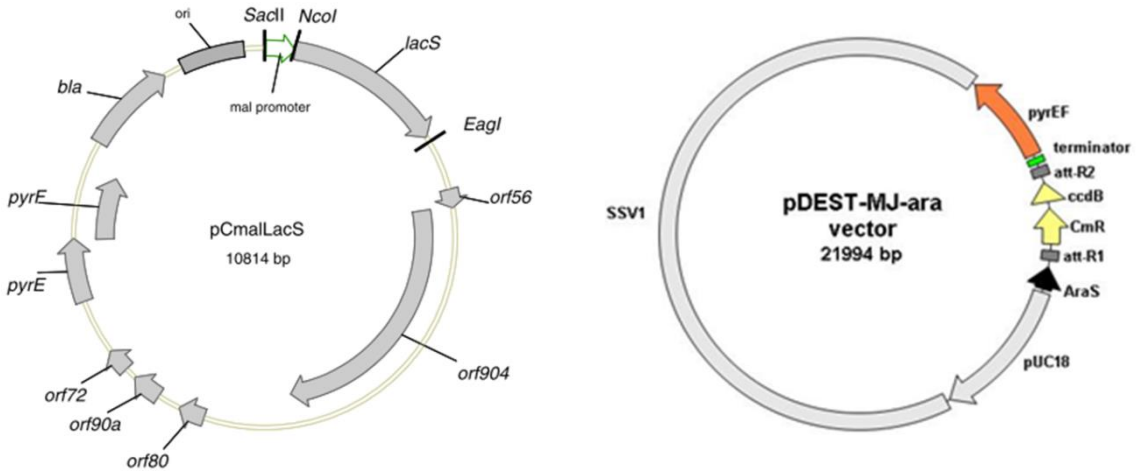


Fig. 5: Vectors used for constructing pCAra-GW. A) Vector map of pCmalLacS is a pRN1 based shuttle vector harboring a *pyrEF* selection marker for *Sulfolobus* auxotrophic mutants, a pBlueskript vector part (ampicillin resistance and ColE1 ori) and a *lacS* reporter gene inducible by maltose. Source: Berkner et al., 2010. **B)** Vector map of pDEST-MJ-ara, a SSV1 based shuttle vector comprising the *pyrEF* selection marker and a Gateway®cassette harboring the *CmR* (chloramphenicol resistance) gene and the *ccdB* suicide gene. Upstream to the attR1 site of the cassette, an arabinose promoter is located. Source: Zebec, submitted.

3.4.2 Removal of the maltose promoter and *lacS* gene via restriction digestion

To remove the *lacS* gene and the maltose promoter, pCmalLacS was first linearized using the restriction enzyme *Sac*II cutting directly upstream of the maltose promoter and leaving a 3' overhang (Fig. 6). The reaction was performed according to the protocol in

Table 4. The product was purified (see *3.3.3.2*) and the plasmid was quantified (see *3.3.2.4*).

500ng of the linearized vector were then digested using *Eag*I having the recogniton site immediately adjacent to the *lacS* gene producing a 5' overhang (Tab. 4). The digest was conducted according to the protocol in table 4 and loaded on 0.8 agarose. The piece removed from the plasmid was 1817 bp in size. The plasmid backbone was purified from the agarose gel (see *3.3.5.2*) and used for further cloning.

3.4.3 Amplification of the Gateway® cassette + arabinose promoter

The fragment comprising the Gateway® cassette + *ara*S to be ligated was amplified from the pDEST-Ara vector using the 406Seq_FW primer binding the last nucleotides of the *pyrEF* gene adjacent to the attR2 site and the 406Seq_RV primer covering the first 20 nucleotides of the pUC18 part downstream of the arabinose promoter, comprising a *Eco*RI recognition site (Tab. 1). The PCR amplification was achieved according to the protocol listed in table 3 for Phusion Polymerase whereby a blunt end PCR product was produced. The elongation time was chosen to be 75 sec, since the amplicon was 2200bp in length and 31 PCR cycles were set. An aliquot of the PCR product was checked on 0.8% agarose and the rest was purified (see *3.3.3.2*).

3.4.4 Ligation of the vector

To generate the destination vector pCGate-Ara, the amplified Gateway® cassette + *ara*S was ligated with the previously digested "pC" vector using a blunt end cloning strategy following the manual for Sticky-End cloning of the CloneJet PCR cloning kit (ThermoScientific). Deviant to the protocol which instructs blunting of the PCR product, 150ng of the linear pC vector were blunt ended following the company`s manual. For this, the overhangs generated by the restriction enzymes *Sac*II and *Eag*I were removed and filled in by the action of the applied blunting enzyme, respectively. After incubation, the PCR amplified Gateway® cassette + *ara*S was added in the molar plasmid-insert ratio 1:3 together with 1µl T4-ligase to the appropriate reaction and incubated over night at 20°C. The ligase was heat inactivated at 65°C and 4µl of the reaction were used for

transformation of 40 µl *E.coli* DB 3.1 cells (gyrA462) (Invitrogen™) resistant to the toxic *ccdB* gene product following the standard heat shock protocol (see 3.3.1). The transformation mixture was plated on LB agar supplemented with 50µg/ml chloramphenicol and incubated at 37°C o/n. Only cells harboring the ligated Gateway®-destination vector (pCAra-GW) were able to grow, since only they comprised the chloramphemincol resistance gene.

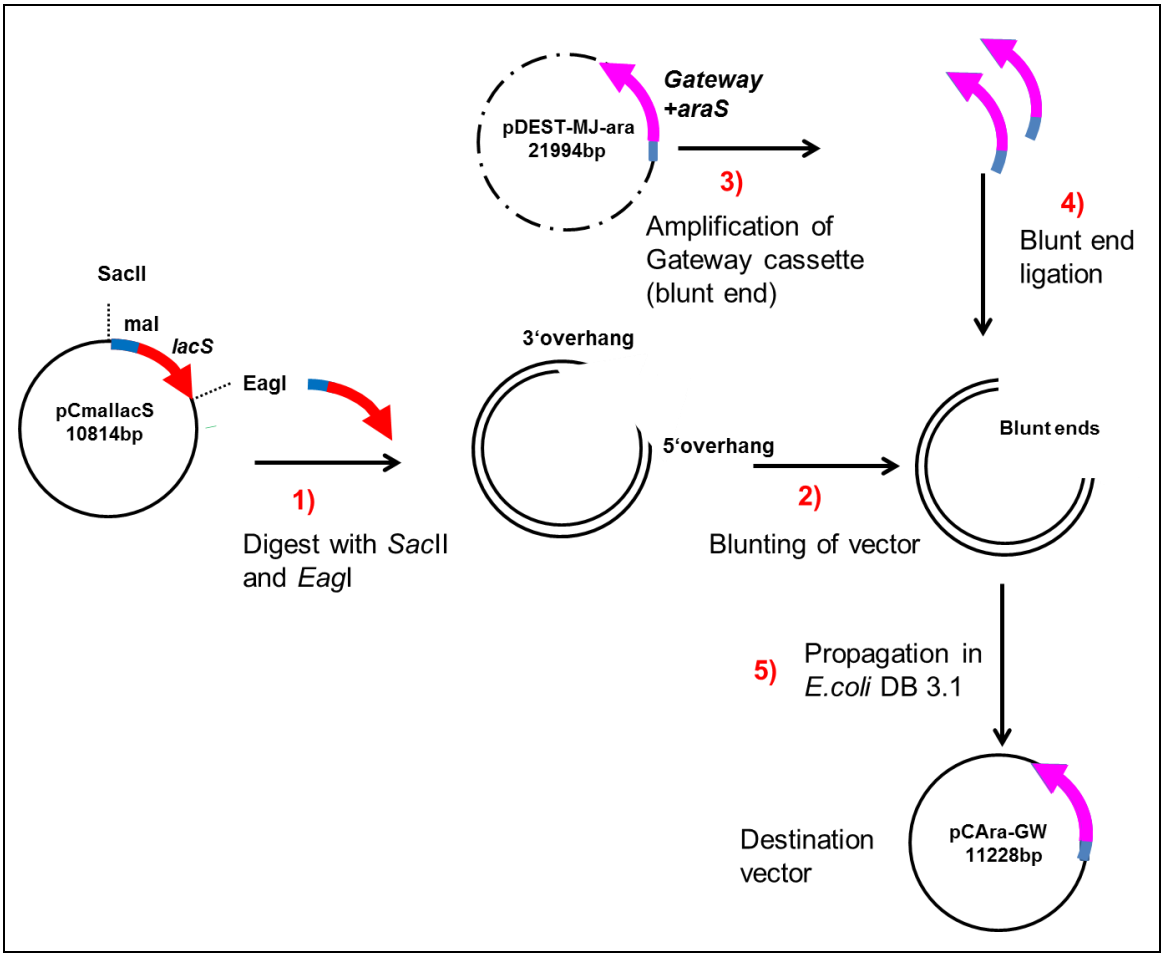


Fig. 6 Scheme of pCAra-GW cloning. pCmalLacS was used as starting vector for construction of the destination vector pCAra-GW, harboring a Gateway® cassette (pink) and an arabinose inducible promoter (*ara*S) (blue) upstream of it.

3.4.5 Verification of pCAra-GW

3.4.5.1 Check PCR analysis

Colonies were picked and a Colony PCR (see *3.3.3.1*) using 406Seq_FW and 406Seq_RV primer (same as used for cloning) was performed as described in table 3, whereby 2 min elongation time were chosen and 30 cycles were set. The colonies were transferred into 5ml liquid LB media treated with the same concentration of chloramphenicol and incubated o/n. The colony PCR was loaded on 0.8% agarose for visualization and positive cultures of positive clones were centrifuged at full speed and further used for plasmid extraction (see *3.3.2.1*). Recovered plasmids were quantified (see *3.3.2.4*).

3.4.5.2 Sequencing of Gateway cassettes

To gain a verification of the gateway cassettes on nucleotide level, one pCAra-GW plasmid was sequenced using the 406Seq_FW and the 406Seq_RV primer, respectively in separate reactions. The sequencing mixture was set up according to the instructions of the LGC genomics sequencing manual. Sanger sequencing was achieved by the eponymous company (LGC Genomics GmbH, Berlin/Germany). To obtain the full length sequence of the cassette, a GW-internal primer (Tab.1) was designed on the basis of the sequence chromatogram gained by previous 406Seq_FW primer sequencing. After proofreading of the three sequence chromatograms and removal of the not well resolved initial sections of the sequences (up to 30nt), contigs were assembled using the forward, internal and reverse sequence. The contig was aligned using ClustalW alignement tool (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and was blasted against the NCBI database.

3.4.5.3 EcoRI and BglII restriction digest

For integrity check of the vector and to determine the orientation of the Gateway® cassette + *ara*S insert, an *Eco*RI restriction digest was conducted according to the instructions in table 4. Also, a *Bgl*II restriction digest was performed to linearize the plasmid (Tab.4).

3.5 Construction of pIZ shuttle vectors via Gateway® recombination

3.5.1 Constructs used for Gateway® cloning

3.5.1.1 The protospacer inserts

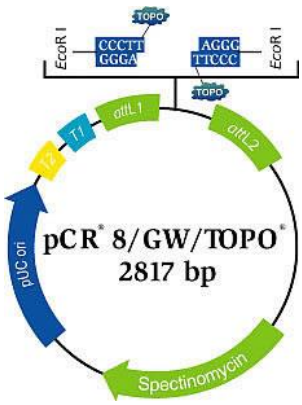
The CRISPR negative control protospacer variant 406-R was generated in earlier studies (Manica et al., 2011) and consists of the ORF406 from the pNOB8 conjugative plasmid which harbors a 37nt sequence showing a perfect match to the CRISPR spacer 53 of locus A in *S. solfataricus* P2, but not in P1 and therefore serves as negative control in this study (Fig.7B). The second CRISPR negative control, WOP (*W*ith*O*ut *P*rotospacer) consists only the ORF406 sequence without any protospacer embedded (Manica et al., 2013) (Fig. 7B).

The protospacer constructs D63-HA and D63-7U matching the spacer 63 of locus D in *S. solfataricus* P1 were recently obtained by exchanging the A53 spacer of ORF406 constructs for the reverse complement of the D63 spacer sequence by Overlapping Extension and were further modified by introducing 7 GU mismatches (Zebec, submitted). The protospacer D63-HA shows a 3' PAS (protospacer adjacent sequence) matching the 5'handle of the according crRNA, whereas D63-7U harbors a not-matching PAS sequence (Fig.7B).

3.5.1.2 pENTRY

The protospacers (embedded in the ORF406) and the *lacS* gene (SSO3019) of *S. solfataricus* P2 had been cloned into pCR8/GW/TOPO plasmids (Invitrogen™) (Fig.7A) to generate pENTRY vectors (Manica et al., 2013 and unpublished; Zebec, submitted). All pENTRY vectors had been verified by sequencing. An ENTRY vector consists of the attL1 and attL2 sites between which the insert of interest had been integrated via TOPO TA cloning (Invitrogen™). Beside a pUC ori necessary for propagation in *E.coli*, the vector harbors a spectinomycin resistance gene making selection on spectinomycin possible (Fig. 7A). ENTRY clones are used for site specific Gateway® recombination.

A)



B)

5'→3' protospacer sequence

HA	**-ATAGTTTGGGTATTGTTGTTTTGGTTTTTATATGAAA***CTTTCAAT-*
7U	**-ATAGTTTGGGTATTGTTGTTTTGGTTTTTATATGAAA***GAGAGCTG-*
406R	**-GAAGGCTGAGGATGAGGTTACACGATGTTGCTATTCA-**
WOP	----------

Fig.7: pENTRY vector and protospacer sequences. **A)** pCR 8/GW/TOPO vector (Invitrogen ™) used for Gateway® in vitro recombination in this study. **B)** Protospacer constructs cloned into PCR 8/GW/TOPO vectors to create different pENTRY vectors used for LR-in vitro recombination. Bolt: protospacer sequence matching the spacer; italic: PAS (protospacer adjacent sequence).

3.5.2 Gateway® *in vitro* recombination

To transfer the attL sites flanked sequence located on the pENTRY vector to the pCAra-GW destination vector, an att site specific Gateway® *in vitro* recombination (attL1 recombines with attL1, etc.) was performed following the manual of the Invitrogen™ Gateway® LR Clonase™ II Enzyme mix whereby pCAra-GW served as destination vector and was added in same amounts as the pENTRY vector. The set up was as follows:

1µl pENTRY (150ng)
1µl pCAra-GW (150ng)
1µl TE buffer, pH=8
1µl 5x buffer
1µl LR clonase (at the end, directly from -80°C)
5µl

The reaction was incubated at 21°C o/n. 2-3µl of the Gateway® reaction were used for transformation of 20-30µl of *E.coli* ER 1821 (NewEngland Biolabs) cells (see *3.3.1*) and plated on LB-agar supplemented with 50µg/ml ampicillin. Cells were incubated at 37°C o/n and screened for colonies the next day. Exclusive growth of expression clones, called pIZ- vectors is ensured since the ampicillin selection marker is situated on the pCAra-GW

backbone excluding growth of pENTRYs, and the expression of the *ccdB* suicide gene product kills the cells harboring a pCAra-GW vector not having exchanged the cassette.

The pIZ vectors were checked via PCR (see *3.3.3.1*) and visualized on 0.8% agarose. Verified vectors were used in transformation studies.

3.6 Transformation of Sulfolobus

3.6.1 Plasmid preparation

3.6.1.1 Methylation of plasmids

All final plasmid propagations of cloning steps before *Sulfolobus* transformation were carried out in the *E.coli* strain ER 1821 harboring the pM.*Esa*BC4I plasmid (New England Biolabs) which expresses the eponymous methylase. M. *Esa*BC4I specifically methylates the N4-position of the first cytosine of GGCC motifs leading to the protection of the plasmids from *Sua*I restriction enzyme degradation.

3.6.1.2 Dialysis of plasmids

Prior to transformation of *Sulfolobus*, the plasmids were dialyzed to remove salts potentially affecting electroporation. 40µl of plasmid DNA were put on a Millipore filter (Millipore, 0.025 µm) previously placed onto the surface of a 1L dH_2O water bath supplemented with 1ml 50mM Tris-HCl. After 1-2 hours incubation at room temperature (RT), the DNA was removed and the concentration was determined (see *3.3.2.4*).

3.6.2 Preparation of competent cells

Sulfolobus auxotrophic mutant cells were sampled at an OD_{600} of 0.2 (see 3.1.1), cooled down in an ice-water bath and centrifuged at 4500rpm, at 4°C for 15 min (8 Universal 320R, *Hettich Zentrifugen*). The pellet was washed in 20ml 20mM sucrose (Serva) three times whereby for final resuspension, an appropriate volume of supernatant was used to yield a concentration of 10^{10} cells per ml.

3.6.3 Electroporation

In this study, two different electroporation procedures were tested for transformation of *S. solfataricus* M16 which differed in the amount of plasmid DNA, the electroporation settings and the recovery of the transformed cells. The Standard method "ST" was conducted according to the protocol of Berkner et al., 2007, where 150ng of dialyzed plasmid DNA were mixed with 50µl of competent cells (10^{10} cells/ml). Contrary, 1µg of plasmid DNA was mixed with the same amount of cells in the "ICE" approach carried out after Deng et al., 2010. For electroporation, cooled 1mm electrode gap cuvettes (PeqLab, LE) were used in both approaches into which the transformation mixture was pipetted directly to the bottom. Immediately afterwards, electroporation was performed using a Gene Pulser Xcell (BioRad) apparatus. Following conditions were used for the ST-method: 1240V, 25µF, 1000Ω, whereby for the ICE-method, 1,5kV, 25µF, 400Ω were set. After electroporation, the ST - approach cells were directly transferred into with 50µl 75°C-pre-heated recovery solution (Berkner et al., 2007) and incubated for 45 min at 75°C under continuous shaking (550rpm). For the ICE method, the transformation suspension was incubated in 1ml ddH_2O for 10min on ice. After recovery, both approaches were directly transferred into 50ml tryptone/sucrose media and incubated at 78°C (see *3.1.1*).

Transformation assays for *S. solfataricus* M18 and *S. tokodaii* were performed according to the ST method.

3.7 Verification of transformation

3.7.1 PCR on total DNA

To verify the plasmid in the transformed culture, a total DNA extraction was conducted (see *Sulfolobus* DNA extraction) whereof 10ng were used as template for PCR amplification using following the Phusion polymerase PCR protocol (Tab. 3). For detection of pIZ-βGal, 406Seq_FW and 406Seq_RV primer were used, whereby the elongation time of the PCR reaction was chosen to be 75 sec. For verification of the pIZ-protospacer

constructs, ORF904_FW and ORF904_RV primer were used in PCR reaction and the elongation time was set to 40 sec. PCR reactions were visualized on 0.8% agarose.

3.7.2 Southern blot

3.7.2.1 Probe design

Two different dsDNA probes were designed in this study: the ORF904 pIZ vector specific probe and the D291 virus specific probe. Both probes were constructed by a 32 cycle Phusion PCR amplification (Tab. 3) using digoxygenin (DIG) labeled nucleotides (DIG DNA Labeling Kit, Roche) together with unlabeled dNTPs in different ratios. For the ORF904 probe amplification, ORF904 primers (Tab. 1) were used and as template pIZ-7U (1ng) was applied. DIG labeled dNTPs and unlabeled dNTPs were added in an 1:2 ratio, and elongation time was set on 45sec, since the amplified fragment was 1030 bp in size. 1ng of the SSV1 based plasmid pMJ0305 was used as template for D291 probe amplification, whereby D291 primers were used (Tab. 1) and DIG labeled dNTPs and unlabeled dNTPS were added in equal amounts. Elongation for the 329bp sized D291 probe was chosen to be 30 sec. Both PCRs were checked on 1.5% agarose and were purified (see *3.3.3.2*).

3.7.2.2 AseI digest and pre-gel

Prior to probe hybridization, 2-18 µg of total genomic DNA of M18 transformants and appropriate controls comprising untransformed pIZ-HA and pIZ-WOP as well as pMJ0305 were digested by the *Ase*I restriction enzyme according to the protocol in table 4. The digest was loaded on a 0.7% agarose for verification and to visualize the intensities of the digested DNA for potential adjusting of DNA amounts for the following blot.

3.7.2.3 Southern hybridization

*Ase*I digested DNA and 10µl of 1kb+ GeneRuler DNA ladder were separated on 0.9% agarose (without Ethidium bromide) by gel electrophoresis at 20V for 16 hours. For the last 20 min, the voltage was increased to 90V. After electrophoresis, the gel was shortly swayed in ddH_2O and the marker lane was cut out, post stained with Ethidium bromide and a picture was taken under UV next to a fluorescent ruler to later localize the hybridization signals on the film. Afterwards, the gel was placed beneath a nitrocellulose

membrane (Hybond-XL, Amersham) in the middle of a typical Southern blot sandwich consisting of three wet Whatman papers beneath the gel connecting the gel/membrane layer with 20xSSC high-salt transfer buffer, and three filter papers on the top of the membrane. Through capillary action, the DNA was blotted onto the membrane o/n (18h). The transferred DNA was fixed on the membrane through incubation at 80°C for 2 hours. The membrane was then incubated in hybridization solution (see below) at 42°C for 2 hours. Both probes were mixed together with 50µl hybridization solution and denatured for 5 minutes at 95°C. The membrane was incubated with the two – probed mixture at 42°C o/n. The hybridization probe solution was removed and the membrane was washed twice in 2xSSC/0.1% SDS for 15 min at RT and afterwards two times with 0.2xSSC/0.1% SDS buffer at 60°C. After washing, the membrane was incubated with anti-DIG-antibodies attached to alkaline phosphatase (Anti-Digoxigenin-AP, Fab fragments, Roche) binding the DIG labeled probes hybridized to the appropriate sequence. To visualize the probe bound DNA fragments, the membrane was incubated 10 min in the Disodium 3-(4-methoxyspiro {1,2-dioxetane-3,2'-(5'-chloro)tricyclo [3.3.1.1^{3,7}]decan}-4-yl)phenyl phosphate (CSPD), a chemiluminescent substrate being dephosphorylated by the alkaline phosphatase which leads to light reaction detected on an X-ray film.

Hybridization solution (200ml):	20x SSC:
100ml 100% Formamid	3M NaCl
50ml 20xSSC	0.3M Sodiumcitrat
40ml 10% blocking reagent	pH=7.0
660µl 30% N-laurosyl-sarcosyl-sodium-salt	
400µl 10% SDS	
9ml dH_2O	

3.7.3 Quantitative real time PCR (qPCR)

To estimate the plasmid copy number per cell, two qPCR reactions were performed using the same template DNA (genomic DNA of “induced” samples of M18 transformants– see Northern blot) but different pPCR primer sets. Q1 406_FW and Q1 406_RV primer (Tab. 1) covering the protospacer D63 located on the pIZ-plasmid were used to quantify the pIZ-plasmid in 25 ng total DNA. To evaluate the numbers of chromosomes in same DNA

extractions, q3194FW and q3194RV primer specifically binding the Sso 3194 chromosomal gene, were applied. Three replicates of DNA extractions per cell sample with three technical qPCR replicates each were analyzed.

The DNA extractions were primarily diluted to 5ng/µl. As positive controls for plasmid quantification, three dilutions of untransformed pIZ-HA plasmids (10^6, 10^5 and 10^4 copies) and as negative controls, four water controls were prepared. As a qPCR standard, 10-fold dilutions from $4x10^7$ copies on of linearized pEntryA53* (Zebec, in prep.) were used.

For accurate pipetting, master mixes (without template DNA) were prepared appropriate to the number of samples. The master mix was transferred into 8 strip tubes (MicroAmp®). Content and amount of each qPCR reaction are given below.

10µl pPCR Sybrgreen mix (Qiagen™)
2µl primer mix (RV and FW) 10µM
5µl genomic DNA (5ng/µl)
3µl DEPC water
20µl reaction

qPCR amplification was carried out in a Eppendorf Mastercycler epgradientS realplex^2 (Eppendorf) as follows:

95°C – 5 min	
95°C – 15 sec	40x
60°C – 20 sec	
72°C – 15 sec	
95°C – 15 sec	
60°C – 15 sec	
MC – 20 min	
95°C – 15sec	

Total amounts of qPCR evaluated copy numbers were analyzed after calculating the average of all triplicates.

3.7.4 Northern analysis

3.7.4.1 Induction of transformants

For testing the arabinose promoter activity, the transformed M18 cells harboring the pIZ-protospacer constructs (4 samples) were grown from glycerol stocks in 50ml selective NZ/S growth media (see *3.1.1*) until late exponential phase. Every culture inoculated as triplicate to an OD_{600} of 0.2 in N-Z-Amine (0.2% w/v) and minimal sucrose (0.04% w/v) media. "Uninduced" cultures were harvested to 7ml in falcon tubes, centrifuged at 4500rpm and 4°C (Universal 320R, *Hettich Zentrifugen*) and pellets were immediately frozen. For induction of the *ara*S promoter, 20% arabinose was added to the residual cultures at 0.16% and incubation was proceeded for four hours. "Induced cultures" were then sampled as described above. The pellets of "uninduced" and "induced" samples were used for further DNA and RNA extraction (see *3.3.2.3*).

3.7.4.2 RNA quality check

1µg of the RNA extracts of one chosen triplicate of "uninduced" and "induced" cultures of M18-pIZ-HA, M18-pIZ-WOP and M18-pIZ-406R cells were added to 9µl RNA mix (see below) and loaded on 1.2% agarose to visualize the amount and integrity check (all 3 ribosomal RNA subunits must be apparent).

RNA mix:

250µl Formamide
83µl 37% Formaldehyde
50µl gelbuffer pH=8
50µl DNA loading buffer

3.7.4.3 Northern hybridization

Total RNA of M18-pIZ-HA (10µg) M18-pIZ-WOP (2.5µg) and 406R (2.5µg) of "uninduced" and arabinose "induced" transformants were mixed with 15µl (WOP, 406R) or 20µl (FA) RNA mix, and were incubated at 65°C for 10 min for denaturation. Additionally, 10µl of RNA high range ladder (RiboRuler, Thermo Scientific) was denatured. The samples and RNA ladder were loaded on 1% unstained formaldehyde gel. Gel electrophoresis was achieved in 1x MOPS-running buffer (see below) at 50V for two h, 80V for 3.4 h and 110V

for the last 5 min. The gel was directly transferred onto a nitrocellulose membrane accordingly to the Southern blot (see Southern blot). The RNA was attached to the membrane by cross-linking via ultraviolet irradiation (1200J/m^2). For visualization of RNA, the membrane was incubated in 0.4% methylene blue (0.04% in 5M NaAC pH=5.5). The RNA fragments of the marker lane were localized and a picture was taken to later identify the size of the hybridized RNA fragments on the X-ray film. The DIG labeled ds DNA probe ORF 406 had been previously constructed in our laboratory by PCR reaction using Pro-FW1 and Pro-RW2 primer (Zebec, submitted.). The probe does not cover the protospacer D63, but binds the RNA of expressed ORF406 at the 5' end (Tab. 2). Equivalent to Southern hybridization, the probe was mixed with 50ml hybridization solution and was denatured at 95°C for 10 min before membrane hybridization. The membrane was incubated with the probe-hybridization mix at 42°C o/n. Subsequent washing steps and X-ray visualization were carried out as described for Southern blot (see *3.7.2.3*)

Hybridization solution (100ml):	10xMOPS running buffer
50ml 100% Formamide	41.8g MOPS in 800ml DEPC, adjust pH to 7
10ml 10% SDS	16.6ml 3M NaAC-DEPC
10g dextrane sulphate	20ml 0.5m EDTA-DEPC (pH=8)
5ml 20xSSPE	Fill up to 1000ml – sterile filtrate
0.5g low fatmilk powder	
24.5ml DEPC-H_2O	

3.7.5 Semi-quantitative PCR

3.7.5.1 Semi-quantitative PCR using ORF904 primer

Semi-quantitative PCR was performed according to the protocol of Phusion PCR (Tab.3) but with the modification of using 25ng template DNA instead of 1ng. The Semi – quantitative approach was designed to simulate a qPCR reaction by choosing cycle numbers according to the cycle thresholds measured for the different samples. ORF904 FW/RV primers were used in this approach to amplify the pIZ-plasmid in same already processed DNA samples (qPCR, Southern blot) and 20 cycles were set with an elongation step of 40 sec. DNA extracts from former M16 cells transformed with pIZ-7U and M18 cells with pIZ-406R which had never been processed before served as controls for the

Semi-quantitative PCR were. The transformation of the controls had been achieved according to the ST method (see Transformation of *Sulfolobus*). The entire PCR reaction volume (25µl) was loaded on 0.8% agarose.

3.7.5.2 Semi-quantitative PCR using pyrEB and ORF904 primers

The PCR was as performed as described above but amounts of the amplicon were analyzed after 15, 21, 24 and 27 cycles. Therefore each sample was examined in quadruplicates in one thermocycler using a program looping back to the annealing step after a short 5 minutes 4°C break after the appropriate cycle number. During the 5 min break, the thermocycler was opened and the sample was taken out.

Two semi-quantitative PCR reactions were conducted in parallel in the "homologous recombination study". 25ng of genomic "induced" DNA of M18 pIZ-protospacer transformants were analyzed following the Phusion PCR protocol (Tab. 3) using pyrB_FW and pyrE_RV primers (Tab.1) with an elongation time of 45 sec. The parallel approach was carried out with ORF904FW/RV (Tab. 1) primer but otherwise using the same settings. Entire PCR products (25µl) were loaded on 0.8% agarose and compared.

4 Results

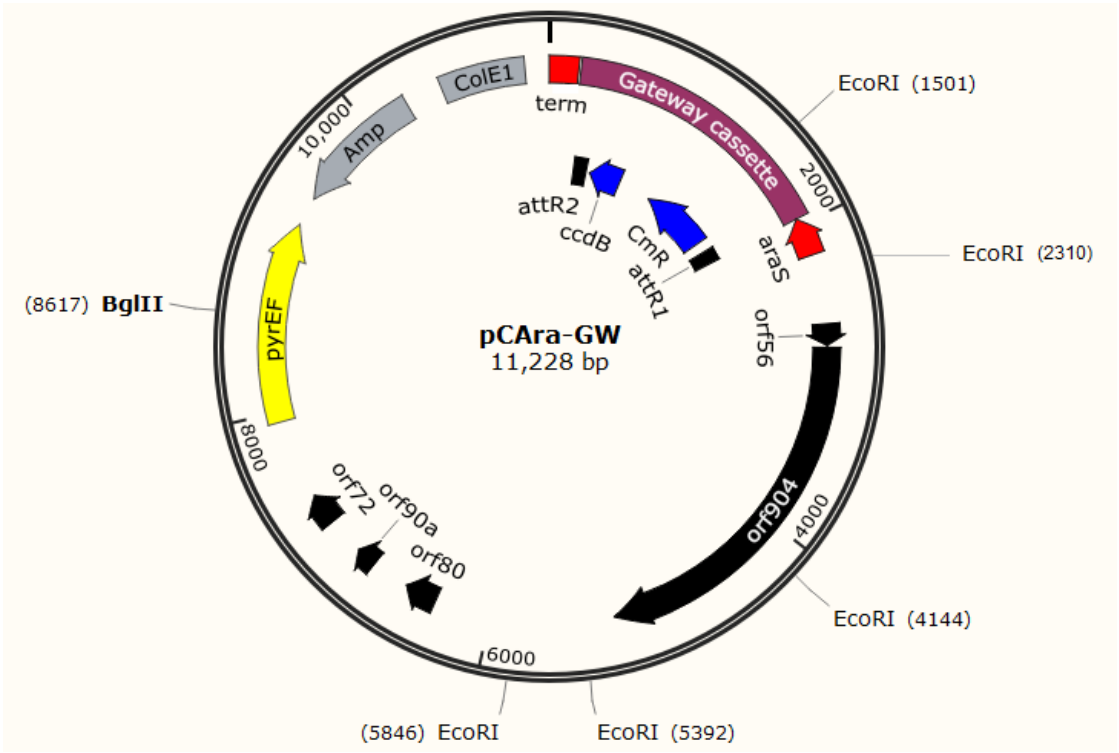
4.1 Construction of vectors used in this study

4.1.1 pCAra-GW

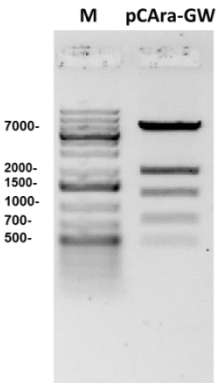
The destination vector serving as backbone for all expression vectors in this study was gained by cloning a Gateway® cassette harboring an arabinose inducible promoter (*araS*) situated upstream of the attR1 site into the pRN1 based shuttle vector pCmalLacS from which the *lacS* and maltose promoter have been removed previously (see Materials and Methods and protocol of the "Großpraktikum").

Cloning results were checked via PCR using the primer pair 406Seq_FW and 406Seq_RV which amplified the expected size of 2.3kb bp (Gateway® cassette + *ara*S) (data not shown). Additionally, the insert was sequenced utilizing the primer 406Seq_FW, 406Seq_RV as well as a Gateway® internal primer in separate reactions. The sequencing results were aligned to public databases reassuring the integrity of the single components (data not shown). Since the Gateway® cassette has been inserted via a blunt end cloning strategy, its orientation was checked via *Eco*RI restriction digest (Fig. 8B). Besides the determination of the orientation, the digest also proves the integrity of the vector. Five fragments of about 6900bp, 2100bp, 1240bp, 800bp and 450bp are apparent on the gel which is in accordance to in silico *Eco*RI digestion of the vector (Nebcutter, New England Biolabs®; http://tools.neb.com/NEBcutter2/). The plasmid is 11228bp in size and the orientation of the insert has been determined to be counterclockwise. Fig. 8C depicts the linearized pCAra-GW vector after *Bgl*II restriction as a strong band above the 10000 bp marker band.

A)



B)



C)

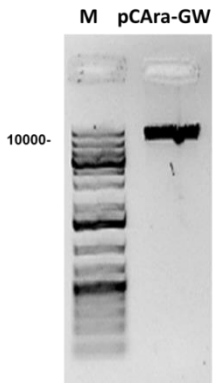


Fig. 8: pCAra-GW vector map and integrity check. A) The vector consists of a pRN1 plasmid part comprising amplification genes necessary for propagation in *Sulfolobus* indicated in black, a pBlueskript vector part made up of a *bla* gene conferring ampicillin resistance (AmpR) and an *E.coli* origin of replicatoin (ColE1) indicated in grey, a *pyrEF* gene cluster of *S. solfataricus* P2 as auxotrophic selection marker in yellow, the Gateway® cassette in purple with the corresponding genes and att sites cited beneath and the arabinose inducible promoter (*ara*S) in red as well as a terminator (term) in red. Restriction sites of *Eco*RI and unique BglII are indicated. **B)** *Eco*RI digest of 500ng pCAra-GW for orientation verification showing 6 bands at 6900bp, 2100bp, 1240bp, 800bp and 450bp . **C)** *Bgl*II digest (one cutter in *pyrEF*) of 500ng pCAra-GW loaded on 0.8% agarose. M= 3µl DNA ladder. GeneRuler 1kb+ (Thermoscientific) DNA ladder was used for all gels.

pCAra-GW can be propagated in *E.coli* strains carrying a mutated gyrase subunit A which is otherwise affected by the toxic gene product of the *ccdB* suicide gene comprised in the Gateway® cassette. Within Gateway® *in vitro* recombination assays, att site specific recombination between entry clones harboring the gene of interest and pCAra-GW leads to the formation of expression vectors comprising the pCAra backbone and the gene of interest located at the position the Gateway® cassette was placed before.

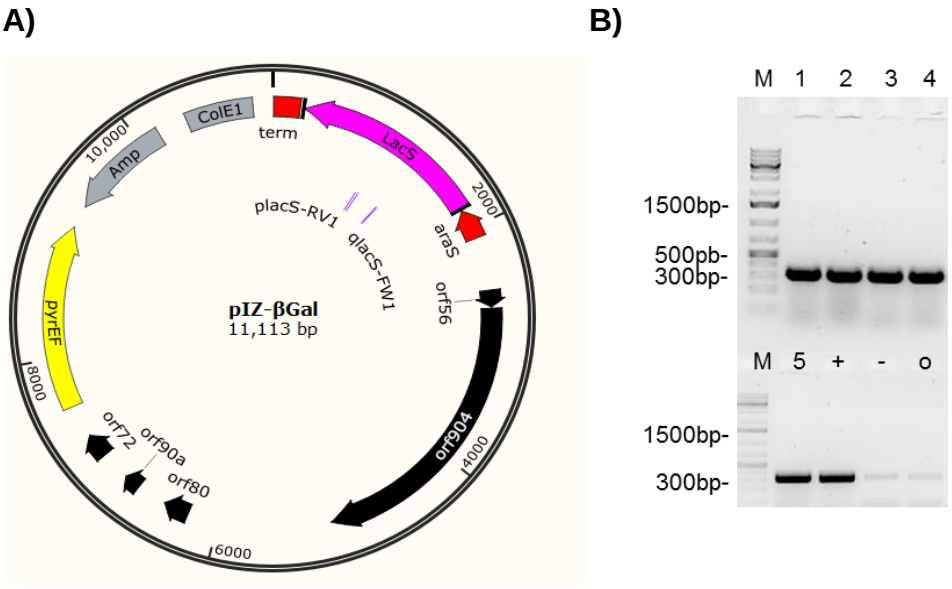
4.1.2 Expression vectors used in transformation assays

4.1.2.1 pIZ- βGal construct

As a trial construct to test under which conditions transformation of *S. solfataricus* P1 using a plasmid is feasible, the *lacS* gene of *S. solfataricus* P2 expressing β-galactosidase was placed into pCAra-GW vector generating the expression vector pIZ-βGal (Fig.9A). For this, a pENTRY vector (pENTRY-lacS) carrying the *lacS* gene bordered by attL sites was used in the Gateway® *in vitro* recombination assay.

In vitro recombination is considered to be accurate, owing to the Gateway® cassette and the ampicillin marker reassuring forced selection. Nevertheless, a colony PCR using qlacS_ FW1/RW1 primer pair proofing the presence of the *lacS* insert was established. Fig. 9B shows the exact size of 300bp of the primer specific band in all picked colonies. As positive control, pCmalLacS was used harboring the *lacS* gene of *S. solfataricus*. The negative control (pMJ plasmid without *lacS* gene) shows a faint band. Given that the same intense band also appears in the water control, the negative control was perceived as credible.

Also the topologies of four pIZ-βGal preparations were determined by loading total plasmid (uncut) on an agarose gel (Fig.9C). For pIZ-βGal no. 1, an additional smaller fragment was apparent at approximately 4kb possibly representing supercoil formation of the plasmid. Other conformations, such as multimers and open circle are visible in ranges between 20000 and 10000bp (Fig.9C).



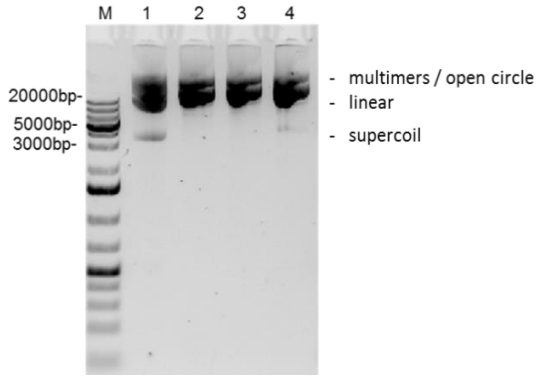


Fig.9: pIZ-βGal vector map and integrity check. A) Map of expression clone pIZ-βGal as result of LR Gateway *in vitro* recombination between pENTRY-lacS and pCAra-GW. qlacS FW1 and RV1 primers sites are indicated. **B)** Check colony PCR on five *E.coli* ER 1821 colonies grown upon transformation with the recombination solution loaded on 0.8% agarose. *lacS* specific primer amplifying 300 bp fragment were used in PCR. 1-5: 5 different *E.coli* clones, M: 2.5µl DNA ladder, (+) positive control: pCmalLacS plasmid, (-) negative control (pMJ plasmid not carrying the *lacS*), (o) water control (water supplemented to PCR reaction instead of DNA). **C)** Topology check of pIZ-βGal plasmids: 100ng of 4 (1-4) independently recovered pIZ-βGal plasmids loaded on 0.8% agarose. Multimeric and open circular forms appear higher on the gel than the linear plasmid. Supercoiled conformations are detected below the linearized plasmid. M= 3µl DNA ladder. GeneRuler 1kb+ DNA ladder (Thermoscientific) was used for all gels.

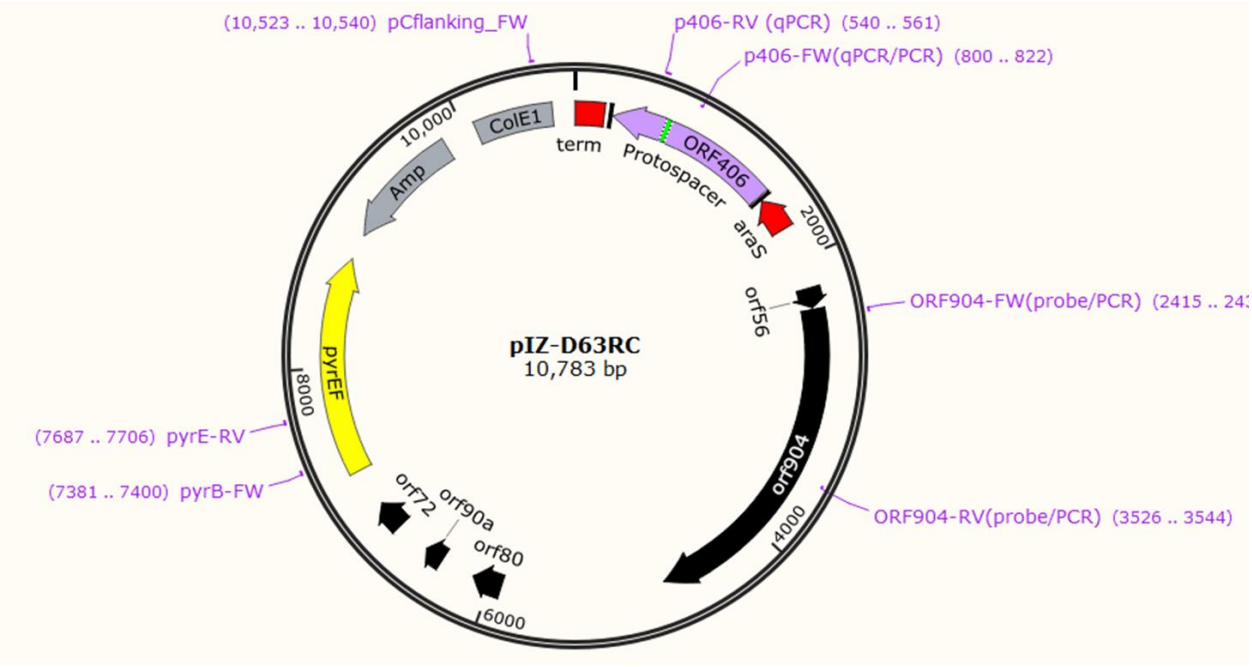
4.1.2.2 pIZ-protospacer constructs

For plasmid based CRISPR studies, the protospacer constructs recently designed in our laboratory by inverse PCR (Manica et al., 2013; Zebec, submitted and see Introduction and Materials and Methods) were used for *in vitro* recombination with pCAra-GW creating the pIZ-protospacer expression vectors (Fig.10A). The recombined protospacers D63-HA and D63-7U have a cognate spacer (spacer 63) in the CRISPR locus D of the *S.*

solfataricus P1 chromosome and had been embedded in the approximately 1000 bp sized ORF406 for cloning reasons (Fig.10A). As described in the introduction, the mRNA of these protospacers was shown to be targeted by the Cmr guided crRNA in virus infection studies. However, DNA interference was reduced due to GU pairing (7U construct) and additional handle complementarity (HA construct). As negative controls, two protospacers were used, namely WOP (*W*ith*O*ut *P*rotospacer) carrying the "empty" ORF406 insert, and 406R matching a spacer in the *S. solfataricus* P2 but not in P1 (Manica, in press; Zebec, submitted).

All four protospacer constructs were available in pENTRY vectors (Zebec, unpublished) and therefore directly used for Gateway® *in vitro* recombination of pCAara-GW yielding pIZ-HA, pIZ-7U, pIZ-WOP and pIZ-406R constructs (summarized here as pIZ-protospacer variants). A map of a pIZ-protospacer vector is depicted in Fig.10A. To verify the presence of the insert, a PCR on recovered plasmids using the primers Q1 406FW binding the 406 ORF-Protospacer insert and pCFlanking_FW binding the vector backbone as indicated on the vector map, yielding the expected size of 1100bp (Fig.10B).

A)



B)

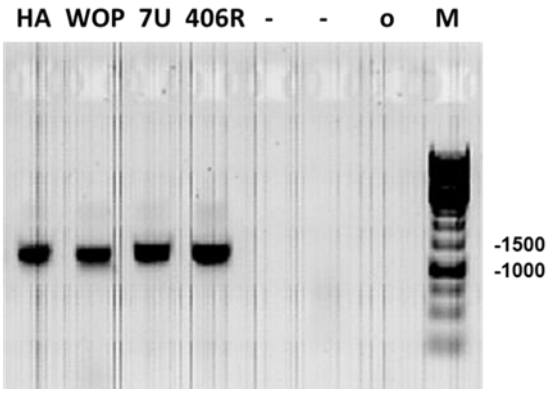


Fig. 10: Vector map of pIZ-protospacer construct with primer sites indicated and verification of the ORF406/D63 insert by PCR on extracted plasmids. A) Map of the pIZ-protospacer vector harboring the ORF406 (purple) sequence stretch interrupted by the modified reverse complement spacer 63 of the CRISPR locus D, i.e. protospacer D63 (green). Primers used within this study are cited together with their position on the vector. **B)** PCR check specific for the ORF406/D63 insert from pIZ-HA/WOP/7U/406R recovered plasmids on 0.8% agarose. Used primers were pCFlankingFW and Q1 406FW amplifying 1100 bp. 31cycles were used for PCR, 5µl of the samples were loaded (1ng DNA was used as template for PCR). Negative control: pIZ-βGal plasmid loaded twice. M= DNA Ladder, 1kb GeneRuler (Thermoscientific).

4.2 Transformation studies in Sulfolobus

4.2.1 Transformation of *S. solfataricus* M16 with pIZ-βGal

Two different transformation assays were prepared each in parallel with a negative control treated the same as the samples but without the plasmid DNA. One assay was the standard assay used in our laboratory, named “ST” (Standard) and the other approach referred to as “ICE” was conducted according to the protocol from Deng et al. (Deng, 2010). The two methods differ in the amount of applied plasmid, the electroporation protocol and the recovery of the transformed cells as illustrated in Fig. 11A. After the recovery step, all cell suspensions were incubated in 50 ml tryptone/sucrose (T/S) Brock media and treated equally.

4.2.1.1 Growth of cultures

Transformed *S. solfataricus* M16 cells were constantly grown in T/S media for 11 days and reached an optical density of about 0.5 each. 2ml of inoculum were transferred into fresh media containing N-Z-Amine (NZ) instead of tryptone. After 8 days, the ST culture reached an OD_{600} of 0.7 and the ICE culture 1.1, whereas for both negative controls an OD_{600} of 0.2 was measured which remained consistent over 7 days of further incubation. ST and ICE inoculum were each set on OD_{600} 0.2 in fresh NZ/S media. After three days, for both cultures an OD_{600} of 1.19 was quantified. Again the liquid cultures were adjusted to 0.2. After four days of incubation, both cultures reached an optical density of 1.1 representing an average generation time of about 24 hours. Growth was followed by every day measurement of the $OD_{600.}$ Optical densities of the different treated cultures of the last incubation round are depicted below as a growth curve indicating that there is no great difference between the two approaches (Fig. 11B).

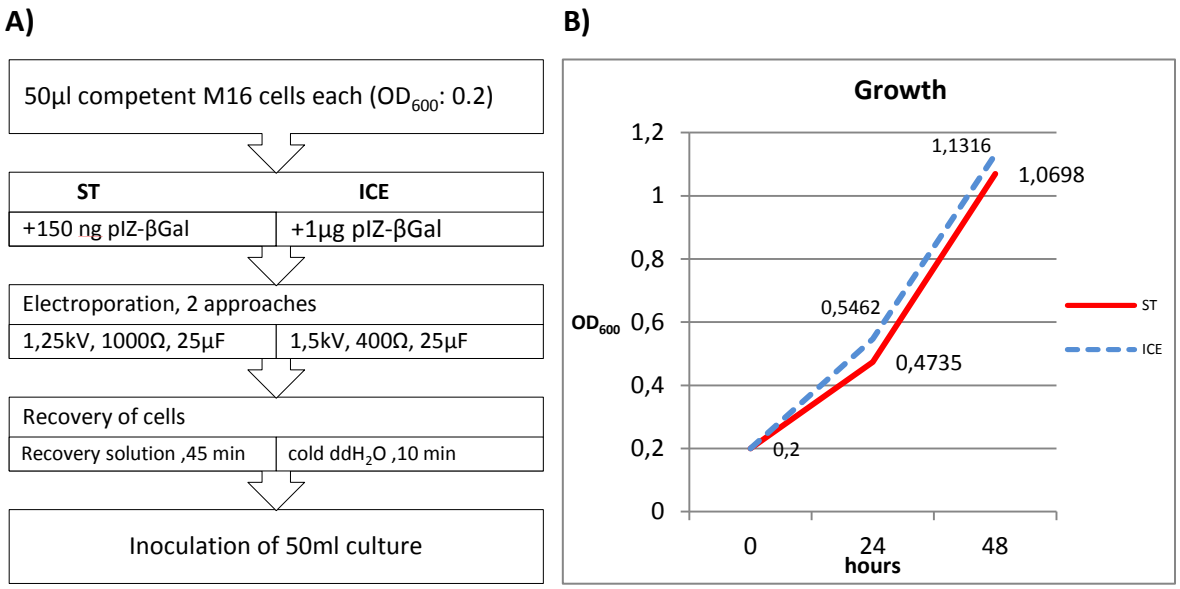


Fig. 11: Transformation approaches for *S. solfataricus* M16 and growth of both transformation mixtures. A) Transformation set up for two different approaches. 50µl of competent *S. solfataricus* M16 cell suspension were used for each transformation. **B)** Optical densities of ST and the ICE cultures measured each day at same time points are diagramed as curve over four days. OD_{600}=optical density spectrophotometrically measured by a wavelength of 600nm; blue line= ICE method, optical densities are indicated above; red line= ST method, optical densities are indicated beneath.

4.2.1.2 Transformation verification via PCR and retransformation

To verify that the growing cells are positive transformants, the presence of the shuttle vector pIZ-βGal had to be shown in the cultures. Therefore, a PCR on the total DNA extraction was performed using the primer pair 406FSeq_FW/RV amplifying 2050bp (Fig.12A). The products of ST and ICE show bands of expected sizes which is in accordance with the positive control represented by the untransformed pIZ-βGal plasmid. Since studies on SSV1 virus and *S. tokodaii* are conducted in our laboratory, the DNA extractions were checked for contamination by doing virus – and *S. tokodaii* specific control PCRs which were both negative (data not shown).

The total DNA was used to retransform *E.coli* TOP10 cells under ampicillin selection. Two colonies and five colonies were detected upon ICE and ST total DNA retransformation, respectively. For technical reasons, a colony PCR was performed only for the ST cells using the same primers as described above showing specific amplification in all five samples (data not shown). Plasmids of three ST colonies were isolated and an *Eco*RI restriction digest was conducted (Fig. 12B). The digests as well as recovered uncut plasmid was

loaded on 0.8% agarose. Bands of 7000-6000 bp, slightly above 2000bp, at 1500bp, 500 bp and a small band at about 300 bp appeared. The pattern apparent in sample Eco1 shows an additional intense band situated between 20000 and 10000 bp probably representing uncut or partially cut plasmids. However, the corresponding uncut DNA sample shows a lower band than the others at about 4000bp indicating that the preparation might contain a mixture of plasmids or was still in the supercoil form. As control, original pIZ-βGal was digested (180ng).

A)

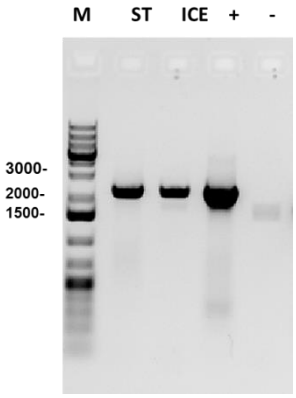
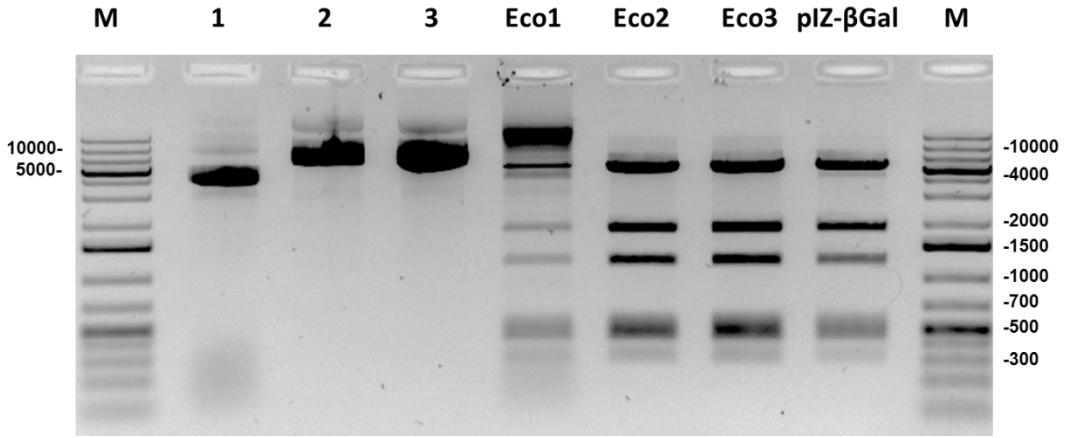


Fig.12: Verification of pIZ-βGal in transformed *S. solfataricus* M16 cells. A) PCR on 10ng total DNA recovered from M16 cells. Used primer pair was 406Seq_FW/RV amplifying 2050bp; cycles for PCR amplification were 29. ST=PCR on Standard approach sample; ICE= PCR sample of ICE approach. Used positive control (+) was pIZ-βGal untransformed vector (18ng as template for PCR amplification), whereas as negative control, pCmalLacS not harboring primer binding sites was applied. 5µl were loaded each. M= 3µl DNA ladder, 1kb+ GeneRuler (Fermentas). **B)** 1-3: Total DNA load of 100ng plasmid DNA extracted from three retransformed *E.coli* TOP10 cells. Eco1-3: *Eco*RI digests of 800ng of the same plasmid DNA resulting from retransformation. Same numbers denote same plasmid extraction. 7µl were loaded each. M= 1kb+ DNA ladder, GeneRuler (Fermentas), 2µl loaded on the left, 3µl loaded on the right side. 0.8% agarose.

B)



4.2.2 Transformation of *S. tokodaii* cells c92

To test whether other *Sulfolobus* strains would be also amenable for pCAra-GW based expression vectors, *S. tokodaii pyrEF* auxotrophic mutant c92 was transformed using the pIZ-βGal plasmid as described with the ST method. Contrary to *S. solfataricus* transformation, *S. tokodaii* cells were raised in N-Z-Amine/Glucose media. It took the cells 12 days to reach an OD_{600} of 0.32, whereas the negative control showed clear growth retardation by reaching a maximum OD of 0.03. The putative transformed culture was transferred into new medium two times and sampled at an OD_{600} of 0.2. After the second inoculation, the cells showed an average generation time of about 10 hours.

To check for successful transformation, again total DNA extract of c92 was used as template for PCR amplification using 406Seq_FW and 406Seq_RV primers enclosing the *lacS* and arabinose promoter outside of the att sites. The gel shows two bands, one at an expected size of 2200 bp and the other located 400 bp below (Fig.13A). As positive control, pIZ-βGal untransformed plasmid was used.

Attempts to retransform the total DNA yielded obvious retardation of *E.coli* colony growth as only after 30 hours at 37°C, clear big colonies were detectable on the plate. Five colonies were chosen for plasmid extraction but only one plasmid recovery in unusual low concentration of 50ng/µl was achieved. 500ng of this sample were digested using *Eco*RI but only a smear was detectable on the gel (data not shown).

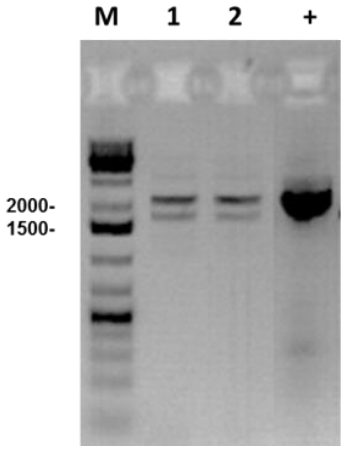


Fig. 13: PCR on *S.tokodaii* c92 total DNA of transformants. Two bands are visible at about 2200 bp and 1800bp. The bigger fragment is specific for expected size. For PCR, 10µg of template were used and the untransformed pIZ-βGal plasmid served as positive control.

4.3 CRISPR-protospacer analysis using pIZ shuttles

4.3.1 Transformation of *S. solfataricus* M18 with pIZ-protospacer constructs

For CRISPR studies, *S. solfataricus* M18 was transformed with the pIZ-protospacer constructs: pIZ-HA, pIZ-7U, pIZ-406R and pIZ-WOP according to the ST method described above where the initial incubation of the cells after recovery was achieved in tryptone/sucrose media and further incubations N-Z-Amine/sucrose solutions. After three transfers, i.e. 6 generations the average growth rate was 24h, whereas the negative control did not show any growth, as the highest OD_{600} measured was 0.08. Glycerole stocks were prepared of the sample. It should be noted here, that later incubation of transformed cells from glycerol stocks showed growth rates of 8 hours.

4.3.1.1 Transformation verifications

4.3.1.1.1 PCR on extracted DNA

For further Southern blot analysis new primers were designed amplifying a 1030 bp stretch of the ORF904 located on the pRN1 backbone (cf.Fig 10A). The primers had been tested first on untransformed pIZ-protospacer plasmids proving the specificity for all four protospacer constructs (data not shown). ORF904_FW and ORF904_RV primers were subsequently used for a first transformation check of M18 cells by doing a 32 cycles PCR on extracted total DNA of the putative transformants, using SSV1 plasmid vector as negative- and untransformed pIZ-HA as positive control. Aliquots were loaded on 0.8% agarose depicted in Fig. 14A. The bands all appeared at the expected size, whereby pIZ-7U shows a lower intensity than the others. To visualize the plasmids on the total DNA extract, Southern blot hybridizations were conducted.

4.3.1.1.2 Probe design for Southern blot analysis

Within the Southern blot a pIZ-vector specific as well as a virus specific probe were used concurrently. For pIZ-specific probe design, the primers ORF904_FW and ORF904_RV (described above) were used in PCR amplification using the pIZ-7U vector, whereas for

the SSV1 probe, D291 specific primers were applied amplifying 329bp of the SSV1 specific ORF D291 present on every SSV1 virus based vector used in the laboratory. Since the ORF904 probe specific for the pIZ-vector spans 1030 bp, the unlabeled and labeled digoxigenin-dNTPs were used in 1:2 ratio whereas for the D291 probe covering a much smaller fragment, they were added in equal amounts. Fig 14B shows the control of the probe design whereby the labeled probe (*) is larger than the unlabeled control (-).

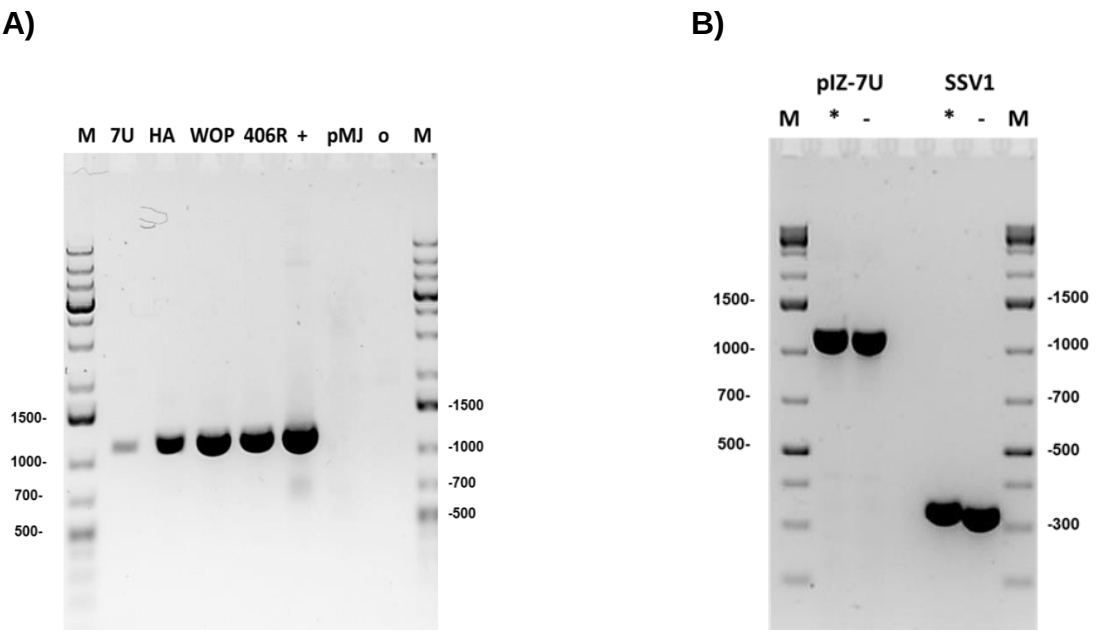
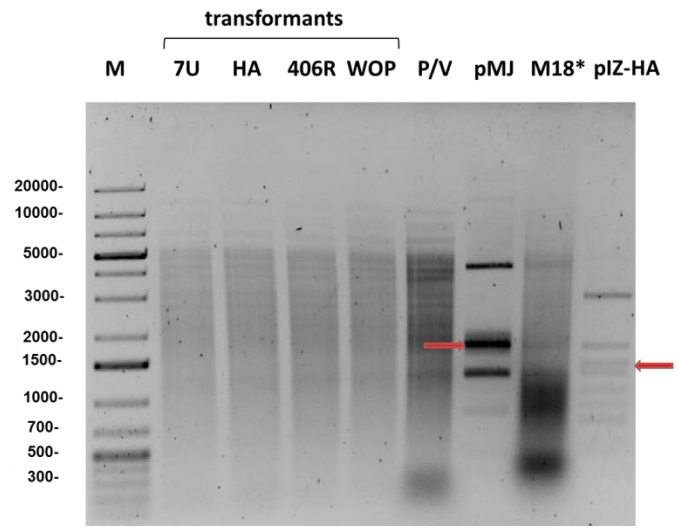


Fig. 14: Check PCR on total DNA of M18-pIZ-protospacer transformants and visualization of the PCR for probe design. A) 32 cycle PCR on total DNA of M18-pIZ-protospacer transformants using primers 904FW and 904RV amplifying 1030bp (10µg were used as template for PCR). Positive control (+): pIZ-HA untransformed plasmid (1µg used in PCR), negative control: pMJ virus shuttle vector (1µg used in PCR);(o): water control. **B)** Aliquots of probe amplification for Southern blot; pIZ-specific probe: 32 cycle PCR with ORF904 primers on pIZ-7U template using labeled and unlabeled dNTPs in 1:2 ratio (*); control PCR on the same template using unlabeled dNTPs (-). SSV1 specific probe: 32 cycle PCR with D291 primers on SSV1 template using labeled and unlabeled dNTPs in 1:1 ratio (*); control PCR on the same template using unlabeled dNTPs (-).

4.3.1.1.3 Southern hybridization

Southern hybridization was designed to answer the following questions: Are the cells transformed with the pIZ-protospacer constructs and not contaminated with the SSV1 virus and if so, does the plasmid amount differ between the transformants as the PCR results indicate. 3µg of total DNA of the putative M18 pIZ-protospacer transformants were used. As controls served 58ng of untransformed pIZ-HA vector as well as 110ng of the SSV1 based vector pMJ0305 which both represent $4.74x10^9$ copies. When considering 3µg of chromosomal DNA from *S. solfataricus* (2.9Mb) comprising $9.2x10^8$ copies under the supposition that a *Sulfolobus* cell harbors 1-2 genomes, the pIZ-HA and pMJ0305 dilutions both represent a copy number of 5-10 vectors per cell. Furthermore, 3 µg of total DNA extract from M18 cells previously transformed with pMJ0305 and a mixture of the latter with M18-pIZ-HA transformants of this study (3µg M18-pMJ0305, 2.5µg M18-pIZ-HA) to visualize plasmid and virus within one sample (V/P, i.e. virus/plasmid control) were used in the hybridization assay. Prior to the Southern blot, samples and controls were restriction digested using *Ase*I and visualized on a pre-gel to check the digestion and intensity of the samples (Fig. 15A). Probe ORF904 and D291 were expected to hybridize to restriction fragments of 1384bp and 1957bp in size, respectively (position indicated with red arrows on the pre-gel) (Fig. 15A). The Southern blot is depicted in Fig. 15B.



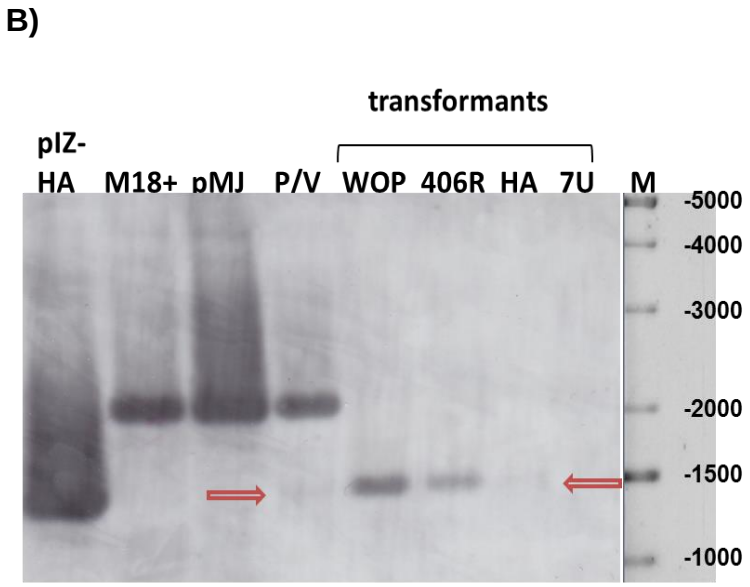


Fig. 15: Pre-gel and Southern hybridization of M18-pIZ-protospacer transformants with pIZ-vector and SSV1-virus specific probes simultaneously. A) 3µg of total DNA from M18-protospacer transformants (WOP,406R, HA,7U), 58 ng of pIZ-HA control (untransformed plasmid), 110ng of untransformed pMJ0305 SSV1 based vector (pMJ), 3µg total DNA of M18 – pMJ0305 transformants (M18*), a mixture of 2.5µg of M18-pIZ-HA DNA and 3µg of M18-pMJ0305 DNA (V/P) were *Ase*I digested and loaded on 0.7% agarose representing the pre-gel as guideline for further Southern hybridization. The red arrow on the left indicates the approximate size of the pIZ-protospacer plasmid visualized by ORF 904 probe hybridization (1384bp) and the arrow on the right points the position of the fragment binding to the D291 probe (1957bp) in further Southern blot hybridization. M: 4µl of DNA ladder, 1kb+ (Thermoscientific). **B)** X-ray film visualizing the Southern blot of the M18-protospacer transformants and appropriate controls conducted with pIZ-plasmid specific ORF 904 and SSV1 virus specific D291 digoxigenin labeled probes. Same digestions and amounts as used in the pre-gel were applied. The red arrows highlight the plasmid pIZ-HA only faintly visible in the transformation sample and the V/P control. pIZ-plasmids are visualized at 1384 bp and SSV1 based vectors at 1957 bp. M: 10µl of DNA ladder, GeneRuler 1kb+ (Thermoscientific).

For M18-pIZ-WOP and -406R transformants triggering no CRISPR degradation, the plasmid is visible at its expected position. No additional band is detectable on the blot for those samples. For M18-pIZ-HA DNA, only a very faint fragment (red arrows) is apparent in the transformation sample, as well as in the P/V control where approximately same amounts of total M18-pIZ-HA DNA were used. For M18-pIZ-7U total DNA, no respective band is detectable. The dilution control of the untransformed plasmid pIZ-HA and pMJ0305 are at expected location whereby their intensity is higher than that of the transformants. Samples harboring the SSV1 based virus vectors are located shortly beneath the 2000bp of the marker lane. No indefinable bands are present and no virus specific band is detected within M18-pIZ-protospacer total DNA.

4.3.2 Arabinose induction of transformants

4.3.2.1 Northern blot analysis upon arabinose induction

For testing the arabinose promoter and detection of possible CRISPR mediated mRNA degradation in samples HA and 7U, M18-pIZ-7U/HA/406R and –WOP cells from the same transformation approach examined above were raised from a glycerin stock to an OD of 1 in N-Z-Amine/sucrose media of which an inoculum was incubated in N-Z-Amine media with reduced sucrose concentration.

At an OD of 0.2, samples were harvested in triplicates referring as “uninduced samples”, whereas the rest was subsequently supplemented with arabinose and sampled as “induced samples” after 4 hours incubation. The DNA as well as RNA was extracted whereby the latter was checked on the gel which affirmed integrity by visualizing all three rRNA subunits at same intensities for the samples used in the Northern blot (data not shown).

Northern hybridization was performed on total RNA of one triplicate of “induced” as well as on the corresponding “uninduced” sample using the probe 406 hybridizing to the 5’ end of the ORF406 not spanning the protospacer D63 (Fig. 16). The probe was already

available in the laboratory as it was used for previous studies on the same protospacers (Zebec, submitted).

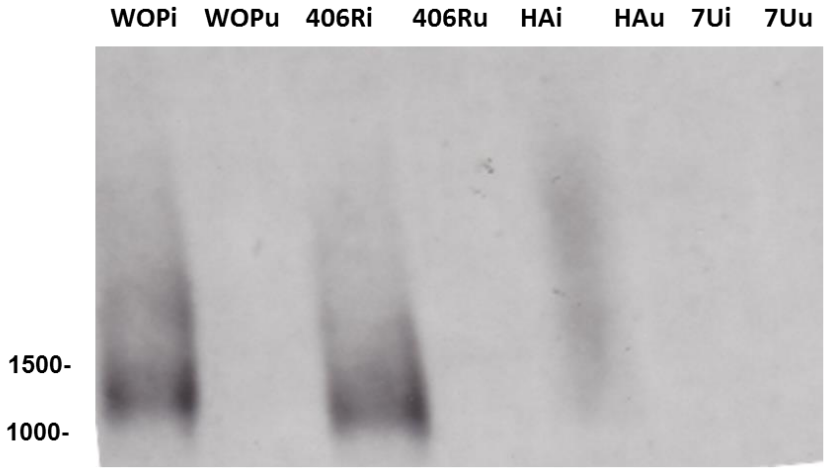


Fig. 16: Northern hybridisation. 2.5µg of total RNA of arabinose induced ("I") and uninduced ("u") samples of M18-pIZ-WOP and M18-pIZ-406R as well as 10µg of M18-pIZ-7U and M18-pIZ-HA induced and uninduced samples, respectively were hybridized to a ORF 406 (protospacer insert of pIZ-vectors) specific probe. The probe binds the 1100bp transcript of the ORF406 comprising the protospacer D63. Only for arabinose induced samples of M18-pIZ-WOP and M18-pIZ-406R the mRNA is apperant.

2.5µg of total RNA of M18-406R and M18-WOP were used as well as 10µg of HA and 7U. The blot visualizes the band shortly above the 1000bp marker lane which correlates with the size of the mRNA of the total ORF406 insert. As expected, for M18-pIZ-WOP induced as well as M18-pIZ-406R induced RNA a clear band is visible, whereas for the uninduced samples no band can be detected. No sharp band is apparent for M18-HA induced RNA and no mRNA is visualized in case of the M18-7U sample which refers to the low amount of plasmid in the cells of those samples (Fig. 16).

4.3.3 Quantification of the plasmids

4.3.3.1.1 qPCR analysis

To elucidate numbers of present pIZ-protospacer plasmids and host chromosomes in 1µg DNA of transformants, a qPCR was performed on total DNA of "induced" samples. Induced samples stem from the same transformation assay as the previous processed DNA samples (Southern blot No.1) but were raised from a glycerin stock of the respective

and incubated for approximately four more generations (c.f. Northern blot). We decided to choose these DNA extracts for further studies due to their progressed growth and for potential comparison to the DNA extracts of shorter incubation.

As a reference gene to measure chromosome copy numbers, Sso 3194 encoding non-phosphorylating GAP dehydrogenase (GAPN) was chosen as PCR target being unique in the *S. solfataricus* genome. Sso 3194 was amplified using q3194FW and RV primer. The qPCR assay specific for pIZ-protospacer plasmids was achieved using 406 Q primers binding the ORF406 embedding the protospacer D63 (c.f. Fig.10A). Fig. 17 depicts a bar diagram of mean values of measured pIZ-protospacer copies and M18 chromosomes per 1µg DNA of transformed M18 cells (analyzed in biological and technical triplicates). Comparisons of plasmid and chromosomal copy numbers allow conclusions of approximate copy numbers per cell. However, it must be taken into account that *Sulfolobales* exhibit a fluctuant genome number changing between one and two per cell. Consequently, copy numbers per cell are potentially higher than the qPCR data convey. The chromosome amounts are relatively stable among all transformants ranging between $3.2x10^6$ (M18-406R) and $3.98x10^6$ (M18-7U) copies per µg total DNA (80µl). The plasmid value for pIZ-WOP shows the highest amount of $3.16x10^6$ followed by pIZ-406R with $2.1x10^6$ copies. The quantity of the pIZ-HA plasmid representing $3.1x10^5$ copies is ten times lower than the CRISPR negative controls. M18-pIZ-7U DNA revealed very few pIZ-7U plasmid amounts of no more than 322 copies.

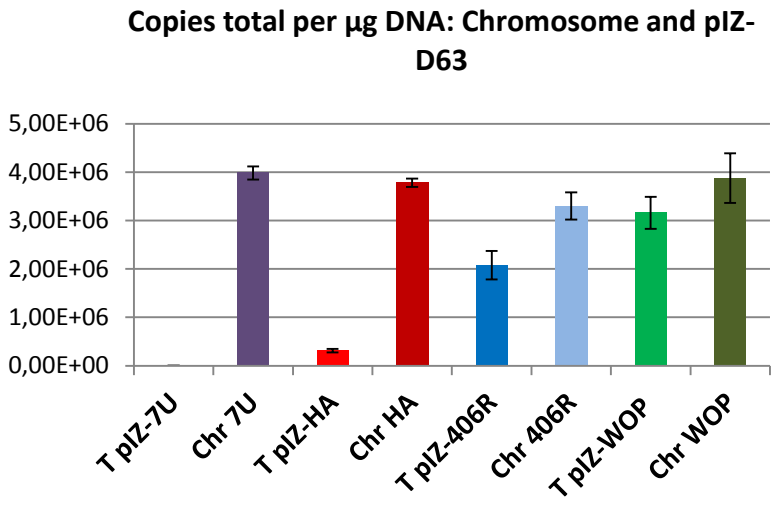
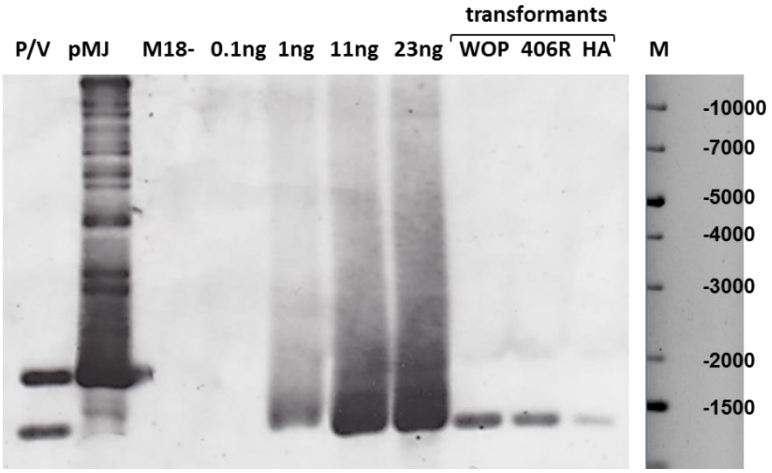


Fig. 17: Mean values of chromosomal and pIZ-plasmid-derived copies of triplicates of 1µg total DNA of induced transformants analysed by qPCR. Chromosomal copies were determined using q3194FW and RV primer amplifying the reference gene Sso 3194 occurring only once in the genome of *S. solfataricus.* Amounts of pIZ-plasmids of the same samples were evaluated via protospacer specific 406 Q FW and RV primer amplification. Bars representing chromosomal and plasmid counts of same samples are depicted side by side and in same ground colours (pIZ-7U transformants: violet; pIZ-HA transformants: red; pIZ-406R transformants: blue; pIZ-WOP transformants: green). Plasmid counts are abbreviated as "T" (transformants) + plasmid name and the chromosomal counts as "Chr" + protospacer acronym. The amounts of chromosomal counts do not differ notably from each other, whereas the plasmids show varying amounts.

4.3.3.1.2 Southern blot analysis according to qPCR data

A second Southern blot was performed using the same DNA extractions for comparing copy numbers gained by qPCR with dilutions of plain pIZ-HA plasmids representing different copy numbers. For M18-pIZ-WOP/406R transformants 6µg were loaded on the blot, whereby 18µg were applied for M18-pIZ-HA. The M18-pIZ-7U sample was omitted due to too low amounts. As virus controls, pMJ0305 vector was used and for comparison of virus and plasmid a 1:5 mixture of M18-pMJ0305 DNA and M18-pIZ-WOP DNA was applied. Additionally, untransformed M18 DNA was loaded as negative control. The copy numbers of all samples as determined by qPCR are given in Fig.18. Total DNA of the samples was digested with *Ase*I whereby dilutions and pMJ0305 digestion were used from stocks prepared for the previous Southern blot. For M18-protospacer transformants, plasmid specific bands are visible at 1384bp and for bands specific for SSV1 virus plasmids appear shortly below 2000bp. The untransformed M18 control does not show any band which is also the case for the lowest dilution of 0.1ng representing $8x10^6$ copies. P/V control, representing a mixture of M18 SSV1 transformants and M18

pIZ-WOP transformants (5:1) displays both plasmid types at expected high and strength. The M18 transformants show bands of varying intensity comparable to the 1ng dilution control representing $8.4x10^7$ copies.



6µg DNA of M18-pIZ-WOP and -406R transformants; 18µg DNA of M18-pIZ-HA transformants	
pIZ-WOP	$1.92x10^7$
pIZ-406R	$1.2x10^7$
pIZ-HA	$3.04x10^6$
pIZ-WOP plasmid dilutions	
0.1 ng dil.	$8.4x10^6$
1 ng dil.	$8.4x10^7$
11ng dil.	$9.26x10^8$
23 ng dil.	$1.94x10^9$
Positive/negative controls	
P/V (5µg/1µg)	$1.3x10^7$/n/a
55 ng pMJ-control	$2.4x10^9$
3µg M18 DNA	n/a

Fig. 18: Southern hybridization using pIZ-plasmid and SSV1 virus specific probes on *Ase*I digested total DNA extracts of pIZ-protospacer M18 transformants. 6µg DNA of M18-pIZ-WOP and M18-pIZ-406R transformants as well as 18µg of M18-pIZ-HA DNA were used in Southern hybridization performed with ORF904 and D291 digoxigenin probes in parallel. ORF904 probes hybridize to 1384bp fragments, whereas virus specific D291 probe binds a 1957bp fragment. Copy numbers of pIZ-HA/WOP/406R within the respective amount of used DNA were determined according to previous qPCR data and listed in the table on the right. pIZ-7U untransformed plasmid dilutions (23ng – 11ng – 1ng – 0.1ng) representing different copy numbers, cited in the table are visualized as well as the P/V control consistent of 5µg M18-pIZ-WOP DNA and 1µg DNA of M18-pMJ0305 transformants. 55ng of untransformed pMJ0305 plasmid is shown (pMJ) and 3µg of untransformed M18 DNA was loaded as negative control.

4.3.4 Semi quantitative PCR assay using ORF904 primer on qPCR samples and DNA extracts of former transformants

To elucidate whether plasmid quantification determined by 406 Q primer specific qPCR analysis remains constant when using a different pIZ-plasmid specific primer set, a semi quantitative PCR was established with primers ORF904FW and RV amplifying 1030 bp. Same DNA replicates analyzed via qPCR were applied as templates for the semi quantitative PCR in same amounts. Furthermore, DNA extracts used in the first Southern blot analysis stemming from the same transformed cell stock as the qPCR samples but grown in media for a shorter period and without supplemented arabinose were analyzed as well. Additionally, DNA extractions from former transformation assays, where M16 was transformed with pIZ-7U and M18 with pIZ-406R never been processed before were also tested. For all samples, 25 ng were used as template and 20 cycles were set for PCR.

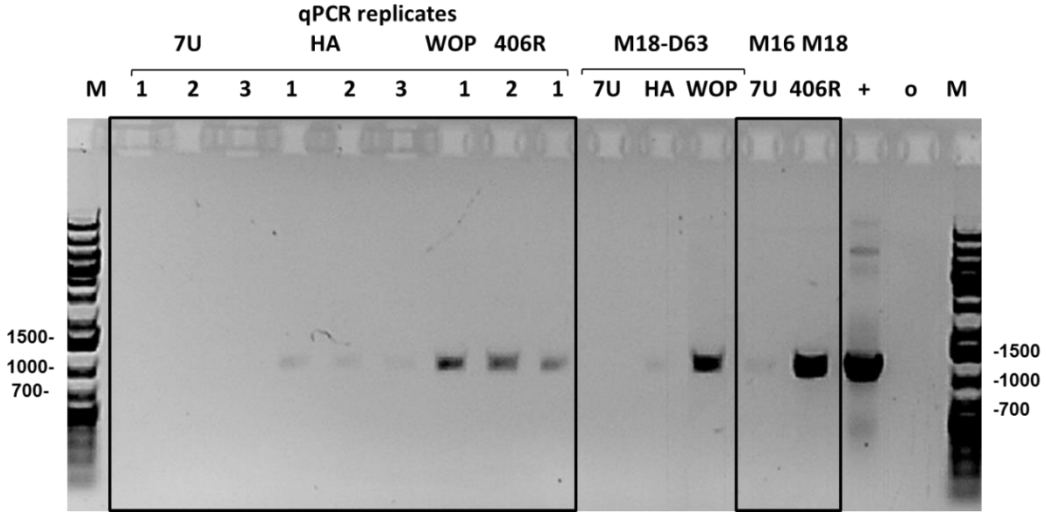


Fig. 19: Semi quantitative PCR on different DNA extractions of M18-pIZ-transformants using ORF904FW and RV primers specifically amplifying 1030bp of the pIZ-plasmids. DNA extracts from "induced" M18-pIZ-7U/HA/WOP/406R replicates used in former qPCR analysis are visible on the left side of the gel (left box). The different numbers refer to different replicates of the same DNA extract, whereby for M18-pIZ-WOP only one replicate was processed (qPCR replicates). Samples cited as M18-D63 represent the same DNA extractions as used for the first Southern blot hybridization stemming from the same transformation approach as the qPCR templates but from DNA extractions of transformants grown without arabinose supplementation. PCR products of M18-pIZ-7U/HA and – WOP were loaded. The box located on the right half of the gel highlights the PCR products of M16-pIZ-7U and M18-pIZ-406R DNA resulting from former transfection studies of M16 and M18 cells using the pIZ-7U and pIZ-406R protospacer, respectively. (+) positive control: untransformed pIZ-HA plasmid, (o): water control. 25ng total DNA served as template for 20 cycles PCR. Total amounts (25µl) of PCR amplification were loaded on 0.8% agarose. M= 2.5µl (right) and 3µl (left) DNA ladder, 1kb+ (Thermoscientific).

The gel shows bands appearing at the expected size of 1030bp specific for ORF904 primer amplification. Regarding the qPCR replicates (left box), M18-pIZ-HA PCR products show faint bands, whereby the PCR fragments of M18-pIZ-WOP and -406R are visible (Fig.19). The samples of the first DNA extract (located in the center of the gel) reflect the same pattern. The PCR product of the control sample M16-pIZ-7U indicates a band of very low intensity, whereby the PCR fragment of M18-pIZ-406R control is clearly visible (Fig.19).

4.3.5 Homologous recombination analysis

Since M18-pIZ-7U and M18-pIZ-HA transformants contained very low plasmid numbers but showed approximately the same cell amounts and growth rates as M18-pIZ-WOP and M18-pIZ-406R harboring the pIZ-plasmids complementing the uracil auxotrophy, the chromosomal *pyrEF* originally disrupted by a 1359bp sized IS element (ISC1359) could potentially be converted to an intact one. To test for such a conversion, semi quantitative PCR was established using newly designed primers. Primer pyrB_FW and pyrE_RV were designed to bind the first nucleotides of the *pyrB* and within the *pyrE* gene, respectively (Fig.10A and Fig.20). Owing to the difference in size of the PCR product being 326bp in case of a wild type *pyrEF* and 1658bp within M18 mutants, intact genes are distinguishable from disrupted ones (Fig.20). Since the primer also bind the *pyrEF* selection marker on the pIZ plasmids (leading to the smaller fragment), all samples were analyzed via a second PCR using ORF904 primers binding the plasmid backbone and therefore determining the plasmid quantities under same conditions. This parallel approach identifies the amplification origin of the 326bp *pyrEF* PCR fragment, since it should show same intensities as the ORF904 amplicon when originating from the pIZ-plasmid.

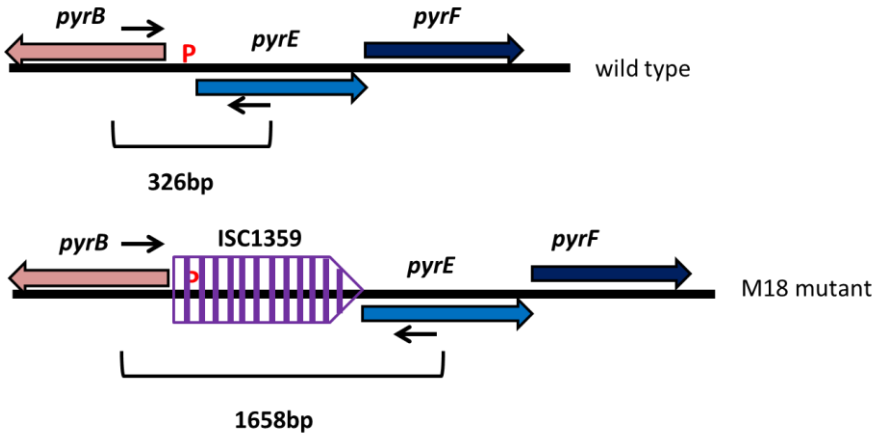


Fig. 20: Schematic depiction of the intact and disrupted *pyrEF* gene cluster and sizes of fragments yielded upon pyrB FW and pyrE RV PCR amplification. Primer pyrB FW and pyrE RV (black arrows) amplify a 326bp fragment in case of the functional *pyrEF* gene cluster present in *S. solfataricus* P1 wild type cells or on the pIZ-plasmid vectors. A PCR product of 1658bp is amplified in PCR in case of IS element 1359 insertion disrupting the promoter region of the *pyrEF* genes which consequently renders the cell uracil auxotrophic (mutant M18). PCR amplification using the primer pair pyrB FW and pyrE RV therefore distinguishes between functional and disrupted *pyrEF.*

As in qPCR, 25ng of total DNA were used as templates and same amount of untransformed pIZ-HA plasmid was used as positive - and "empty" pENTRY vector as negative control within semi quantitative PCR analysis. Entire amounts of PCR reactions were visualized on agarose gels depicted below (Fig.21) which are discussed right away in this section for retaining lucidity. For simplicity, the PCR product of functional *pyrEF* either originating from the pIZ-plasmid or from the chromosomal intact gene clusters is named *pyrEF300* and the disrupted is designated as *pyrEF1600* indicating the approximate size of the fragments.

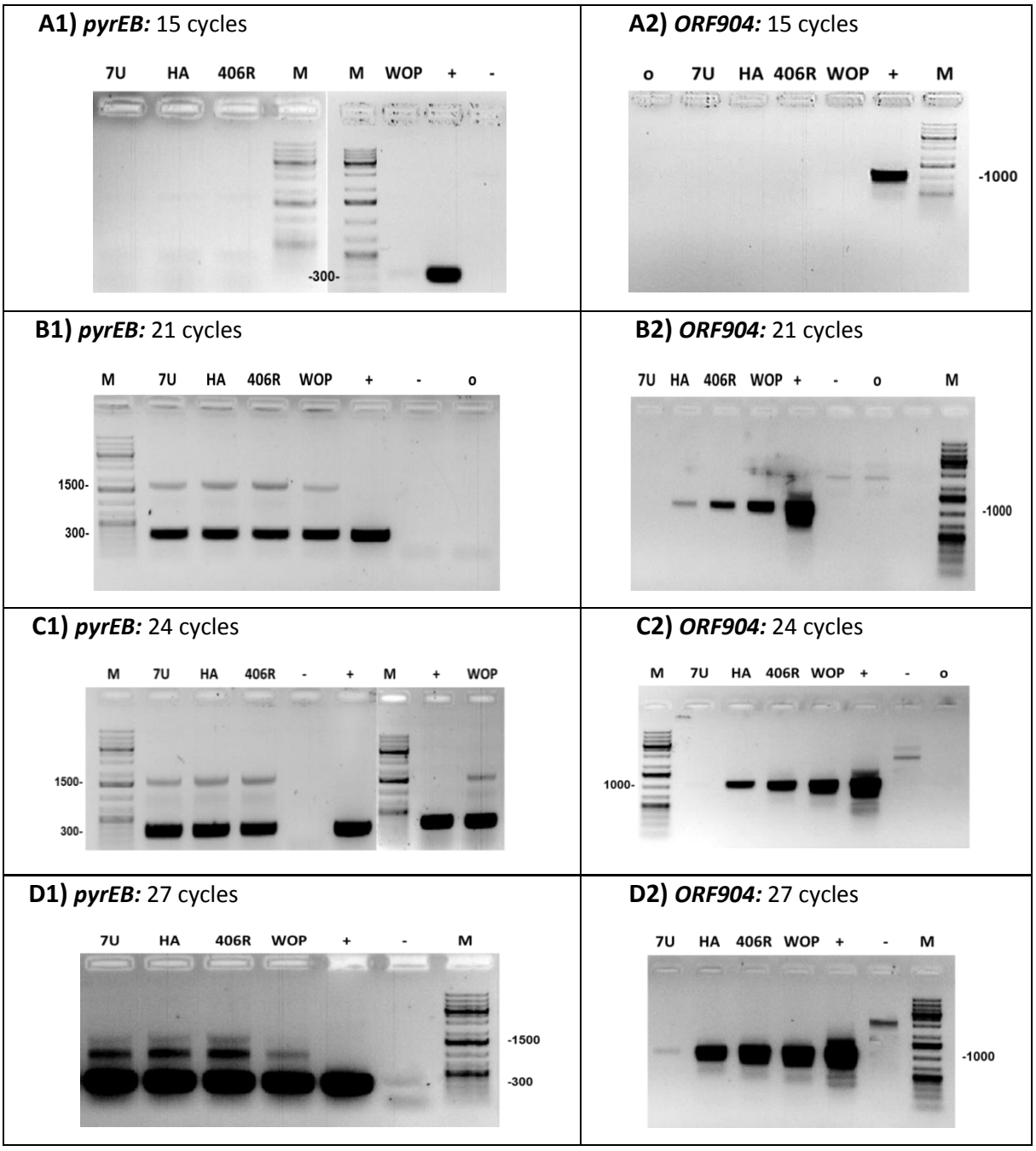


Fig. 21: Semi quantitative PCR of M18-pIZ-protospacer total DNA. PCR was obtained for 15, 21, 24 and 27 cycles whereby same templates were used (25ng). pyrEB primers were used for the recombination PCR (left) and ORF904 primers were used for the plasmid control PCR (right). Entire PCR samples (25µl) were loaded on 0.8% agarose. M= 3µl marker lane: 1kb+ GeneRuler (Thermo Scientific).

After 15 cycles, indications of very faint bands of both, the *pyrEF1600* as well as for the *pyrEF300* are visible (Fig.21 A1), whereby the corresponding plasmid control gel representing ORF904 PCR products used as benchmark for plasmid quantity does not

show any band beside the positive control (Fig.21 A2). Upon six additional duplication rounds, important changes have occurred within the different samples visualized in Fig.21 B1,B2. M18-pIZ-protospacer samples show the *pyrEF1600* and *pyrEF300* bands at 1658bp and 326bp, respectively, whereby in any case, *pyrEF300* seems to be more intensive than the bigger one representing the defective *pyrEF* (Fig.21 B1). The ORF904 plasmid control gel for 21 cycles depicts intense bands at expected position (1030 bp) for M18-pIZ-WOP and - 406R samples, whereby only a faint band is visible for M18-pIZ-HA and no band is apparent for the M18-pIZ-7U sample (Fig.21 B2).

These results strongly indicate that *pyrEF300* in case of M18-pIZ-7U DNA stems from an intact chromosomal *pyrEF* as well as the major part of the *pyrEF300* from M18-pIZ-HA does, since no plasmid is detected for M18-pIZ-7U and a weaker plasmid band signal (Fig.21 B2) than that of the *pyrEF300* band (Fig.21 B1) is visualized for M18-pIZ-HA transformants (Fig.21 B2).

The two CRISPR negative controls, M18-pIZ-WOP and M18-pIZ-406R show a similar intensity of the *pyrEF300* band as the ORF904 plasmid specific fragment, leading to the assumption that the *pyr300* originates manly from the pIZ-vectors (Fig. 21 B1,B2). Certainly, revertants within WOP and 406R culture cannot be excluded.

The *pyrEF1600* also appears within all samples in a far lower quantity than the *pyrEF300*. In case of M18-pIZ-WOP,-406 and -HA this band is referred to the chromosomal *pyrEF* still being mutated to allow the presence of the complementing plasmid. Nevertheless, the intensity of disrupted *pyrEF1600* was expected to be approximately as high as that of the ORF904 band and *pyrEF300* band, respectively, when correlating the amount of detected plasmid directly to the amount of chromosomes. Only the M18-pIZ-HA correlates with this conception, whereby M18-pIZ-WOP and -406R show a less intense *pyrEF1600*. A possible explanation for this is that PCR primer have been absorbed by the more abundant shorter *pyrEF* fragment not allowing the bigger fragment to amplify as much as the smaller one leading to an almost consistent band intensity of *pyr1600* in all samples. Also for M18-pIZ-7U, a 1600bp fragment is detectable indicating the presence of some remnant *pyrEF* disrupted chromosomes within the culture.

The same pattern is visible for 24 cycles (Fig.21C), whereby the 300bp band within *pyrEF* amplification shows a higher intensity than the 1658bp band (Fig.20C1). When comparing the *pyrEF1600* bands of cycle 21 and cycle 24, no considerable increase of product is obvious, whereas the 326bp fragment seems to gain intensity (Fig.21 B1, C1).

After 27 cycles, the pIZ-7U plasmid is visible as very faint band within the ORF904 control PCR (Fig.21 D2). The gel depicting *pyrEF* amplification shows an additional band beneath the *pyrEF1600* supposed to be a chimera of the 1600bp *pyrEF* band owing to the technique used for semi quantitative PCR, where additional cooling steps were juxtaposed to take out the samples after corresponding cycles (see Materials and Methods). This is underpinned by the decrease of intensity of the 1600bp band and was also observed in previous PCRs using the same method (data not shown).

To surely verify the quantity of pIZ plasmids and exclude unspecific amplification or recombination within the plasmid, again a 21 cycle PCR was conducted using primer pair ORF904 FW and pyrE RV spanning 5100 bp of the whole pIZ-D63 plasmid. The gel shows specific bands for all constructs with higher intensities for pIZ-406R and –WOP as for –HA. pIZ-7U is not visible (Fig. 22).

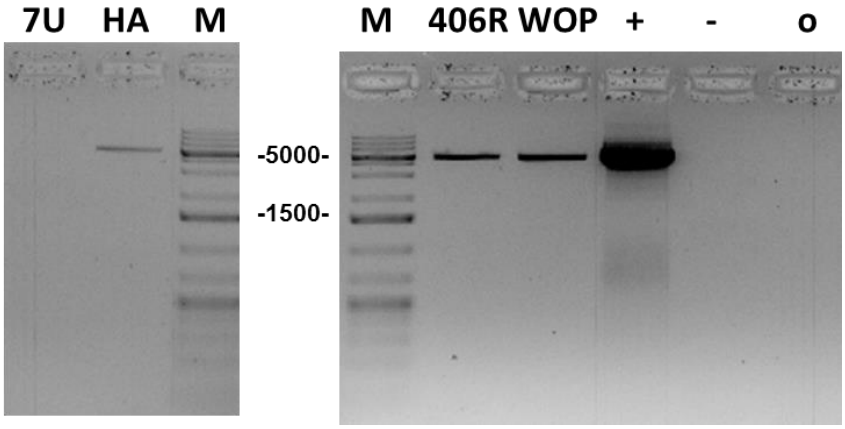


Fig. 22: Semi quantitative PCR on M18-pIZ-protospacer total DNA. Primer combination of pyr904 FW and pyrE RV were chosen and 21 cycles were set. M= 2.5µl marker lane: 1kb+ GeneRuler (Thermo Scientific), 0.8% agarose.

5 Discussion

5.1 Development of a genetic toolbox for S. solfataricus P1

In this study a plasmid – based shuttle vector for *Sulfolobus* was constructed on the basis of pRN1. pRN1 was chosen since pRN family based shuttle vectors had already been successfully used to transform *S. islandicus, S. acidocaldarius*, *S. tokodaii* and recently also in *S. solfataricus* P2 (pRN2 derivate) (Berkner et al., 2007, 2010a; Deng et al., 2009; Gudbergsdottir et al., 2011, Zebec, in prep.). Additionally, pRN protein family motifs were found integrated in the genome of *S. solfataricus* P2 and *S. tokodaii* testifying *S. solfataricus* being a possible host for this plasmid group (Kawarabayasi et al., 2001; X. Peng, Holz, Zillig, Garrett, & She, 2000). It is worth to mention here, that during this study, we never found any integrated pRN1 plasmid or any chromosomal background signals of residual pRN sequences (at least in the range of our detection methods) in the *S. solfataricus* P1 (c.f. Southern blot and PCR data).

The shuttle vector pCmalLacS (Berkner et al., 2010b) comprising the pRN1 plasmid backbone replicating in *Sulfolobus* and a part of the pBlueskript vector for propagation in *E.coli* was modified rendering the pCAra-GW vector harboring a Gateway® cassette for quick cloning procedures via *in vitro* recombination. In addition, an arabinose promoter enabled induction of transcription of a further recombined genetic gene upon arabinose supply (Fig.16). The Gateway® cassette – arabinose promoter insert is situated counterclockwise within the vector which was proven by restriction analysis (Fig.8B). The destination vector pCAra-GW can only be propagated in special *E.coli* strains, such as DB 3.1. carrying a mutation within the gyrase therefore being unaffected by the toxic CcdB expressed from the Gateway® cassette (Bernard & Couturier, 1992). After *in vitro*

recombination, i.e. substitution of the Gateway® cassette with a sequence of interest, expression vectors can be propagated in any *E.coli* strain. Replication in *E.coli* is possible due to the ColE1 origin of replication and selection for positive *E.coli* transformants is achieved on ampicillin drop out media owing to the *ampR* gene expressing ampicillin hydrolyzing β-lactamase. Propagation of the plasmid within *S. solfataricus* is ensured by the presence of pRN1 specific ORF904 and ORF56 shown to be essential for proper replication within *Sulfolobus* (Berkner & Lipps, 2007; Lipps, 2009). The sole open reading of the plasmid pRN1 not comprised entirely in the pCAra-GW is ORF90b has been reported to be dispensable for proper propagation (Berkner et al., 2007). The expression vectors based on the pCAra-GW construct complement uracil auxotrophy in *pyrEF* defective mutant strains by the *pyrEF* selection marker of *S. solfataricus* P2. Therefore, cells transformed with such expression vectors are grown in uracil-free media.

5.1.1.1 Transformation studies in the S. solfataricus β-galactosidase mutant M16

For the first transformation study, pCAra-GW was recombined with the entry clone pENTRY-*lacS* rendering the 11113 bp sized expression vector pIZ-βGal carrying the *lacS* gene of *S. solfataricus* P2 coding for β-galactosidase (Fig.9A). The recombination solution was directly used to transform *E.coli* E.R 1812 (NewEngland Biolabs) cells conferring C5 methylation of CCGG sequence stretches in order to protect the plasmids from restriction digestion by enzymes specifically recognizing this motif. This methylation step is generally conducted prior to transformation of *S. acidocaldarius* with any vector since the Archaeon was shown to harbor the R.*Sua*I restriction barrier efficiently degrading foreign DNA within CCGG patterns (Grogan & Hansen, 2003; Kurosawa & Grogan, 2005). For *S. solfataricus* no restriction methylation barrier has been described so far, but pIZ-βGal was methylated anyway for being on the safe side.

The *S. solfataricus* M16 mutants which harbors insertional elements within its *pyrEF* and *lacS* genes (Martusewitsch et al., 2000) was transformed using the pIZ-βGal vectors in two different approaches. The M16 ST (standard) approach was performed according to the protocol generally used for pRN1 plasmid transformation of *S. acidocaldarius* (Berkner et

al., 2007). The ICE method was chosen since it was successfully applied for *S. islandicus* transformation using the pRN2 based plasmid pHZ2 (Deng et al., 2009). After a first incubation of both cultures in tryptone/sucrose uracil selection media for 11 days, an inoculum was transferred in N-Z-Amine/sucrose media. The switch of carbon source possibly displays a crucial step in this transformation study, since strong background growth has been reported for *S. solfataricus* auxotrophic mutants on tryptone containing uracil free media (Berkner et al., 2007). In contrast to *S. acidocaldarius* not showing background growth of *pyrEF* mutants in uracil selective media, *S. solfataricus* was elucidated to contain homologs of uracil permeases which possibly enable mutant growth on very low amounts of uracil (Berkner et al., 2007). Use of acid hydrolyzed N-Z-Amine instead of tryptone has been shown to reduce background growth, whereas complete prevention has not been achieved (C. Zhang & Whitaker, 2012b). However, the selective pressure acting on the cells was sufficient in our study, since transformation of pIZ-βGal plasmids was successful for all samples initially positively identified by PCR on total DNA extracts (Fig.12B). Both transformation methods seemed to have similar transformation efficiencies and led to similar growth of cultures since both different treated cultures have reached almost identical optical densities at same time points (Fig. 12B). Stationary phase was reached after 4 days incubation from starting cultures with an OD_{600} of 0.2 (Fig. 12B). Negative controls for both methods did not show continuing growth after incubation with N-Z-Amine solution underpinning the enhanced selection force due to the chemical component. It is worth to mention, that raise of pIZ – transformed M18 cells from glycerin stocks showed 8 hours generation time after one additional transfer in new media (i.e. after four generations). These findings indicate both methods to be constructive for transformation of *S. solfataricus* P1, whereby further transformation studies were conducted using the ST method.

In earlier attempts of plasmid transformation only low amounts of transformed plasmids were verified by PCR reactions in total DNA, but neither were the authors able to visualize positive transformants via different detection methods nor did they achieve retransformation of total DNA of transformants in *E.coli* (Berkner & Lipps, 2008a; Leigh et al., 2011). However, within this study, the pIZ-βGal plasmid was successfully recovered

from *E.coli* TOP10 upon retransformation of total DNA from pIZ-βGal transformed M16 cells. Due to technical reasons (loss of samples during the plasmid extraction), only the plasmids extracts of the ST method were restriction digested using *Eco*RI for integrity and verification check of the plasmid. Beside the digest, also untreated plasmids were loaded on agarose. Two out of three recovered plasmids showed exactly the same restriction pattern as the original pIZ-βGal plasmid used as control (Eco2,3, Fig.12B). Thus, the retransformation proved the successful transformation of M16 cells using the vector pIZ-βGal and the maintenance of the respective within the cells (Fig.12B). Beside the two identical patterns of retransformed plasmids, lane no.1 (Eco1) shows an additional intense band between 10000 and 20000bp (Fig.12B). The apparent band resembles a linearized but not further digested plasmid since its position can be assigned the size of pIZ-βGal. The other bands of this sample reflect the same restriction digest pattern as visible for the samples 2 and 3, even though at lower intensity. When inspecting the total plasmid load of extract no. 1, the most intense band is detected at about 4 kb which compared to the topology test of untransformed pIZ-βGal plasmids (Fig. 8C) corresponds to the suggested supercoil conformation. Entire plasmids of samples Eco2 and Eco3 appear identical and do not show the 4000kb band. In general, supercoiled elements are not as amenable for restriction digest as untangled DNA. Higher concentrations of restriction enzymes are even recommended for restriction reactions on supercoiled targets (NewEngland BioLabs® guidance for cleavage of supercoiled DNA: https://www.neb.com/tools-and-resources/selection-charts/cleavage-of-supercoiled-dna). Therefore we speculate the supercoil conformation of plasmids in sample Eco1 being responsible for incomplete digest of some pIZ-βGal plasmids resulting in an additional undigested band.

Furthermore, also *S. tokodaii* cells were transformed with pIZ-βGal using the ST strategy and N-Z-Amine/glucose media directly in the first incubation. Transformed cultures set to an OD_{600} of 0.2 needed 10 hours to reach an OD_{600} of 0.4, whereas the negative control did not achieve an OD higher than 0.03. PCR on total c92 DNA resulted in two bands instead of one indicating either changes of some plasmids within the cells such as

sequence deletions, partly integrations or homologous recombinations or they may result from contamination.

The impossibility to recover the pIZ-βGal plasmids from retransformation of total c92 transformants indicates a possible integration event of the plasmid. Since it was shown that *S. tokodaii* is transformable using the pCmalLacS plasmid (Zebec, in prep) we assume the cells to be transformed but we could not elucidate the origin of the second band resulted from PCR amplification. To clarify these results, the transformation should be repeated in further studies.

5.2 Application of the pIZ plasmid system for CRISPR studies in S. solfataricus

Differences in plasmid maintenance between plasmids carrying matching protospacers and plasmids without

The protospacer constructs D63-7U and D63-HA are reverse complements of the spacer 63 of the chromosomal CRISPR locus D of *S. solfataricus* P1 leading to complementary binding of their mRNA (transcribed in sense direction) to the D63-crRNA mediating CRISPR interference. These two protospacer constructs show 7 mismatches to the spacer where cytosine was substituted with thymine via inverse PCR. Furthermore, the HA protospacer has an 8 nucleotide PAS sequence at its 3'end matching the 5' 8nt – handle of the D63-crRNA. DNA interference is supposed to be achieved by the CRISPR/Cas system, where an enzyme complex (CASCADE or Csm in case of *S. solfataricus* P1) is guided by an incorporated crRNA to a complementary invader DNA sequence which in turn is bound by the crRNA (crRNA:ssDNA hybrid) and subsequently cleaved by the complex. However, *in vivo* studies in *S. epidermidis* which were recently confirmed in our laboratories for *S. solfataricus* P1 have shown that at least 4 matching basepairs between the 5' handle derived from the 5'repeat of the chromosomal spacer of the crRNA and the 3' PAS of the protospacer leads to protection from DNA interference, since the CRISPR system recognizes the protospacer as "self" preventing the chromosomal spacers from

degradation (Manica et al., 2013; L. a Marraffini & Sontheimer, 2010a). Recent *in vivo* transfection studies of the auxotrophic *S. solfataricus* mutant M18 using SSV1 virus based shuttle vectors carrying the protospacer constructs indicated these features (mismatches and handle matching PAS) to represent effective protections from CRISPR mediated DNA interference (Zebec, submitted.). The mismatches of the protospacer were shown to be tolerated within the mRNA:crRNA hybrid (GU pairing) whereas they restricted the binding of the crRNA to the protospacer DNA and consequently 68% of DNA interference. The additional handle binding in case of construct D63-HA protected almost 100% of DNA interference. This protection led to the degradation of the protospacer mRNA possibly mediated by the CRISPR/Cmr complex which was revealed by qPCR and also visualized on a Northern blot (Zebec, submitted). This study has proven a silencing effect of the CRISPR system in *S. solfataricus* P1.

The protospacers D63-7U and D63-HA together with the construct 406R containing a protospacer not matching a spacer in P1 and therefore not triggering CRISPR interference (Zebec, submitted) and construct WOP not containing any protospacer were placed into pCAra-GW yielding pIZ-7U/pIZ-HA/pIZ-406R and pIZ-WOP plasmid construct (Fig.7B, 10A). The initial idea was to find out whether the above mentioned results from infection studies and therefore RNA interference could be detected by using a non-infectious plasmid as genetic shuttle instead of a viral vector. Like in previous pIZ-βGal transformation studies, M18-pIZ-protospacer transformants showed an average generation time of 24 hours for all four transformants, whereby the negative controls did not grow. After six generations, DNA was extracted and transformation was verified by ORF904FW/RV plasmid specific primer amplification, whereby the M18-pIZ-7U sample and to a lesser extent also M18-pIZ-HA show a PCR product of much fainter than the M18-pIZ-WOP and M18-pIZ-406R samples, respectively indicating these transformants to contain less plasmid than the others (Fig. 14A). To verify this result and to get an idea of the plasmid amount and its cellular state (integrated or extrachromosomal), a Southern blot was conducted on 3µg digested total DNA samples using a plasmid specific probe (ORF904) and a virus probe (D291) in parallel. The results of the Southern hybridization reflect the PCR reaction insofar that no plasmid was detected for M18-pIZ-7U and only

very low amounts were discovered in M18-pIZ-HA total DNA. Contrary, for the CRISPR negative controls M18-pIZ-WOP and M18-pIZ-406R a band at ORF904 specific size of 1384bp was clearly visible. These results strongly indicate a difference of plasmid maintenance or transformation efficiency between the CRISPR system affecting and non-affecting protospacers. Since Southern blot hybridization is not as sensitive as PCR amplification, those differences appear even more dramatic on the blot leading to a non-visible 7U plasmid whereas after 32 cycle PCR the presence of 7U was verified (c.f.14A, 15B).

These findings differ enormously from the results of the above mentioned study using pMJ0305 SSV1 based vectors, where maintenance of the protospacer constructs HA was equal to the controls, since CRISPR mediated DNA interference was abolished (Zebec, submitted). Thus, Southern blot and PCR results indicate that pRN1 based pIZ-plasmids harboring protospacers matching a spacer of the chromosome are less stable than the pIZ plasmids without a matching protospacer even when carrying a GU and handle which normally leads to protection from CRISPR mediated DNA degradation (L. a Marraffini & Sontheimer, 2010b; Zebec, submitted.).

When comparing the intensities of the CRISPR negative samples M18-pIZ-WOP and M18-pIZ-406R to the pIZ-HA plasmid dilutions (controls) representing 5-10 copies per cell (rough calculation according to chromosomal copy numbers) it is obvious that plasmid numbers are not as abundant in 3µg total DNA indicating a low copy number of pIZ-plasmids. However no exact copy number determination can be obtained on basis of this comparison. Beside the verification of M18 cells to carry the pIZ plasmids, even if in varying quantities, this Southern blot provides evidence that no virus contamination was present within the transformants since the virus specific probe hybridized to a band size of 1957bp as apparent in the pMJ (untransformed SSV1 viral based vector), M18+ (M18 transformed with pMJ) and the V/P control. The latter is a mixture of M18-pIZ-HA total DNA and M18+ where the specificity of both probes is proven by the visualization of the virus and the plasmid within one sample. Unfortunately, the M18-pIZ-HA DNA was used for this control containing only low amounts of the plasmid which explains the barely

visible plasmid band in the control sample (red arrows). The second Southern blot (Fig.18) displays a better implementation of this control.

Given that only one band at expected size is visible in total DNA preparations of transformants, an integration of the plasmid into the chromosome or other alterations in plasmid conformation can be excluded. This is in accordance to other pRN1 shuttle vector constructs reported to be sustained extrachromosomally in *S. acidocaldarius* cells (Berkner et al., 2007).

Inserts are inducible upon arabinose supplementation

The arabinose induction of the transformants was performed to answer the following questions: i) is the arabinose promoter working and ii) is CRISPR specific mRNA degradation in case of HA and 7U detectable, as observed in previous studies (Zebec, submitted). The arabinose promoter was reported to yield a six – fold induction of insert transcription and contrary to other promoters, low basal activity (Lubelska et al., 2006; N. Peng, Xia, Chen, Liang, & She, 2009). The Northern blot probe visualized the whole mRNA of ORF406 (comprising the protospacer D63) of the induced samples in case of M18-pIZ-WOPi and M18-pIZ-406Ri (samples treated with arabinose) whereby the RNA of uninduced samples was not detectable which confirms the stringency of the promoter. For M18-pIZ-HAi, only a smear is visible and no band is apparent in case of M18-pIZ-HAu, M18-pIZ-7Ui and u, respectively (Fig. 16). The absence of the RNA for these samples can clearly be referred to the low amount of plasmid present in those cells. The integrity test of all samples yielded the same expected pattern and same intensity of all rRNA subunits excluding a general loss of RNA possibly due to technical reasons in case of these 7U and HA samples. The fail to visualize the HA and 7U transformants because of quantity reasons and detection limits of the Northern blot was not surprising. The question whether there is mRNA interference which would have led to an additional band beneath the 1000 originating from the cut within the D63 protospacer mRNA, in case of the few residual cells carrying the pIZ-HA of pIZ-7U plasmid therefore remains unanswered.

Correlation between chromosomal numbers and transformed plasmids

For a better valuation of the amounts of the pIZ constructs within the M18 transformants, a qPCR was established. Total amounts of chromosomes and plasmids comprised in 1µg DNA of the same transformants were determined and as expected demonstrated a clear underrepresentation of pIZ-7U and pIZ-HA plasmids within the cell culture whereas pIZ-WOP and pIZ-406R are present in much higher amounts (Fig.17). However, chromosomal counts do not vary considerably among the different samples indicating similar cell amounts for each sample (Fig.17). When comparing chromosomal copies to plasmid reads, the pIZ-HA plasmid represents only 10% of the total chromosomal amount, whereas the pIZ-7U not even makes up 1% of the chromosomal content (Fig.17). When considering these findings together with the observation that all cultures showed the same growth rate and approximately same chromosome contents it becomes obvious that the M18 cells initially transformed with pIZ-7U and pIZ-HA plasmids, respectively, have found a way to grow without the uracil auxotrophy complementing plasmid leading to the assumption that the cells managed to restore the disrupted *pyrEF* to a functional one. Before addressing this topic in detail, the question of the copy number of pIZ-WOP and pIZ-406R plasmids per cell has to be answered. The CRISPR negative controls seem to be “normally” maintained within the culture and therefore can be considered as benchmark for copy number determination of expression vectors with pCAra backbone.

It was shown that the cell cycle of *Sulfoblobus* cells residing in stationary phase arrests in the G2 stage prohibiting further cell division consequently leaving the cells with two chromosomes (Bernander, 1997). Also, within the dividing exponential phase, cells were found to often harbor two genomes at once (Bernander, 1997). Therefore, chromosome counts cannot be directly correlated to cell amounts making an exact determination of the copy number impossible since the amount of the cells residing in the G2 stage cannot be determined using qPCR. In case of the CRISPR negative control plasmids pIZ-WOP and pIZ-406R, all cells at least carry one plasmid when considering as many as 20% and 45% of the M18-pIZ-WOP and M18-pIZ-406R transformants to carry two chromosomes, respectively. However, lower numbers of cells harboring two genomes cannot be excluded meaning that residual cells in CRISPR negative controls not harboring the

plasmid but able to grow because of restoration of their *pyrEF* are present, albeit in low amounts. Taken all together, pIZ vectors are low copy number plasmids showing 1-2 copies per cell when considering two chromosomes being ubiquitous per cell which is expected since low copy numbers were also reported for other pRN1 plasmids derivates used in transformation studies of *S. acidocaldarius* (Berkner et al., 2007).

We decided to perform another Southern blot using the same probes as in the first one but calculated the amounts of applied DNA on the basis of the qPCR data. The blot revealed that the DNA extracts of pIZ-protospacer transformants used for qPCR ("induced") were also free from virus contamination since the D291 probe visualizing any SSV1 derivate was visible in the virus containing controls (pMJ, P/V) but not in the transformants samples (Fig. 18). Furthermore, untransformed M18 DNA control did not show any band which excluded any unspecific binding of the probes or natural occurring pRN1 plasmid or SSV1 virus within M18 cells (Fig.18). The intensities of pIZ-WOP and pIZ-406R in 6µg of DNA of M18 transformants could be correlated to the 1ng dilution of the pIZ-HA plain plasmid control representing $8.4x10^7$ copies. This agreed with the qPCR data from which $1.92x10^7$and $1.2x10^7$were calculated for 6µg DNA of M18-pIZ-WOP and M18-pIZ-406R, respectively. 18µg M18-pIZ-HA DNA representing $3.04x10^6$ copies showed a lighter band than the 1ng which is expected. The 0.1ng control sample which should have represented the 10^6 dilution did not reveal any band maybe due to technical reasons because of the low dilution rate. The V/P virus control showed two specific bands for the ORF904 plasmid probe and SSV1 probe resulting from a mixture of M18-pMJ0305 and M18-pIZ-WOP DNA.

Semi-quantitative PCR assay and reproducibility of results

To verify the qPCR data directly with different primers, a semi quantitative PCR assay was performed using ORF 904FW/RV primers and 20 cycles (Fig.19). pIZ-WOP and pIZ-406R show bands of approximately the same intensities as expected since they both reached the qPCR cycle threshold after 18-19 cycles, whereas pIZ-HA showed only a very faint band which is also in accordance to qPCR since it reached the threshold after 21 cycles. Samples from previous DNA extraction (used in Southern blot no.1) expectedly showed

the same intensities. For this analysis, DNA was extracted from a previous transformation assay, where M16 was transformed with pIZ-7U and M18 with pIZ-406R. The pIZ-7U M16 control revealed fragments of expected size fainter than that of the M18-pIZ-HA transformants whereas the pIZ-406R M18 control showed a band as intense as the other CRISPR negative controls used in this study. These samples represent very important controls since they proof the reproducibility of the outcomes gained in the transformation studies in that pIZ-7U shows far more diminished plasmid maintenance than the CRISPR negative controls. This underpins the assumption that the loss of the plasmid pIZ-7U is caused by the presence of the matching protospacer and therefore by CRISPR interference. The question is now, what happened to the plasmids where CRISPR seems to have acted on. The qPCR and also growth analysis showed that there were no losses in cell number and growth, therefore the cells must have undergone changes leading to prototrophy upon infection with protospacer carrying plasmids. To study this phenomenon, homologous recombination studies were established.

5.3 Homologous recombination as rescue effect upon CRISPR interference?

S. solfataricus mutant M18 is a P1 strain containing the insertional element ISC1359 within the promoter region of the gene cluster *pyrEF* situated 12bp upstream of the *pyrE* (Martusewitsch et al., 2000). This promoter disruption leads to breakdown of uridine *de novo* synthesis and renders the cells auxotrophic for uracil (Martusewitsch et al., 2000). Since the amounts of chromosomal copies evaluated by qPCR range from $2x10^6$ in case of M18-pIZ-406R up to $3.9x10^6$ for M18-pIZ-7U per µg DNA and therefore are quite similar for all transformants, whereas only the inactive protospacer constructs M18-pIZ-WOP and -406R efficiently harbor a pIZ vector complementing the *pyrEF* function within the cell, the chromosomal *pyrEF* was supposed to be reverted to an intact copy being the reason for growth of the M18-pIZ-7U and most of M18-pIZ-HA cells.

Therefore, all samples were checked via PCR using different cycle numbers representing a semi quantitative method when comparing the intensities of products derived from different cycle numbers on the gel. *pyrEB* primers amplified a 1658 bp fragment in case of a disrupted *pyrE* within mutant cells and a 326 bp amplicon was produced for an intact gene cluster not containing the IS element present in reverted cells as well as transformed cells, since the *pyrEF* of the pIZ plasmid is also covered by the primer pair. To distinguish between chromosomal - and plasmid derived 326bp product, the same PCR approach was applied for ORF904 plasmid specific primer whereby intensities of intact *pyrEF* fragments could be compared to those of ORF904 PCR products within same cycles. Comparing PCR products gained in 15,21,21 and 27 cycle PCRs, it is obvious that the intact *pyrEF* in case of M18-pIZ-7U and M18-pIZ-HA mainly originates from the chromosome, since no plasmids/lower amounts of plasmid are detected in plasmid verification PCRs using ORF904 primer. Contrary, in the CRISPR negative controls M18-pIZ-WOP and M18-pIZ-406R approximately same intensities of the *pyrEF* and the plasmid are visible allowing the suggestion that the major part of the amplified intact *pyrEF* comes from the plasmid, as expected.

All in all, these data imply that the majority of M18-pIZ-7U and M18-pIZ-HA cells which have been initially transformed with a pIZ plasmid containing a full size protospacer matching the CRISPR spacer D63, have restored their functional *pyrEF.* However this, if at all happened only to a minor extent to M18-pIZ-WOP and M18-pIZ-406R transformants still harboring the pIZ plasmids not matching a spacer. Also, for previous transformation of M16 with the protospacer construct pIZ-7U, very low levels of the plasmid could be detected in semi-quantitative PCR (Fig.19), whereas liquid cultures showed growth like the control culture of M18-pIZ-406R (data not shown) proven to contain the plasmid. Although the *pyrEF* of M16-pIZ-7U total DNA was examined, it is quite possible that also here a conversion to a functional gene cluster had happened.

Unsuccessful previous attempts to transform *S. solfataricus* P1 were mainly attributed to the revision of the *pyrEF* and therefore the loss of the plasmid. People assumed either a recombination to occur or a reversion through the active removal of the IS element itself

probably through stress triggered by electroporation (Berkner & Lipps, 2008a; Christa Schleper, personal communication).

The results of our study imply that the reversion to intact *pyrEF* in case of M18-pIZ-7U and M18-pIZ-HA happened via homologous recombination between the plasmid and the chromosome and not upon overgrowth of revertants in consequence of electroporation. Fundamental arguments allowing this assumption are the plasmid maintenance and absence of an obvious revision within the samples pIZ-WOP and pIZ-406R for which the same competent cells and electroporation settings were used and samples were treated equally (e.g. incubation in the same shaker, same flasks, etc.). Also, plasmid maintenance was observed in further transformation studies using pIZ vectors, again with exception of M16-pIZ-7U. Moreover, the negative control that had been treated the same way as the samples but without addition of a plasmid, did not show any growth over a period of three weeks with continuous supply of nutrients, excluding an immediate "natural" rise of revertants (data not shown). Furthermore, it is very likely that the plasmid derived *pyrEF* recombines with a disrupted chromosomal one in case of mutant cells carrying an insertional element within the mutated *pyrEF* and not a full length deletion. Such recombinant cells arose as byproducts within the first targeted gene disruption study in *S. solfataricus* 98/2 derivate PBL 2002 (*lacS*::IS*1217*) performed by Worthington and colleagues, where they used a knock out cassette consisting of a recombinant but functional *lacS* gene flanked by sequences of the chromosomal α amylase gene. Beside the desired genotype showing the gene knock out upon recombination of the flanking sites of the cassette, they also obtained cells having exchanged the disrupted chromosomal *lacS* with the intact one (Worthington et al., 2003). Contrary, such cells were not observed when using the same approach in *lacS* deletion cells not displaying homologous stretches (Schelert et al., 2004). Owing to the usage of the *S. solfataricus* P2 *pyrEF* as selection marker on the pIZ plasmid similar to that of M18 cells, identical gene stretches are offered for homologous recombination.

The main question to be answered is why only pIZ-7U and pIZ-HA constructs undergo recombination, whilst the others are stably maintained within the cells. Since pIZ-7U and

pIZ-HA are the only constructs carrying a protospacer matching the spacer 63 of the CRISPR locus D we deduce the CRISPR system to target the protospacer which in turn triggers homologous recombination.

Considering D63RC- 7U and -HA protospacers being targeted by the CRISPR/Cas system on the DNA level would lead to nucleolytic cleavage of the protospacer ssDNA-crRNA hybrid and consequently to disruption of the pIZ-plasmid. Given that the plasmid harbors the *pyrEF* essential for survival, the cells might have subsequently undergone homologous recombination (HR) as a rescue effect. Studies in *S. acidocaldarius pyrE* deletion strains reported a considerable increase of homologous recombination events when transforming unmethylated plasmids being targets for the *Sua*I restriction barrier over methylated (Kurosawa & Grogan, 2005b). The authors elucidated linearized or fragmented DNA to be more favorable used for HR which is according to findings in bacterial, eukaryotic cells and other Archaea such as *Thermococcus kodakariensis*, *Haloferax volcanii* and *Methanococcus voltae* (Kurosawa & Grogan, 2005b; Patel et al., 1994; Sato et al., 2003). In the study of Sato and coworkers only double–crossover based homologous recombination events were observed even when transforming circular plasmids designed for single crossover, i.e. integration of the whole plasmid. Double-crossover can only occur between linear and circular DNA, inferring that plasmids were linearized prior to HR (Sato et al., 2003). These findings intensify the assumption of the CRISPR/Cas system to trigger homologous recombination, since participant enzymes were proven to cut within the protospacer at GC regions which consequently would lead to a linearization of the plasmid. The pIZ-plasmid has not been found integrated in the chromosomes for neither cells which leads to the assumption of a double-crossover event within the M18-pIZ-7U and most of –HA cells (see above).

Recent *in vivo* studies in *S. solfataricus* demonstrated the enormous impact of the CRISPR/Cas on transfection efficiencies when using pMJ0305 virus based vectors carrying a protospacer covering a chromosomal spacer. Total immunity, i.e. no plaques/no transformants were observed for protospacer constructs showing a 100% match to the chromosomal spacers as result of degradation of the virus (Manica et al., 2011). This was

also observed within another study performed *in S. solfataricus* P2 and *S. islandicus*, where plasmids were used as shuttle vectors carrying a perfect matching protospacer. Different to the viral approach, transformation exclusively relied on *pyrEF* selection equal to our study. They observed only few transformants upon growth on selective media whereof the majority showed broad deletions of whole CRISPR loci in order to maintain the plasmid complementing uracil auxotrophy (Gudbergsdottir et al., 2011). An auxotrophic mutant strain of *S. solfataricus* P2 carrying an IS element disrupting the *pyrEF* was used in that study, whereby the plasmid harbored a *pyrEF* derived from *S. acidocaldarius* to omit possible recombination events. Thus, this study greatly contributes to the understanding of ours because on the one hand, just as the virus studies, it demonstrates the enormous pressure of the CRISPR/Cas system leading to invariable elimination of the matching invader even if promoting cell death and on the other hand that cells counteract the CRISPR system in rather deleting whole sequences to keep the plasmid in the cell than getting rid of the IS element within their *pyrEF.* This once more underpins our assumption of HR to have occurred instead of reversion through IS removal.

The protospacers D63RC-7U and -HA were originally designed to circumvent DNA interference by GU pairing in case of both constructs and through an addition of a 8nt sequence at the 3'end of HA, perfectly matching the "autoimmunity" handle of the crRNA, as depicted and described above. Virus based *in vivo* transfection studies of M18 using the protospacer construct D63-7U omitted 68% of DNA interference and D63RC-HA was protected to 95% determined by transfection efficiency (Zebec, submitted.). Prevention of DNAi enabled examination of RNA interference on the protospacer performed by the CRISPR/Cmr complex. Indeed, degradation of the 7U and HA mRNA could be detected via qPCR and even by Northern blot analysis (Zebec, submitted.).

Within our studies, no stable maintenance was detected for the 7U protospacer carrying plasmid and only low amounts, i.e. 10% of plasmid (according to total chromosomal amounts, qPCR data) were found in HA samples making an examination of RNAi impossible (c.f. Northern blot). Obviously, there is more plasmid maintained in case of the

pIZ-HA vector transformants than in 7U transformants reflecting previous studies in that HA is not as strongly affected by DNA interference as 7U. Nevertheless, the proportions should be higher, since approximately 95% of cells in case of HA carried the intact virus. These are reasons to expect that DNAi is possibly not adequately prevented in our study indicating a fundamental difference to the previous findings using recombinant viral pMJ0305 vectors. In protospacer studies targeting family II spacers (locus A and B) using pMJ0305 vector derivates, the CRISPR/Csm (typeIIIA) was supposed to be the DNA attacking complex in *S. solfataricus* P1, since a PAM seemed to be irrelevant for interference (Manica et al., 2013, 2011). Contrary, plasmid based protospacer studies of Gudbergsdottir et al. revealed a CC PAM for family I CRISPR loci being necessary for proper DNAi (comprising locus D of *S. solfataricus*) (Gudbergsdottir et al., 2011) suggesting that the CASCADE (type I) is the targeting agent. Following these assumptions, the 7U and HA protospacers used in our study should be prevented from type IA cleavage due to the lack of a typical PAM motif and of course, the GU pairing mismatches. However, the dependence of the different Cas complexes on a PAM and its general role in CRISPR interference has not been completely solved yet.

The D63 crRNA was shown to be most prominent within the dsRNA degrading Cmr complex which is neither dependent on a PAM nor is it sensitive to a paired handle on the protospacer whereas a flap at the 3' end is essential (Hale & Duff, 2010; J. Zhang et al., 2012, Zebec in revision). The Cmr complex of *S. solfataricus* P2 was postulated not to cut DNA (J. Zhang et al., 2012). This was reinforced by recent studies in *S. islandicus* containing Cmr-α which is distantly related to the Cmr of *S. solfataricus* showing antisense transcription dependent DNA interference, whereby Cmr-β closely resembling the CMR of *S. solfataricus* was again not susceptible for DNAi (Deng et al., 2013). According to these findings it can be excluded that the Cmr actively is involved in plasmid degradation, since this would require DNA cleaving activity.

Even so, the complex could be the driving force for homologous recombination since RNA interference, which was proven to be active on the transcripts of the protospacers (HA: 45% and 7U: 25%) (Zebec,in prep.) is an energy draining process. Therefore,

recombination would be a beneficial solution for the cells shedding the plasmid without any losses. Although pMJ0305 vectors harbor the *pyrEF* selection marker, they are less dependent on the auxotrophic selection owing to their spreading nature. The infectiousness of those virus plasmids together with their high copy number per cell and their integration into the chromosome would therefore keep them in the culture even upon *pyrEF* recombination of the chromosomes. Considering this, a virus based vector is not affected by recombination and CRISPR targeting therefore is continuing.

Our data clearly point to the possibility that the different impacts of the CRISPR system on the same protospacers within the same host can be led back to the nature of the genetic vehicle harboring the respective and therefore it can be argued that the CRISPR system affects low copy number plasmids differently than scattering viruses.

6 Conclusion

The establishment of a plasmid-based shuttle vector system for the Archaeon *S. solfataricus* P1 has long lagged behind due to instability of the plasmids after transformation. PyrEF disruption mutants of *S. solfataricus* P1, being uracil auxotrophs, were reported to lose the non-spreading vectors in consequence of leaky *pyrEF* selection (Berkner et al., 2007). However, our study showed that transformation of *S. solfataricus* P1 mutants with a pRN1-based cryptic plasmid is possible by increasing selection pressure from the media by using N-Z-Amine instead of tryptone. Altogether, five pIZ-plasmid variants were transformed at least once in this study and their presence in transformed culture was proven by different methods. Beside PCR reactions verifying two plasmid-specific ORFs and once the half plasmid, intact pIZ-lacS vectors were successfully recovered from *E.coli* cells upon retransformation of total *S. solfataricus* DNA. Moreover, two Southern hybridizations on two different total DNA extracts from M18-pIZ-WOP - and M18-pIZ-406R transformants (CRISPR negative controls) proved the presence of the plasmid and its exclusive episomal state. Additionally, qPCR quantified the pIZ plasmids and revealed an average copy number of one to two plasmids per cell.

The second part of this study once more underlined the complexity and variability of the CRISPR system. When providing SSV1 based viral shuttles and pIZ-vector variants with identical protospacers having a cognate CRISPR spacer in the *S. solfataricus* chromosome, the CRISPR responses vary significantly with respect to the replicon: Modified protospacers that contain complementary sequences at their 3’ end to the 5’ autoimmunity handle of the interfering crRNA to circumvent DNA interference were not targeted when provided on a SSV1 vector (Manica et al., 2013; Zebec, submitted), but seemed to be degraded when carried with the plasmid. Since the loss of the plasmid is only detected in pIZ-vectors harboring matching protospacers but not in equally treated negative controls, general failure of transformation was excluded. Furthermore, the same phenomenon was observed in PCR assays, when screening *S. solfataricus* M16 cells

transformed with the same plasmids. We therefore argue that the modifications to circumvent CRISPR/Cas mediated DNA interference in SSV1 infection, are not sufficient to protect plasmid DNA from degradation. Cells which have lost the plasmid had probably undergone homologous recombination to recover their defective *pyrEF*. Literature shows linearized plasmids to be more susceptible to homologous recombination than circular ones. Since Cas complexes cleave matching sequences, this might even specifically trigger recombination events. Homologous recombination of cells carrying the intact plasmid without protospacer was not observed in this experiment.

Altogether, the toolbox of *S. solfataricus* P1 was extended by a plasmid based shuttle vector system which can be used to do genetic studies in the host. The non-infectious nature of the plasmid is advantageous over the uncontrollable spreading nature of the SSV1 virus, but seems to evocate different CRISPR responses than the virus does. According to these facts, the plasmid is not suitable to replace the virus vector in doing CRISPR studies, but it reveals the system in a different light. It seems that not only the sequence directly interacting with the crRNA determines the CRISPR reaction, but that also the nature of the extrachromosomal element influences the outcome.

7 References

Aagaard, C., I. Leviev, R. N. Aravalli, P. Forterre, D. Prieur, and R. A. Garrett. 1996. General vectors for archaeal hyperthermophiles: strategies based on a mobile intron and a plasmid. FEMS Microbiol. Rev. 18:93–104.

Albers, S., Jonuscheit, M., Dinkelaker, S., Urich, T., Tampé, R., Driessen, A. J. M., … Tampe, R. (2006). Production of Recombinant and Tagged Proteins in the Hyperthermophilic Archaeon Sulfolobus solfataricus Production of Recombinant and Tagged Proteins in the Hyperthermophilic Archaeon Sulfolobus solfataricus. doi:10.1128/AEM.72.1.102

Albers, S.-V., & Driessen, A. J. M. (2008). Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome. Archaea (Vancouver, B.C.), 2(3), 145–9.

Allers, T., & Mevarech, M. (2005). Archaeal genetics - the third way. Nature reviews. Genetics, 6(1), 58–73. doi:10.1038/nrg1504

Aravalli, R.N., and Garrett, R.A. (1997) Shuttle vectors for hyperthermophilic archaea. Extremophiles 1: 183–191.

Aucelli, T., Contursi, P., Girfoglio, M., Rossi, M., & Cannio, R. (2006). A spreadable, non-integrative and high copy number shuttle vector for Sulfolobus solfataricus based on the genetic element pSSVx from Sulfolobus islandicus. Nucleic acids research, 34(17), e114. doi:10.1093/nar/gkl615

Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., … Yakunin, A. F. (2011). A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. Molecular microbiology, 79(2), 484–502. doi:10.1111/j.1365-2958.2010.07465.x

Balch, W. E., Fox, G. E., Magrum, L. J., Woese, C. R., & Wolfe, R. S. (1979). Methanogens: reevaluation of a unique biological group. Microbiological reviews, 43(2), 260–96.

Barns, S. M., Delwiche, C. F., Palmer, J. D., & Pace, N. R. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. Proceedings of the National Academy of Sciences of the United States of America, 93(17), 9188–93.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., … Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science (New York, N.Y.), 315(5819), 1709–12. doi:10.1126/science.1138140

Berkner, S., Grogan, D., Albers, S.-V., & Lipps, G. (2007). Small multicopy, non-integrative shuttle vectors based on the plasmid pRN1 for Sulfolobus acidocaldarius and Sulfolobus solfataricus, model organisms of the (cren-)archaea. Nucleic acids research, 35(12), e88. doi:10.1093/nar/gkm449

Berkner, S., & Lipps, G. (2007). Characterization of the transcriptional activity of the cryptic plasmid pRN1 from Sulfolobus islandicus REN1H1 and regulation of its replication operon. Journal of bacteriology, 189(5), 1711–21. doi:10.1128/JB.01586-06

Berkner, S., & Lipps, G. (2008a). Genetic tools for Sulfolobus spp.: vectors and first applications. Archives of microbiology, 190(3), 217–30. doi:10.1007/s00203-008-0392-4

Berkner, S., & Lipps, G. (2008b). Mutation and reversion frequencies of different Sulfolobus species and strains. Extremophiles : life under extreme conditions, 12(2), 263–70. doi:10.1007/s00792-007-0125-7

Berkner, S., Wlodkowski, A., Albers, S.-V., & Lipps, G. (2010a). Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius. Extremophiles : life under extreme conditions, 14(3), 249–59. doi:10.1007/s00792-010-0304-9

Berkner, S., Wlodkowski, A., Albers, S.-V., & Lipps, G. (2010b). Inducible and constitutive promoters for genetic systems in Sulfolobus acidocaldarius. Extremophiles : life under extreme conditions, 14(3), 249–59. doi:10.1007/s00792-010-0304-9

Bernander, R. (1997). Cell cycle characteristics of thermophilic Cell Cycle Characteristics of Thermophilic Archaea, 179(16).

Bernard, P., & Couturier, M. (1992). Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. Journal of Molecular Biology, 226(3), 735–745. doi:http://dx.doi.org/10.1016/0022-2836(92)90629-X

Bertani, G. (1951): *Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli*. In: *J. Bacteriol.* Bd. 62, Nr. 3, S.293-300

Bertani, G., and L. Baresi. 1987. Genetic transformation in the methanogen Methanococcus voltae PS. J. Bacteriol. 169:2730- 2738

Bitan-banin, G., Ortenberg, R., & Mevarech, M. (2003). Development of a Gene Knockout System for the Halophilic Archaeon Haloferax volcanii by Use of the pyrE Gene

Development of a Gene Knockout System for the Halophilic Archaeon Haloferax volcanii by Use of the pyrE Gene, 185(3). doi:10.1128/JB.185.3.772

Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. (2005). Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. Microbiology (Reading, England), 151(Pt 8), 2551–61. doi:10.1099/mic.0.28048-0

Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F., & Forterre, P. (2005). Nanoarchaea: representatives of a novel archaeal phylum or a fast-evolving euryarchaeal lineage related to Thermococcales? Genome biology, 6(5), R42. doi:10.1186/gb-2005-6-5-r42

Brochier-Armanet, C., Boussau, B., Gribaldo, S., & Forterre, P. (2008). Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. Nat Rev Micro, 6(3), 245–252. Retrieved from http://dx.doi.org/10.1038/nrmicro1852

Brock, T. D., K. M. Brock, R. T. Belly & R. L. Weiss, (1972) Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol 84: 54-68.

Brouns, S. J. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J. H., Snijders, A. P. L., … van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defense in prokaryotes. Science (New York, N.Y.), 321(5891), 960–4. doi:10.1126/science.1159689

Bryant, M. P. (1972). Commentary of anaerobic on the Hungate technique for culture, 1324–1328.

Cannio, R., Contursi, P., Rossi, M., & Bartolucci, S. (1998). An autonomously replicating transforming vector for Sulfolobus solfataricus. Journal of bacteriology, 180(12), 3237–40. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=107829&tool=pmcentrez&rendertype=abstract

Carroll, D. (2004) Using nucleases to stimulate homologous recombination. Meth. Mol. Biol. 262, 195–207.

Carte, J., Wang, R., Li, H., Terns, R. M., & Terns, M. P. (2008). Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes, 3489–3496. doi:10.1101/gad.1742908.archaea

Cline, S. W., & Doolittle, W. F. (1987). Efficient transfection of the archaebacterium Halobacterium halobium. Journal of bacteriology, 169(3), 1341–4. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=211943&tool=pmcentrez&rendertype=abstract

Clore, A. J., & Stedman, K. M. (2007). The SSV1 viral integrase is not essential. Virology, 361(1), 103–11. doi:10.1016/j.virol.2006.11.003

Deng, L., Garrett, R. a, Shah, S. a, Peng, X., & She, Q. (2013). A novel interference mechanism by a type IIIB CRISPR-Cmr module in Sulfolobus. Molecular microbiology, 87(5), 1088–99. doi:10.1111/mmi.12152

Deng, L., Zhu, H., Chen, Z., Liang, Y. X., & She, Q. (2009). Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon Sulfolobus islandicus. Extremophiles : life under extreme conditions, 13(4), 735–46. doi:10.1007/s00792-009-0254-2

Cubellis,M.V., Rozzo,C., Monteucchi,P. and Rossi,M. (1990) Isolation and sequencing of a new beta-glycosidase-encoding archaebacterial gene. Gene, 94, 89–94.

Deveau, H., Barrangou, R., Garneau, J. E., Labonté, J., Fremaux, C., Boyaval, P., … Moineau, S. (2008). Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. Journal of bacteriology, 190(4), 1390–400. doi:10.1128/JB.01412-07

Elferink MG, Schleper C, Zillig W. Transformation of the extremely thermoacidophilic archaeon Sulfolobus solfataricus via a self-spreading vector. FEMS Microbiol Lett. 1996 Mar 15;137(1):31-5. PubMed PMID: 8935654. combination. Meth. Mol. Biol. 262, 195–207.

Elkins, J. G., Podar, M., Graham, D. E., Makarova, K. S., Wolf, Y., Randau, L., … Stetter, K. O. (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. Proceedings of the National Academy of Sciences of the United States of America, 105(23), 8102–7. doi:10.1073/pnas.0801980105

Erauso, G, Marsin, S., Baucher, M. F., Erauso, G., Marsin, S., Benbouzid-rollet, N., … Prieur, D. (1996). Sequence of plasmid pGT5 from the archaeon Pyrococcus abyssi : evidence for rolling-circle replication in a hyperthermophile . Sequence of Plasmid pGT5 from the Archaeon Pyrococcus abyssi : Evidence for Rolling-Circle Replication in a Hyperthermophile.

Erauso, Gaël, Stedman, K. M., van de Werken, H. J. G., Zillig, W., & van der Oost, J. (2006). Two novel conjugative plasmids from a single strain of Sulfolobus. Microbiology (Reading, England), 152(Pt 7), 1951–68. doi:10.1099/mic.0.28861-0

Erdmann, S., & Garrett, R. a. (2012). Selective and hyperactive uptake of foreign DNA by adaptive immune systems of an archaeon via two distinct mechanisms. Molecular microbiology, 85(6), 1044–56. doi:10.1111/j.1365-2958.2012.08171.x

Felgner, P. L., Gadek, T. R., Holm, M., Roman, R., Chan, H. W., Wenz, M., … Danielsen, M. (1987). Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proceedings of the National Academy of Sciences of the United States of America,

84(21), 7413–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=299306&tool=pmcentrez&rendertype=abstract

Forterre, P., Brochier, C., & Philippe, H. (2002). Evolution of the Archaea. Theoretical Population Biology, 61(4), 409–422. doi:10.1006/tpbi.2002.1592

Garneau, J. E., Dupuis, M.-E., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P., … Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature, 468(7320), 67–71. Retrieved from http://dx.doi.org/10.1038/nature09523

Grabowski, B., & Kelman, Z. (2003). ARCHAEAL DNA REPLICATION: Eukaryal Proteins in a Bacterial Context. Annual Review of Microbiology, 57(1), 487–516. doi:10.1146/annurev.micro.57.030502.090709

Greve, B., Jensen, S., Brügger, K., Zillig, W., & Garrett, R. a. (2004). Genomic comparison of archaeal conjugative plasmids from Sulfolobus. Archaea (Vancouver, B.C.), 1(4), 231–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2685578&tool=pmcentrez&rendertype=abstract

Grogan, D. W. (1997). Rates of spontaneous mutation in an archaeon from geothermal environments . Rates of Spontaneous Mutation in an Archaeon from Geothermal Environments, 179(10).

Grogan, D.W. (1991) Evidence that beta-Galactosidase of Sulfolobus solfataricus is only one of several activities of a thermostable beta-D-glycosidase. Appl Environ Microbiol 57: 1644–1649. Grogan,

Grogan, D. W. & R. P. Gunsalus, (1993) Sulfolobus acidocaldarius synthesizes UMP via a standard de novo pathway: results of biochemical-genetic study. J Bacteriol 175: 1500-1507.

Grogan, D. W., & Hansen, J. E. (2003). Molecular Characteristics of Spontaneous Deletions in the Hyperthermophilic Archaeon Sulfolobus acidocaldarius, 185(4), 1266–1272. doi:10.1128/JB.185.4.1266

Gudbergsdottir, S., Deng, L., Chen, Z., Jensen, J. V. K., Jensen, L. R., She, Q., & Garrett, R. a. (2011). Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. Molecular microbiology, 79(1), 35–49. doi:10.1111/j.1365-2958.2010.07452.x

Hale, C. R., & Duff, M. O. (2010). NIH Public Access, 139(5), 945–956. doi:10.1016/j.cell.2009.07.040.RNA-Guided

Holmes, M., Pfeifer, F., & Dyall-Smith, M. (1994). Improved shuttle vectors for Haloferax volcanii including a dual-resistance plasmid. Gene, 146(1), 117–121. doi:http://dx.doi.org/10.1016/0378-1119(94)90844-3

Huber, H., Hohn, M. J., Rachel, R., Fuchs, T., Wimmer, V. C., & Stetter, K. O. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. Nature, 417(6884), 63–67. Retrieved from http://dx.doi.org/10.1038/417063a

Huet, J., Schnabel, R., Sentenac, a, & Zillig, W. (1983). Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. The EMBO journal, 2(8), 1291–4. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=555274&tool=pmcentrez&rendertype=abstract

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, a. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. Journal of bacteriology, 169(12), 5429–33. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=213968&tool=pmcentrez&rendertype=abstract

Jain, R., Rivera, M. C., & Lake, J. a. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. Proceedings of the National Academy of Sciences of the United States of America, 96(7), 3801–6. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=22375&tool=pmcentrez&rendertype=abstract

Jansen, R., Embden, J. D. a Van, Gaastra, W., & Schouls, L. M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. Molecular microbiology, 43(6), 1565–75. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11952905

Jimenez, A. (1985). The Mechanism of Resistance to Puromycin and to the Puromycin-precursor 0-Demethyl-puromycin in, 2877–2883.

Jonuscheit, M., Martusewitsch, E., Stedman, K. M., & Schleper, C. (2003). A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector. Molecular Microbiology, 48(5), 1241–1252. doi:10.1046/j.1365-2958.2003.03509.x

Kate, M. (1993). Chapter 9 Membrane lipids of archaea. In D. J. K. and A. T. M. B. T.-N. C. B. M. Kates (Ed.), The Biochemistry of Archaea (Archaebacteria) (Vol. Volume 26, pp. 261–295). Elsevier. doi:http://dx.doi.org/10.1016/S0167-7306(08)60258-6

Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., … Kikuchi, H. (2001). Complete Genome Sequence of an Aerobic Thermoacidophilic Crenarchaeon , Sulf olobus tokodaii strain7, 140, 123–140.

Konneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B., & Stahl, D. A. (2005). Isolation of an autotrophic ammonia-oxidizing marine archaeon. Nature, 437(7058), 543–546. Retrieved from http://dx.doi.org/10.1038/nature03911

Kozubal, M. A., Romine, M., Jennings, R. deM, Jay, Z. J., Tringe, S. G., Rusch, D. B., … Inskeep, W. P. (2013). Geoarchaeota: a new candidate phylum in the Archaea from high-temperature acidic iron mats in Yellowstone National Park. ISME J, 7(3), 622–634. Retrieved from http://dx.doi.org/10.1038/ismej.2012.132

Kurosawa, N., & Grogan, D. W. (2005a). Homologous recombination of exogenous DNA with the Sulfolobus acidocaldarius genome: properties and uses. FEMS microbiology letters, 253(1), 141–9. doi:10.1016/j.femsle.2005.09.031

Kurosawa, N., & Grogan, D. W. (2005b). Homologous recombination of exogenous DNA with the Sulfolobus acidocaldarius genome: properties and uses. FEMS microbiology letters, 253(1), 141–9. doi:10.1016/j.femsle.2005.09.031

Langer, D., Hain, J., Thuriaux, P., & Zillig, W. (1995). Transcription in archaea: similarity to that in eucarya. Proceedings of the National Academy of Sciences of the United States of America, 92(13), 5768–72. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=41582&tool=pmcentrez&rendertype=abstract

Leigh, J. a, Albers, S.-V., Atomi, H., & Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. FEMS microbiology reviews, 35(4), 577–608. doi:10.1111/j.1574-6976.2011.00265.x

Lillestøl, R. K., Redder, P., Garrett, R. a, & Brügger, K. (2006). A putative viral defence mechanism in archaeal cells. Archaea (Vancouver, B.C.), 2(1), 59–72. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2685585&tool=pmcentrez&rendertype=abstract

Lillestøl, R. K., Shah, S. a, Brügger, K., Redder, P., Phan, H., Christiansen, J., & Garrett, R. a. (2009). CRISPR families of the crenarchaeal genus Sulfolobus: bidirectional transcription and dynamic properties. Molecular microbiology, 72(1), 259–72. doi:10.1111/j.1365-2958.2009.06641.x

Lin, C. C., & Casida, L. E. (1984). GELRITE as a Gelling Agent in Media for the Growth of Thermophilic Microorganisms. Applied and environmental microbiology, 47(2), 427–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=239688&tool=pmcentrez&rendertype=abstract

Lintner, N. G., Kerou, M., Brumfield, S. K., Graham, S., Liu, H., Naismith, J. H., … Lawrence, C. M. (2011). Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for

antiviral defense (CASCADE). The Journal of biological chemistry, 286(24), 21643–56. doi:10.1074/jbc.M111.238485

Lipps, G. (2009). Molecular biology of the pRN1 plasmid from Sulfolobus islandicus. Biochemical Society transactions, 37(Pt 1), 42–5. doi:10.1042/BST0370042

Lubelska, J., Jonuscheit, M., Schleper, C., Albers, S.-V., & Driessen, A. M. (2006). Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon Sulfolobus solfataricus. Extremophiles, 10(5), 383–391. doi:10.1007/s00792-006-0510-7

Lucas, S., Toffin, L., Zivanovic, Y., Charlier, D., Forterre, P., & Prieur, D. (2002). Construction of a Shuttle Vector for , and Spheroplast Transformation of , the Hyperthermophilic Archaeon Pyrococcus abyssi, 68(11), 5528–5536. doi:10.1128/AEM.68.11.5528

Lundberg, K. S., Shoemaker, D. D., Adams, M. W. W., Short, J. M., Sorge, J. A., & Mathur, E. J. (1991). High-fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus. Gene, 108(1), 1–6. doi:http://dx.doi.org/10.1016/0378-1119(91)90480-Y

Makarova, K. S., Aravind, L., Wolf, Y. I., & Koonin, E. V. (2011). Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. Biology direct, 6(1), 38. doi:10.1186/1745-6150-6-38

Manica, A., Zebec, Z., Steinkellner, J., & Schleper, C. (2013). Unexpectedly broad target recognition of the CRISPR-mediated virus defence system in the archaeon Sulfolobus solfataricus. Nucleic acids research, 1–9. doi:10.1093/nar/gkt767

Manica, A., Zebec, Z., Teichmann, D., & Schleper, C. (2011). In vivo activity of CRISPR-mediated virus defence in a hyperthermophilic archaeon. Molecular microbiology, 80(2), 481–91. doi:10.1111/j.1365-2958.2011.07586.x

Marraffini, L. a, & Sontheimer, E. J. (2010a). Self versus non-self discrimination during CRISPR RNA-directed immunity. Nature, 463(7280), 568–71. doi:10.1038/nature08703

Marraffini, L. a, & Sontheimer, E. J. (2010b). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. Nature reviews. Genetics, 11(3), 181–90. doi:10.1038/nrg2749

Marraffini, L. A., & Sontheimer, E. J. (2008). CRISPR Interference Limits Horizontal Targeting DNA, 322(December), 1843–1845.

Martusewitsch, E., Sensen, C. W. and Schleper C., (2000). High Spontaneous Mutation Rate in the Hyperthermophilic Archaeon Sulfolobus solfataricus Is Mediated by Transposable Elements †, 182(9), 2574–2581.

Matsumi, R., Manabe, K., Fukui, T., Atomi, H., & Imanaka, T. (2007). Disruption of a Sugar Transporter Gene Cluster in a Hyperthermophilic Archaeon Using a Host-Marker System Based on Antibiotic Resistance Disruption of a Sugar Transporter Gene Cluster in a Hyperthermophilic Archaeon Using a Host-Marker System Based on Ant. doi:10.1128/JB.01692-06

Metcalf, W. W., Zhang, J. K., Apolinario, E., Sowers, K. R., & Wolfe, R. S. (1997). A genetic system for Archaea of the genus Methanosarcina: liposome-mediated transformation and construction of shuttle vectors. Proceedings of the National Academy of Sciences of the United States of America, 94(6), 2626–31. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=20139&tool=pmcentrez&rendertype=abstract

Mojica, F J M, Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology (Reading, England), 155(Pt 3), 733–40. doi:10.1099/mic.0.023960-0

Mojica, Francisco J M, Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. Journal of molecular evolution, 60(2), 174–82. doi:10.1007/s00239-004-0046-3

Kary B. Mullis: The polymerase chain reaction. Birkhäuser, Boston 1994

Nakata, A., Amemura, M., & Makino, K. (1989). Unusual nucleotide arrangement with repeated sequences in the Escherichia coli K-12 Unusual Nucleotide Arrangement with Repeated Sequences in the Escherichia coli K-12 Chromosome, 171(6).

Nunoura, T., Takaki, Y., Kakuta, J., Nishi, S., Sugahara, J., Kazama, H., … Takami, H. (2011). Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic acids research, 39(8), 3204–23. doi:10.1093/nar/gkq1228

Offre, P., Spang, A., & Schleper, C. (2012). Archaea in Biogeochemical Cycles. Annual Review of Microbiology, 67(1), 130628184403000. doi:10.1146/annurev-micro-092412-155614

Palm, P., Schleper, C., Grampp, B., Yeats, S., McWilliam, P., Reiter, W.-D., & Zillig, W. (1991). Complete nucleotide sequence of the virus SSV1 of the archaebacterium Sulfolobus shibatae. Virology, 185(1), 242–250. doi:http://dx.doi.org/10.1016/0042-6822(91)90771-3

Patel, G. B., Nash, J. H., Agnew, B. J., & Sprott, G. D. (1994). Natural and Electroporation-Mediated Transformation of Methanococcus voltae Protoplasts. Applied and environmental microbiology, 60(3), 903–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=201408&tool=pmcentrez&rendertype=abstract

Peck, R. F., DasSarma, S., & Krebs, M. P. (2000). Homologous gene knockout in the archaeon Halobacterium salinarum with ura3 as a counterselectable marker. Molecular Microbiology, 35(3), 667–676. doi:10.1046/j.1365-2958.2000.01739.x

Peng, N., Xia, Q., Chen, Z., Liang, Y. X., & She, Q. (2009). An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Molecular microbiology, 74(4), 928–39. doi:10.1111/j.1365-2958.2009.06908.x

Peng, X., Holz, I., Zillig, W., Garrett, R. a, & She, Q. (2000). Evolution of the family of pRN plasmids and their integrase-mediated insertion into the chromosome of the crenarchaeon Sulfolobus solfataricus. Journal of molecular biology, 303(4), 449–54. doi:10.1006/jmbi.2000.4160

Perez-Rodriguez, R., Haitjema, C., Huang, Q., Nam, K. H., Bernardis, S., Ke, A., & DeLisa, M. P. (2011). Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in Escherichia coli. Molecular microbiology, 79(3), 584–99. doi:10.1111/j.1365-2958.2010.07482.x

Possot, O., Gernhardt, P., Klein, a, & Sibold, L. (1988). Analysis of drug resistance in the archaebacterium Methanococcus voltae with respect to potential use in genetic engineering. Applied and environmental microbiology, 54(3), 734–40. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=202533&tool=pmcentrez&rendertype=abstract

Pourcel, C., Salvignol, G., & Vergnaud, G. (2005). CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. Microbiology (Reading, England), 151(Pt 3), 653–63. doi:10.1099/mic.0.27437-0

Prangishvili, D. A., Vashakidze, R. P., Chelidze, M. G., & Gabriadze, I. Y. (1985). A restriction endonuclease SuaI from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius. FEBS Letters, 192(1), 57–60. doi:http://dx.doi.org/10.1016/0014-5793(85)80042-9

Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., & Wagner, R. (2010). Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. Molecular microbiology, 75(6), 1495–512. doi:10.1111/j.1365-2958.2010.07073.x

Reeks, J., Naismith, J. H., & White, M. F. (2013). CRISPR interference: a structural perspective. The Biochemical journal, 453(2), 155–66. doi:10.1042/BJ20130316

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Har- bor, N.Y.

Sakofsky, C. J., & Grogan, D. W. (2013). Endogenous Mutagenesis in Recombinant Sulfolobus Plasmids. Journal of bacteriology, 195(12), 2776–2785. doi:10.1128/JB.00223-13

Sato, T., Fukui, T., Atomi, H., & Imanaka, T. (2003). Targeted Gene Disruption by Homologous Recombination in the Hyperthermophilic Archaeon Thermococcus kodakaraensis KOD1, 185(1), 210–220. doi:10.1128/JB.185.1.210

Schelert, J., Dixit, V., Hoang, V., Drozda, M., Blum, P., & Simbahan, J. (2004). Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon Sulfolobus solfataricus by Use of Gene Disruption Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon Sulfolobus solfataricus by Us. doi:10.1128/JB.186.2.427

Schelert, J., Drozda, M., Dixit, V., Dillman, A., & Blum, P. (2006). Regulation of mercury resistance in the crenarchaeote Sulfolobus solfataricus. Journal of bacteriology, 188(20), 7141–50. doi:10.1128/JB.00558-06

Schleper, C., Kubo, K., & Zillig, W. (1992). The particle SSV1 from the extremely thermophilic archaeon Sulfolobus is a virus: demonstration of infectivity and of transfection with viral DNA. Proceedings of the National Academy of Sciences of the United States of America, 89(16), 7645–9. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=49767&tool=pmcentrez &rendertype=abstract

Semenova, E., Jore, M. M., Datsenko, K. a, Semenova, A., Westra, E. R., Wanner, B., … Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proceedings of the National Academy of Sciences of the United States of America, 108(25), 10098–103. doi:10.1073/pnas.1104144108

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., … Van der Oost, J. (2001). The complete genome of the crenarchaeon Sulfolobus solfataricus P2. Proceedings of the National Academy of Sciences of the United States of America, 98(14), 7835–40. doi:10.1073/pnas.141222098

Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P., & Siksnys, V. (2011). Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. The EMBO journal, 30(7), 1335–42. doi:10.1038/emboj.2011.41

Söllner, S., Berkner, S., & Lipps, G. (2006). Characterisation of the novel restriction endonuclease SuiI from Sulfolobus islandicus. Extremophiles : life under extreme conditions, 10(6), 629–34. doi:10.1007/s00792-006-0019-0

Stahl, D. a, & de la Torre, J. R. (2012). Physiology and diversity of ammonia-oxidizing archaea. Annual review of microbiology, 66, 83–101. doi:10.1146/annurev-micro-092611-150128

Stedman, K. M., Schleper, C., Rumpf, E., & Zillig, W. (1999). Genetic requirements for the function of the archaeal virus SSV1 in Sulfolobus solfataricus: construction and testing of viral shuttle vectors. Genetics, 152(4), 1397–405. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1460719&tool=pmcentrez&rendertype=abstract

Thia-Toong, T. L., M. Roovers, V. Durbecq, D. Gigot, N. Glansdorff & D. Charlier, (2002) Genes of de novo pyrimidine biosynthesis from the hyperthermoacidophilic crenarchaeote Sulfolobus acidocaldarius: novel organization in a bipolar operon. J Bacteriol 184: 4430-4441.

Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., & Urich, T. (2011). archaeon from soil, 1–6. doi:10.1073/pnas.1013488108/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1013488108

Tumbula, D. L., R. A. Makula, and W. B. Whitman. 1994. Transformation of Methanococcus maripaludis and identification of a PstI-like restriction sys- tem. FEMS Microbiol. Lett. 121:309–314.

Van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., & Brouns, S. J. J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. Trends in biochemical sciences, 34(8), 401–7. doi:10.1016/j.tibs.2009.05.002

Waege, I., Schmid, G., Thumann, S., Thomm, M., & Hausner, W. (2010). Shuttle Vector-Based Transformation System for Pyrococcus furiosus Shuttle Vector-Based Transformation System for Pyrococcus furiosus □, 76(10). doi:10.1128/AEM.01951-09

Wiedenheft, B., E. van Duijn, J. B. Bultema, S. P. Waghmare, K. Zhou, A. Barendregt, W. Westphal, A. J. Heck, E. J. Boekema, M. J. Dickman and J. A. Doudna (2011). "RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions." Proc Natl Acad Sci U S A 108(25): 10092-10097.

Wagner, M., Berkner, S., Ajon, M., Driessen, A. J. M., Lipps, G., & Albers, S.-V. (2009). Expanding and understanding the genetic toolbox of the hyperthermophilic genus Sulfolobus. Biochemical Society transactions, 37(Pt 1), 97–101. doi:10.1042/BST0370097

Woese, C R, & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proceedings of the National Academy of Sciences of the United States of America, 74(11), 5088–90. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=432104&tool=pmcentrez&rendertype=abstract

Woese, Carl R, Kandlert, O., & Wheelis, M. L. (1990). Towards a natural system of organisms : Proposal for the domains, 87(June), 4576–4579.

Wolfe, R. S. (2011). Chapter one - Techniques for Cultivating Methanogens. In A. C. R. and S. W. R. B. T.-M. in Enzymology (Ed.), Methods in Methane Metabolism, Part A (Vol. Volume 494, pp. 1–22). Academic Press. doi:http://dx.doi.org/10.1016/B978-0-12-385112-3.00001-9

Worthington, P., Hoang, V., Perez-pomares, F., & Blum, P. (2003). Targeted Disruption of the α -Amylase Gene in the Hyperthermophilic Archaeon Sulfolobus solfataricus Targeted Disruption of the ⌴ -Amylase Gene in the Hyperthermophilic Archaeon Sulfolobus solfataricus. doi:10.1128/JB.185.2.482

Yeats, S., McWilliam, P., & Zillig, W. (1982). A plasmid in the archaebacterium Sulfolobus acidocaldarius. The EMBO journal, 1(9), 1035–8. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=553158&tool=pmcentrez&rendertype=abstract

Yosef, I., Goren, M. G., & Qimron, U. (2012). Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. Nucleic acids research, 40(12), 5569–76. doi:10.1093/nar/gks216

Zhang, C., & Whitaker, R. J. (2012a). A broadly applicable gene knockout system for the thermoacidophilic archaeon Sulfolobus islandicus based on simvastatin selection. Microbiology (Reading, England), 158(Pt 6), 1513–22. doi:10.1099/mic.0.058289-0

Zhang, C., & Whitaker, R. J. (2012b). A broadly applicable gene knockout system for the thermoacidophilic archaeon Sulfolobus islandicus based on simvastatin selection. Microbiology (Reading, England), 158(Pt 6), 1513–22. doi:10.1099/mic.0.058289-0

Zhang, J., Rouillon, C., Kerou, M., Reeks, J., Brugger, K., Reimann, J., … White, M. F. (2012). Europe PMC Funders Group Structure and mechanism of the CMR complex for CRISPR- mediated antiviral immunity, 45(3), 303–313. doi:10.1016/j.molcel.2011.12.013.Structure

Zheng et al., H. (2012). Development of a Simvastatin Selection Marker for a. doi:10.1128/AEM.06095-11

Zillig, W.; Palm, P.; Klenk, H.-P. (1992). A Model of the Early Evolution of Organisms: The Arisal of the Three Domains of Life from the Common Ancestor. In The Origin and

Evolution of the Cell. (Hartman, H. and Matsuno, K., eds). World Scientific Publishing Co. Pte. Ltd. Singapore 163 - 1982.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., … Kristjansson, J. K. (1993). Screening for Sulfolobales, their Plasmids and their Viruses in Icelandic Solfataras. Systematic and Applied Microbiology, 16(4), 609–628. doi:http://dx.doi.org/10.1016/S0723-2020(11)80333-4

8 Zusammenfassung

Sulfolobus solfataricus, ein Vertreter der dritten Domäne des Lebens – den Archaeen, ist ein hyperthermophiler Mikroorganismus, der sich zu einem wichtigen Modellorganismus etabliert hat. Er gehört zu den Crenarchaeota, und war der erste seiner Art, der mit einem vollständigen Transformationssystem auf Basis des Sulfolobus-shibatae Virus 1 (SSV1) ausgestattet wurde. Vektoren, die auf der DNA des SSV1 Virus basieren, verbreiten sich nicht nur über Zellteilung, sondern können sich eigenständig über Zellinfektion verteilen. Diese Eigenschaft ist einerseits von Vorteil, da das Transformationssystem nicht an ein stringentes Selektionssystem gebunden ist. Andererseits bringt der infektiöse Charakter stark schwankenden Virus Zahlen und auch Veränderungen der frei-vorliegenden Form des Virus mit sich, was zur Beeinträchtigung wichtiger genetischer Studien, wie beispielsweise Gen - Überexpression oder - Ausschaltung (Gen-Knockout), führt. Frühere Versuche, ein nicht-infektiöses Transformationssystem für *S. solfataricus* zu entwerfen, stellten sich als schwierig heraus.

Das Ziel dieser Studie war es, einen Transformationsvektor basierend auf einem Plasmid zu konstruieren, welcher in niedriger Kopienzahl in Uracil – auxotrophen *S. solfataricus* P1 Mutanten stabil erhalten bleibt.

In den bereits bestehende Vektor pCmalLacS, der auf der Nukleinsäuresequenz des pRN1 Plasmids aus *S. islandicus* basiert, wurde eine Gateway® Rekombinationskassette mit einem davorstehenden Arabinose induzierbaren Promoter innerhalb mehrerer Klonierungsschritte eingebracht. Durch diese Kassette konnten in den resultierten Zielvektor pCAra-GW gewünschte Sequenzen über die etablierte Gateway®-Klonierung eingebracht werden, wodurch verschiedene pIZ-Expressionsvektoren konstruiert wurden. Die Transkription der eingebrachten Sequenzen ist durch Zugabe von Arabinose in das Nährmedium durch den Promoter induzierbar.

Fünf verschiedene pIZ-Expressionsvektoren, jeweils eine andere integrierte Sequenz enthaltend, wurden im Laufe dieser Studie erstellt und erfolgreich innerhalb drei

verschiedener Transformationsansätzen durch Elektroporation auf die *S. solfataricus* Zellen übertragen. Die Zellen wurden anschließend in selektivem N-Z-Amine/Saccharose Medium inkubiert.

Im Gegensatz zu früheren Studien, war es uns möglich, genomische DNA der *Sulfolobus* Transformanten in *E.coli* zu retransformieren und das intakte pIZ-Plasmid aus *E.coli* wieder zu extrahieren. Über Southern Hybridisierung konnten wir nicht nur die Präsenz der Plasmide nachweisen, sondern auch, dass sie nicht – integriert, in freier Form im Zytoplasma vorlagen. Die Kopienzahl des Vektors pro µg DNA wurde durch die quantitative PCR Methode bestimmt, indem ein chromosomales Gen als Referenz herangezogen wurde. Die Anzahl der Chromosomen und des transformierten pIZ-WOP Expressionsvektors pro µg DNA waren annähernd ident (um die $3x10^6$ Kopien pro µg DNA), was bedeutet, dass die Kopienzahl des Plasmids eins – bis zwei pro Zelle betrug.

Des Weiteren wurden die Plasmide praktisch angewandt um die Aktivität des CRISPR (*C*lustered *R*egularly *I*nterspaced *S*hort *P*alindromic *R*epeats) Systems zu untersuchen. CRISPR ist ein kürzlich entdecktes Immunsystem-Äquivalent vieler Prokaryoten, welches wirkend durch spezifische RNA Fragmente, extrazellulären Eindringlingen durch Nukleinsäuresequenz - Erkennung abwehrt. Ein CRISPR Lokus liegt im Chromosom und besteht aus einer Ansammlung kurzer, unterschiedlicher Spacer-Sequenzen, welche von meist uniformen, Palindromseuqenzen, sogenannten “repeats“ voneinander abgegrenzt werden. Spacer sind DNA Sequenzen, die von Viren oder Plasmiden aus einer früheren Infektion der Zelle oder der Vorgängerzellen stammen. Transkription dieser Loki bringt kurze crRNAs (CRISPR RNA) hervor, welche aus einem Spacer und Teilen der angrenzenden „repeats“ bestehen, in Proteinkomplexe eingebaut werden und komplementäre fremd-DNA oder – RNA (sogenannte „protospacer“) binden und durch Enzymaktivität abbauen. In unserem Labor wurden vor kurzem solche Protospacer, welche dem Spacer 63 des Lokus D im *S. solfataricus* Genoms entsprechen, so verändert, dass CRISPR- vermittelter DNA Abbau (DNA-Interferenz) durch Fehlpaarung in der Nukleinsäuresequenz (GU pairing - Konstrukt D63-7U), und noch effektiver, durch Basen-Komplementarität des Protospacers zum Autoimmunität - 5‘ Endes (handle) der crRNA

(Konsrukt D63-HA), verhindert wurde. Transformationsstudien von *S. solfataricus* M18 mit SSV1 Virus Vektoren, welche diese Protospacer integriert trugen, führten zu stabilen Viren in den Zellen, aber zum Abbau der Protospacer mRNA. Gegenteilig, deuten unsere Ergebnisse auf eine DNA-Inteferenz welche zum Verlust der Plasmiden in den Kulturen führte (wenn auch in unterschiedlichem Maße), wenn dieselben Protospacer in pIZ-Plasmiden transformiert worden waren. Nur geringe Kopienzahlen für pIZ-HA Plasmid ($3.1x10^5$ copies/µg DNA) und weniger als 400 Kopien des pIZ-7U Vektors konnten nachgewiesen werden, während negativ Kontrollen, die keinen Protospacer enthielten, mit Kopienzahlen von $3.18x10^6$ per µg DNA stabil waren. Überraschenderweise, wurde für alle Kulturen ähnliches Wachstum und ähnliche Chromosomen Anzahl (durch qPCR) gemessen, was darauf hindeutete, dass einige Zellen, die mit dem Protospacer Plasmiden transformiert worden waren, ihr defektes *pyrEF* Operon, höchstwahrscheinlich über Homologe Rekombination als Reaktion auf den CRISPR DNA-Angriff, wiederhergestellt hatten. Alles in allem, zeigen unsere Resultate, dass das CRISPR System verschieden auf die Art des extrachromosomalen Elements, welches einen Protospacer trägt, reagieren.

9 Curriculum vitae

Isabelle Anna Zink

Ausbildung:

seit September 2012	Diplomarbeit im Department "Ecogenomics and Systems Biology", Wien
seit 2009	Diplomstudienzweig Mikrobiologie/Genetik, Universität Wien
2006 – 2009	Diplomstudium Biologie, Universität Wien, Abschluss 1. Abschnitt
Juli 2006	Matura
1998 – 2006	Neusprachliches Gymnasium
Fremdsprachen:	Englisch in Sprache und Schrift, Französisch (6 Jahre Gymnasium), Latein (4 Jahre Gymnasium)