# Responses of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV-light

vom Fachbereich Biologie
der Technischen Universität Darmstadt zur
Erlangung des akademischen Grades eines
*Doctor rerum naturalium*
genehmigte

**DISSERTATION**

Vorgelegt von

**Dipl. Biol. Sabrina Fröls**

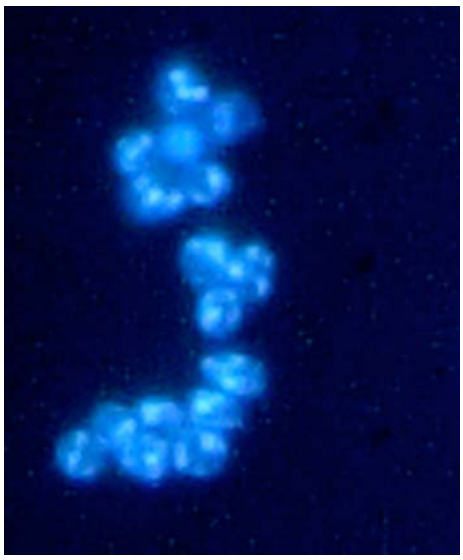
aus Frankfurt am Main

Darmstadt 2008

D17

Tag der Einreichung: 09. April 2008
Tag der mündlichen Prüfung: 17. Juni 2008

1. Berichterstatterin extern : Prof. Dr. Christa Schleper
2. Berichterstatterin: Prof. Dr. Felicitas Pfeifer
3. Berichterstatter: Dr. habil. Arnulf Kletzin



Cellular aggregation of *S. solfataricus*
after UV-irradiation.

(DAPI stained, fluorescence micrograph)

“Nothing in an organism makes sense except
in the light of the functional context.”

(Jan-Hendrik S. Hofmeyr)

# Index

# 1. Summary

In nature UV-light is the most DNA-damaging factor. Photoproducts, that are not removed, result e.g. in the inhibition of replication or can cause lethal mutations. Compared to Eukarya and Bacteria, the DNA-damage response mechanisms in Archaea are not well understood. In particular hyperthermophilic and acidophilic archaea might have to deal with an additional constant challenge to maintain their genomic stability due to their life in harsh environments. This makes them interesting objects to study DNA damage response mechanisms.
In this study the transcriptional and cellular reactions of the hyperthermophilic archaeon *Sulfolobus solfataricus* and its UV-inducible virus (SSV1) to UV-light were investigated by applying a post-genomic approach.

SSV1 showed a co-ordinately regulated transcriptional cycle from 0.5 h to 8.5 h after UV-treatment. By using a high-density DNA-microarray approach, the transcripts could be classified into three categories: immediate early (T-ind), early (T5, T6 and T9) and late transcripts (T3, T1/2, T4/7/8 and Tx). The latter were up-regulated upon the onset of viral genome replication. This tightly regulated transcriptional pattern of SSV1 has not been described before for any archaeal virus and is reminiscent of those of many bacteriophages and some eukaryotic viruses. Six host genes were exclusively regulated in an infected strain upon UV-treatment indicating specific virus-host interactions. Among these were genes encoding topoisomerase VI, which probably plays an essential role in the replication of SSV1. All 34 gene products of SSV1 were tested for protein-protein interactions in a yeast two-hybrid approach. Some of the eight observed interactions suggested new putative protein functions involved in the regulation or involved in the particle assembly of SSV1.

*S. solfataricus* exhibited a complex transcriptional and cellular reaction to UV-light. The UV-dependent transcriptional reactions were investigated by a genome-wide DNA-microarray analysis that extended over 10 time points of two strains. 55 UV-dependently regulated genes that clustered into three major groups were identified. These genes indicated an immediate arrest of replication (*cdc6-2, cdc6-1*) and a stop in the cell cycle (e.g. *soj, ssh7*), during the UV-dependent reaction from 1.5 h to 5 h after UV-treatment. In addition potential transcription factors (e.g. *tfb-3*) were identified, which might be

involved in secondary UV-dependent reactions. The induction of an operon involved in homologous recombination (*rad50*/*mre11*) indicated the formation of DNA double-strand breaks (DSB). Consistent with this, DNA DSB were observed by pulse-field gel electrophoresis between 2 h and 8 h after UV-treatment, probably as a result of replication stops due to unrepaired photoproducts. Another, rather unexpected finding was the induction of an operon encoding a potential type II/type IV pili biogenesis system (*sso0117* through *sso0121*) for secretion or pili formation. In support of this, a statistical microscopic analysis demonstrated that at least 50-70% of the cells formed aggregates, particularly between 6 h and 8 h after UV-exposure. In addition the study of a deletion mutant verified that the pili are encoded by the potential pili biogenesis operon and that they are essential for mediating the UV-dependent cellular aggregation. Aggregate formation was stimulated by chemically induced DSB in DNA, but not by other environmental stressors, indicating that this reaction is UV-specific. Furthermore, it was shown that UV-light strongly stimulated the conjugation activity of *S. solfataricus* (with a frequency of up to $10^{-2}$), whereas no conjugative activity was observed without UV-irradiation.

The data of this thesis open new opportunities towards an understanding of the complex mechanisms involved in DNA-repair after UV-damage in Archaea, and provide supporting evidence to a link between recombinational repair *via* cellular aggregation and subsequent conjugation as major response of *S. solfataricus* to UV-light damage.

## List of publications

The following publications and manuscripts are included in this thesis. The manuscripts were reformatted to generate a uniform style, but the original content of the text and the figures of the publications have been maintained:

**Frölś, S.**, Gordon, P.M., Panlilio, Schleper, C. and Sensen, C.W. (2007a) Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. Virology, 365:48-59.

**Frölś, S.**, Gordon, P.M., Panlilio, M.A., Duggin, I.G., Bell, S.D., Sensen, C.W. and Schleper, C. (2007b) Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. J Bacteriol, 189: 8708-8718.

**Frölś, S.**, Zolghadr, B., Wagner, M., Ajon, M., Folea, M., Egbert J. Boekema, E.J., Arnold J.M. Driessen, A.J.M., Schleper C. and Albers, S.V. UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. Submitted April 2008

S. V. Rajagopala, S.V. **Frölś, S.**, Schleper, C. and Uetz, P. Protein-protein interactions of the *Sulfolobus shibatae* virus 1 (SSV1). Manuscript in preparation

Manuscript and patent not included in this thesis:

C. Schleper, C., Jonuscheit, M., Eck, J., Niehaus, F., Albers, S.V., **Frölś, S.** Patent: ARCHAEON EXPRESSION SYSTEM International publication number: WO2004/106527, since 2004

Teichmann, D., **Frölś, S**., and Schleper, C. *In vitro* and *in vivo* studies of UV-regulated promoters from the hyperthermophilicarchaeon *Sulfolobus solfataricus*. Manuscript in preparation

# 2. General introduction

The ability to react to changing environmental conditions is one of the central characteristics of living systems. A quick and efficient response and adaptation to non-equilibrium conditions, i.e. stress, is necessary for the survival of all organisms, yet the mechanisms that underlie these reactions are quite variable both in the microbiotic as well as macrobiotic world. A full comprehension of the molecular mechanisms of adaptations requires both, the characterization of single molecules as well as the characterization of their complex cellular networks (Kitano, 2002).
Whereas the detailed study of single molecules is indispensable to define the biochemical and structural properties of basic components of the system, their role for the complex cell can only be understood in the context of the cellular networks. Such networks are found on different hierarchic levels in biology, ranging from the complexity of metabolic pathways, to the protein-protein interactions in single cells and to the networks between multicellular structures (Kitano, 2002). They are studied by post-genomic and systems biology approaches.

This work was directed at the global study of the transcriptional and cellular reaction of the hyperthermophilic archaeal model organism *Sulfolobus solfataricus* to the stressor UV-light. All organisms exposed to sunlight have to deal with a constant damage of their DNA by the natural UV-irradiation, but they developed different types of reactions to secure their survival. In addition, a detailed analysis of the reactions of a UV-inducible *S. solfataricus* virus were performed and integrated with earlier studies to gain new insights into the virus regulation. The transcriptional dynamics and the global cellular reactions to UV-exposure of *S. solfataricus* and SSV1 were studied by applying a post-genomic biology approach.

## 2.1 Archaea - the third domain of life

The studies of Carl Woese which were based on the analysis of ribosomal RNA catalogues firstly discovered that the Archaea represent a separate, third domain of life besides Eukarya and Bacteria, (Woese & Fox, 1977; Woese *et al.*, 1990). This separation was further supported by specific features of archaea, like their unique cytoplasmic

membranes with phytanyl ether lipids instead of fatty acid ester lipids, the special cell wall composition and their restriction to extreme environments (Woese *et al.*, 1978). The first characterized species of the archaea were isolated from extreme habitats with high salinity, high temperature and acidity, or strict anoxia (Zillig *et al.*, 1990; Stetter, 2006). Until today, the archaea are famous for their specialization to extreme environments like hot springs of volcanic origin or hydrothermal vents (thermophiles), salt lakes and sea water evaporation ponds (halophiles) or acidic mine drains (acidophiles). The domain Archaea was later phylogenetically divided into the two kingdoms of the Crenarchaeota and Euryarchaeota based on 16S rDNA analysis, (Woese *et al.*, 1990). This original characterisation is still valid for most of the cultivated strains. The Euryarchaeota represent a physiological heterogenic and globally distributed group with the extremely halophilic, the strictly methanogenic and some (hyper)thermophilic organism. Crenarchaeota were originally thought to be the oldest evolutionary group and only representing hyperthermophilic archaea. Most have sulfur-dependent and autotrophic metabolisms, perhaps reminiscient to organisms that might have existed on the early earth (Fig. 2.1), (Zillig *et al.*, 1990; Stetter, 2006).

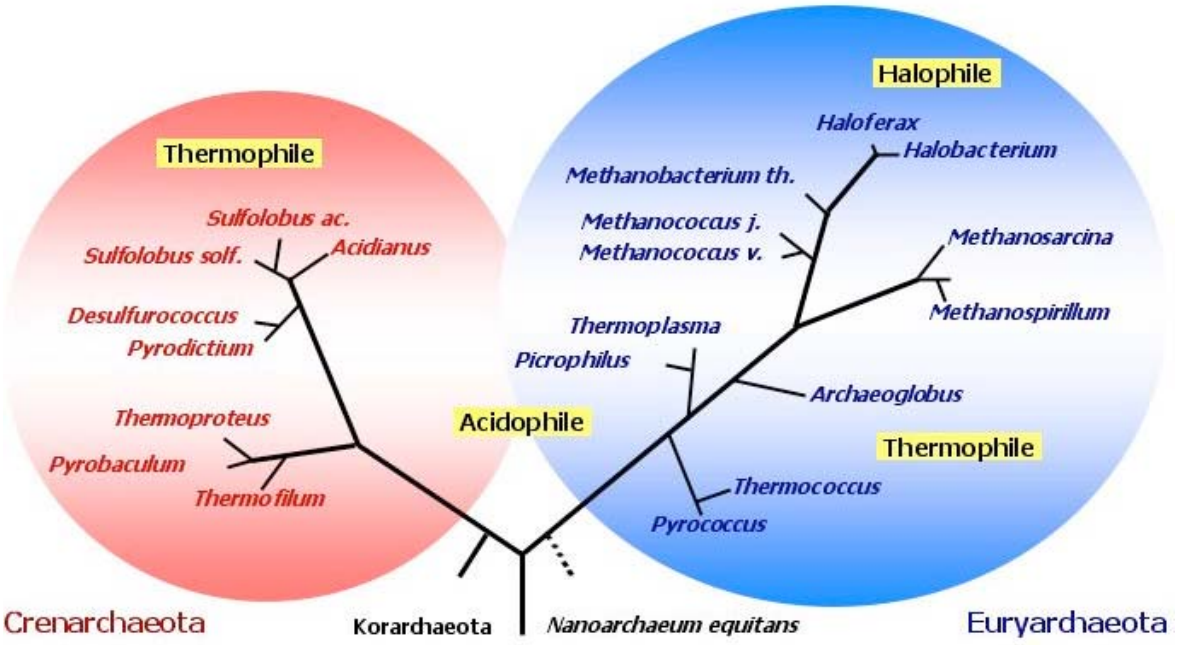



**Figure 2.1:** The domain of Archaea with its kingdoms. Only branches of cultivated species are shown. The schematic tree is based on 16S rDNA phylogenetic analysis. Korarchaeota (of which only enrichments exist) might form a separate kingdom. The placement of the taxon *Nanoarchaeum equitans* is currently debated (modified from C. Schleper)

However, subsequent molecular environmental studies have shown that Archaea, and particularly Crenarchaeota are not exclusively restricted to extreme environments. They were found in various moderate common place environments like soil, freshwater or the ocean. Although this group of moderate archaea seems to be big and widespread, only two cultivated isolates and a few enrichments have so far been described (Delong & Pace, 2001; Könnecke *et al.*, 2005; Leininger *et al.*, 2006; Wuchter *et al.,* 2006; de la Torre *et al.*, 2008). It is currently debated, if moderate archaea even form a separate kingdom within the Archaea (Brochier-Armanet *et al.*, 2008). Similarly, two further hyperthermophilic groups have been proposed to from separate kingdoms of so-called Nanoarchaeaota and Korarchaeota (Auchtung *et al.*, 2006; Di Giulio, 2007; and see Fig 2.1)

## 2.2 Archaea as models for the central information processing in Eukarya

Molecular studies of Archaea are particularly interesting, because these microorganisms exhibit many fundamental similarities to Eukarya. The overall cellular architecture and physiology of Archaea is, however, more similar to that of Bacteria. Archaea are prokaryotic organisms missing the cellular nucleous and cellular organelles. The single cells are small, of approximately 1-2 μm in diameter, diverse in morphology (from coccoid to rods) and contain usually a single small circular chromosome like most Bacteria.

Even though the genomes of the Archaea are hundred to thousand-fold smaller, compared to those of Eukarya, their basal central information processing machineries, namely replication, transcription and translation, exhibit astonishing similarities to Eukarya (see Tab. 2.1). In direct comparison, the archaeal central information processes are highly simplified and represent a minimal set of the complex eukaryotic processes. For instance, the basal transcription apparatus in human cells is composed of > 42 general transcription factors (Nikolov & Burley, 1997), whereas the archaeal apparatus requires only 3 factors (Geiduschek & Ouhammouch, 2005; Thomm, 2006). Therefore, the identification of these factors and the study of their interactions in Archaea are of special interest in order to investigate which solutions of minimal sets are found and to create simplified models for the complex eukaryotic systems.

The first similarities between the transcription apparatus of the Archaea and Eukarya were discovered through the characterisation of an archaeal DNA-dependent RNA-polymerase (Zillig *et al.*, 1979). Additional studies revealed that 15 components of the single archaeal RNA polymerase are homologous to the shared subunits of the three eukaryotic RNA polymerases (Huet *et al.*, 1983). More recently, it could be demonstrated that also the three-dimensional structures of these polymerases are similar to each other (Hirata *et al.*, 2008). The archaeal promoter structure consists of two central regions, an AT-rich TATA box sequence at around 25 bp upstream of the transcriptional start site, flanked upstream by the TFB recognition element (BRE), both representing homologues of the core promoter elements of Eukarya (Palmer & Daniels, 1995; Qureshi & Jackson, 1998). The TATA box binding protein (TBP) is also highly conserved between both domains of life and TBPs from human and yeast can functionally replace those of mesophilic Archaea (Wettach *et al.*, 1995; Bell & Jackson, 1998). The archaeal transcription factor TFB is essential to direct the initiation of transcription and mediates, together with the TBP-DNA complex, the binding of the RNA–polymerase (Gohl *et al.*, 1995; Hausner *et al.*, 1996; Quershi *et al.*, 1997). At least one TBP and one TFB is found in each archaeal genome, but some species have multiple copies, suggesting that different combinations of TBP-TFB interactions may regulate specific gene sets (Baliga *et al.*, 2000; Shockley *et al.*, 2003).

The DNA replication machinery of Archaea is described as a stripped down or ancestral version of the eukaryotic one (Dionne *et al.*, 2003; Barry & Bell, 2006). The ORC/cdc6 homologues are probably involved in the regulation of the initiation of replication and bind specifically to the origin recognition boxes of the genome (Robinson *et al.*, 2004). The copy numbers of the genes encoding for ORC/cdc6 homologues vary in Archaea but at least one homologue has been found in all genomes (Barry & Bell, 2006). At least seven more factors of the basal replication machinery are shared between Eukarya and Archaea. Interestingly, some Archaea possess multiple origins of replication for their genome like it is the case in Eukarya. The simultaneous replication might allow a faster replication and the multiple origins may be involved in regulating the process itself (Robinson *et al.*, 2004; Lundgren *et al.*, 2004; Robinson *et al.*, 2007).

**Table 2.1:** Comparison of basal machineries involved in central information processing in the three domains of life, Archaea, Eukarya and Bacteria.

| Feature | | Archaea | Eukarya | Bacteria |
|---|---|---|---|---|
| **Replication**[a] | Initiator | Orc1/Cdc6 | ORC | DnaA |
| | Helicase loader | Orc1/Cdc6 | Ccd6 +Cdt1 | DnaC |
| | Helicase | MCM | MCM complex | DnaB |
| | Single-strand binding | RPA like SSB | RPA | SSB |
| | Primase | Primase | Primosome | DnaG |
| | DNA polymerase | DNA Pol B familiy | DNA Pol B familiy | DNA Pol III |
| | Sliding clamp | PCNA | PCNA | ß-Clamp |
| | Clamp loader | RF-C | RF-C | γ-Complex |
| | Ligase | DNA lig 1 | DNA lig 1 | Ligase |
| | Primer removal | FEN/Rad2 | FEN/Rad2 | DNA Pol I |
| **Promoter-structure**[b] | Assembling of the initiation complex | TATA Box<br>-25 to -30 | TATA Box<br>-25 to -30 | Pribnow Box<br>-10 |
| | Determination of orientation and strength | BRE<br>> - 30 | BRE<br>> -30 | -35 Box sequence |
| **Transcription**[b] | Catalytic enzyme | RNA polymerase<br>(10-14 subunits) | RNA polymerase II<br>(12 subunits) | RNA polymerase<br>(4 subunits) |
| | General transcription factor | TBP | TBP | Sigma factor |
| | TBP, DNA, RNAP interaction | TFB | TFIIB | |
| | Stabilisation of the transcription complex | TFE | TFIIE | |
| **Translation** [C] | Ribosomal proteins | 15 SSU and 19 LUS are universal in all domains<br>13 SSU and 20 LSU are shared | | |
| | Translation initiation factors | eIF5B | IF2 | |
| | | a/eIF2 | eIF2 | |
| | | aIF6 | aIF6 | |
| | | aIF2B | eIFB | |

[a] Data complied from Lao-Sirieix *et al.,* 2007

[b] Data complied from Thomm, 2007

[C] Data complied from Londei, 2007

Even the last step of the central information process, the translation of the mRNA into proteins, shows some close similarities to that of Eukarya. Several translation elongation factors are exclusively shared by Archaea and Eukarya (see Tab 2.1; Londei, 2005). Whereas the ribosomal subunits exhibit similar sedimentation coefficients as those in Bacteria (Londei *et al.*, 1986), homologues to nearly half of the archaeal ribosomal proteins are only found in Eukarya (Lecompte *et al.*, 2002).

## 2.3 *Sulfolobus solfataricus* and its virus SSV1

*Sulfolobus solfataricus*, the object of this study, belongs to the hyperthermophilic Crenarchaeota. *S. solfataricus* and other species of the genus *Sulfolobus* were isolated from terrestrial solfataras, which are small steam-heated pools of volcanic origin (Fig. 2.2 A) (Zillig *et al.*, 1980). They are found in geothermally active regions all over the world, like in Yellowstone National Park (USA), Japan, Iceland, New Zealand, El Salvador, Italy and Kamchatka/Russia (Rice *et al.*, 2001).

The order Sulfolobales (Stetter, 1989) is represented by the genera *Sulfolobus*, *Stygiolobus*, *Metallosphaera*, *Acidianus* and *Sulfurisphaera* (Fuchs *et al.*, 1996). Common properties are a coccoid morphology, the acidophily (pH 2-3), hyperthermophily (75-88°C) and a low genomic G+C content (31-45 mol%), (Stetter *et al.*, 1990).

The best characterised species of the genus *Sulfolobus* is *S. solfataricus*. It grows optimally at a temperature of 80°C and a pH of 3 (Brock *et al.*, 1972; Zillig *et al.*, 1980). *S. solfataricus* is an aerobic heterotrophic organism, using different carbohydrates or amino acids as growth substrates (Grogan, 1989), which allows easy cultivation. The flagellated, single cells exhibit an irregular coccoid shape of about 1 to 2 μm in diameter (Fig. 2.2 B) (Zillig *et al.*, 1980). Since their isolation, different *Sulfolobus* species have attracted the attention of researchers due to their unique hyperthermophilic properties but also because of their eukaryote-like features.

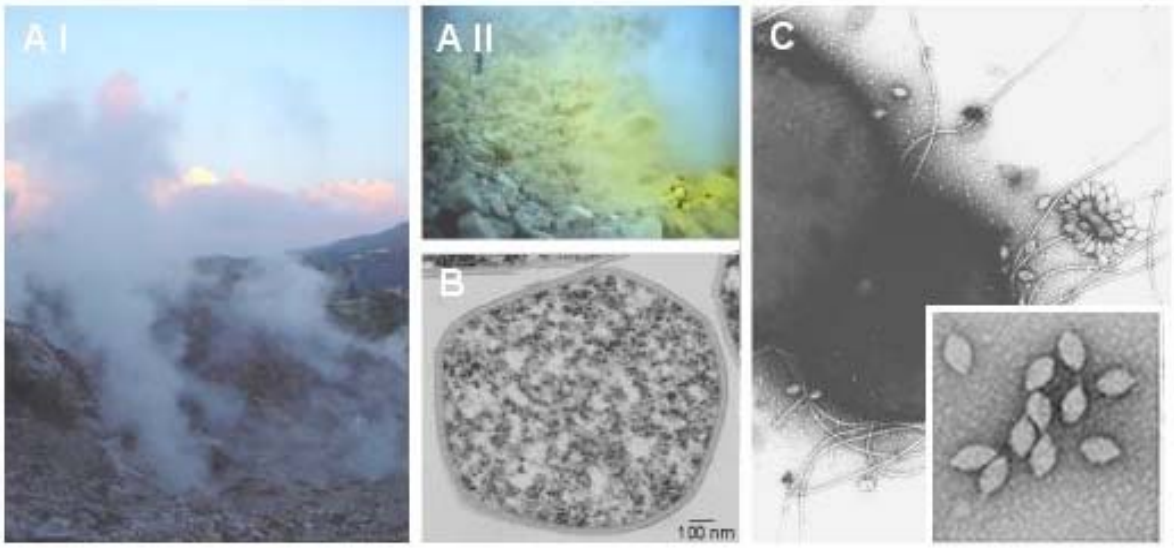



**Fig. 2.2:** (**A) I and II**; Terrestrial solfataras, habitats of the hyperthermophilic Crenarchaeon species *Sulfolobus ssp.* (source W. Zillig, C. Schleper), (**B**) Transmission electron microscopic picture of a single *Sulfolobus solfataricus* cell (source W. Zillig). (**C**) Electron microscopic picture of virions from the *Sulfolobus shibatae* virus 1 (SSV1), a viral species in the familiy Fuselloviridae. The spindle-shaped particles are 60 x 100 nm in size. Samples were negatively stained with 3% uranyl acetate (source C. Schleper).

During the course of isolation of the different *Sulfolobus* species and of other hyperthermophilic Crenarchaeota, many DNA viruses and conjugative plasmids were discovered and characterized (Zillig *et al.,* 1988, 1996). The isolated crenarchaeal viruses are highly diverse as is reflected by their very unusual morphotypes: flexible or rigid rods, filaments, spindle-shapes and spherical virions (Zillig *et al.*, 1996; Prangishvilli & Garrett, 2004, 2005). All of the isolated viruses contain double-stranded (ds) DNA and they are, in contrast to dsDNA viruses of Bacteria, in general non-lytic to their host. These attributes were used to study the basic central information processing and to develop genetic systems for the hyperthermophilic archaea.

One of the first isolated and best-studied crenarchaeal viruses is the *Sulfolobus shibatae* virus 1 (SSV1; Fig. 2.2 C) (Martin *et al.*, 1984), which was early used as a model system to study basic transcription and its regulation in Archaea (Reiter *et al.*, 1987, 1988a, 1988b; Zillig *et al.*, 1992) (see also chapters 3.2 and 4.2). After infection the virus DNA is integrated site-specifically into the host genome (Reiter *et al.*, 1990). UV-light or replication inhibitors, like mitomycin C, were able to induce the virus proliferation (Martin *et al.*, 1984; Reiter *et al.*, 1988c). The promoters are similar to those of the RNA polymerase II in Eukaryotes and represent the first examples for the TATA-box motif in

Archaea (Reiter *et al.*, 1988a, 1988b). The only deviation from the promoter consensus was found for a single transcript that appear early after induction by UV-light. While the induction of central transcripts after UV-light was shown (Reiter *et al.*, 1988a), the full transcriptional cycle of the virus was unknown.

## 2.4 *Sulfolobus solfataricus* - an archaeal model system

Recent progress has made *S. solfataricus* an excellent candidate for comprehensive studies and systems biology approaches. The genome sequences of three species of the genus *Sulfolobus* are available: *S. solfataricus* P2 (She *et al.*, 2001), *S. acidocaldarius* (Chen *et al.*, 2005) and *S. tokodaii* strain 7 (Kawarabayasi *et al.*, 2001), allowing comparative genomic analyses and the development of post-genomic techniques.

Based on the genome of SSV1, an integrative shuttle vector was developed, which, together with the selectable uracil auxotrophic mutant strain *S. solfataricus* PH1-M16, led to the development of a transformation system, applicable for promoter and gene expression studies (Fig 2.3, Martusewitsch *et al.*, 2000; Jonuscheit *et al.*, 2003; Albers *et al.*, 2006).

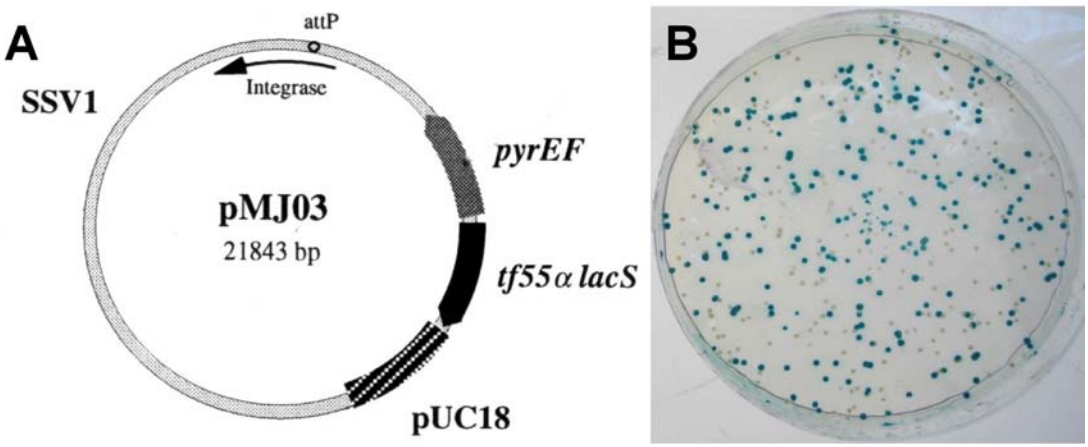



**Figure 2.3:** (**A**) SSV1-based shuttle vector pMJ03; the genome of SSV1 is displayed in light gray; the 5´-FOA selection marker *pyrEF* in gray; the reporter gene *lacS* (ß-galactosidase) in fusion with the promoter of the heat-shock gene *tf55* in black. The arrow indicates the SSV1 ORF D335 encoding a site-specific recombinase. (**B**) Blue-white screening of colonies from *S. solfataricus* colonies on solid media after transfection with pMJ03

In addition, a genetic system allowing targeted gene disruption for the analysis of defined genotypes was developed (Albers & Driessen, 2007). It is based on strain *S. solfataricus* 98/2, an isolate from Yellowstone National Park (Schelert *et al.,* 2004) that differs from *S. solfataricus* P2 by carrying a natural deletion of ca. 58 kb in its genome. The two genetic systems are now widely applied tools for genetic manipulations aimed to verify hypotheses that arise from integrative approaches.

Beside the genetic tools that are available for *Sulfolobus*, it is equally important that some basic knowledge about fundamental cellular processes, such as e.g. the cell cycle and basic transcription, are available, that provide a framework for the interpretation of data.

### 2.4.1 Cell cycle properties

The dynamics of growth and vitality of the cell are regulated by the cell cycle, which requires a complex coordinated interaction of molecules. In addition, the cellular reactions towards DNA-damage, e.g. caused by UV-light, interfere with the cell cycle.

Among the Archaea, the cell cycle of *Sulfolobus* has been studied most intensively and has led to remarkable findings. *Sulfolobus* represents the first non-eukaryotic organism for which more than one replication origin has been identified (Robinson & Bell, 2004). Equally unexpected was the finding that during the exponential growth phase, more then half of the cells carry two copies of the chromosome (Poplawski & Bernander, 1997).

*S. solfataricus* needs approximately 6 to 7 h for one cell division, under optimal growth conditions (Bernander & Polawski, 1997; Poplawski & Bernander, 1997). The cell cycle starts with a very short G1 phase, which is probably needed for the preparation of the replication, indicated by a simultaneous expression of the putative initiators of replication, Cdc6-1 and Cdc6-3 (Bernander, 2000; Robinson *et al.*, 2004). These proteins bind to the origin recognition boxes (ORB) at the origin of replication (ORC) and mediate the loading of the MCM helicases (Jiang *et al.,* 2007). Three ORCs were identified in *S. solfataricus* and remarkably, simultaneous replication initiations at the three ORC were revealed by a microarray based marker frequency analysis (Lundgren *et al.*, 2004). The amplification from 1N to 2N in the S-phase requires 37% of the cell cycle and with termination of replication the cells pass over into the G2 phase. At the same time the two Cdc6 proteins 1 and 3 disappear and a third protein Cdc6-2 is present during the whole G2 phase

(Robinson *et al.*, 2004). The first part of the G2 phase (60% of the cell cycle) is dominated by major cell growth (Bernander & Poplawski, 1997). During the mitosis, the second part, a visible separation of the nucleoids takes place, probably mediated by coiled-coil proteins belonging to the family of structural maintenance of chromosomes proteins (Smc) (Poplawski & Bernander 1997; Elie *et al.,* 1997; Bernander, 1998; Robinson *et al.*, 2007).

### 2.4.2 UV-light induced DNA-damage and repair

In contrast to the comparably well-studied cell cycle, much less is known about the DNA repair mechanisms present in the Archaea (see also chapter 5.2). Although homologous genes encoding enzymes of repair systems have been found in Bacteria and Eukarya, the active repair mechanisms are more or less unknown (Aravind *et al.,* 1999; Kelman & White, 2005).

Sunlight represents the most severe DNA-damaging factor and induces mainly cis-syn-cyclobutane pyrimidine dimers (CPD) on dipyrimidine sequences. Non-repaired CPD blocks the replication and transcription or causes frameshift mutations. The natural UV-radiation reaching the ground is composed of 94% UV-A and 6% UV-B. The daily dose of DNA-damaging UV-B light on a sunny day is measured with an average of 2 $kJ/m^2$ per day (www.temis.nl). Both, the technically used UV-C and the natural UV-B induce CPD in the DNA, with UV-C (254 nm) being about 100-fold more effective than UV-B radiation (Kuluncsics *et al.,* 1999).

Different mechanisms were identified in *S. solfataricus,* which seem to be involved in the reversion, bypass and remove of CPDs in the DNA:

(I) The DNA-photolyase is able to revert CPDs directly and its catalytic activity is enhanced in the presence of UV-light. It was demonstrated for *S. solfataricus* that the removal of the CPDs was 2-fold higher compared to the control incubation in the dark, which suggested a DNA-photolyase activity (Dorazi *et al.*, 2007).

(II) The trans-lesion DNA-polymerase Dpo4 of *S. solfataricus* belongs to the Y-family of polymerases, with lesion bypass properties. It was demonstrated *in vitro* that the Dpo4 is able to insert bases opposite to CPD lesions, suggesting a close relation to the eukaryotic

polη involved in translesion DNA-synthesis (Boudsocq *et al.*, 2001).

(III) An active NER system to remove photoproducts was postulated for *S. solfataricus,* indicated by an efficient repair of CPD in the dark (Salerno *et al.,* 2003). Four homologous components of the eukaryotic nucleotide excision repair system (NER) were identified in the genome of *S. solfataricus*: the nucleases XPF and Fen1/XPG and the helicases XPB and XPD (Kelman & White, 2005). However, the functional activity of the archaeal NER system has not been shown and homologous proteins of the eukaryotic NER system involved in the damage recognition and verification are lacking in the archaeal genomes (Shuck *et al.,* 2008).

## 2.5 Aims of this study

At the beginning of this work in December 2004, no studies using post-genomic approaches had been reported for the kingdom Crenarchaeota. The recent development of DNA-microarrays for *S. solfataricus* enabled us to study global cellular reactions on the transcriptional level. We were interested in analysing the response of *S. solfataricus* to UV-light in order to identify and characterize the factors involved in DNA-repair, which were not well understood in *S. solfataricus* and in Archaea in general. Another aim of the study was to reveal if complex regulons, encoding repair functions, are specifically induced upon UV-damage, comparable to the SOS response in bacteria. The fact that SSV1 is inducible by UV-light suggested the presence of such regulons. Furthermore, the study of UV-response enabled the characterization of the transcriptional program of the virus SSV1 and corresponding host interactions, a system of reduced complexity, which was comparably well characterized in its single components before (Reiter *et al.,* 1987; 1988a; 1988b; 1990). The data should be analysed by applying two strategies. A bottom-up approach was used to incorporate given functional information about genes of interest, e.g. of the virus SSV1. In addition, a more generic top-down approach should be used to identify UV-regulated genes in a genome-wide analysis. The results of both strategies should be used to generate new hypotheses about the responses of *S. solfataricus* to UV-light and to enable the design of experiments for their subsequent detailed verification.

The results are summarized in the following chapters 3 to 6:

The first two chapters address the transcriptional regulation and protein interactions of SSV1 and compare these data with available knowledge. In chapter 3 the transcriptional cycle of SSV1 upon UV-treatment and the virus-host interactions are investigated. Furthermore, some of the SSV1 reactions are complemented by detailed molecular analyses. In the fourth chapter, the protein-protein interactions of the viral proteins are studied in a yeast two-hybrid approach to gain more information about the protein characteristics of SSV1.

The second part of this work includes the characterization of the genome-wide UV-dependent transcriptional reactions of *S. solfataricus*. The aim was to describe the global cellular reaction and to identify genes most prominently involved in it. This study is described in chapter 5. Based on these data and additional biochemical and cellular analyses, new insights into the cellular reaction to UV-light were obtained. The last study described in chapter 6 focuses on one particular aspect of UV-reaction that was discovered in the global analyses. Genetic and biochemical methods were used to investigate a particular UV-inducible operon that mediates cellular aggregation and potentially conjugation in *Sulfolobus*.

Finally, in chapter 7 all data obtained in this work are summarised and discussed with regard to the UV–response of *S. solfataricus* as well as SSV1 and in the context of other relevant post-genomic approaches.

Chapter 3

# Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA-microarrays

Published in:

**Virology**

Volume 365, Issue 1, 15 August 2007, Pages 48-59

Elsevier

# Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA-microarrays

**Sabrina Frölsa** , Paul M.K. Gordon , Mayi Arcellana Panlilio , Christa Schleper and Christoph W. Sensen

## 3.1 Abstract

The spindle shaped *Sulfolobus* virus SSV1 was the first of a series of unusual and uniquely shaped viruses isolated from hyperthermophilic archaea. Using whole-genome microarrays we show here that the circular 15.5 kb DNA genome of SSV1 exhibits a chronological regulation of its transcription upon UV-irradiation, reminiscent to the life cycles of bacteriophages and eukaryotic viruses. The transcriptional cycle starts with a small UV-specific transcript and continues with early transcripts on both its flanks. The late transcripts appear after the onset of viral replication and are extended to their full lengths towards the end of the approximately 8.5 hour cycle. While we detected only small differences in genome-wide analysis of the host *Sulfolobus solfataricus* comparing infected versus uninfected strains, we found a marked difference with respect to the strength and speed of the general UV-response of the host. Models for the regulation of the virus cycle, and putative functions of genes in SSV1 are presented.

## 3.2 Introduction

Newly discovered viruses of the thermoacidophilic archaeal genera *Sulfolobus* and *Acidianus* have recently led to the definition of many novel viral families with unusual morphologies and genomes (Haring *et al.*, 2005a, 2005b; Prangishvili *et al.*, 2006; Zillig *et al.*, 1996). These archaeal viruses, which are not reminiscent of any known bacteriophage or eukaryotic virus, exhibit various types of lysogeny or carrier states that allow them to reside and propagate within their hosts. In contrast, purely lytic viruses have thus far hardly been isolated from extremophilic organisms (Zillig *et al.*, 1996; Prangishvili *et al.*, 2001), perhaps due to the difficulty in sustaining the integrity or infectivity of free virus particles under high temperature and acidity. One of the first and best studied viruses in hyperthermophilic archaea is SSV1, which was originally isolated from *Sulfolobus shibatae* (Martin *et al.*, 1984). It is particularly well known as it served as an early model for the study of basic transcription and regulation in archaea (Reiter *et al.*, 1987, 1988a, 1988b; Zillig *et al.*, 1992). SSV1 infects different strains of *Sulfolobus* (Schleper *et al.*, 1992), that grow optimally around 80°C and pH 3. The lemon-shaped virus particles of 100 nm in length harbour a double-stranded circular DNA genome of 15.5 kb (Palm *et al.*, 1991) that is highly positively supercoiled and covered by a nucleoprotein (Nadal *et al.*, 1986; Reiter *et al.*, 1987).

The capability of SSV1 to infect hosts was long overlooked, because the virus does not produce significant numbers of particles upon infection (Schleper *et al.*, 1992). Instead, the genome is rapidly and site-specifically integrated into a tRNA gene of the host (Reiter *et al.*, 1990) paralleled by a short slow-down of growth (Schleper *et al.*, 1992). Even after infection with an excess of virus particles the host recovers well and often grows even better than before, as if the presence of the viral genome confers some advantage (Schleper *et al.*, 1992) (and own observations). Interestingly, besides the integrated copy, the circular SSV1 genome resides in the cells in a plasmid form with 3 to 4 copies per host chromosome. Upon irradiation of host cells with UV-light, a strong replication of the viral DNA is induced and large amounts of particles (up to 100 per cell) are released into the culture medium, without apparent lysis of the host cells (Martin *et al.*, 1984; Schleper *et al.*, 1992). Ten transcripts of various lengths of SSV1 have been mapped, that start from seven different promoters (Reiter *et al.*, 1987). The identification of their starts and stops led to the first definition of a consensus for promoters and terminators in Archaea, with the promoters being reminiscent of those for RNA polymerase II in eukaryotes (Reiter *et al.*, 1988a, 1988b). With one exception, all transcripts of SSV1 (from T1 through T9)

seem to be produced at low levels in the 'latent' state. A short transcript (T-ind), which lacks the canonical TATA-box typical of archaeal promoters, is expressed solely upon UV-irradiation of the host cell (Reiter *et al.*, 1988a). The genome contains 34 predicted ORFs, but the function of only 4 proteins are known. Besides three structural proteins (Reiter *et al.*, 1987), the integrase, which belongs to the family of site-specific tyrosine recombinases, has been well characterized (Serre *et al.*, 2002; Muskhelishvili *et al.*, 1993). Related recombinases and integration mechanisms as typified through SSV1 have been found widespread in the *Sulfolobus* genomes and in other archaea (She *et al.*, 2001a). Two proteins have been recently crystallized from SSV1, with their sequence and structure suggesting potential roles as transcriptional regulators (Kraft *et al.*, 2004a, 2004b).

SSV-like viruses seem to be widespread and ubiquitous in hot springs, as they have been isolated from various places throughout the world and from different *Sulfolobus* strains (Rice *et al.*, 2001; Stedman *et al.*, 2003; Wiedenheft *et al.*, 2004; Zillig *et al.*, 1994). Comparative genomic and structural studies have elucidated conserved functions in these fuselloviridae (Wiedenheft *et al.*, 2004). The study of archaeal viruses has also led to the identification of features that are conserved in viruses or phages of all three domains of life, the Bacteria, Archaea and Eukarya (Rice *et al.*, 2004) and has inspired theories about the evolution of viruses and their hosts (Forterre *et al.*, 2006).
Beside comparative evolutionary studies, another application of SSV1 is its use as a transformation vehicle (Stedman *et al.*, 1999; Cannio *et al.*, 1998). Our laboratory has recently established a transformation vector, on the basis of the complete viral genome (Jonuscheit *et al.*, 2003), that allows the high level expression of genes and purification of recombinant and tagged proteins (Albers *et al.*, 2006) as well as promoter studies in *S. solfataricus* (Lubelska *et al.*, 2006).

Although the ecology, structure and evolution of SSV1 has been studied for many years and have made SSV1 an important model virus of the Archaea, detailed knowledge about its life cycle and about the function of specific genes was relatively scarce. In order to gain better insights into the biology of this virus, its life cycle, putative gene functions and their effect on the host, we have conducted a genome-wide transcriptional study of SSV1 and the host *S. solfataricus* using a 70-mer oligonucleotide microarray for the analysis. We show that SSV1 exhibits a temporal regulation of its transcription upon UV-induction, and the chronological order of transcription in the SSV1 and *Sulfolobus* genome allows us

to hypothesize about certain aspects of gene function and regulation in this hyperthermophilic virus-host system.

## 3.3 Results

### 3.3.1 UV effects on growth of host and virus

Production of *SSV1* virus particles does not result in lysis of the host, but cells show a marked growth retardation over a time period of ca. 10 hours after UV treatment, while virus particles are produced and expelled into the medium. (Schleper *et al.*, 1992; Martin*et al.*, 1984). Fig. 3.1 displays an example of such a growth retardation for cultures that were used in the microarray experiments of this study. We have used strain PH1, a beta-galactosidase mutant of *S. solfataricus* P1 as well as a lysogen thereof, PH1(SSV1), that has been isolated from a single plated colony after infection of a culture with SSV1. Both strains showed an apparent growth arrest up to 4 h after UV-treatment, which was independent of the virus and probably due to cell damages. By contrast, control cultures of both strains resumed growth starting 0.5 h after mock treatment.

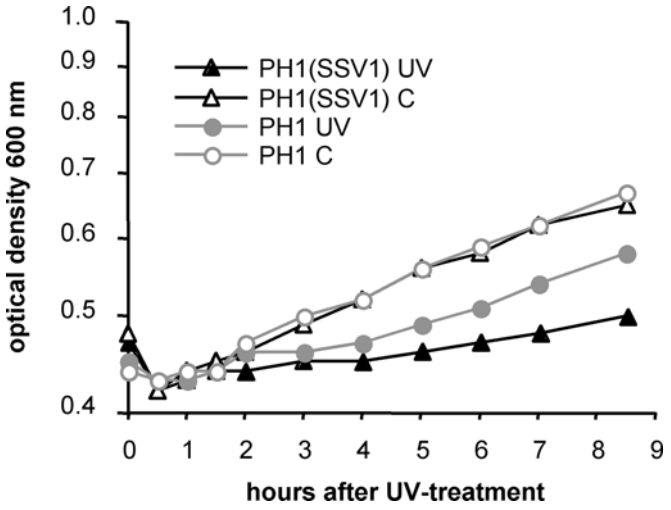



**Figure 3.1:** Growth of *S. solfataricus* culture PH1 (circles) and the SSV1 infected *S. solfataricus* culture PH1(SSV1) (triangles) after UV-treatment. The cultures were split into two halfs, one of which was subjected to UV-treatment and the second half was mock treated (exact same treatments but without exposure to UV light, labelled C). Cultures were immediately re-cultivated, after the UV-treatment at time point 0 h. The time points of sampling for DNA and RNA isolation are indicated by the symbols.

The initial drop in the optical density of the lysogenic cultures and the slow growth of all cultures (including mock treated controls) was probably due to the transient transfer of the cells to room temperature (cold shock), which was required for the UV-treatment. The lysogenic strain PH1(SSV1) was inhibited for another 5 h or even longer, apparently due to virus production. Replication of the viral DNA started approximately 5 to 6 h after UV-treatment as seen in total DNA preparations (not shown) and Southern analyses (Fig. 3.2).

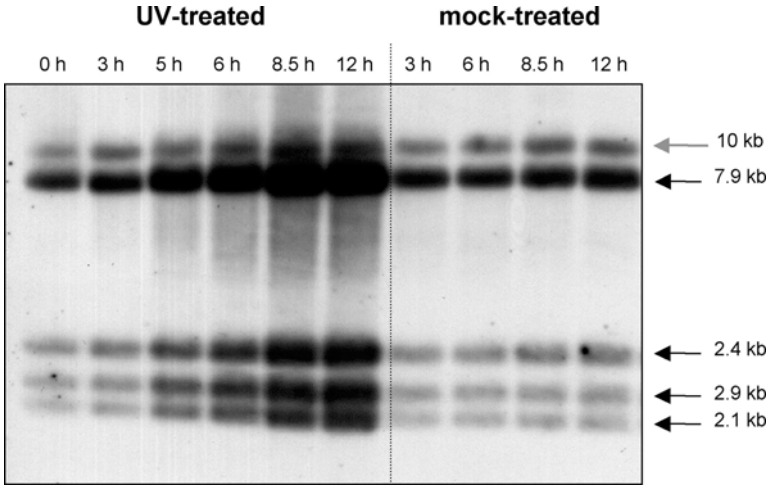



**Figure 3.2:** Southern analysis of *EcoRI*-cut total DNA from strain PH1(SSV1) that was hybridized with a randomly labelled SSV1 DNA. Samples were taken from 0 to 12 h after UV-treatment and mock control. The arrows indicate the *EcoRI* restriction pattern of SSV1 episomal DNA, which appears upon the onset of viral replication, as well as a 10-kb fragment that contains one of the virus flanks in the chromosome. The second flank of the integrated viral DNA is not visible.

Furthermore, we determined the percentage survival of cells on the basis of colony forming units directly before and after UV-treatment. While the mock treated control cultures (which were cooled down during the procedure like UV-treated cultures) showed no effect, i.e. exhibited the same plating efficiency as “pre-mock” cultures, the survival rate of UV-treated PH1 was 40% (compared to pre-UV cultures) indicating that the UV-dose resulted in severe cell damage. The lysogen PH1(SSV1) had a plating efficiency of only 10% after UV-treatment, which must additionally be caused by the stress imposed on the cells when producing virus particles.

### 3.3.2 Analysis of the transcription cycle

The RNA isolated from UV-treated and mock-treated cells was analyzed in Northern hybridizations (Fig. 3.3) to verify successful induction of the viral cycle and to evaluate the quality of the isolated nucleic acids.

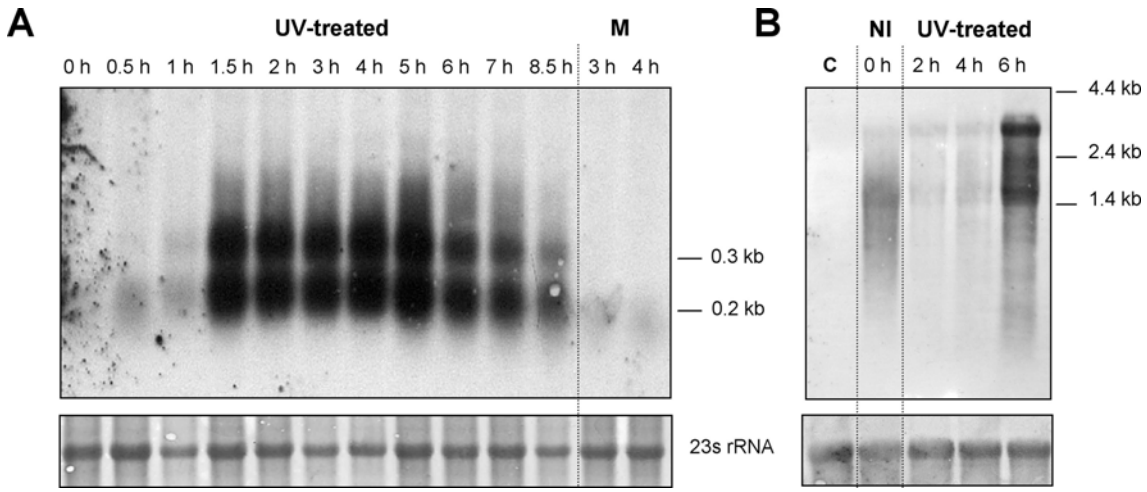



**Figure 3.3:** Northern Analyses with transcript-specific DIG-labelled RNA-probes (**A**) of the UV-inducible transcript T-ind (0.3 kb). PH1(SSV1) total RNA was isolated from 0 to 8.5 h after UV-treatment. For control PH1(SSV1) total RNA from 3 h and 4 h after mock treatment was used (M). (**B**) Northern analysis of the transcript T5 (3 kb), in the non-induced stage (NI), and 2 h, 4 h, 6 h after UV-treatment. Total RNA of the non-infected strain PH1 (C) was used as control. Each lane was loaded with 5 µg of total RNA. Methylenblue-stained 23 rRNA is shown underneath to estimate the relative amounts of blotted RNA.

For microarray hybridizations, the total RNA was reverse transcribed and dually labelled with fluorescent dyes. Whole genome arrays with 70-mer oligonucleotides specific to the open reading frames (ORFs) of the *S. solfataricus* genome, as well as SSV1's ORFs and various other genetic elements, were used in the hybridization experiments (see Materials and Methods). Our data are based on multiple independent experiments and controls (see Materials and Methods). At 0.5 h after UV-treatment no significant change in the mRNA level was observed between experiment and control culture. The first reaction of SSV1 was detected 1 h after UV-treatment with a high increase of the UV-inducible transcript T-ind, that rose 16-fold after 2 h, with no other transcript being induced (Fig. 3.4). The 16-fold increase is relative to the microarray's T-ind probe background noise in the control, as T-ind transcripts do not exist in the control sample. The high level of T-ind was observed until 5 h, after which it started to decrease. These results were also confirmed

by Northern-analyses, using a T-ind specific DIG-labelled mRNA probe (Fig. 3.3 A). While no T-ind transcripts were detectable before UV-treatment in PH1 (SSV1), the maximal transcript levels were found between 1.5 h and 5 h and no T-ind was detected in the mock-treated control. The Northern analysis displayed different amounts of a shorter (0.2 kb) and the full-length (0.3 kb) T-ind transcript that could not be resolved in the microarray study. The relative amount of both transcripts seemed to be similar at 1.5 h, but the long transcript dominated at the maximal expression level (at 5 h).

One hour after the appearance of T-ind, the 5´ located genes of the transcripts T5 and T6 were first detected. Both transcripts flank the T-ind region (though on opposite strands) and both promoters contain an inverted repeat sequence, which is unique in SSV1 (Reiter *et al.*, 1988a) (see Fig. 3.6 for promoter sequences). The full-length transcripts were detected 5 h after UV-treatment (Fig. 3.4). We cannot yet distinguish whether the different lengths of transcripts are caused by mRNA degradation, or by early transcriptional termination/antitermination processes. The sporadic expression of ORF D-355, which encodes the integrase of SSV1 and is located at the 3´-end of T5, suggests that the polycistronic transcript is at least partially degraded. Fig. 3.3 B displays a Northern analysis with a T5-specific DIG-labelled RNA probe which was designed to hybridise to the transcript region of E-178 and F-93 (for ORF numbers see Fig. 3.9) located in the middle of T5. In contrast to T-ind, which is not detectable in samples without UV-treatment, a short and low-copy form of T5 was observed at 0 h (Fig. 3.3 B). Similarly, transcriptional activity from all transcripts except T-ind was seen in the microarray analysis of the non-induced stage, where we compared an infected to a non-infected strain (see NI in Fig. 3.4), confirming the results of earlier studies (Reiter *et al.*, 1988a). However, at 2 h and 4 h after UV-treatment no T5 was detectable (Fig. 3.3 B), suggesting that it is either downregulated and/or quickly degraded upon UV induction, a scenario likely affecting other transcripts of SSV1 as well.

The last of the early transcripts, T9, appeared at 5 h after UV-treatment. This is shortly before the onset of SSV1 replication, which starts between 5 h and 6 h after induction, as seen in Southern analyses (Fig. 3.2) and confirmed by DNA microarray experiments with genomic DNA (data not shown).

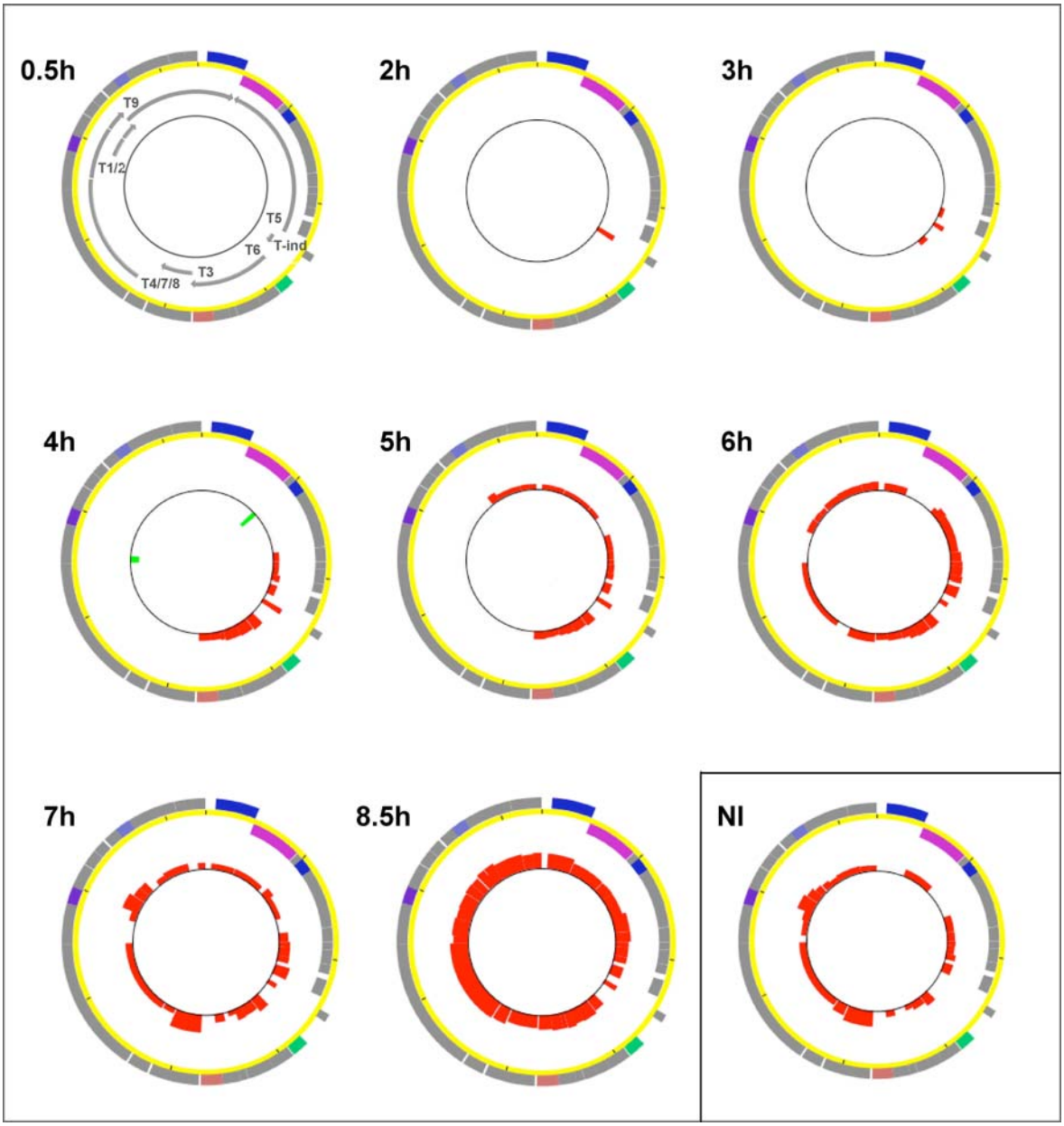



**Figure 3.4:** Relative change in transcript levels of SSV1 genes between 0.5 h and 8 h after UV-treatment (change is relative to transcript levels in mock-treated SSV1 infected culture). The NI (Non-Induced) stage shows the constitutive expression level of SSV1 genes in culture (relative to background signal obtained from non-infected culture). The microarray data are graphically displayed in the genomic context with the program Bluejay (Turinsky *et al.*, 2005), with red bars showing transcript levels elevated (and green reduced) relative to the control sample. The detailed expression ratios for every gene and timepoint are listed in the Supplementary data (Tab. S.3.1). The outermost circle shows the genes of SSV1 with their strand orientation (clockwise in outer ring, counter-clockwise in inner ring). The arrows at time point 0.5 h represent the 10 transcripts as determined earlier for SSV1 (Reiter *et al.*, 1987, 1988a). Each dataset was generated from 2-3 completely independent UV-experiments starting from new cultures and each hybridization was performed as a dye-swap to reduce technical variability.

All other transcripts (T1/2, T3, T4/7/8) were first observed at 6 h after UV-treatment. The levels of the mRNAs of the two coat proteins VP1, VP2 and the nucleoprotein VP3 (Reiter *et al.*, 1987) increased continuously to a maximal induction of 12-fold at 8.5 h after UV-treatment. Slightly higher mRNA levels of VP1 at later stages and in the non-treated lysogenic host were probably due to earlier transcription termination of T8 and T2, resulting in transcripts T7 and T1, respectively (Reiter *et al.*, 1988b).

The monocistronic transcript T3 increased to a nearly 15-fold level at 7 h and it was also present in a comparably high amount at the non-induced stage. Together with the VP1 mRNA, T3 represented the most abundant transcript at 7 h and in the non induced stage.

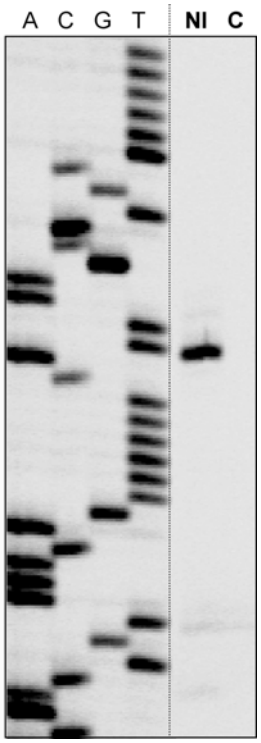



**Figure 3.5:** Primer extension analysis showing transcript start of ORF C-124, that appeared as a late transcript in the microarray analysis. Lanes ACGT: sequence ladder of C-124. Total RNA from a SSV1 lysogen was used from the non induced stage, (NI, lane 1), and from a non infected PH1 strain (C, lane2).

A transcript of ORF C-124 had not been previously described in the literature. It was first observed in the microarrays 7 h after UV-treatment and was also present in the non-induced stage (termed transcript Tx in Figs. 3.6 and 3.9). The simultaneous expression of C-792 and C-124 with the two coat proteins and the nucleoprotein at the late stage in the SSV1 cycle suggests a structural function.

The start of transcript C-124 was mapped by primer extension (Fig. 3.5). Its upstream region revealed some (but limited) similarity to the canonical consensus sequence of archaeal and most SSV1 promoters with a BRE element and a TATA box (Fig. 3.6). While

direct repeats are found in the T5 and T6 promoter regions (underlined in Fig. 3.6, Reiter *et al.*, 1988a), we have not discovered any further putative common cis-regulatory elements in the SSV1 promoters.

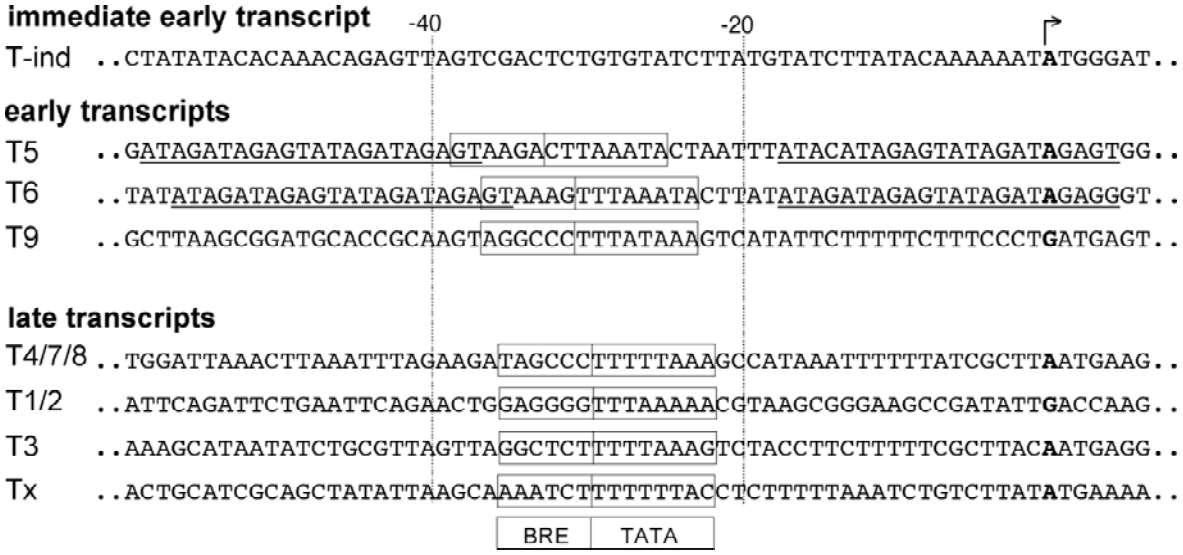



**Figure 3.6:** Comparison of all mapped promoters of SSV1 (Reiter *et al.*, 1988a and this study). Sharp bent arrow and bold letter indicate mapped transcription starts (+ 1), represent the putative BRE and TATA box elements, respectively, and direct repeats in promoter regions of T5 and T6 are underlined.

### 3.3.3 Transcriptional activity in the non-induced state and effects on the host

Considering the tight temporal regulation of transcription upon UV-treatment of the host, it is interesting to note that a low transcriptional activity of SSV1 was also observed when the cells had not been treated with UV-light (except for T-ind). While the early transcripts were incomplete (T5, T6, T9), the late transcripts were fully represented (T2, T3, Tx, T4, Fig. 3.4 NI). Notably, the mRNA of the site-specific recombinase/integrase (ORF D-335) was visible, while the upstream located parts of the respective transcript T5 were not. The absolute strengths of transcripts seen in the non-induced state (Fig. 3.4 NI) is however not comparable to the others states of our analysis, because in the former we used a non-infected strain as a control, while the UV-experiments show relative increases in comparison to mock treated SSV1-carrying controls. The Northern analysis of transcript T5 (Fig. 3.3) shows that the absolute amount of mRNA in the non-induced state is much lower than after UV-treatment.

As noted earlier, growth of the host cells does not seem to be affected by the presence of SSV1 (Grogan *et al.*, 1990; Schleper *et al.*, 1992). Concomitantly, we did not see a

considerable difference on the genome-wide transcriptional level of the host, *S. solfataricus*, without UV-treatment, when we compared an infected (lysogenic) strain with the SSV1-free strain.

### 3.3.4 Differences in the transcriptional host reaction after UV-treatment

While the overall genomic UV-response was comparable with respect to the transcriptional changes (Fröls *et al.*, submitted), a major difference was found with respect to the strength and speed of the host reaction.

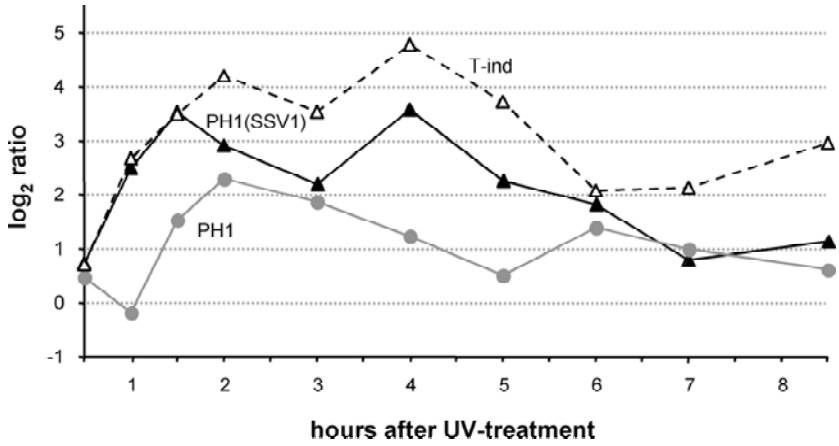



**Figure 3.7:** Average relative mRNA levels ($log_2$) of the 19 most strongly up-regulated host genes upon UV-treatment as well as $log_2$ ratios of T-ind (dashed line). The nature of the up-regulated host genes is exhaustively discussed in Fröls *et al.* (submitted for publication). Data were generated by hybridizing cDNA of a UV-treated culture in competition with cDNA of a mock-treated culture of the same time point. Instantaneously, and was not only faster but also stronger, while the UV-specific reaction of strain PH1 was delayed by ca. 1.5 hours under the same conditions. This finding indicated that an infected strain is apparently more sensitive to UV-light than an uninfected strain.

This is demonstrated in Fig. 3.7, where we compare the average curve of the 19 most strongly up-regulated genes of both strains. For comparison the figure also displays expression levels of the strongest SSV1 transcript T-ind. The SSV1-containing strain PH1(SSV1) reacted instantaneously, and was not only faster but also stronger, while the UV-specific reaction of strain PH1 was delayed by ca. 1.5 h under the same conditions. This finding indicated that an infected strain is apparently more sensitive to UV-light than an uninfected strain.

Genes showing the most pronounced differences in expression levels after UV-treatment between the infected and the non-infected strain are listed in Table 3.1. Most of them exhibited a much stronger down or up regulation in the infected strain, like *cdc6*-1 or *bcp*-

2 (4- to 8-fold higher). But some genes, like the three 7-kDa DNA-binding proteins of *S. solfataricus*, showed a stronger downregulation in the non-infected strain. In order to mimic the stronger reaction of the SSV1-infected strain, we compared its data with those of the non-infected strain treated with a higher UV-dose of 200 J/m$^2$, instead of 75 J/m$^2$ (data not shown). According to this analysis, only six of the 22 genes listed in Table 3.1 showed an exclusive reaction in the infected strain. Their transcriptional pattern is displayed in Fig. 3.8.

**Table 3.1:** Differently regulated genes between strains PH1 and PH1(SSV1) after UV-treatment

| PH1 | PH1 (SSV1) | Gene ID | Operon [a] | Homology [b] | Predicted |
|---|---|---|---|---|---|
| | | SSO0968 | 1 (1/2) | ABE | DNA topoisomerase VI subunit B |
| | | SSO0969 | 1 (2/2) | ABE | DNA topoisomerase VI subunit A |
| | | SSO1460 | | ABE | Penicilin amidase / acylase |
| | | bac18_n0002 | 2 (1/2) | | Hypothetical protein |
| | | bac18_n0001 | 2 (2/2) | S | Conserved hypothetical protein |
| | | SSO0823 | | S | Conserved hypothetical protein |
| | | SSO2121 | | ABE | Peroxiredoxin (*bcp-2*) |
| | | SSO2750 | 3 (1/2) | S | Conserved probable ATPase |
| | | SSO2751 | 3 (2/2) | S | Conserved hypothetical kinase |
| | | SSO0257 | | AE | Cell division control 6/orc1 protein homolog (*cdc6-1*) |
| | | SSO0034 | | AB | ATPase involved in chromosome partitioning (*soj*) |
| | | SSO1210 | | | Hypothetical protein, (coiled coil region) |
| | | SSO0858 | 4 (1/2) | C | Conserved hypothetical protein |
| | | SSO6687 | 4 (2/2) | S | Conserved hypothetical protein |
| | | SSO3207 | | S | Conserved hypothetical kinase |
| | | SSO2200 | | AB | Conserved hypothetical probable ATPase |
| | | SSO0048 | | C | Transcription regulator (Lrs14 homolog) |
| | | SSO3066 | | AB | Arabinose binding protein, ABC Transporter |
| | | SSO9180 | | S | 7 kDa DNA-binding protein |
| | | SSO10610 | | S | 7 kDa DNA-binding protein |
| | | SSO9536 | | S | 7 kDa DNA-binding protein |
| | | SSO0271 | | AE | 26S proteasome regulatory subunit |

Highly induced
Induced
Not regulated
Slightly repressed
Repressed

[a] Number of operon (1-4), and in brackets: position in operon and total no. of genes in operon

[b] Homologues (blastp e-value <$10^{-40}$) in:
S, Sulfolobaceae; C, Crenarchaeaota; A, Archaea; B, Bacteria; E, Eukarya; V, Viruses

Five genes were highly induced and four of these genes were co-transcribed in two operons (Tab. 3.1). The first operon encodes the DNA-topoisomerase VI which might play a role in the viral replication and/or topology of SSV1 DNA (see Discussion). The second operon encodes two proteins of unknown function. SSO1210 represented the only notably downregulated gene. It is noteworthy that the upregulated chromosomal genes reacted during the transcriptional induction of the early genes of SSV1, indicating a potential co-regulation.

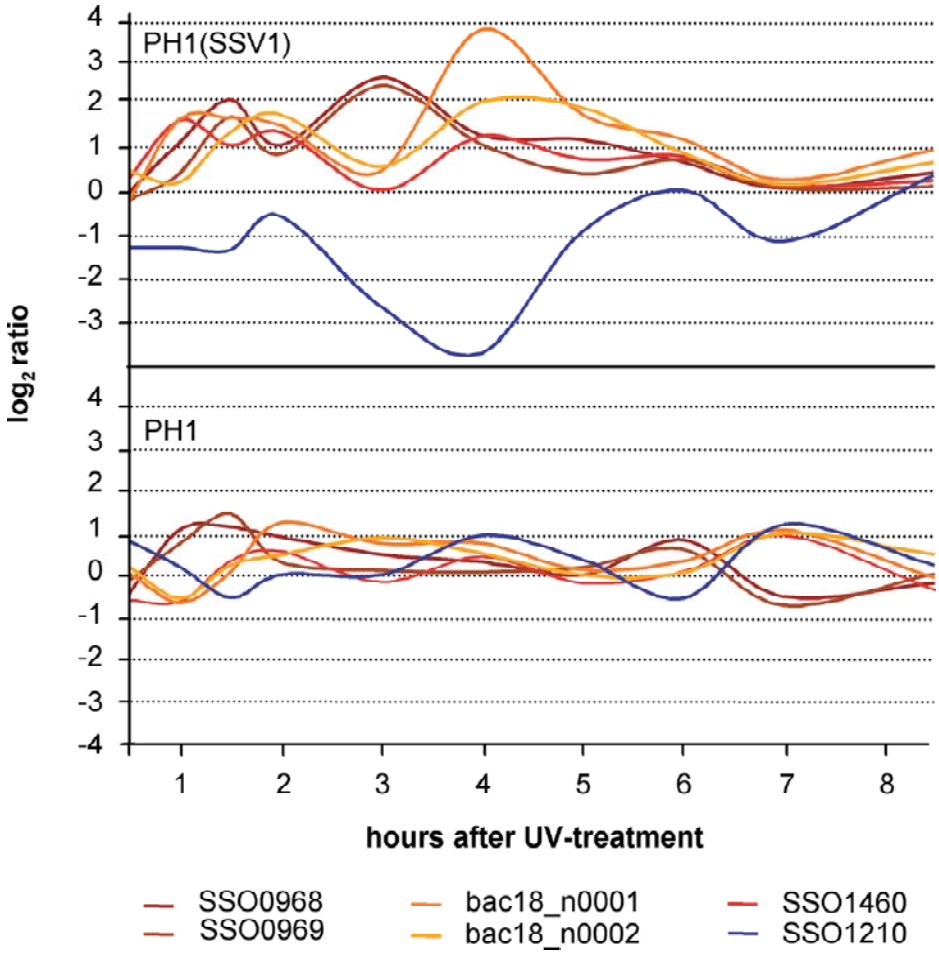



**Figure 3.8:** Host genes specifically reacting to the presence of SSV1. The graphs display expression patterns of the genes with the most distinguished reaction between the two strains PH1(SSV1) and PH1 after UV-treatment. Detailed $log_2$ ratios and statistical evaluations are listed under Supplementary data Table S.3.2.

## 3.4 Discussion

### 3.4.1 SSV1 exhibits a chronological transcription cycle

Initial transcription analyses of SSV1 had revealed some regulatory elements, including the unusual promoter of the UV-inducible transcript T-ind and inverted repeat structures in the promoter elements of what were identified as potentially early transcripts, i.e. T5 and T6 (Reiter *et al.*, 1987, 1988a). Our study shows a complete and detailed scenario of a viral cycle demonstrating that SSV1 exhibits a tight chronological transcriptional regulation, reminiscent to that of many bacterial and eukaryal viruses. This cycle is very different from the *Sulfolobus* virus SIRV, for which a rather simple and barely chronological pattern of transcription has been described after infection (Kessler *et al.*, 2004). Starting with the UV-inducible very early transcript (T-ind), whose promoter deviates from the canonical archaeal consensus in lacking a TATA box (Reiter *et al.*, 1988a), the polycistronic transcripts of SSV1 can be further categorized by their time of appearance into early (T5, T6, T9) and late (T1/2, T3, Tx and T4/7/8) (see Fig. 3.9). These results correlate well with the known or putative functions of the corresponding genes, as far as these have been revealed. For example, four genes in transcript T5 which are conserved among other SSV viruses show features typical of regulatory proteins or transcription factors, i.e. D-63 with a leucine-zipper motif (Kraft *et al.*, 2004a) F-93, a putative DNA-binding protein (Kraft *et al.*, 2004b) and E-51, a member of the CopG family with a Helix-Turn-Helix motif. Furthermore, the conserved gene for the integrase is encoded on the early transcript T5 (D-335). However, many other genes encoded in the early transcript T5 are not conserved among the SSV1 viruses, indicating specific regulatory functions for the UV-inducible virus SSV1. The products of the delayed early transcript T9 that appears 5 hours after UV-treatment, shortly before the onset of replication, might have central functions in this process. Six of the seven ORFs encoded by T9 are highly conserved among all known SSV types (Wiedenheft *et al.*, 2004) and two of them are also homologous to ORFs of the satellite particle pSSVx (Arnold *et al.*, 1999). Furthermore, ORF B251 shows similarity to a DnaA-like protein (Koonin *et al.*, 1992). Together these observations indicate that products of T9 could be involved in replication.

As expected, the late transcripts T2, T7/8 cover the three structural proteins of SSV1 VP1-3. Other genes in these transcripts might encode factors involved in the assembly of the virion.

Interestingly, B-115 (of transcript T7/8), that shows similarity to an ArsR-like putative transcriptional repressor appears as the very last gene that is significantly upregulated at the end of the transcription cycle and also in the non-induced stage. It could be involved in downregulation of the early SSV1 genes towards the end of the viral cycle.

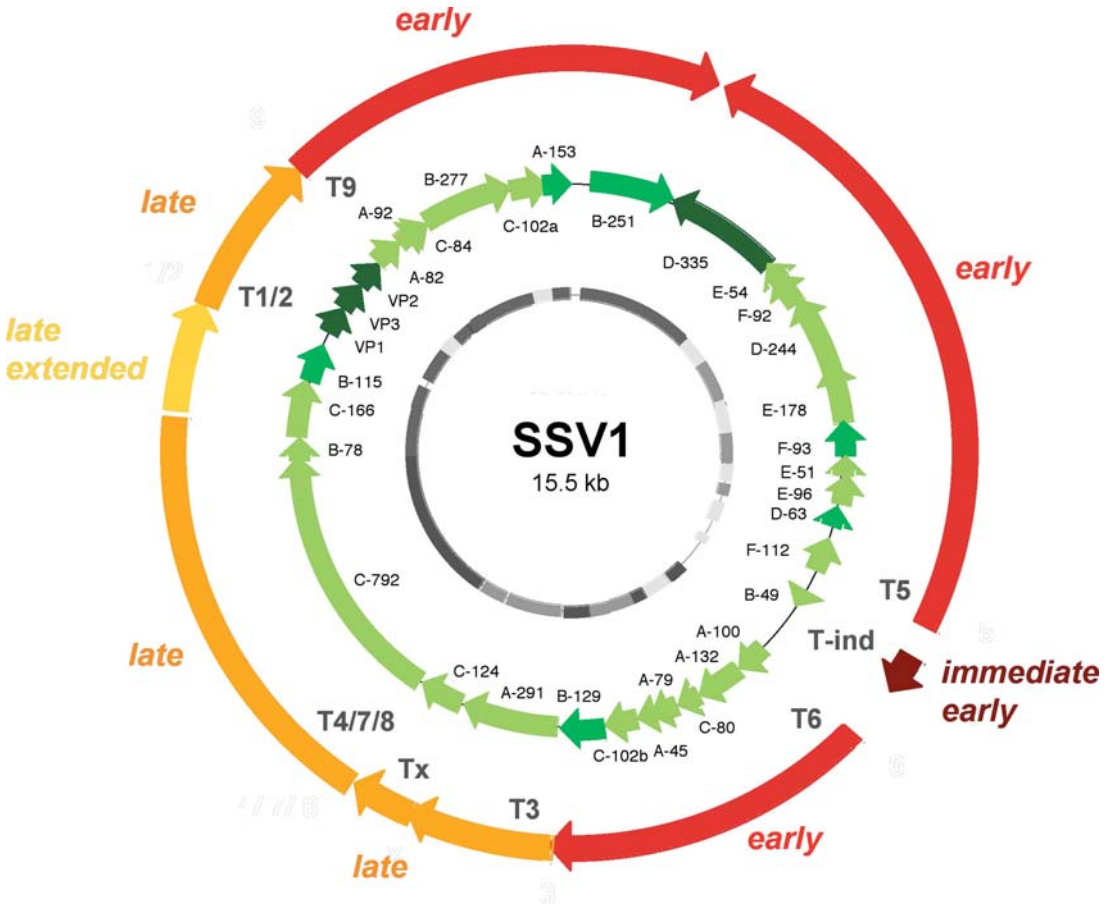



**Figure 3.9:** Summary of microarray and transcript mapping analyses. Similar to those of other viruses and phages of bacteria and eucaryotes, the transcripts of SSV1 can be categorized according to their time of appearance and (putative) functional roles into immediate early, early, late (for further discussion see text). Genes are labelled in dark green (known function), green (predicted function) or light green (unknown function). The inner circle shows the conservation of the genes among the four sequenced SSV types (Wiedenheft *et. al.*, 2004); dark grey: genes conserved in all four viruses; grey: genes conserved in three SSV types; light grey: genes conserved only between two SSV types or unique for SSV1.

### 3.4.2 Regulation of SSV1

The following speculative scenario for the regulation of the SSV1 cycle is compatible with the microarray data.

After UV-treatment the host cells stop growth and express a specific set of approx 40 genes, including transcription factors and a potential replication inhibitor, but no genes that are thought to be involved in direct DNA-repair (Fröls *et al.*, submitted for puplication). A UV-specific transcriptional regulator or factor allows transcription from the T-ind promoter on SSV1. We have found possible candidates in the genome that are induced immediately after UV-treatment (not shown). This putative UV-factor can also be activated upon treatment with mitomycin C or in stationary phase, as these conditions also lead to SSV1 induction (Martin *et al.*, 1984; Zillig, *et al.*, 1998). T-ind transcripts accumulate over 2 hours, including the longer T-ind which may represent or may encode an activator for transcripts T5 and T6. Two inverted repeats that flank the TATA box in these promoters may represent a binding site for a repressor, or alternatively for an activating factor such as the T-ind product. A similar mechanism has been described for the activator E1A of adenovirus, of which a shorter and a longer transcript are produced by a splicing mechanism. In E1A, only the longer transcript contains the CR3 domain, which acts as an activation domain and is essential for the strong activation of the early viral promoters by interaction with the cellular transcription factors (Flint *et al.*, 1997; Berk *et al.*, 2005).

In eukaryotic viruses, many proteins that specifically activate transcription can also stimulate origins of DNA replication (DePamphilis *et al.*, 1993; Heintz *et al.*, 1992). Preliminary results suggest that the origin of replication of SSV1 lies in the vicinity of the T-ind promoter (I. Dugin, S. Bell and S. Fröls, unpublished). Since the T-ind protein (B-49) was found to form a dimer in a yeast-2-hybrid screen (S. Fröls, C. Schleper and P.Uetz, unpublished), and since it is produced between 1.5 and 5 h after UV-treatment, with replication starting around 5 to 6 h after UV-treatment, the T-ind product could help in stimulating replication by sequence-specific binding to DNA. However, it is probably not essential to natural propagation of the virus, since it is not present in the closely related, non-UV-inducible viruses SSV2 and SSVK1 and SSVRH.

### 3.4.3 Reaction of the host to virus induction

While a SSV1-infected host is not visibly impaired in growth compared to a non-infected strain, we observed a difference in growth and on the transcriptional level upon UV-treatment. With respect to transcription, the infected strain reacted faster and

considerably more strongly to the treatment as if the presence of virus or some viral proteins changed the sensitivity of the strain to DNA-damage (Fig. 3.7).

Among the few genes that showed a differential response were those encoding the two subunits of topoisomerase VI. This enzyme could play an essential role in replication of SSV1, because it is the only topoisomerase that can relax positive superturns produced during DNA-replication. It might yet exhibit another function for the virus: The strong positive supercoil of SSV1 DNA in viral particles that was originally thought to be produced by the activity of reverse gyrase (Nadal *et al.*, 1986), could as well be formed through the action of other proteins. SSV1 DNA could be wrapped by a DNA-binding protein into a positive sense and a topoisomerase (like topoVI) could subsequently relax the compensating negative superturns. The involvement of topoVI in the SSV1 cycle could be tested by the use of a specific inhibitor for this protein (Gadelle *et al.*, 2005).

In conclusion, the chronological transcription of SSV1 after UV-treatment is reminiscent to that of well-known bacteriophages or viruses. However the basis for its regulation still needs to be resolved and it is not linked to an SOS-like response as known for bacteria. Similarly, most of the gene functions of SSV1 as well as its replication mechanism are not yet elucidated. Our microarray analysis will help to shed light on the biology of SSV1 and related viruses of archaea.

## 3.5 Materials and Methods

*Growth of Sulfolobus strains*

Strains *S. solfataricus* PH1 (Schleper *et al.*, 1994) and *S. solfataricus* infected with SSV1 (a.k.a. PH1(SSV1)) (Martusewitsch *et al.*, 2000) were grown at 78°C and pH 3 in Brock´s medium (Grogan *et al.*, 1989), with 0.1% (w/v) tryptone and 0.2% (w/v) D-arabinose. The optical density of liquid cultures was monitored at 600 nm. For surviving rate and UV-dose determination, solid media was prepared by adding gelrite to a final concentration of 0.6% and $Mg^{2+}$ and $Ca^{2+}$ to 0.3 and 0.1 M, respectively. Plates were incubated for five days at 78°C.

*UV-treatment*

Freshly inoculated 400 ml cultures were grown at 78°C until log phase was reached ($OD_{600\ nm}$ 0.3 - 0.5). Culture flasks were transferred to a dark room under red light and were divided into two halves. For UV-treatment, aliquots of 50 ml were transferred to a plastic container (20 cm x 10 cm x 4 cm) and irradiated with UV-light at room temperature for 45 s at 254 nm (W20, Min UVIS; DEGESA), while shaking the culture carefully. The used UV-energy was approximately 75 $J/m^2$. It was determined by comparing the plating efficiency (cfu/ml) using a UV-light with known strength and determining survival rates (Fröls *et al.*, submitted for puplication). The second half of the culture was equally cooled and treated, but the UV-light was shut off (mock treatment). For the microarray experiments cDNA from UV-treated cells was competitively hybridized to cDNA of mock-treated cells. Only for NI in Figure 3.4, we used cDNA of a lysogen versus cDNA of a non-infected strain.

*DNA preparation and analysis*

Extrachromosomal and genomic DNA was prepared from *Sulfolobus* using standard procedures (Martusewitsch *et al.*, 2000). DNA was cut with *EcoRI* and used for Southern blotting by standard techniques. Southern blots were probed with randomly digoxigenin-labelled SSV1 DNA (Roche Biochemicals).

*RNA preparation and analysis*

Total RNA was extracted using standard procedure (Chirgwin *et al.*, 1979). RNA quality was determined by agarose gel electrophoresis and by determination of the ratio of absorption at 260 nm and 280 nm. Only RNA samples with a ratio between 2.1 and 1.9 were used for further experiments. For Northern analysis, 5 µg of RNA were separated on denaturing, formaldehyde-containing 1.2% (w/v) agarose gels, followed by transfer to nylon membranes.

Strand-specific T-ind (B-49) and T5 (flanking ORFs F-93 to E-178) RNA digoxigenin-labelled probes were synthesized with the T3/T7 in vitro transcription system (Fermentas) using the inserted genomic regions of SSV1 in pBluescript and pCR4-TOPO as linearized template.

*Labelling of cDNA*

5-(3-aminoallyl)-dUTP (Ambion) labelled cDNA was generated by using the RevertAidTM H Minus First Strand cDNA Synthesise Kit (Fermentas) according to the manufacturer's directions, with the following modifications: 10 µg total RNA was added to 16 µl deionized nuclease-free water and 400 ng random hexamer primer. The mixture was denaturated at 70°C over 10 min and chilled on ice. 5 x reaction buffer, 4 µl (20 mM) dNTP/aa-UTP labelling mix 3:2 (5 mM dATP / dGTP / dCTP, 3 mM dTTP, 2 mM aa-dUTP), 20 units ribonuclease inhibitor and 200 units RevertAid™ H Minus M-MuLV reverse transcriptase were added. The reaction mixture was incubated at 25°C for 10 min then shifted to 42°C for 1.5 h. The RNA template was removed by adding 10 µl 1M NaOH and incubating at 70°C for 10 min before adding 10 µl 1M HCL. The cDNA was precipitated and the pellet was dried at room temperature.

*Microarray design and fabrication*

The microarray design targeted the open reading frames (ORFs) of the Sulfolobus solfataricus P2 genome (She *et al.*, 2001b) and various endogenous and exogenous genetic elements of Sulfolobus species. The genomic elements include 3057 "large" ORFs of 300 bases or more, including 305 transposon-related genes and 395 genes of undecided function, 46 tRNAs and 4 rRNAs. 616 "small" ORFs of less than 300 bases were included, consisting of 107 transposon-related genes, 178 hand-annotated or conserved hypothetical genes, and 176 genes with no functional evidence, but whose presence is

predicted by both Glimmer (Delcher *et al.*, 1999) and GeneMark (Besemer *et al.*, 2005. After filtering completely redundant ORFs (269 in transposons-related and 285 other ORFs), and including unique segments of repeated sequences, 3352 uniquely identifiable genomic elements were left to be targeted in the microarray. Extrachromosomal targets included ORFs from SSV1 (34), SSV2 (32), pSSVx (8), pRN1 (8) and pRN2 (6). 16 large intergenic regions of the *S. solfataricus* P2 genome were targeted as potential negative controls.

The 3456 70-mer oligonucleotides for the microarray were designed using the Osprey software (Gordon *et al.*, 2004) with an optimal theoretical melting temperature of 78°C under 0.1 M NaCl, and ordered from Qiagen Inc. with a 5' C6 amino linker. The full list of probes is available at http://osprey.ucalgary.ca/sulfolobus. The oligonucleotides were spotted using a VersArray Chip Printer (BioRad) onto Corning UltraGAPSTM Amino-Silane 25x75mm Coated Slides. The quality of the microarrays was determined by using 5-(3-aminoallyl)-dUTP labelled cDNA generated from genomic DNA.

*cDNA labelling and microarray hybridisation*

The coupling of the cyanine-3 or cyanine-5 (Cy-3/Cy-5) fluorescent molecules was performed by using a slightly modified protocol for the FairPlay Microarray labelling kit (Stratagene catalog #252002) with 10 µl of 2x coupling buffer and 5 µl of dye. The dye-coupled cDNA was purified using a four EtOH wash processes, and then eluted using a 3-cycle fibre matrix recovery in a microspin cup. The final volume was reduced to between 2.5 and 3.0 µl by vacuum. A hybridization solution of 90 µl DIG Easy Hyb, 5 µl yeast tRNA and 5 µl fish sperm was incubated at 65°C for 2 minutes and allowed to cool to room temperature, then it was added in sufficient quantity to the labelled cDNA to total 25 µl and incubated the same way again. The cooled solution was pipetted onto the slide and allowed to incubate at 37°C for 18 hours under controlled humidity, then washed three times at room temperature: 2 x SSC & 0.2% SDS, 0.2 x SSC, and 0.1 x SSC.

Dye-swapping is used to reduce possible data bias due to different inherent fluorescence levels of the Cy-3 and Cy-5 dyes. In the dye swap, one microarray hybridization was performed with the control cDNA sample labelled with Cy-3, and the experiment cDNA sample labelled with Cy-5. Another microarray slide is used to perform a hybridization with control cDNA sample labeled with Cy-5, and the experiment cDNA sample labeled

with Cy-3. The difference in Cy-5/Cy-3 fluoresence ratios in the two microarray hybridizations contributes to the technical variability measure for the up- or downregulation.

Each slide hybridisation experiment was repeated as a dye-swap, and each time point was analysed by combining results of 2 to 3 hybridizations from independent UV-experiments. This resulted in a total of 8 to 12 data points (with two exceptions of 6 data points) for each gene at each time point, as the basis for the quantitative and statistical analysis. In total, 4 independent UV-experiments and 62 successful hybridizations were performed in order to obtain the 8.5 h time-series. Slide scanning was performed by the Scan Array 5000TM (Perkin-Elmer), and spot brightness quantified using QuantArrayTM v. 3 (Perkin-Elmer). The scanning laser intensity was manually adjusted for each chip to optimize the dye-signal acquired.

*Microarray data analysis*

The QuantArray result files for each microarray hybridization slide were run through a set of Perl programs to first determine the quality of the readings intra-chip (spot replicates), and inter-chip (dye-swap replicates). Poor quality slides were excluded from the analysis, and experiments were re-performed as required. The signal measurements on each slide were scaled to provide an overall average fold-change of 0 for the 5th through 95th percentiles of the chip.

For each spot, the signal intensity was calculated as the minimum of the 65% confidence interval of the signal, minus the maximum of the 65% confidence interval for the background. Where dye-swaps were successful for an experiment timepoint, the resulting expression ratio was averaged across both slides (4 spots), inverting ratios where appropriate. Where only one slide succeeded in an experiment timepoint, the ratio was simply the average of the intra-chip replicates (2 spots). In either case, the standard deviation of the measurement replicates was also noted for each gene, along with percentile rank and order of magnitude of the signal intensity. All of these data provide an overview of the quality of the gene expression measurement within the sample provided (technical reproducibility).

Replicate experiments were combined for each timepoint, providing an overview of the reproducibility of gene expression measurements across biological samples. A mean (average of the ratios p) was calculated for each gene at each timepoint, and the standard deviation of the p ratios was also computed. Signal percentile ranks were recalculated, and min/max order-of-magnitude for raw signal intensities noted. The end product for each timepoint in an experimental condition was a table providing: (1) a mean expression ratio for the gene, (2) a measure of the biological variability of the mean, (3) a measure of the variability between technical replicates, (4) a percentile rank of the mRNA abundance, and (5) the order of magnitude range of the mRNA abundance. All of these factors were important in considering the reliability and interpretation of the microarray results.

The calculation of p-values (chance of false positive) was then performed using t-tests. Two t-tests were performed for each gene at each timepoint, the first using the biological replicate variability and degrees of freedom, the second using the technical replicate variability and degrees of freedom. The result yielding the higher p-value ("worse-case scenario") was assigned to the gene in the tables described in the last paragraph.
For the study described in this paper, data for the non-induced state (NI) were calculated by competitive microarray hybridization of samples from a lysogenic and from a non-lysogenic strain, whereas all time-series gene expression values resulted from the comparison of UV-treated and non-UV-treated samples of the lysogenic strain.

*Primer extension analysis for transcript Tx*

Fifteen picomoles of an ID800 5´prime labelled gene-specific primer (5'-caa tta ctt ttc cgt tat aca ata ctt tc-3') was incubated with 5 µg of total RNA at 70°C for 5 min and chilled on ice. Extensions were made in a total volume of 20 µl using reverse transcriptase (Fermentas) by adding the manufacturer's reaction buffer, 20 M dNTPs, 20 U RNase inhibitor before incubating at 42°C for 1 h. The enzyme was inactivated at 70°C for 10 min. The products were analysed by PAGE on a Li-cor machine (DNA Sequencer 4000, MWG-Biotech) using corresponding sequencing reactions as a size marker.

## 3.6 Supplementary data

**Table S.3.1**

**Transcription levels of SSV1 non induced and after UV-treatment [#]**

| Gene ID | | ORF | Length in nt | Transcript | log$_2$ ratios* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NCBI SSV_1p.. | Magpie SSV1_0.. | | | | NI non induced stage | time after UV-treatment | | | | | | | | |
| | | | | | | 0.5 h | 1 h | 1.5 h | 2 h | 3 h | 4 h | 5 h | 6 h | 7 h | 8.5h |
| 02 | 05 | D-355 | 1008 | T5 | **1.53** | 0.37 | -0.33 | -0.27 | 0.45 | 0.29 | 0.14 | **1.00** | 0.85 | **1.21** | **2.33** |
| 03 | 09 | E-54 | 165 | T5 | 0.75 | -0.34 | -3.21 | -0.05 | -0.18 | 0.15 | -3.79 | **1.04** | **1.36** | 0.35 | **2.32** |
| 04 | 10 | F-92 | 279 | T5 | 0.90 | -0.49 | -0.86 | -0.08 | 0.30 | 0.80 | 0.93 | **1.04** | **2.20** | **1.90** | **2.46** |
| 05 | 12 | D-244 | 735 | T5 | 0.69 | -0.42 | -0.38 | 0.14 | 0.15 | 0.91 | 0.79 | 0.85 | **1.95** | 1.19 | **2.41** |
| 06 | 13 | E-178 | 537 | T5 | 1.69 | -0.57 | -0.92 | 0.49 | 0.34 | 0.51 | 0.88 | **1.41** | **1.88** | 0.63 | **3.17** |
| 07 | 14 | F-93 | 282 | T5 | 1.71 | -0.42 | -0.71 | 0.53 | 0.51 | 0.95 | **1.35** | **1.38** | **2.39** | **1.99** | **3.21** |
| 08 | 16 | E-51 | 156 | T5 | 1.87 | -0.43 | -0.29 | 0.54 | 0.14 | **1.08** | **1.34** | **1.45** | **2.62** | **2.40** | **2.79** |
| 09 | 17 | E-96 | 291 | T5 | 1.70 | -0.42 | -0.75 | 0.81 | 0.32 | **1.23** | **1.35** | **1.58** | **2.70** | **2.45** | **2.80** |
| 10 | 20 | D-63 | 192 | T5 | **2.20** | -0.44 | -0.37 | 0.77 | 0.34 | 1.14 | **1.83** | **1.68** | **2.92** | **2.50** | **2.75** |
| 11 | 23 | F-112 | 339 | T5 | **2.16** | 0.04 | -0.07 | 0.92 | 0.33 | **1.25** | **1.74** | **1.87** | **2.87** | **3.19** | **2.63** |
| 12 | 27 | B-49 | 150 | T-ind | 0.69 | 0.71 | **2.68** | **3.50** | **4.21** | **3.54** | **4.79** | **3.73** | **2.08** | **2.13** | **2.96** |
| 13 | 31 | A-100 | 303 | T6 | **2.69** | 0.20 | 0.27 | 0.84 | 0.80 | **1.46** | **3.17** | **2.68** | **2.97** | **3.25** | **2.90** |
| 14 | 33 | A-132 | 399 | T6 | **1.78** | 0.17 | 0.39 | 0.88 | 0.45 | **1.14** | **2.72** | **2.31** | **2.77** | **2.35** | **3.40** |
| 15 | 34 | C-80 | 243 | T6 | **1.18** | -0.11 | -0.07 | 0.52 | 0.63 | 0.99 | **2.48** | **2.07** | **2.52** | **2.54** | **3.22** |
| 16 | 37 | A-79 | 240 | T6 | 0.86 | -0.20 | 0.06 | 0.78 | 0.30 | 0.66 | **2.40** | **1.82** | **1.93** | **1.08** | **3.55** |
| 17 | 39 | A-45 | 138 | T6 | 0.85 | -0.29 | -0.63 | 0.60 | 0.60 | 0.76 | **2.00** | **1.53** | **1.79** | 0.70 | **3.33** |
| 18 | 40 | C-102b | 309 | T6 | 1.47 | -0.29 | 0.07 | 0.42 | 0.34 | -0.26 | 1.69 | **1.61** | **1.74** | **1.80** | **3.28** |
| 19 | 41 | B-129 | 390 | T6 | 0.95 | 0.00 | -0.11 | 0.29 | 0.74 | -0.77 | 1.63 | **1.80** | **1.46** | -0.25 | **2.98** |
| 20 | 43 | A-291 | 876 | T3 | **3.38** | 0.62 | 0.94 | 0.96 | 0.00 | 0.41 | 0.79 | 0.97 | **1.88** | **3.73** | **2.71** |
| 21 | 46 | C-124 | 375 | Tx | **2.42** | 0.48 | 0.78 | 0.87 | 0.50 | 0.64 | 0.31 | 0.79 | 0.55 | 1.14 | **3.40** |
| 22 | 48 | C-792 | 2379 | T4/7/8 | 1.52 | -0.21 | 0.11 | 0.28 | 0.32 | 0.34 | 0.08 | 0.62 | **1.15** | 1.39 | **3.66** |
| 23 | 51 | B-78 | 237 | T4/7/8 | 0.85 | -0.69 | -0.23 | 0.03 | 0.98 | -0.54 | -1.91 | 0.64 | 0.55 | 0.05 | **2.97** |
| 24 | 52 | C-166 | 501 | T7/8 | 1.16 | -0.89 | -0.41 | -0.02 | -0.32 | -0.80 | 0.22 | 0.49 | 0.73 | 0.46 | **3.29** |
| 25 | 54 | B-115 | 347 | T7/8 | **1.61** | -0.42 | 0.38 | -0.08 | -0.43 | -0.57 | 0.49 | 0.65 | 0.85 | **1.32** | **3.04** |
| 26 | 56 | VP1 | 222 | T1/2 T7/8 | **3.35** | 0.32 | -0.21 | 0.26 | -0.35 | 0.60 | 0.54 | 0.91 | **1.37** | **3.47** | **3.24** |
| 27 | 59 | VP3 | 279 | T2/8 | **2.67** | -0.24 | -0.16 | 0.19 | -0.11 | 0.55 | 0.38 | 0.85 | **1.23** | **2.56** | **3.51** |
| 28 | 61 | VP2 | 225 | T2/8 | **1.92** | -0.22 | -0.45 | -0.33 | -0.76 | -0.66 | 0.45 | 0.61 | **1.07** | **2.40** | **3.41** |
| 29 | 62 | A-82 | 249 | T9 | 1.56 | -0.21 | 0.65 | 0.14 | 0.44 | -0.83 | -0.02 | 0.87 | **1.36** | 0.31 | **3.83** |
| 30 | 64 | C-84 | 278 | T9 | **1.05** | 0.06 | 0.25 | 0.25 | -0.09 | 0.51 | 0.20 | 0.67 | **1.35** | 1.04 | **2.44** |
| 31 | 65 | A-92 | 255 | T9 | 1.06 | 0.07 | 0.47 | 0.54 | -0.44 | -0.84 | 0.16 | **1.86** | **1.27** | **1.04** | **3.85** |
| 32 | 67 | B-277 | 834 | T9 | 1.38 | -0.06 | 0.29 | 0.11 | 0.33 | -0.60 | 0.77 | 1.10 | **1.54** | 1.48 | **3.33** |
| 33 | 69 | C-102a | 309 | T9 | 1.21 | -0.27 | 0.66 | 0.20 | 0.20 | -0.74 | 0.42 | **1.12** | **1.62** | 0.73 | **3.00** |
| 34 | 72 | A-153 | 465 | T9 | 1.15 | -0.32 | 0.66 | 0.27 | 0.29 | -0.56 | 0.48 | 1.09 | **1.79** | **1.25** | **3.18** |
| 01 | 02 | B-251 | 756 | T9 | 0.92 | -0.08 | -1.61 | -0.07 | 0.44 | 0.00 | 0.16 | 1.06 | **1.51** | **1.14** | **3.07** |

[#] corresponding values displayed in Fig. 3.4

* values in red are significant (i.e. < or > 1 log$_2$ ratio), highlighted bold: statistically supported p-value of < 0.1

**Table S.3.2**

**Different regulated genes between PH1(SSV1) and PH1 after UV-treatment***

| **PH1(SSV1)** | $log_2$ ratios at time in hours after UV-treatment | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| gene ID | 0.5 h | 1 h | 1.5 h | 2 h | 3 h | 4 h | 5 h | 6 h | 7 h | 8.5 h |
| SSO0968 | -0.03 | **1.08** | **2.07** | 1.02 | **2.57** | 1.25 | **1.14** | 0.63 | 0.04 | 0.40 |
| SSO0969 | -0.21 | 0.39 | **1.67** | 0.81 | **2.38** | 1.04 | 0.38 | 0.69 | 0.04 | 0.08 |
| SSO1460 | 0.28 | **1.59** | **1.01** | **1.31** | 0.01 | **1.22** | 0.70 | 0.76 | 0.11 | 0.22 |
| bac18_n0001 | -0.24 | **1.61** | **1.63** | **1.47** | 0.43 | 3.67 | **1.72** | **1.15** | 0.23 | 0.90 |
| bac18_n0002 | 0.42 | 0.19 | 1.31 | **1.72** | 0.53 | **2.00** | **1.88** | 0.84 | 0.14 | 0.61 |
| SSO0823 | -0.31 | 0.99 | **1.22** | **1.54** | **1.56** | **1.73** | **1.83** | 0.83 | 0.59 | 0.69 |
| SSO2121 | 0.85 | **1.54** | **1.64** | 0.93 | **1.68** | 1.54 | **1.14** | 0.15 | -0.23 | **1.53** |
| SSO2751 | -0.25 | **-2.07** | **-1.02** | **-2.04** | -0.53 | **-2.21** | -0.52 | 0.06 | 0.15 | -0.68 |
| SSO2750 | -0.43 | **-1.76** | **-1.86** | **-1.74** | -0.27 | **-1.80** | 0.02 | 0.22 | 0.06 | -0.88 |
| SSO0257 | -0.01 | -0.90 | **-2.28** | **-1.94** | -0.30 | -2.01 | -0.61 | -0.40 | -0.19 | 0.31 |
| SSO0034 | -0.41 | -0.84 | **-1.79** | **-1.48** | **-1.16** | **-1.76** | -0.61 | -0.24 | -0.12 | -0.66 |
| SSO1210 | -1.33 | -1.33 | -1.38 | -0.60 | **-2.68** | -3.73 | -0.95 | 0.00 | -1.18 | 0.34 |
| SSO0858 | 0.15 | **-1.00** | **-1.36** | **-1.83** | -0.86 | **-2.04** | -0.92 | -0.69 | -0.08 | -0.13 |
| SSO6687 | 0.12 | -0.85 | **-1.65** | **-1.76** | -0.89 | **-2.00** | -0.76 | -0.73 | -0.14 | -0.37 |
| SSO3207 | -0.33 | -1.35 | **-2.17** | **-1.28** | -0.91 | **-1.60** | -0.22 | -0.28 | -0.34 | 0.12 |
| SSO2200 | -0.84 | -0.25 | **-1.53** | **-1.17** | -1.41 | -0.47 | -0.99 | -0.45 | -0.06 | 0.49 |
| SSO0048 | 0.40 | -0.27 | -0.71 | -1.60 | -0.24 | -1.10 | -1.01 | -1.30 | -0.72 | 0.46 |
| SSO3066 | -0.26 | -0.34 | 0.19 | -1.30 | 0.05 | -0.58 | -0.25 | -0.33 | 0.40 | 0.37 |
| SSO9180 | 0.00 | -0.66 | -0.81 | -1.69 | -0.91 | -0.39 | -0.29 | -0.85 | -0.60 | 0.91 |
| SSO10610 | 0.05 | -0.66 | -0.68 | -1.90 | -0.79 | -0.45 | -0.38 | -0.89 | -0.41 | 0.39 |
| SSO9536 | -0.10 | -0.76 | -0.99 | -1.84 | -0.66 | -0.69 | -0.42 | -1.14 | -1.48 | 1.01 |
| SSO0271 | -0.28 | -0.71 | -0.59 | -1.50 | 0.22 | -1.30 | -0.64 | -0.47 | -0.70 | -0.39 |
| **PH1** | $log_2$ ratios at time in hours after UV-treatment | | | | | | | | | |
| gene ID | 0.5 h | 1 h | 1.5 h | 2 h | 3 h | 4 h | 5 h | 6 h | 7 h | 8.5 h |
| SSO0968 | -0.43 | **1.08** | **1.14** | 0.90 | 0.48 | 0.30 | -0.01 | 0.83 | -0.52 | -0.21 |
| SSO0969 | -0.12 | 0.76 | **1.45** | 0.31 | 0.10 | 0.06 | 0.17 | 0.61 | -0.73 | 0.03 |
| SSO1460 | -0.62 | -0.64 | 0.33 | 0.56 | -0.17 | 0.43 | -0.21 | 0.07 | 0.94 | -0.36 |
| bac18_n0001 | 0.03 | -0.67 | 0.07 | **1.24** | 0.74 | 0.75 | 0.11 | 0.32 | 1.06 | -0.08 |
| bac18_n0002 | 0.17 | -0.57 | 0.24 | 0.47 | 0.88 | 0.51 | 0.00 | 0.05 | 0.99 | 0.48 |
| SSO0823 | 0.20 | -0.53 | 0.54 | 0.60 | **1.21** | -0.01 | -0.28 | 0.54 | 0.58 | 1.08 |
| SSO2121 | 0.54 | 0.77 | 0.50 | 0.73 | 0.54 | 0.75 | **1.05** | 1.22 | -0.56 | -0.68 |
| SSO2751 | 0.21 | 0.34 | 0.06 | -0.42 | -0.29 | -0.34 | -0.04 | -0.27 | **-1.47** | -0.26 |
| SSO2750 | 0.70 | 0.14 | -0.09 | -0.67 | -0.18 | 0.09 | -0.06 | -0.92 | -0.77 | 0.04 |
| SSO0257 | -0.84 | -0.03 | -0.23 | **-1.07** | -1.05 | -0.92 | -0.18 | -0.63 | -0.41 | -0.01 |
| SSO0034 | 0.79 | -0.37 | -0.04 | -0.83 | -0.78 | -0.38 | -0.45 | -0.99 | -0.33 | -0.08 |
| SSO1210 | 0.80 | 0.18 | -0.55 | 0.00 | 0.00 | 0.94 | 0.36 | -0.57 | **1.20** | 0.23 |
| SSO0858 | 0.31 | 0.53 | -0.57 | **-1.27** | -0.83 | -0.74 | -0.09 | -0.45 | -0.94 | -0.40 |
| SSO6687 | 0.04 | 0.15 | -0.50 | **-1.40** | -0.73 | **-1.11** | 0.02 | -0.32 | **-1.01** | -0.37 |
| SSO3207 | 0.05 | -0.04 | -0.57 | -1.20 | -0.08 | -0.23 | -1.47 | -0.69 | -0.13 | 0.55 |
| SSO2200 | -0.49 | 0.10 | 0.30 | -0.73 | -0.96 | **-1.05** | 0.00 | -0.49 | -1.28 | -0.16 |
| SSO0048 | 0.12 | 0.42 | -0.25 | -0.79 | -0.98 | -0.85 | -0.70 | 0.14 | -1.41 | -0.81 |
| SSO3066 | -0.60 | 1.57 | -0.84 | -1.34 | -1.23 | -0.45 | -0.78 | -0.61 | -1.49 | -0.54 |
| SSO9180 | -0.86 | -0.10 | -0.38 | -0.89 | -1.50 | -1.55 | -0.65 | 0.65 | -1.41 | -0.86 |
| SSO10610 | -0.85 | 0.03 | -0.34 | -1.02 | -1.65 | -1.87 | -0.70 | 0.59 | -1.89 | -0.67 |
| SSO9536 | -0.50 | -0.20 | -0.71 | -1.01 | -1.47 | -1.23 | -0.49 | 0.71 | -1.43 | -0.72 |
| SSO0271 | -0.62 | 0.44 | 0.09 | -1.01 | -1.58 | -1.94 | -0.81 | -0.24 | -1.67 | -0.65 |

*estimate $log_2$ ratios of the different regulated genes in PH1(SSV1) and PH1 after UV-treatment (correspond to table 3.1).All ratios are listed: non significant in grey, significant up regulated and down regulated values in red respectively green. Statistically supported $log_2$ ratios with a p-value < 0.0500 are highlighted bold.

## 3.7 References


Albers, S. V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampe, R., Driessen, A. J., and Schleper, C. (2006). Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Appl Environ Microbiol 72(1), 102-11.

Arnold, H. P., She, Q., Phan, H., Stedman, K., Prangishvili, D., Holz, I., Kristjansson, J. K., Garrett, R., and Zillig, W. (1999). The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. Mol Microbiol 34(2), 217-26.

Berk, A. J. (2005). Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. Oncogene 24(52), 7673-85.

Besemer, J., and Borodovsky, M. (2005). GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33(Web Server issue), W451-4.

Cannio, R., Contursi, P., Rossi, M., and Bartolucci, S. (1998). An autonomously replicating transforming vector for *Sulfolobus solfataricus*. J Bacteriol 180(12), 3237-40.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18(24), 5294-9.

Delcher, A. L., Harmon, D., Kasif, S., White, O., and Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27(23), 4636-41.

DePamphilis, M. L. (1993). Origins of DNA replication that function in eukaryotic cells. Curr Opin Cell Biol 5(3), 434-41.

Flint, J., and Shenk, T. (1997). Viral transactivating proteins. Annu Rev Genet 31, 177-212.

Forterre, P. (2006). Three RNA cells for ribosomal lineages and three DNA viruses to replicate their genomes: a hypothesis for the origin of cellular domain. Proc Natl Acad Sci U S A 103(10), 3669-74.

Fröls, S., Gordon, P.M.K., Arcellana Panlilio, M., Duggin, I.G., Bell, S.D., Sensen, C.W., and Schleper, S. (2007) Cellular and Transcriptional Responses of the hyperthermophilic archaeon Sulfolobus solfataricus to UV irradiation. submitted

Gadelle, D., Bocs, C., Graille, M., Forterre, P. (2005). Inhibition of archaeal growth and DNA topoisomerase VI activities by the Hsp90 inhibitor radicicol. Nucleic Acids Res. 33(7):2310-7.

Gordon, P. M., and Sensen, C. W. (2004). Osprey: a comprehensive tool employing novel methods for the design of oligonucleotides for DNA sequencing and microarrays. Nucleic Acids Res 32(17), e133.

Grogan, D., Palm, P., and Zillig, W. (1990). Isolate B12, which harbours a virus-like element, represents a new species of the archaebacterial genus *Sulfolobus, Sulfolobus shibatae*, sp. nov. Arch Microbiol 154(6), 594-9.

Grogan, D. W. (1989). Phenotypic characterization of the archaebacterial genus *Sulfolobus*: comparison of five wild-type strains. J Bacteriol 171(12), 6710-9.

Haring, M., Rachel, R., Peng, X., Garrett, R. A., and Prangishvili, D. (2005a). Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, *Acidianus* bottle-shaped virus, from a new family, the *Ampullaviridae*. J Virol 79(15), 9904-11.

Haring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R. A., and Prangishvili, D. (2005b). Virology: independent virus development outside a host. Nature 436(7054), 1101-2.

Heintz, N. H., Dailey, L., Held, P., and Heintz, N. (1992). Eukaryotic replication origins as promoters of bidirectional DNA synthesis. Trends Genet 8(11), 376-81.

Jonuscheit, M., Martusewitsch, E., Stedman, K. M., and Schleper, C. (2003). A reporter gene system for the hyperthermophilic archaeon *Sulfolobus solfataricus* based on a selectable and integrative shuttle vector. Mol Microbiol 48(5), 1241-52.

Kessler, A., Brinkman, A. B., van der Oost, J., and Prangishvili, D. (2004). Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. J Bacteriol 186(22), 7745-53.

Kraft, P., Kummel, D., Oeckinghaus, A., Gauss, G. H., Wiedenheft, B., Young, M., and Lawrence, C. M. (2004a). Structure of D-63 from sulfolobus spindle-shaped virus 1: surface properties of the dimeric four-helix bundle suggest an adaptor protein function. J Virol 78(14), 7438-42.

Kraft, P., Oeckinghaus, A., Kummel, D., Gauss, G. H., Gilmore, J., Wiedenheft, B., Young, M., and Lawrence, C. M. (2004b). Crystal structure of F-93 from *Sulfolobus* spindle-shaped virus 1, a winged-helix DNA binding protein. J Virol 78(21), 11544-50.

Koonin, E. V. (1992). Archaebacterial virus SSV1 encodes a putative DnaA-like protein. Nucleic Acids Res 20(5), 1143.

Lubelska, J. M., Jonuscheit, M., Schleper, C., Albers, S. V., and Driessen, A. J. (2006). Regulation of expression of the arabinose and glucose transporter genes in the thermophilic archaeon *Sulfolobus solfataricus*. Extremophiles 10(5), 383-91.

Martin, A., Yeats, S., Janekovic, D., Reiter, W. D., Aicher, W., and Zillig, W. (1984). SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium *Sulfolobus acidocaldarius* isolate B12. Embo J 3(9), 2165-2168.

Martusewitsch, E., Sensen, C. W., and Schleper, C. (2000). High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements. J Bacteriol 182(9), 2574-81.

Muskhelishvili, G., Palm, P., and Zillig, W. (1993). SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. Mol Gen Genet 237(3), 334-42.

Nadal, M., C. Jaxel, C. Portemer, P. Forterre, G. Mirambeau,, and Duguet, a. M. (1986). Positively supercoiled DNA in a virus like particle of an archaebacterium. Nature (London) 321, 256-258.

Palm, P., Schleper, C., Grampp, B., Yeats, S., McWilliam, P., Reiter, W. D., and Zillig, W. (1991). Complete nucleotide sequence of the virus SSV1 of the archaebacterium *Sulfolobus shibatae*. Virology 185(1), 242-50.

Prangishvili, D., Garrett, R. A., and Koonin, E. V. (2006). Evolutionary genomics of archaeal viruses: unique viral genomes in the third domain of life. Virus Res 117(1), 52-67.

Prangishvili, D., Stedman, K., and Zillig, W. (2001). Viruses of the extremely thermophilic archaeon *Sulfolobus*. Trends Microbiol 9(1), 39-43.

Reiter, W. D., and Palm, P. (1990). Identification and characterization of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium *Sulfolobus* sp. B12. Mol Gen Genet 221(1), 65-71.

Reiter, W. D., Palm, P., Henschen, A., Lottspeich, F., Zillig, W., and Grampp, B. (1987). Identification and characterisation of the genes encoding three structural proteins of the *Sulfolobus* virus-like particle SSV1. Mol Gen Genet 206, 144-153.

Reiter, W. D., Palm, P., and Zillig, W. (1988a). Analysis of transcription in the archaebacterium *Sulfolobus* indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res 16(1), 1-19.

Reiter, W. D., Palm, P., and Zillig, W. (1988b). Transcription termination in the archaebacterium *Sulfolobus*: signal structures and linkage to transcription initiation. Nucleic Acids Res 16(6), 2445-59.

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T., and Young, M. J. (2001). Viruses from extreme thermal environments. Proc Natl Acad Sci U S A 98(23), 13341-5.

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J. E., Douglas, T., and Young, M. (2004). The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. Proc Natl Acad Sci U S A 101(20), 7716-20.

Schleper, C., Kubo, K., and Zillig, W. (1992). The particle SSV1 from the extremely thermophilic archaeon *Sulfolobus* is a virus: demonstration of infectivity and of transfection with viral DNA. Proc Natl Acad Sci U S A 89(16), 7645-9.

Schleper, C., Roder, R., Singer, T., and Zillig, W. (1994). An insertion element of the extremely thermophilic archaeon *Sulfolobus solfataricus* transposes into the endogenous beta-galactosidase gene. Mol Gen Genet 243(1), 91-6.

Serre, M. C., Letzelter, C., Garel, J. R., and Duguet, M. (2002). Cleavage properties of an archaeal site-specific recombinase, the SSV1 integrase. J Biol Chem 277(19), 16758-67.

She, Q., Peng, X., Zillig, W., and Garrett, R. A. (2001a). Gene capture in archaeal chromosomes. Nature 409(6819), 478.

She, Q., Singh, R. K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M. J., Chan-Weiher, C. C., Clausen, I. G., Curtis, B. A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P. M., Heikamp-de Jong, I., Jeffries, A. C., Kozera, C. J., Medina, N., Peng, X., Thi-Ngoc, H. P., Redder, P., Schenk, M. E., Theriault, C., Tolstrup, N., Charlebois, R. L., Doolittle, W. F., Duguet, M., Gaasterland, T., Garrett, R. A., Ragan, M. A., Sensen, C. W., and Van der Oost, J. (2001b). The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci U S A 98(14), 7835-40.

Stedman, K. M., Schleper, C., Rumpf, E., and 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.

Stedman, K. M., She, Q., Phan, H., Arnold, H. P., Holz, I., Garrett, R. A., and Zillig, W. (2003). Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. Res Microbiol 154(4), 295-302.

Turinsky, A. L., Ah-Seng, A. C., Gordon, P. M., Stromer, J. N., Taschuk, M. L., Xu, E. W., and Sensen, C. W. (2005). Bioinformatics visualization and integration with open standards: the Bluejay genomic browser. In Silico Biol 5(2), 187-98.

Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A. K., Zoeller, L., Snyder, J., Douglas, T., and Young, M. (2004). Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. J Virol 78(4), 1954-61.

Zillig, W., Palm, P., Langer, D., Klenk, H.P., Lanzendorfer, M., Hudepohl, U., Hain, J. (1992). RNA polymerases and transcription in archaebacteria. Biochem Soc Symp. 1992;58:79-88.

Zillig, W., Arnold, H. P., Holz, I., Prangishvili, D., Schweier, A., Stedman, K., She, Q., Phan, H., Garrett, R., and Kristjansson, J. K. (1998). Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. Extremophiles 2(3), 131-40.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janekovic, D., Hain, J., Lanzendörfer, M. and Kristjansson J.K. (1994). Screening for *Sulfolobales*, their plasmids and their viruses in Icelandic solfataras. Syst. Appl. Microbiol. 16, 609-628.

Zillig, W., Prangishvilli, D., Schleper, C., Elferink, M., Holz, I., Albers, S., Janekovic, D., and Gotz, D. (1996). Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea. FEMS Microbiol Rev 18(2-3), 225-36.

**Acknowledgements**

This work was supported by grant Schl410/6-1 of the Deutsche Forschungsgemeinschaft, the BiotechGenoMik Netzwerk Göttingen project 4.1 of the German BMBF as well as by initial funding to CS from the University of Bergen. The work in Canada was supported by Genome Canada through Genome Alberta's Integrated and Distributed Bioinformatics Platform Project, as well as by The Alberta Science and Research Authority, Western Economic Diversification, The Alberta Network for Proteomics Innovation and the Canada Foundation for Innovation. We are grateful to two anonymous reviewers for important comments on the manuscript.

Chapter 4

# Protein-protein interactions of the *Sulfolobus shibatae* virus 1 (SSV1)

Manuscript

In preparation

# Protein-protein interactions of the *Sulfolobus shibatae* virus 1 (SSV1)

Seesandra V. Rajagopala , **Sabrina Fröls** , Christa Schleper and Peter Uetz

## 4.1 Abstract

SSV1 is the type virus of the Fuselloviridae family of hyperthermophilic Archaea and was the first high-temperature virus to be characterized in detail. 28 of its 34 open reading frames (ORFs) have no known function. We have cloned all 34 ORFs into two-hybrid vectors and tested them in all pairwise combinations for interactions. Nine interactions were found, including 5 homotypic interactions (C-84 – C-84, C-102a – C-102a, D-335 – D-335, F-93 – F-93, B-49 – B-49, B-129 – B-129) and 3 heterotypic interactions (VP1-E-178, C-84 – E-178, B-78 – A-132). These results confirm that D-335 and F-93 bind to DNA as dimers and suggest new functions for E-178 and C-84 in capsid assembly.

## 4.2 Introduction

The lemon-shaped virus SSV1 has originally been isolated from the hyperthermophilic archaeon *Sulfolobus shibatae* which grows aerobically and heterotrophically at 80°C and pH 3 (Martin *et al.,* 1984). SSV1 was among the first viruses to be studied from hyperthermophilic Archaea. It was shown to infect the closely related strain *S. solfataricus* (Schleper *et al.,* 1992) and novel variants of this virus were recently isolated from hot springs in Yellowstone, Iceland and Kamchatka/Japan (Stedman *et al.,* 2003; Wiedenheft *et al.,* 2004). The double-stranded circular DNA genome of SSV1 contains 15,456 base pairs and has been completely sequenced (Palm *et al.,* 1991). SSV1 has been used for intensive transcriptional studies, that led to a first definition of archaeal transcriptional promoter and termination sites (Reiter *et al.,* 1988a; 1988b) and the virus also serves as a vector to transform *S. solfataricus* (Jonuscheit *et al.,* 2004). The genome of SSV1 encodes 34 open reading frames (ORFs). Upon infection, the circular DNA is site-specifically integrated into a tRNA gene of its host genome and carried as a prophage in the linear integrated form but also as a low copy plasmid (Reiter *et al.,* 1989, Schleper *et al.,* 1992). After UV-irradiation of the lysogenic host the phage DNA is packaged into lemon-shaped viral particles, which are released into the culture medium without apparent lysis of the host cells (Martin *et al.,* 1984; Schleper *et al.,* 1992). Besides these basic features of the viral life cycle little is known about the details of SSV1 biology. In particular, the function of most of the 34 virus proteins remains mysterious. The function of 6 proteins is known from experimental data or has been predicted to some extend based on homology and crystal structures: The three structural proteins, i.e. the capsid proteins VP1 and VP3, as well as the DNA binding protein VP2 were the first to be characterized (Reiter *et al.,* 1987). Furthermore, D-335 was found to encode the viral site-specific recombinase, which is able to catalyze both proviral integration and excision (Muskhelishvili *et al.,* 1993; Serre *et al.,* 2002). The crystal structure of F-93 was recently solved and shows homology to the SlyA and MarR subfamilies of winged-helix DNA binding proteins (Kraft *et al.,* 2004b). This strongly suggests that F-93 functions as a transcription factor that recognizes a (pseudo) palindromic DNA target sequence. Similarly, the crystal structure of D-63 reveals a helix-turn-helix motif that dimerizes to form an antiparallel four-helix bundle. Based on the conserved surface features Kraft *et al.* (2004a) suggested that it may bind small molecules and/or possibly functions as an adaptor protein in macromolecular assembly. However, a more specific function could not be inferred from the structure.

We decided to study the functions of SSV1 proteins more systematically, and screened its proteome for protein-protein interactions. We have cloned all ORFs into two-hybrid bait and prey vectors and tested each of the 34 x 34 = 1156 protein pairs for interactions. Although this does not involve host-virus interactions, we expected a number of interactions that should provide some evidence for possible functions. For example, proteins that interact with one of the 6 experimentally studied proteins mentioned above are likely to be involved in related functions. This principle has been shown to be an efficient way to study protein function in a number of other systems including viruses (Uetz *et al.,* 2000; 2004).

## 4.3 Results and Discussion

Altogether, we tested 34 x 34 = 1156 protein combinations. These tests resulted in nine interactions, which is much less than we expected based on our previous screens of other eukaryotic and prokaryotic organisms and viruses (e.g. Uetz *et al.,* 2000, Rain *et al.,* 2002, and Uetz *et al.,* unpublished data), which usually yielded more than one interaction per protein on average.

The low recovery of protein-protein interactions in the 2-hybrid screen might be due to the huge difference in growth temperature between yeast and the hyperthermophilic virus-host system. The interactions found involve proteins that are conserved in all four SSV variants (shaded black in Fig. 4.1) as well as proteins that are less conserved (grey) or even unique to SSV1 (white).

### 4.3.1 Interactions involving proteins with known or putative function

D335, the viral tyrosine recombinase was found to interact with itself. This confirms an observation by Muskhelishvili (1994) who showed that D-335 binds to a palindromic DNA -sequence which is part of the attP site, suggesting that the protein binds as dimer with each of the two subunits recognizing one half of the palindrome. Furthermore, Letzelter *et al.* (2004) showed by in vitro complementation studies of mutant proteins that the active site is shared by two different monomers of the protein, as found for the yeast Flp recombinase (Letzelter *et al.,* 2004). F-93 was also found to interact with itself. This confirms the previous finding that F-93 exists as a homodimer in solution and that a tight

dimer is also present in the 2.7-Å crystal structure (Kraft *et al.,* 2004b). Further, the crystal structure revealed a fold that is homologous to the SlyA and MarR subfamilies of winged-helix DNA binding proteins. This strongly suggests that F-93 functions as a transcription factor that recognizes a palindromic DNA target sequence.

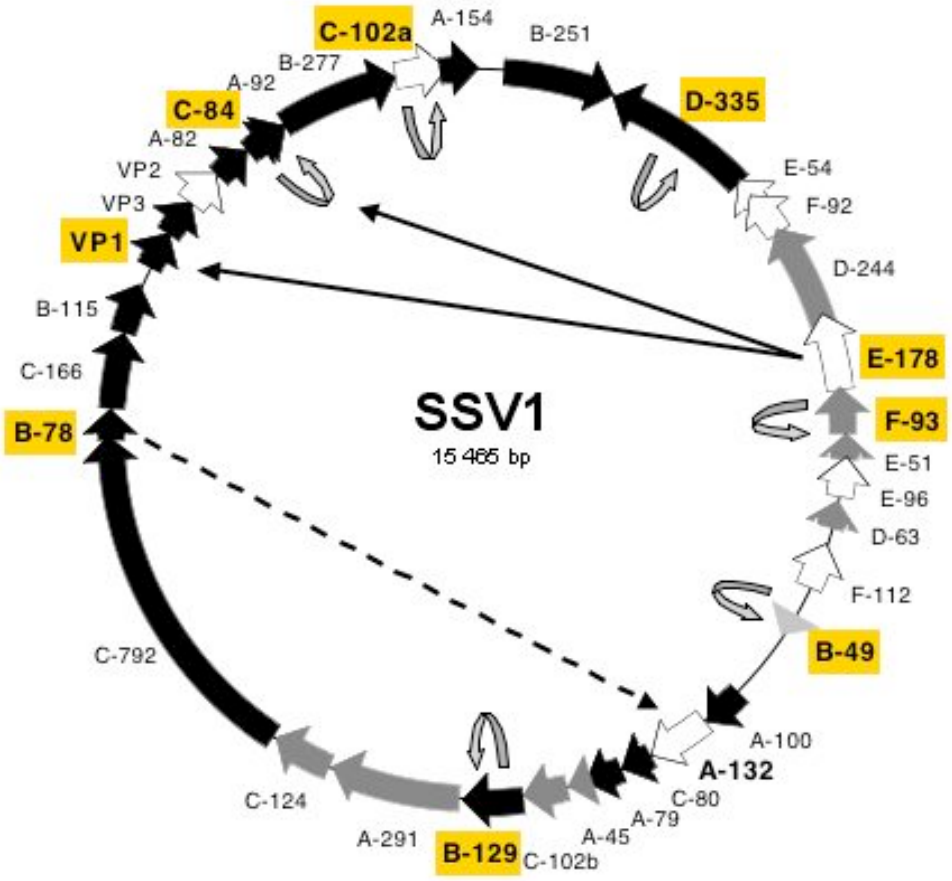



**Figure 4.1:** Genome map of SSV1 and its intragenomic protein-protein interactions. All ORFs involved in interactions are highlighted bold. Homotypic interactions are marked with curved arrows, heterotypic interactions with straight arrows (bait • prey), dashed arrow indicates weak interactions (as defined by slow growth of the yeast expressing the protein pair). ORFs shared between all SSV variants (SSV1, SSV2, SSV RH and SSV K1) are shown in black (for details see Stedmann *et al.,* 2004), conservation among 3 viruses in dark grey, among 2 in light grey. ORFs unique to SSV1 are shown in white.

We found that the capsid protein VP1 interacted with ORF E178 whose function is unknown. This interaction suggests that E178 is involved in packaging or assembly of the viral particle. Alternatively, E178 may be involved in processing of VP1, given the observation of Reiter et al. (1987) that VP1 is expressed as a pro-protein and processed

to yield the mature protein. Although VP1 and VP3 are 56% identical, VP3 was not found to interact with E178 in our assay, suggesting that the interaction may be mediated by the less conserved N-terminal half of VP1.

### 4.3.2 Interactions involving proteins with unknown function

Four proteins of unknown function, B-49, B-129, C-84, and C-102a, were found to form dimers (or possibly multimers). The functional relevance of these dimerizations are not known. Finally, another interaction involved ORFs B-78 and A-132.

Interestingly, homodimerizations were the predominant form of interaction found in our screens. Surprisingly, the homodimerization of D-63 that has been detected through crystallization studies (Kraft *et al.,* 2004a) was not detected in our screen. The reason for this false negative is unknown. It might indicate that a number of interactions have not been detected in our screen because they might not be stable enough under the growth conditions of yeast. Other known reasons for false negatives involve temperature-dependent incorrect folding of proteins or steric hindrances due to the Gal4 fusion proteins that may prevent interactions. This may also explain why most two-hybrid interactions are only detected in one direction, that is, bait B-78 detects prey A-132 but not vice versa. Therefore it is also likely that some interactions have not been detected in our screen.

In order to uncover the function of the SSV1 proteins it will be helpful to also study their interactions with the Sulfolobus host proteins. Given the huge diversity and novelty of the viruses from hyperthermophilic archaea (Prangishvili *et al.,* 2004), more detailed studies on their gene and protein functions have the potential to uncover novel features as well as similarities to bacterial or eukaryotic viruses (Rice *et al.,* 2004) which will shed light on the evolutionary origin of these unique forms of life.

## 4.4 Materials and Methods

*Cloning of SSV1 ORFs into Gateway vectors*

34 viral genes were selected from the published genome sequence (Palm *et al.,* 1991); the ORFs were amplified by PCR with primers containing attB1 and attB2 sites of the Gateway system (Invitrogen) (Tab. S.4.1.) As PCR templates we used SSV1 plasmid DNA isolated from *Sulfolobus shibatae* strain B12. The PCR-products were cloned into the entry vector pDONR201 by BP clonase recombination as recommended by the manufacturer (Invitrogen). The SSV1 ORFs from pDONR201 were cloned into the yeast two hybrid prey vector pGADT7g (which encodes the activation domain of yeast transcription factor Gal4) by LR clonase recombination (Invitrogen). The pGADT7g vector was created by inserting a Gateway conversion cassette (Invitrogen) into the *SmaI* restriction site of pGADT7 vector (Clontech). SSV1 baits were constructed using homologous recombination in yeast. All viral ORFs were amplified by PCR with common primers that bind to the pDONR201 vector 24 base pair upstream of the insert. In addition, the primers contained a 20 base pair specific common tail which was homologous to the two-hybrid vector pOBD2 (Forward primer: 5 AATTCCAGCTGACCACCATG CCA AGT TTG TAC AAA AAA GCA GGC 3 , Reverse primer: 5 GATCCCCGGGAATTGCCATGATC TTG TGC AAT GTA ACA TCA GAC 3 ). A second PCR step with 70-mer primers was used to add another 50 nucleotides of homology to the PCR products to be cloned (Second round Forward primer: 5' C TAT CTA TTC GAT GAT GAA GAT ACC CCA CCA AAC CCA AAA AAA GAG ATC GAA TTC CAGCTG ACCACCATG3, Second round Reverse primer: 3 GTACCGTTAAGGGCCCCTAGGC AGCTGGACGTCTCTAGATACTTAGCATCT ATGACTTTTTGGGGCGTTC 5 ). The PCR product and pOBD2 (linearized with *PvuII* and *NcoI*) was used for yeast homologous recombination (see Cagney *et al.,* 2000 for details).

*Yeast strains and transformation*

Prey clones were transformed into *Saccharomyces cerevisae* strain Y187 (MAT•, ura3- 52, his3-200, ade2-101, trp1-901, leu2-3,112, gal4•, gal80•, URA3::GAL1UAS -GAL1TATA — lacZ met—, Harper *et al.,* 1993),and baits (i.e. PCR product and linearized pOBD2 vector) into haploid yeast strain AH109 (MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4•, gal80•, LYS2::GAL1UAS-GAL1TATA-HIS3, GAL2UAS-GAL2TATA-ADE2, URA3::MEL1UAS-MEL1 TATA-lacZ, James *et al.,* 1996) by the yeas. Transformation protocol described by Cagney *et al.* (2000). The positive transformants were selected by the growth of yeast on

synthetic plates lacking either leucine (pGADT7g transformants) or tryptophan (pOBD2 transformants) (Sherman, 1991).

*Yeast two-hybrid matrix screen*

Prey colonies were arrayed in quadruplicate onto omnitrays (Nunc) filled with rich solid media (YEPD). Liquid cultures of each bait were independently mated with the prey array using a 384-pin high density replica tool as described (Cagney *et al.,* 2000, Uetz *et al.,* 2004). After 1 day of growth at 30°C, diploids containing both bait and prey construct were selected by transfer to synthetic medium lacking leucine and tryptophan and incubated for 3 days. Two-hybrid selection was performed by replicating the diploid array onto medium lacking leucine, tryptophan, and histidine (reporter gene) and supplemented with different concentrations (0 mM, 3 mM, 10 mM) of 3amino-1,2,4-triazole to suppress self activation of certain baits. Yeast twohybrid protein-protein interaction were scored after 5 and 10 days of incubation at 30°C.

## 4.5 Supplementary data

**Table S.4.1.:** List of used primers to ampilfiy the SSV1 ORF´s. The forward primers containing the attB1 site, the reverse primers the attB2 site (Invitrogen), followed by 13 sequence specific nucleotides.

| | Forward primer (attB1 -- sequence specific) | Reverse primer (attB2 -- sequence specific) |
|---|---|---|
| **att sites** | 5´-GGGG-ACA-AGT-TTG-TAC-AAA-AAA-GCA-GGC-TTA-- | 5`-GGGG-AC-CAC-TTT-GTA-CAA-GAA-AGC-TGG-GTT-- |
| **Nr. ORF** | | |
| **1 B-251** | --ATGGTAAGGAACA-´3 | --CTATCTTGTTACG-´3 |
| **2 D-335** | --ATGACGAAAGATA-´3 | --TCAGACCCCTTTT-´3 |
| **3 E-54** | --ATGGCTAGAATAC-´3 | --TTATGGTTGCCAA-´3 |
| **4 F-92** | --ATGATATTCAACT-´3 | --TCATTTGTCCACC-´3 |
| **5 D-244** | --ATGGATTGTAATA-´3 | --TCATTTTCCGCCC-´3 |
| **6 E-178** | --ATGGATCAAAGAA-´3 | --TCACCATATTGAA-´3 |
| **7 F-93** | --ATGAAAAGTACAC-´3 | --TCACCTCTTTTTG-´3 |
| **8 E-51** | --ATGATAGAGAAGA-´3 | --TCACTCCTCAGCC-´3 |
| **9 E-96** | --ATGATGGTGCAAG-´3 | --TTAGGTCACCTTG-´3 |
| **10 D-63** | --ATGAGTAAAGAAG-´3 | --TCAGCTCACCTTA-´3 |
| **11 F-112** | --ATGGCACAAACTC-´3 | --TCACTGCTTTGCC-´3 |
| **12 B-49** | --ATGGGATGTGCAA-´3 | --TTAGAACAAATCA-´3 |
| **13 A-100** | --ATGGTTTCACCCC-´3 | --TCAGAAGTCAAAA-´3 |
| **14 A-132** | --ATGAAGGCTGAGG-´3 | --TCATTGCCCCTCA-´3 |
| **15 C-80** | --ATGAAGGCTAGGG-´3 | --CTAAACATTTTCT-´3 |
| **16 A-79** | --ATGTTTAGATGCC-´3 | --TCACTCAACCTCT-´3 |
| **17 A-45** | --ATGTATCAATGTC-´3 | --TCATTGTCCTCTC-´3 |
| **18 C-102b** | --ATGAACCTAATTG-´3 | --TCATCCTCTAACG-´3 |
| **19 B-129** | --ATGACGGAGTCAG-´3 | --CTAACTAACGCAG-´3 |
| **20 A-291** | --ATGAGGAAGTCCC-´3 | --TTAATATAGCTGC-´3 |
| **21 C-124** | --ATGAAAAAACTGT-´3 | --CTATCTTCTAAAT-´3 |
| **22 C-792** | --ATGAAGTGGGGAC-´3 | --TCATTCCTCCCTC-´3 |
| **23 B-78** | --ATGACGGACGCAA-´3 | --TTAGTCCCCATCC-´3 |
| **24 C-166** | --ATGGGGACTAAGT-´3 | --TCATTCCGACCCC-´3 |
| **25 VP3** | --ATGGAAATCAGTT-´3 | --TCACTCCTCCTTA-´3 |
| **26 VP2** | --ATGAAGTGGGTGC-´3 | --CTACTTGCGGTGC-´3 |
| **27 A-82** | --ATGAGTGCGTTAG-´3 | --TCATATCCTTTCC-´3 |
| **28 C-84** | --ATGAGGTGGGGTA-´3 | --TTAAGAGATAAGC-´3 |
| **29 B-277** | --ATGTCTGATGGGA-´3 | --TTAGCTCACCCCT-´3 |
| **30 C-102a** | --ATGGTCTCAGTAA-´3 | --TTAGCCATTCAGC-´3 |
| **31 A-153** | --ATGGCTAAAAAGA-´3 | --TTACTCTTCTTCA-´3 |
| **32 B115** | --ATGACGGAGTATA-´3 | --TTAGCCTCTTTGA-´3 |
| **33 VP1** | --GAAGCAACCAACA-´3 | --TCAGTCTTTGTAT-´3 |
| **34 A91** | --ATGATAGGGATAC-´3 | --TCAGACATGCAAA-´3 |

## 4.6 References

Cagney, G., Uetz, P. & Fields, S. (2000). High-throughput screening for protein-protein interactions using two-hybrid assay. Methods Enzymol 328, 3-14.

Cagney, G., Uetz, P. & Fields, S. (2001). Two-hybrid analysis of the Saccharomyces cerevisiae 26S proteasome. Physiol Genomics 7, 27-34

Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K., & Elledge, S. J. (1993) The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75, 805-816.

James, P., Halladay, J., &Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics 144, 1425-1436.

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. Mol Microbiol 48(5):1241-1252.

Kraft, P., Kummel, D., Oeckinghaus, A., Gauss, G. H., Wiedenheft, B., Young, M. & Lawrence, C. M. (2004a). Structure of D-63 from Sulfolobus spindle-shaped virus 1: surface properties of the dimeric fourhelix bundle suggest an adaptor protein function. J Virol 78, 7438-42.

Kraft, P., Oeckinghaus, A., K mmel, D., Gauss, G. H., Gilmore, J., Wiedenheft, B., Young, M. & Lawrence, C. M. (2004b). Crystal Structure of F-93 from Sulfolobus Spindle-Shaped Virus 1, a Winged-Helix DNA Binding Protein. J Virol 78, 11544—11550.

Letzelter, C., Duguet, M., & Serre, M.C. (2004). Mutational analysis of the archaeal tyrosine recombinase SSV1 integrase suggests a mechanism of DNA cleavage in trans. J Biol Chem 279, 28936-28944.

Martin, A., Yeats, S., Janekovic, D., Reiter, W. D., Aicher, W. & Zillig, W. (1984). SAV-1 a temperate UV inducible DNA virus-like particle from the archaebacterium Sulfolobus acidocaldarius isolate B-12. EMBO J 3, 2165—2168.

Muskhelishvili, G., Palm, P. & Zillig, W. (1993). SSV1-encoded site-specific recombination system in Sulfolobus shibatae. Mol Gen Genet 237, 334-342.

Muskhelishvili, G. (1994). The archaeal SSV integrase promotes intermolecular excisive recombination in vitro. Syst Appl Microbiol 16, 605—608.

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, 242-50.

Prangishvili, D. & Garrett, R.A. (2004). Exceptionally diverse morphotypes and genomes of crenarchaeal hyperthermophilic viruses. Biochem Soc Trans 32, 204-208.

Rain, J. C., Selig, L., De Reuse, H., & 10 other authors (2001). The proteinprotein interaction map of Helicobacter pylori. Nature 409, 211-215.

Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E., Johnson, J.E., Douglas, T. & Young, M. (2004). The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. Proc Natl Acad Sci USA 101, 771620.

Reiter, W. D., Palm, P., Henschen, A., Lottspeich, F., Zillig, W. & Grampp, B. (1987). Identification and characterization of the genes encoding three structural proteins of the Sulfolobus virus-like particle SSV1. Mol Gen Genet 206, 144—153.

Reiter, W.D., Palm, P. & Zillig, W. (1988a) Transcription termination in the archaebacterium Sulfolobus: signal structures and linkage to transcription initiation. Nucleic Acids Res 16, 2445-2459.

Reiter, W.D., Palm, P. & Zillig, W. (1988b). Analysis of transcription in the archaebacterium Sulfolobus indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. Nucleic Acids Res 16, 1-19.

Reiter, W.D., Palm, P. & Yeats, S. (1989). Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. Nucleic Acids Res 17, 1907-1914.

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. Proc Natl AcadSci USA 89, 7645-7649.

Serre, M.C., Letzelter, C., Garel, J.R. & Duguet, M. (2002). Cleavage properties of an archaeal site-specific recombinase, the SSV1 integrase. J Biol Chem 277, 16758-67.

Sherman, F., & Wakem, P. (1991). Mapping yeast genes. Methods Enzymol 194, 38-57.

Stedman, K. M., She, Q., Phan, H., Arnold, H. P., Holz, I., Garrett, R. A. & Zillig, W. (2003). Relationships between fuselloviruses infecting the extremely thermophilic archaeon Sulfolobus: SSV1 and SSV2. Res Microbiol 154, 295-302.

Uetz, P., Giot, L., Cagney, G., & 17 other authors (2000). A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature 403, 623-627.

Uetz, P., Rajagopala, S. V., Dong, Y. A., and Haas, J. (2004). From ORFeomes to protein interaction maps in viruses. Genome Res 14, 2029-2033.

Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A.K., Zoeller, L., Snyder, J., Douglas, T. & Young, M. (2004). Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. J Virol 78, 1954-61.

## Acknowledgments

This work was supported by DFG grant Ue 50/2-1 to P.U. and by DFG grant SCHL410-3 to C.S.

Chapter 5

# Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV-damage

Published in:

**Journal of Bacteriology**

Volume 189, Number 23, December 2007, Pages 8708–8718

American Society for Microbiology

# Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV-damage

**Sabrina Fröls** , Paul M.K. Gordon , Mayi Arcellana Panlilio , Iain G. Duggin , Stephen D. Bell , Christoph W. Sensen and Christa Schleper

## 5.1 Abstract

In order to characterize the genome-wide transcriptional response of the hyperthermophilic, aerobic crenarchaeote *Sulfolobus solfataricus* to UV-damage, we used high-density DNA-microarrays which covered 3368 genetic features encoded on the host genome, as well as the genes of several extrachromosomal genetic elements. While no significant up-regulation of genes potentially involved in direct DNA-damage reversal was observed, a specific transcriptional UV-response involving 55 genes could be dissected. Although flow cytometry showed only modest perturbation of the cell cycle, strong modulation of the transcript levels of the Cdc6 replication initiator genes were observed. Up-regulation of an operon encoding Mre11 and Rad50 homologues pointed to induction of recombinational repair. Consistent with this, DNA double-strand breaks were observed between 2 and 8 hours after UV-treatment, possibly resulting from replication fork collapse at damaged DNA-sites. The strong transcriptional induction of genes which are potentially encoding functions for pilus formation suggested that conjugational activity might lead to enhanced exchange of genetic material. In support of this, a statistical microscopic analysis demonstrated that large cell aggregates formed upon UV-exposure. Together, this provided supporting evidence to a link between recombinational repair and conjugation events.

## 5.2 Introduction

Most organisms meet the challenge of maintaining their genome integrity and assuring correct replication of their genetic material while protecting themselves against the DNA-damaging effects of UV-light. This is reflected in the large number of proteins involved in DNA-repair pathways, which are found in all three domains of life; the Bacteria, Eukaryota and Archaea. For hyperthermophilic organisms, like many archaea, that dwell at the upper temperature limit of life (Stetter, 2006), this challenge might be even more demanding. Studies on mutation frequencies and repair in Archaea have been inspired by the expectation that extremophiles growing under conditions which accelerate spontaneous DNA-damage should be particularly proficient in DNA-repair (Crowley *et al.*, 2006; Grogan, 2000; Raychaudhuri *et al.*, 2003). Archaea have also gained special interest because of their unique evolutionary position and their relationship to eukaryotes. Homology in many factors in the systems responsible for transcription and replication have been observed. The homologous, yet simpler, archaeal systems provide a powerful tool for the study of cellular evolution and more complex systems in the eukaryotic nucleus (Klenk, 2006). The homology between the eukaryotic and archaeal domains also exists in DNA-repair systems (Aravind *et al.*, 1999; Kelman and White, 2005). For example, potential factors involved in nucleotide excision repair (NER) of UV-induced DNA-lesions are - in most archaea - exclusively constituted by homologs of the eukaryotic proteins XPF/XPB/XPD/Fen-1. The in vivo function of this system in archaea has not yet been elucidated, and the system also seems to be incomplete (Kelman & White, 2005; Romano *et al.*, 2006). However, Salerno *et al.* (2003) have shown that *Sulfolobus* can efficiently conduct the repair of photoproducts in the dark, suggesting the presence of an active NER system that is perhaps completed by an as-yet-uncharacterized set of genes. By contrast, in the archaeon *Halobacterium* the bacterial uvr system is additionally present, and in that case it seems to be solely responsible for repair of DNA photoproducts in the dark (Crowley *et al.*, 2006).

Notably, some proteins involved in DNA-repair systems in bacteria and eukaroytes are absent in most archaea, such as the mutL/mutS mismatch repair machinery (Grogan, 2004; Kelman & White, 2005), indicating that alternative systems might be present (Makarova *et al.*, 2002). Whereas most repair mechanisms directly act on the damaged DNA, unrepaired lesions can also be overcome during replication. A lesion bypass polymerase (Dpo4) (Kulaeva *et al.*, 1996) has been found in those archaea that contain

photolyases, such as the halophiles, which are exposed to strong solar radiation. Some thermophiles from terrestrial hot springs also contain these enzymes (see Kelman & White, 2005).

*Sulfolobus spp.,* which reside in solfataras (mud pots) all over the globe, have emerged as important model organisms for biochemical and genetic studies of hyperthermophilic archaea, including analyses on genome integrity and DNA-repair. In *S. acidocaldarius* the rate of spontaneous mutation frequencies was found to be comparable to that of other microorganisms, indicating that hyperthermophiles are able to maintain genomic stability, despite the extreme growth conditions (Grogan *et al*., 2001). The anaerobic hyperthermophilic euryarchaeote *Pyrococcus furiosus* has an astonishingly high resistance to gamma-irradiation, and a highly efficient repair mechanism for double-strand DNA breaks (Robb *et al.*, 2001; Williams *et al.*, 2006); by contrast, the sensitivity of *S. acidocaldarius* to gamma-irradiation was found to be comparable to that of *E. coli* (Reilly & Grogan, 2002). Similarly, mutational analyses after exposure to short wave-length UV-light revealed that *Sulfolobus* was as sensitive and equally UV-mutable as *E. coli* and exhibited effective photoreactivation under visible light (Wood *et al.*, 1997). In line with these findings, Salerno *et al.* (2003) identified cis-syn-cyclobutane pyrimidine dimers (CPDs) in *Sulfolobus solfataricus* after treatment with UV-light, which together with pyrimidine 6-4 pyrimidone photoproducts (6-4PP), are known to be direct consequences of UV-induced damage. The same authors demonstrated repair of CPDs in the dark, suggesting the presence of an active NER pathway in *Sulfolobus* (Salerno *et al.*, 2003). However, unlike in other organisms, in *Sulfolobus* it seems to act with the same efficiency on both DNA-strands, lacking a transcription-coupled activity (Dorazi *et al.*, 2007; Romano *et al.*, 2006). Interestingly, an increased rate of exchange of genetic markers was observed with *S. acidocaldarius* mutants upon treatment with UV-light and it was hypothesized that DNA-lesions and double strand breaks stimulate this process (Schmidt *et al.*, 1999; Wood *et al.*, 1997).

*Sulfolobus solfataricus* is a host for the virus SSV1, which contains a 15.5 kb double stranded circular DNA-genome that site-specifically integrates into the host chromosome. Viral replication and propagation is strongly inducible by UV-light (Martin *et al.*, 1984; Reiter *et al.*, 1988; Schleper *et al.*, 1992), which led to the early speculation that Archaea might have an SOS-like system (Martin *et al.*, 1984). Two recent genome-wide transcriptional studies in *Halobacterium sp.*, however, did not reveal a concerted induction

of the genes involved in excision repair or other damage-reversal genes (Baliga *et al.*, 2004; McCready & Marcello, 2003). Instead, only genes involved in homologous recombination seemed to be induced when cells were exposed to low doses of UV-light (McCready *et al.*, 2005).

Here we describe a genome-wide transcriptional analysis of the response of *Sulfolobus solfataricus* to UV-irradiation aimed to investigate the general UV-response of a hyperthermophilic archaeon. Our studies are complemented by analyses of double-strand break formation, the cell cycle and cell physiology.

## 5.3 Results

### 5.3.1 Survival and growth of *S. solfataricus* cells after exposure to UV-light

To evaluate the impact of UV-irradiation on *S. solfataricus*, we first analysed cell viability after UV-doses in a range from 25 to 200 J/m$^2$ of UV-C (254 nm) (Tab. 5.1).

**Table 5.1:** Survival fraction of *S. solfataricus* cells after treatment with UV-light

| Survival of *S. solfataricus* after treatment with different doses of UV-light * | | | |
|---|---|---|---|
| UV-doses (J/m$^2$) | absolte numbers (cfu/ml) | standard deviation (cfu/ml) | survival fraction (in % of cfu frome 0 J/m$^2$) |
| 0 J/m$^2$ | 2.41 x 10$^8$ | 1.40 x 10$^8$ | 100 % |
| 25 J/m$^2$ | 1.35 x 10$^8$ | 9.43 x 10$^7$ | 56 % |
| 50J/m$^2$ | 9.40 x 10$^7$ | 5.17 x 10$^7$ | 40 % |
| 75 J/m$^2$ | 2.47 x 10$^7$ | 1.07 x 10$^7$ | 11 % |
| 100 J/m$^2$ | 5.81 x 10$^6$ | 3.10 x 10$^6$ | 2.5 % |
| 150 J/m$^2$ | 5.33 x 10$^4$ | 4.11 x 10$^4$ | 0.03 % |
| 200 J/m$^2$ | 6.67 x 10$^3$ | 9.43 x 10$^3$ | 0.0003 % |

* an exponentially grown culture was treated with different UV-doses from 25 J/m$^2$ to 200 J/m$^2$ and colony forming units were determined after plating of the cells (cfu/ml), means of n=3 are given.

The UV-dose of 75 J/m$^2$, which was used in all further experiments, yielded a plating efficiency of approximately 10 to 40% compared to non-treated cells. Fig. 5.1 displays representative growth rates of UV-irradiated (75 J/m$^2$) and control cultures of *S. solfataricus* PH1. The strain showed, a growth retardation after UV-treatment compared to the mock-treated control culture.

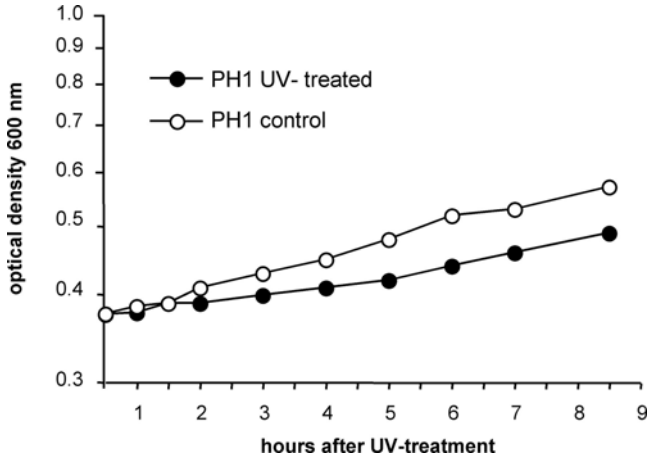



**Figure 5.1:** Representative growth curve of *S. solfataricus* PH1 after UV-treatment. A UV-dose of approximately 75 J/m$^2$ (254 nm) was used for the treatment and the cultures were re-cultivated at time point 0 h. The time points of sampling for DNA and RNA extractions are indicated by the symbols.

### 5.3.2 UV-exposure induces the formation of cell aggregates

Microscopic examination of cells revealed the formation of cell aggregates between 3 and 10 hours after UV treatment, with the greatest level of aggregation appearing on average 6 hours after the UV treatment (Fig. 5.2). On average, two to five cells were found in the early aggregates, while bigger complexes tended to form later. The formation of cell aggregates was similar in both strains, i.e. was independent of the virus SSV1, and was highly reminiscent of the formation of aggregates observed in the context of plasmid-mediated conjugation (Schleper *et al.*, 1995).

We saw much less or even no aggregate formation at all, when the cells were exposed to higher doses of UV-light (200 J/m$^2$), indicating that the cell clumping did not represent a non-specific aggregation of dead cells (not shown). Furthermore, differential staining (see *Materials and Methods*) indicated that at least 50% of the cells within the aggregates were metabolically active (data not shown). We therefore conclude that the aggregation of the cells most probably represents a regulated cellular reaction to the UV-treatment.

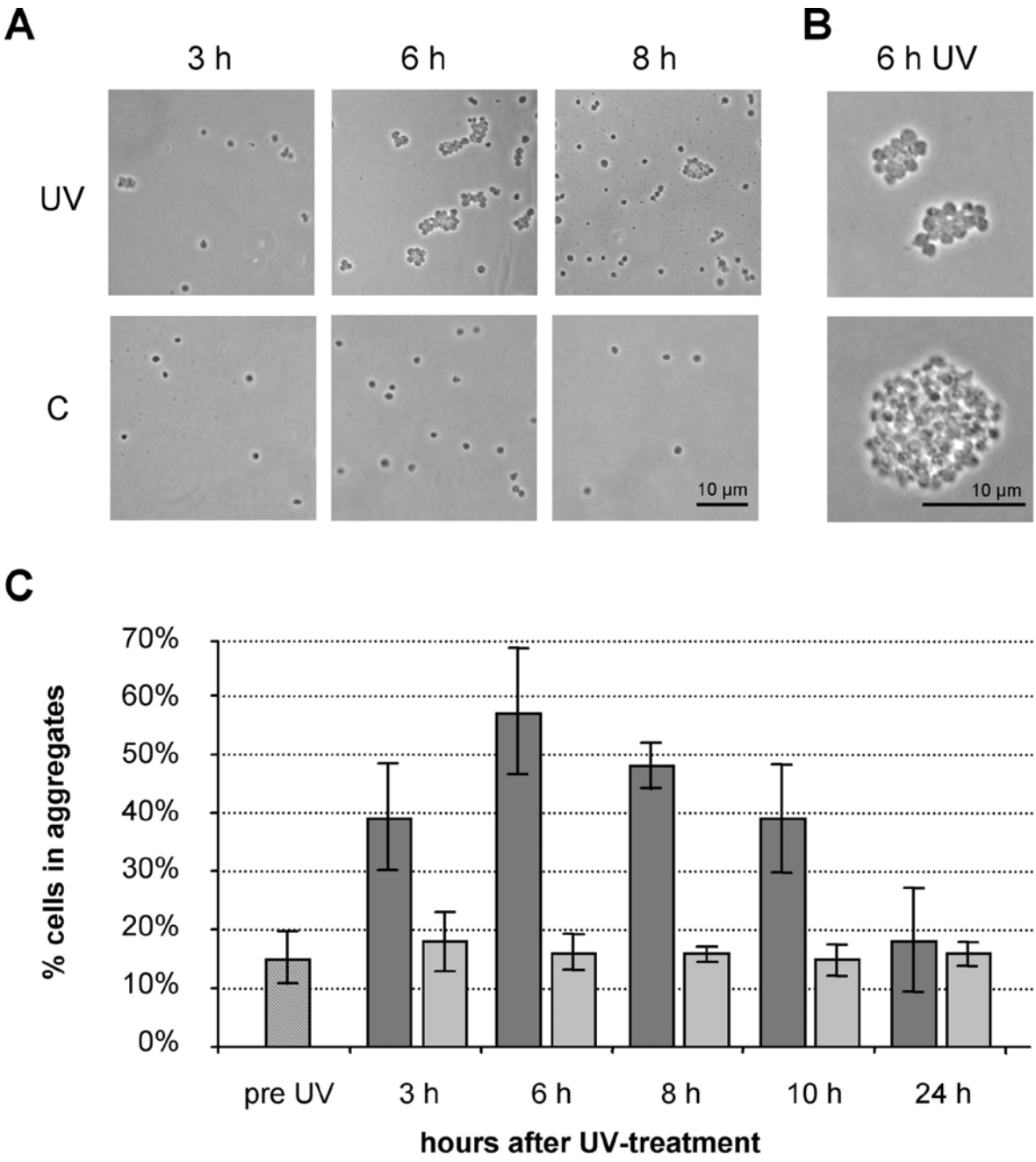



**Figure 5.2:** Aggregation of *S. solfataricus* cells after UV-treatment (**A**) Micrographs (phasecontrast) of *S. solfataricus* cells fixed on a gelrite-coated microscopy slide. Representative pictures of cells and/or aggregates at 3h, 6h and 8h after UV-treatment are shown. C= control culture without UV-treatment. An increasing amount of cell aggregates was observed in UV-treated cultures after 3h. (**B**) Top: typical cell-aggregates as observed around 3 to 6h after UV-treatment. Bottom: typical "maxi" cell aggregate, mostly discovered at 6h after UV-treatment, excluded from statistics in C, because the number of cells was uncountable. (**C**) Quantitative analysis of cell aggregate formation at different time points after UV-treatment. The pre-UV culture was split into a UV-treated culture (dark grey) and a control culture (light grey). The amount of cells in and out of aggregates was counted until 1000 single cells were found in total. Means and standard deviation (error bars) are shown from three independent UV-experiments. The amount of cells found in aggregates is an underestimate, because cells in the large aggregates were not countable (see B).

### 5.3.3 Analysis of cellular DNA content by flow cytometry

During exponential growth, *Sulfolobus* cells remain in a G2/M phase of the cell cycle for a relatively long time period, with only a very short G1 phase (Hjort & Bernander, 1999). Therefore, most cells of an unsynchronized, exponentially growing *Sulfolobus* culture contain two genomes, while a considerably smaller fraction contains only one. Fig. 5.3 shows the DNA-content distribution of an exponential culture of *S. solfataricus* PH1, acquired by flow cytometry (Fig. 5.3, left row, control). After UV-treatment of an exponential culture, we observed an initial modest accumulation of cells with DNA contents that coincide with cells in the G1 and S phases of the normal cell cycle (Fig. 5.3, see panels representing 0.5 h to 5 h post-UV treatment). Five hours after the UV-treatment, this effect became more obvious, and in addition, a heterogeneous population of cells containing greater than two chromosome equivalents became apparent, which was most obvious after 8.5 hours post-UV treatment. A similar phenotype has been observed after treatment of *S. solfataricus* with hydroxyurea, a likely DNA damaging agent (IGD and SDB, unpublished data). Similar data to those shown in Fig. 5.3 were obtained with the PH1(SSV1) strain (see Fig. S.5.1 in suppl. material).

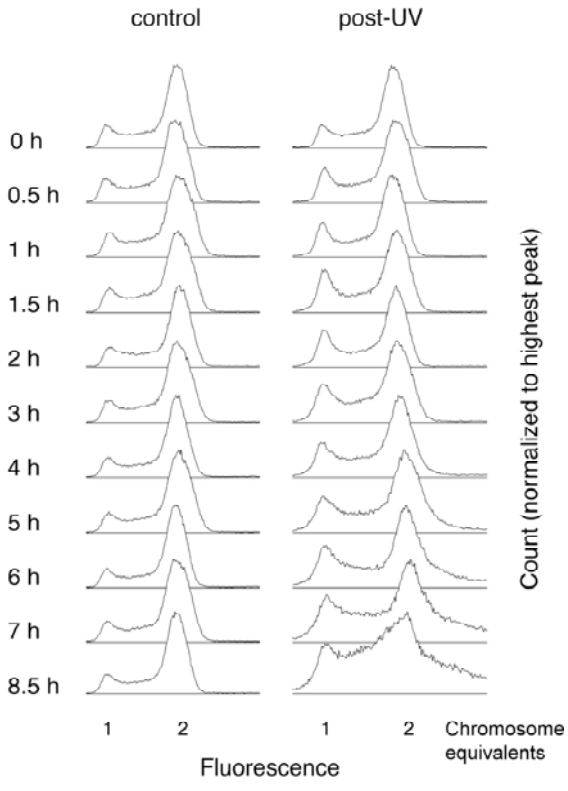



**Figure 5.3:** Flow cytometry analysis of *S. solfataricus* cells after UV-treatment (strain PH1). Cells were fixed in 80% ice-cold ethanol and the DNA contents were measured by fluorescence (see Fig. S.5.1 for strain PH1 (SSV1)). Samples from the UV-treated culture (post-UV) and mock-treated control culture (control) were analysed from 0 h to 8.5 h. A chromosome content of 1N is found in G1 phase, between 1N and 2N represent S phase (DNA synthesis) and 2N is G2 phase of the cell cycle stages (Hjort & Bernander, 1999).

### 5.3.4 Formation of double-strand breaks

While CPDs have been demonstrated to occur in *Sulfolobus* after UV-treatment (Salerno *et al.*, 2003), it has not been investigated if double strand breaks (DSB) are formed as a result of unrepaired lesions during replication, similar to those observed in *E. coli* (Bonura & Smith, 1975), yeast (Haber, 2006) or more recently described in mouse cells (Garinis *et al.*, 2005). In order to analyse the formation and extent of double strand breaks, we analysed DNA prepared after UV-treatment using pulsed field gel electrophoresis (Fig. 5.4). A considerable accumulation of chromosomal fragments of smaller size than in the control samples was observed, peaking at 2 hours after UV-treatment, but visible until 8 to 10 hours after UV-treatment. The most abundant fraction of fragments captured in this analysis ranged from 100 kb up to 600 kb in size, mostly because the electrophoresis conditions were chosen such that all fragments of 600 kb and bigger were compressed in the upper part of the gel.

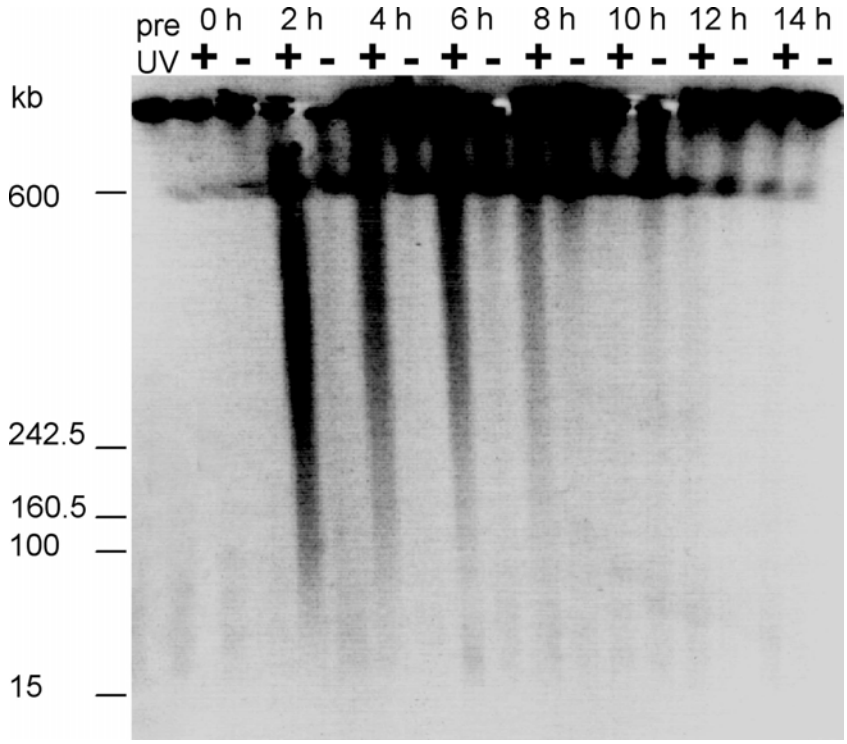



**Figure 5.4:** PFGE-analysis of total DNA from *S. solfataricus* PH1 after UV-treatment to determine the extent of double-strand breaks (DSBs). DNA from UV-treated (+) and control (-) cells were analysed from 0 h to 14 h after UV-treatment and pre-UV. Fragmentation of DNA is visible as a smear in the area of the gel below the compression zone (at 600 kb), from 2 hours to 8 hours mostly in the UV-treated samples.

No DSBs were observed at time point 0 (cells harvested immediately after treatment which might indicate that their formation was not a direct result of the UV-treatment *per se*, but rather, a result of subsequent cellular processes. However less material seemed to have been separated in those lanes. By contrast, we observed CPD formation at time point 0 (data not shown), consistent with previous findings (Salerno *et al.*, 2003), suggesting that CPDs are a direct result of the UV-treatment.

### 5.3.5 General transcriptional response

For each of the 4 independent UV time series experiments, RNA was isolated from UV-treated and control cells and analyzed using Northern hybridizations to evaluate the quality of the isolated nucleic acids and to verify induction of the viral cycle in the lysogenic strain PH1(SSV1) (Fröls *et al.*, 2007) as well as of some UV-responsive chromosomal genes (not shown). For microarray hybridizations, the total RNA was reverse-transcribed and dual-labelled with fluorescent dyes.

We identified 55 UV-responsive genes in *S. solfataricus* that exhibited a pronounced change in mRNA copy numbers over an extended period of time. Among these were 19 genes that we categorized as being strongly induced genes based on k-means clustering (KMC) analysis (group Ia in Tab. 5.2, see also suppl. material Fig S.5.2). Another 14 genes showed a similar expression pattern as group Ia but had smaller amplitudes in mRNA level changes in KMC analysis (group Ib Tab. 5.2). A third group of 22 genes represented the most pronounced down-regulated genes (group II in Tab. 5.2). The averaged expression profiles of these groups are presented in Fig. 5.5 and suppl. Fig. S.5.3 (single ratios are listed in Tab. S.5.1 and S.5.2). The figure shows that the UV-dependent response over time lasted from ca. 1.5 hours after UV-treatment until 5 hours. The start of the transcriptional response in the lysogenic strain PH1(SSV1) was observed considerably earlier and the response was generally stronger, with an average maximal induction level of group Ia genes of 12 fold ($\log_2$ of 3.5) versus 6-fold ($\log_2$ of 2.5) for strain PH1 (see suppl Fig. S.5.3). Therefore the use and comparison of data from both strains helped in dissecting those genes that showed a significant UV-dependent response. Immediately after the UV-treatment, a large number of genes seemed to be induced over only a short time period (1 to 1.5 hours after UV) in strain PH1 and to a lower extent in the infected strain PH1(SSV1), respectively. These genes are represented by the blue dotted line in Fig. 5.5.

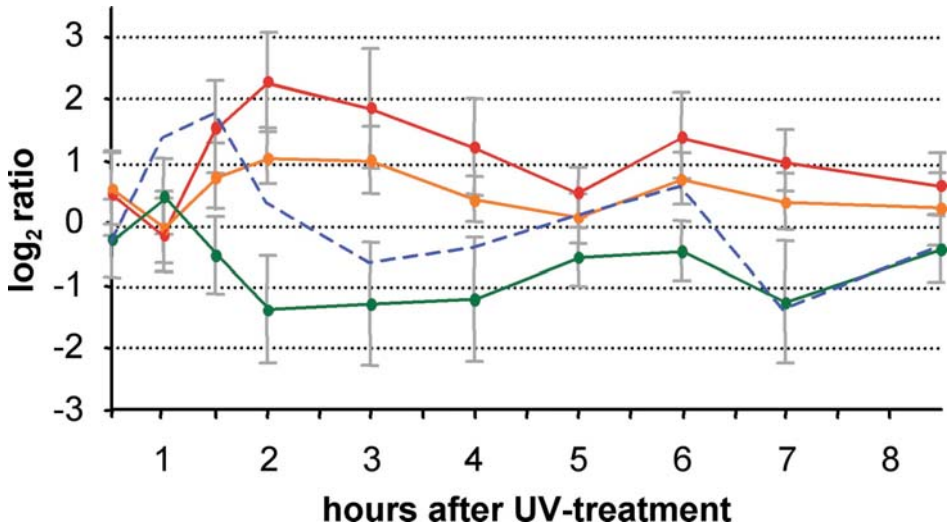



**Figure 5.5:** Expression profiles of the general transcriptional response after UV-treatment of strain PH1. The curves display the mean average of the three identified UV-dependent regulated gene groups as displayed in Table 5.2, the highly induced group of 19 genes (red curve), induced gene group of 14 genes (orange) and the down-regulated group of 22 genes (green). The blue dashed curve was generated from the averages of 11 genes, but represents qualitatively the pattern of approximately 400 genes that are mostly involved in the transcription and translation processes. Errors bars do not represent standard deviations, but express the range of gene expression of the different genes that are "summarized" in each line.

They mostly encode factors involved in translation and transcription, as well as house keeping proteins involved in central metabolic pathways or information processing. A significant increase ($p < 0.05$ in 3 experimental replicates) in the mRNA levels of these approximately 400 genes in strain PH1 was found immediately after UV-treatment until 1.5 h after. This spike occurs before, and is distinct from the long-duration regulation (up or down) of genes in the 1.5 h to 7 h range (compare the dotted blue lines and red lines in Fig. 5.5). We postulate that the genes which show a significant change in mRNA levels over this very short time period appear up-regulated because (1) they are strongly transcribed genes and (2) the cell cycle was modestly perturbed due to the UV-teatment. In order to get insights into the effect of UV-treatment on central processes within the cell, we looked into the datasets for all genes involved in information processing (e.g. replication, transcription, and translation, based on COGs). We found that most of these genes followed the general expression pattern (blue dotted line in Fig. 5.5), although not all peaked considerably (> or < than 2-fold change).

**Table 5.2:** UV-dependent regulated genes of *S. solfataricus*

**UV-responsive genes in *Sulfolobus solfataricus***

| Group [a] | P | S | Gene ID [b] | | COG | Operon [c] | Predicted function | Homology [d] |
|---|---|---|---|---|---|---|---|---|
| **Ia** | | | | SSO0691 | | | conserved hypothetical membrane protein, 7 TMD, N-terminal signal peptide | C |
| | | | | SSO0280 | COG1405K | | conserved hypothetical, potential transcription factor | C |
| | | | | SSO3146 | | | conserved hypothetical membrane protein, 7 TMD, N-terminal signal peptide | C |
| | | | | SSO1501 | | | conserved hypothetical | C |
| | | | *cdc6-2* | SSO0771 | COG1474L | 1 (1/3) | cell division control protein, same operon order in *S. acido.* and *S. tokodaii* | AE |
| | | | | c54_n0003_1 | | 2 (1/2) | hypothetical | |
| | | | *gspE-1* | SSO0120 | COG0630N | 3 (2/5) | putative ATPase of typeII/IV secretion system protein | AB |
| | | | | SSO0152 | COG0433R | | conserved hypothetical, (virD4 like ATPase) | S |
| | | | | SSO2338 | COG0477G | | conserved hypothetical transport protein, 5 TMD | A |
| | | | *dpo2* AT | SSO8124 | | 4 (1/3) | DNA polymerase B2 amino-end, same operon order in *S. acido.* and *S. tokodaii* | S |
| | | | | SSO0283 | COG0433R | | conserved hypothetical probable ATPase, N-terminal signal peptide | S |
| | | | | SSO0117 | | 3 (5/5) | conserved hypothetical, with pilin-like signal peptide | S |
| | | | | SSO0118 | | 3 (4/5) | conserved hypothetical, with pilin-like signal peptide | S |
| | | | | SSO1053 | | | conserved hypothetical, N-terminal signal peptide | S |
| | | | | SSO0037 | | | hypothetical, N-terminal signal peptide | |
| | | | | SSO2395 | | | conserved hypothetical | S |
| | | | | SSO1458 | | 4 (3/3) | conserved hypothetical | C |
| | | | *dpo2* CT | SSO1459 | COG0417L | 4 (2/3) | DNA polymerase B2 carboxy-end | A |
| | | | *insB* | SSO1436<br>SSO1654<br>SSO2320<br>SSO2950 | COG1662L | | Transposase and inactivated derivatives, IS1 family | S |
| **Ib** | | | | SSO0001 | COG1468L | | conserved hypothetical nuclease, RecB-family | A |
| | | | | SSO3177 | | | conserved hypothetical, N-terminal SP | S |
| | | | *herA* | SSO2251 | COG0433R | 5 (1/4) | 3´- 5´ssDNA helicase, same operon order in *S. acido. and S. tokodaii* | AB |
| | | | | c43_n0005 | | | hypothetical | |
| | | | *bcp-2* | SSO2121 | COG0450O | | peroxiredoxin | ABE |
| | | | | SSO0823 | | | conserved hypothetical | S |
| | | | *moaC* | SSO0770 | COG0315H | 1 (3/3) | molybdenum cofactor biosynthesis C | ABE |
| | | | | SSO1823 | | | conserved hypothetical | S |
| | | | | c44_n0017 | | | hypothetical | |
| | | | | bac03_n0001 | | | hypothetical | |
| | | | | lam04_n0008 | | 1 (2/3) | conserved hypothetical transcription regulator | C |
| | | | | bac18_n0006 | | | hypothetical | |
| | | | | SSO0121 | | 3 (1/5) | conserved hypothetical, same operon order in *S. acido.* and *S. tokodaii* | S |
| | | | | SSO0119 | COG2064N | 3 (3/5) | conserved hypothetical membrane protein, 9 TMD, N-terminal SP | S |
| **II** | | | | SSO0909 | COG0464O | 7 (3/3) | conserved archaeal AAA$^+$ ATPase | ABE |
| | | | | SSO2288 | COG0477G | | conserved hypothetical membrane protein, 9 TMD and N-terminal SP | S |
| | | | | SSO3242 | COG1321K | | conserved hypothetical transcriptional regulator | S |
| | | | | SSO2750 | | 8 (2/2) | conserved hypothetical (ATPase) ,same operon order in *S. acido.* and *S. tokodaii* | S |
| | | | | SSO2751 | | 8 (1/2) | conserved hypothetical kinase | S |
| | | | *cdc6-1* | SSO0257 | COG1474L | | cell division control 6/orc1 protein homolog | AE |
| | | | *soj* | SSO0034 | COG1192D | | conserved ATPases involved in chromosome partitioning | AB |
| | | | | SSO0910 | COG5491N | 7 (2/3) | conserved hypothetical, VPS24 domain | C |
| | | | | SSO0881 | COG5491N | | conserved hypothetical, VPS24 domain | C |
| | | | | SSO0858 | | 9 (2/2) | conserved hypothetical | C |
| | | | | SSO5826 | COG3609K | | conserved protein, potential transcription regulator | A |
| | | | | SSO6687 | | 9 (1/2) | conserved hypothetical, same operon order in *S. acido.* and *S. tokodaii* | S |
| | | | | SSO10704 | | | conserved hypothetical | A |
| | | | | SSO3207 | | | conserved hypothetical kinase | S |
| | | | | SSO0451 | COG5491N | | conserved hypothetical, VPS24 domain | C |
| | | | | SSO2200 | COG0433R | | conserved hypothetical probable ATPase | AB |
| | | | | SSO0911 | | 7 (1/3) | conserved hypothetical , same operon order in *S. acido. and S. tokodaii* | C |
| | | | | SSO3066 | COG1653G | | arabinose binding protein, ABC Transporter | AB |
| | | | *ssh7A* | SSO9180 | | | 7 kD DNA-binding protein (Sso7d-2) | S |
| | | | *ssh7A* | SSO10610 | | | 7 kD DNA-binding protein (Sso7d-1) | S |
| | | | *ssh7A* | SSO9535 | | | 7 kD DNA-binding protein (Sso7d-3) | S |
| | | | | SSO0271 | COG1222O | | 26S proteasome regulatory subunit | AE |

highly induced
induced
slightly induced
slightly repressed
repressed

Abbreviations: Ia, highly induced gene group; Ib induced gene group; II, repressed gene group
P, strain PH1; S, strain PH1(SSV1); TMD, transmembrane domain;

[a] UVdependent regulated gene groups were identified by a KMC-cluster analysis (Fig.S.5.2)

[b] genes without SSO Number represent intergenic small ORFs of less than 300 nucleotides

[c] number of operon (1-9), and in brackets: position in operon and total no. of genes in operon

[d] homologues (blastp e-value >$10^{-40}$).S, Sulfolobaceae; C, Crenarchaeota; A, Archaea; B, Bacteria; E, Eucarya

### 5.3.6 Differential reaction of the three *cdc6* genes in *Sulfolobus*

One of the most pronounced transcriptional reactions after UV-treatment was a rise in the mRNA level of the *cdc6-2* gene in both *Sulfolobus* strains and a down regulation of *cdc6-1*, while the *cdc6-3* gene remained essentially unaffected (Fig. 5.6).

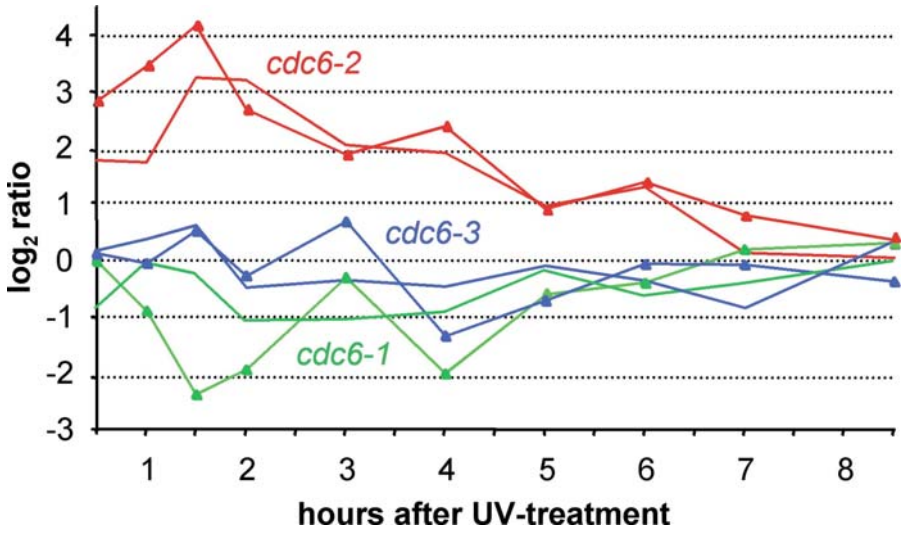



**Figure 5.6:** Expression profiles of the three *cdc6* genes, both strains PH1 (sleek lines) and PH1(SSV1) (lines with triangles) showed a strong upregulation of the potential repressor of replication *cdc6-2* (red), shortly after UV-treatment, while the potential main initiator of replication *cdc6-1* (green) is repressed. The data represent means of 2 to 3 experiments, but the display of standard deviations have been omitted for clarity.

The *cdc6-2* gene was found to be co-transcribed with two other strongly induced ORFs, one of which encoded a hypothetical transcription regulator that could play a role in UV-effected transcriptional responses and the second representing a *moa*C gene, an accessory protein for molybdenum cofactor biosynthesis (Hanzelmann & Schindelin, 2006) (see Fig. S.5.4 II in suppl. material).

### 5.3.7 UV-induced transcriptional response in *S. solfataricus* is limited to 55 genes

Beside the *cdc6-2* operon, we found 30 further up-regulated genes that react most strongly upon UV-treatment (groups Ia and Ib in table 5.2), 13 of which are organized in operons (see column 7 in table 5.2). Among these was a large group of genes encoding hypothetical membrane proteins or proteins with signal peptides, like the two strongly up-regulated genes *sso691* and *sso3146*. One operon encoded homologs of a putative type

II/IV secretion and/or type IV pilus system (Fig. 5.7), with an ATPase (SSO0120), a putative transmembrane protein (SSO0119), and two small proteins (SSO118 and SSO117) with a type IV pilin-like signal peptide (Albers & Driessen, 2005). SSO0120 which, when compared to Hidden Markov Models of the NCBI's COGs database (unpublished), was identified as an ATPase involved in archaeal flagellar biosynthesis, and also matches the central domain of the Flp pilus assembly protein. Also supported by the context of other observations, this operon is likely involved in the synthesis of conjugation pili (see discussion), instead of encoding a secretion system.

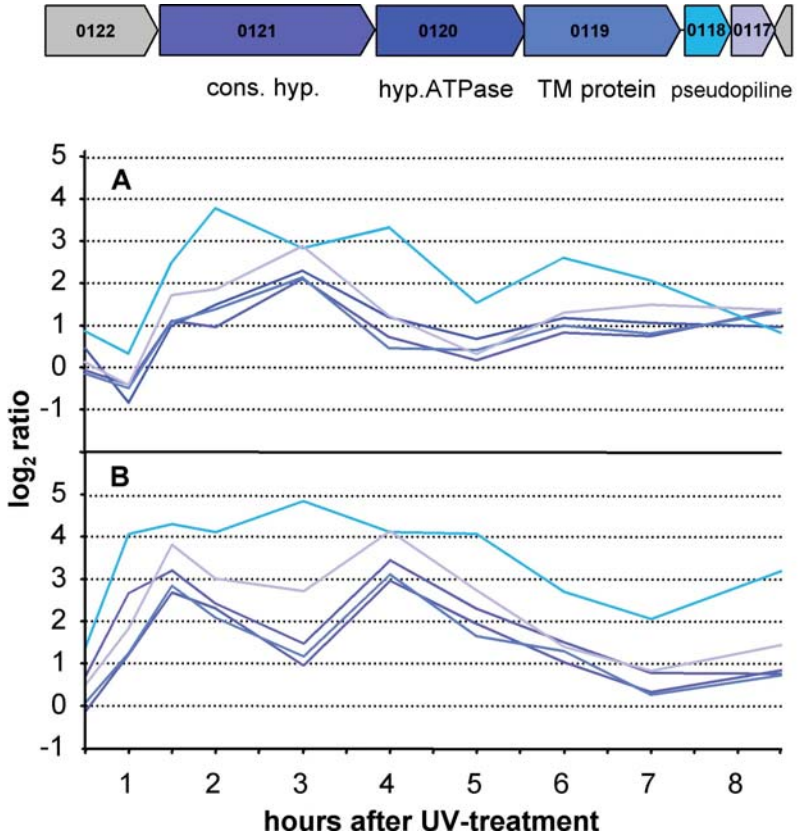



**Figure 5.7:** Expression profile of a strongly UV-induced operon encoding homologues of a putative pilus or secretion system (type II/IV). All five genes show a high induction in both strains PH1 (**A**) and PH1(SSV1) (**B**). The genes flanking the operon (SSO0122 and SSO5209) showed no effect to UV-light (not shown). Predicted gene functions are based on bioinformatics analysis (see also Table 5.2).

Other strongly induced genes encode potential transcription factors (*sso0280* and *Iam04_n0008*), which could be involved in regulating the UV-induced transcriptional response, and two genes encoding $AAA^+$ ATPases (*sso0152*, *sso0283*). Genes putatively involved in DNA-repair processes are discussed in a separate paragraph below.

In group Ia, which contains the highest induced genes, we found also the immediate early transcript T-ind of SSV1 in strain PH1(SSV1), which can be considered as a positive control in this dataset as its UV-dependence and transcriptional pattern has been well characterized (Fröls *et al.*, 2007; Reiter *et al.*, 1988).

Among the 22 prominent down-regulated genes (including 3 operons) we found those encoding potential regulators (*sso3242*, *sso5826*), kinases (*sso2751*, *sso3207*), transporters (*sso2288*, *sso3066*) and diverse ATPases (*sso0909*, *sso2750*, *sso2200*). Among the latter group is the Soj protein, which may be involved in chromosome segregation. Three genes, which can also be found in other crenarchaota (*sso0910*, *sso0881*, *sso0451*), have a conserved VPS24/SNF7 domain, which in eukaryotes is involved in the transport of cellular or transmembrane proteins between the endosomes and lysosomes for degradation events (Hayashi *et al.*, 2005). Of these, *sso0910* belongs to an operon with three genes (operon 7, Fig. S.5.4 I) that shows the strongest down regulation after UV-treatment (*sso0911*, *sso0910*, *sso0909*).

Interestingly, the three genes for SSO7d, encoding one of the two types of chromatin proteins in *Sulfolobus* (*sso10610*, *9535* and *9180*) are also down-regulated.

### 5.3.8 Proteins potentially involved in the repair of DNA-damage

All genes potentially involved in repair systems that have been specifically inspected with respect to their UV-response are listed in the supplementary material, Table S.5.3 Almost none of the genes supposedly involved in damage reversal, including photoreversal and base excision repair (BER), were found to be significantly induced upon UV-light. A few data points of genes from the NER system reached relative ratios above 2-fold (i.e. $\log_2>1$), but since they followed the temporal pattern of the general cell cycle-dependent response, we assume that this is not an indication for a specific transcriptional reaction to UV-damage, but rather reflects a basic and partly synchronized activity of highly transcribed genes. Genes of operon 5 (O5 in table 5.2) encoding the HerA, Mre11, Rad50 and NurA homologs of *Sulfolobus* showed a moderate (two to three-fold) UV-dependent transcriptional upregulation (Fig. 5.8).

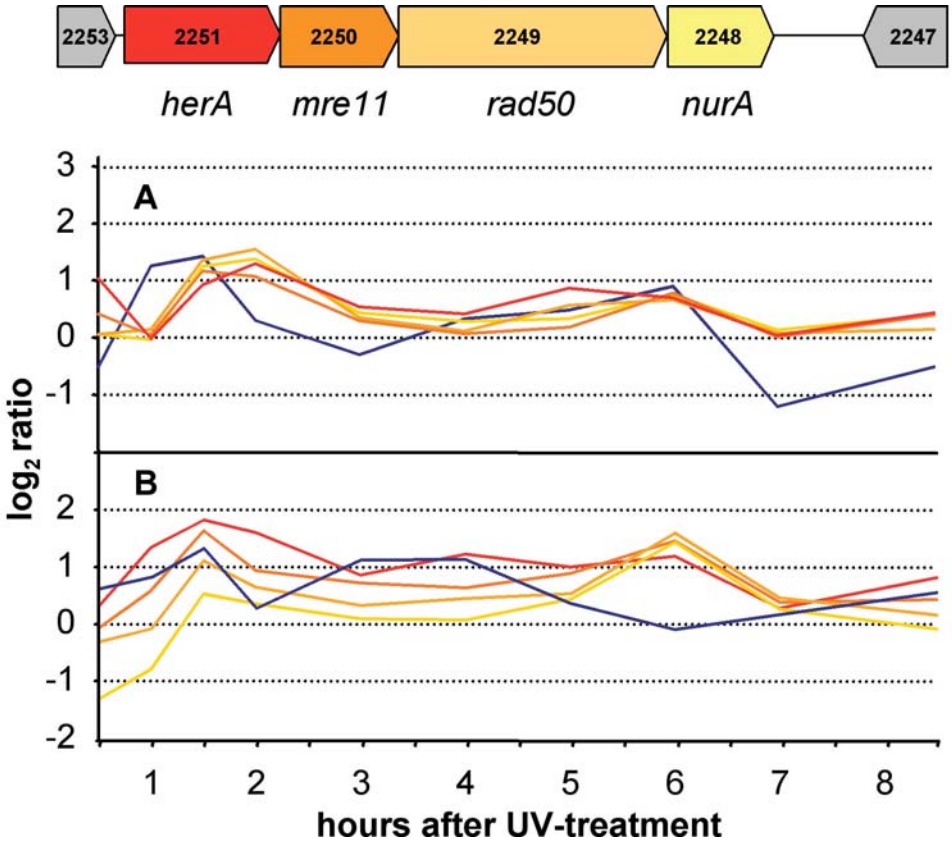



**Figure 5.8:** Expression profile of the archaeal *rad50/mre11* operon after UV-treatment. The transcriptional activity was firstly detected in *S. acidocaldarius*. *herA:* archaeal helicase, encodes a new class of bipolar DNA-helicases (Constantinesco *et al.*, 2004), *mre11*: ssDNA endonuclease and 3´ to 5´ ds DNA-exonuclease, *rad50:* ATPase and *nurA*: nuclease of archaea, a 5´ to 3´ exonuclease. The four genes, which are supposedly involved in homologous recombination as part of the putative Recombination Repair System (RER-System) show a weak but significant UV-dependent response, while *radA* (SSO0250, blue curve) follows the pattern of highly transcribed genes (see. Fig. 5.5 blue line). (**A**) strain PH1 and (**B**) strain PH1(SSV1).

These factors are supposed to be involved in recombinational repair (Constantinesco *et al.*, 2004; Hopfner *et al.*, 2002). Other genes involved in the recombination processes, like the *rad*A gene (*sso0250*), the Holliday junction resolvases (*sso0575*, *sso1176*) or integrase (*sso0375*), however, did not show a UV-dependent expression pattern, (see the blue line in Fig. 5.8 for *rad*A). We found an induction of a RecB-like nuclease (*sso0001*) that could play a role in homologous recombination and recombinational DNA-repair. Like bcp-2 (*sso2121*), it may react due to oxidative stress damage (Stohl & Seifert, 2006).

Among the three identified type B polymerases of *S. solfataricus*, only polymerase II was found to be significantly induced (Fig. S.5.4 III). This B-type polymerase is encoded by three co-transcribed genes (*sso1459*, *sso1458*, *sso8124*), that are proposed to generate a

full length DNA-polymerase by programmed frameshifting (Genbank record AAK41686), as B-family type polymerases usually contain only one polypeptide chain. However, the same triple gene arrangement is found in the genomes of *Sulfolobus tokodaii* and *Sulfolobus acidocaldarius*. From its transcriptional pattern we propose that this polymerase should be involved in DNA-repair/replication after UV-damage. A reaction of the translesion repair polymerase (*dpo4*) is only seen in the uninfected culture, where a significant upregulation from 4 h to 5 h is observed.

## 5.4 Discussion

When discussing the data of this study in the light of other investigations, it is important to note that the UV-dose of 75 J/m$^2$ we applied was lower than that used in an earlier study of *Sulfolobus solfataricus* and many other bacteria (200 J/m$^2$) (Salerno *et al.*, 2003), but similar to that used by McCready *et al.*, 2005 (30–70 J/m$^2$) for genome-wide transcription studies in *Halobacterium sp.* NRC-I. Our survival rates of 10 to 40% for the *S. solfataricus* strain devoid of SSV1 and ~10% for the lysogenic strain are comparable to those found for *S. acidocaldarius* under similar conditions (Schmidt *et al.*, 1999) but are considerably lower than those of *Halobacterium sp.* NRC-I, that showed over 80% survival after exposure to 70 J/m$^2$ (McCready *et al.*, 2005).

The fact that we used two strains (a lysogen and a wild-type non-lysogen), that reacted with almost identical gene sets on the transcriptional level but in a slightly time shifted manner, proved helpful for dissecting the UV-dependent response in *Sulfolobus*. In particular, it helped to distinguish UV-dependent genes from a large number of highly transcribed, but not UV-induced genes, that appeared significant at certain time points, possibly because of a slight alteration of the cell cycle distribution of the population after UV-treatment (see Fig. 5.5 blue dotted line). Another aid in dissecting UV-dependent genes was the relatively long time period over which we did the transcriptional analysis (8-12 hours) and the relatively large number of sampling points.

From the experiments presented here, we can discern three outcomes from UV-light-induced damage in *Sulfolobus*:

### 5.4.1 Growth inhibition, cell death and cell cycle perturbation

From ca. 1.5 to 2 hours after UV-treatment, the optical densitiy of the UV-damaged *S. solfataricus* culture decreased over 5 hours reflecting growth retardation of damaged cells and/or the effect of cell death (Fig. 5.1). We also noted a modest accumulation of cells with a single copy of the chromosome (Fig. 5.3). This presumably reflects the accumulation of cells that fail to progress into the S phase of the cell cycle. It is possible that this could represent a checkpoint-like response in *Sulfolobus*. However, in light of the persistence of this minority population over the 8.5 hour time course and the high levels of mortality caused by UV-treatment, it is perhaps more likely that this reflects an increased sensitivity of G1 and early S-phase cells to UV-induced damage.

With regard to the lack of strong cell cycle responses, it is interesting that we observe clear modulation of the levels of the Cdc6 transcripts. Previous work has revealed that treatment of *S. acidocaldarius* with acetic acid perturbed the cell cycle leading to an accumulation of cells in the G2 phase of the cell cycle (Robinson *et al.*, 2004). This was associated with the presence of high levels of Cdc6-2 and almost undetectable levels of Cdc6-1 and Cdc6-3 proteins. Comparing our dataset with that of (Lundgren & Bernander, 2007) who have investigated cell-cycle dependent transcriptional responses, we find more indications for cell cycle disturbance after UV-treatment, since 18 out of our 22 genes that are significantly down-regulated have been classified as being transcribed and up-regulated in a cell-cycle dependent manner ("cyclic") by Lundgren *et al.* (2007). Nevertheless, the UV-dependent transcriptional response is clearly distinguishable from the genes identified by Lundgren *et al.* (2007), as only 4 of the up-regulated genes are also found to be cell-cycle-dependent (*cdc6-2*, *dpo2*, *sso0152*, *sso1823*).

Perhaps the most obvious effect observed in the flow cytometry is the appearance of cells with greater than 2N contents at late timepoints. This could either be due to additional rounds of replication occurring inappropriately in G2 cells, to lacking cell division after mitose or, as discussed below, due to uptake of DNA from other cells during conjugation processes.

### 5.4.2 Formation of double-strand breaks and induction of the recombinational repair system

Similar to the results of the transcriptome studies in halobacteria (Baliga *et al.*, 2004; McCready *et al.*, 2005), we did not find any indication of a concerted UV-dependent regulation, as would be expected in an SOS-like response, nor did we find a significant induction of the repair genes which are involved in direct DNA-damage removal, such as for example photolyase or components of the putative nucleotide excision repair system. Most probably, these systems are constantly present for instantaneous reaction to DNA-damage and therefore do not react dramatically on a transcriptional level. Alternatively, some factors might be posttranslationally modified for activation and therefore would not appear in a transcriptome analysis. While Salerno *et al.* (2003) described some induction for the NER system (after a UV-dose of 200J/$m^2$), these genes followed the pattern of constitutive, but highly transcribed genes in our study. While *rad*A was induced in *Halobacterium* after a low UV-dose, we did not see a significant induction in *Sulfolobus*, confirming earlier results for this organism (Sandler *et al.*, 1996). Interestingly, we saw a relatively weak, but UV-dependent response of genes from the putative recombinational repair system of *Sulfolobus*, the Mre11 operon, which can also be involved in the repair of double-strand breaks. These results inspired us to analyse the occurrence of double-strand breaks in *Sulfolobus* upon UV-treatment. We observed DSBs between 2 to 8 hours after UV-treatment and hypothesize that the unrepaired fraction of CPDs, which were observed from 0 h to 2 h after UV-treatment (not shown) led to the formation of double-strand-breaks, which are then processed for recombinational repair, which involves factors of the MRE11 complex (Hopfner *et al.*, 2002; Jazayeri *et al.*, 2006).

Furthermore, we observed a considerable accumulation of cells in the S-phase after 8.5 hours (flow cytometry data, Fig. 5.3 and S.5.1) which might represent the fraction of the culture that resume replication after the double-strand breaks have been repaired (see Fig. 5.4). As early as 30 years ago, the occurrence of double-strand breaks upon UV-treatment had been demonstrated for *E. coli* (Bonura & Smith, 1975). Recently, it was shown using mouse skin cells that unrepaired CPDs provoke an accumulation of single and double-strand breaks (SSBs and DSBs) during DNA-replication, which represents a major cause of UV-mediated cytotoxicity (Garinis *et al.*, 2005). Furthermore, CPDs, rather than other DNA-lesions or damaged macromolecules, represented the principal mediator of the cellular transcription response to UV (Garinis *et al.*, 2005). The most prominent

repair pathways that were induced by CPDs were associated with DNA double-strand break signalling and repair, including also Mre11a and Rad50, the two eukaryotic homologues of the genes found in the Mre11 operon of *Sulfolobus* (Garinis *et al.*, 2005).

### 5.4.3 Formation of cell-to-cell contacts – an indication for conjugation?

The strong induction of a typeII/IV system of secretion or pili formation (Table 5.2), with genes that potentially encode type IV pilin-like signal peptides (Albers & Driessen, 2005) inspired us to microscopically investigate if cell aggregates indicative of conjugation were formed. We have reproducibly observed a considerable UV-induced formation of aggregates (Fig. 5.2) as well as the formation of pili (not shown). At least 90% of the cells were found in aggregates, particularly between 3 and 6 hours after the treatment (not all of them are included in the quantitative statistics of Fig. 5.2, because of the huge size of the aggregates). The cell clumps resemble those observed in plasmid-mediated conjugation (Schleper *et al.*, 1995). This finding strongly supports earlier observations of an enhanced exchange of genetic markers upon UV-treatment (Hansen *et al.*, 2005; Reilly and Grogan, 2001; Schmidt *et al.*, 1999; Wood *et al.*, 1997). Wood *et al.* performed experiments with strains that were mutated in the *pyr*D or *pyr*B genes, which are involved in *de novo* uracil formation. The exposure of 70 J/m$^2$ seemed to yield the highest rate of exchange (Wood *et al.*, 1997), which is similar to the UVdose under which we found the most aggregates (75 J/m$^2$).

In light of these observations, we propose that beside well-known DNA repair mechanisms, *Sulfolobus* might use conjugational DNA exchange and subsequent homologous recombination to repair its DNA, since *Sulfolobus* (Hansen *et al.*, 2005; Schleper *et al.*, 1995), like Halobacteria (Papke *et al.*, 2004; Rosenshine *et al.*, 1989), seems to be quite active in conjugaton.

### 5.4.4 Conclusion

In conclusion, our results demonstrate, concurring with other studies, that DNA damage triggers a complex set of events in the cells, and that these events involve many different biological processes. Beside direct damage removal, homologous recombination - and perhaps conjugation - might play a considerable role in this network. The majority of genes that were most prominently down- or up-regulated in this study are of unknown

function and only have orthologues in other *Sulfolobales* or in crenarchaeota, indicating that the UV-dependent and other stress-related responses and repair mechanisms in Archaea are highly diverse and very little understood.

## 5.5 Materials and Methods

*Growth of Sulfolobus strains and UV-treatment*

*S. solfataricus* strains PH1 (Schleper *et al.*, 1994) and PH1(SSV1) (Martusewitsch *et al.*, 2000) were grown at 78°C and pH 3 in Brock´s medium (Grogan, 1989), with 0.1% (w/v) tryptone and 0.2% (w/v) D-arabinose under moderate agitation (ca. 150 rpm in NewBrunswick shaker). The optical density of liquid cultures was monitored at 600 nm. For surviving rate and UV-dose determination, solid media were prepared by adding gelrite to a final concentration of 0.6% and $Mg^{2+}$ and $Ca^{2+}$ to 0.3 and 0.1 M, respectively. Plates were incubated for five days at 78°C. For UV-treatment, aliquots of 50 ml ($OD_{600nm}$ 0.3-0.5) were transferred to a plastic container (20 cm x 10 cm x 4 cm) and irradiated with UV light for 45 sec at 245 nm (W20, Min UVis, DEGESA), while shaking the culture carefully. The treated cultures were stored in the dark at RT for 15 min and were re-incubated at 78°C.

*Surviving rate of S. solfataricus after different UV-doses*

The plating efficiency (cfu/ml) of *S. solfataricus* cells after treatment with different UV-doses (UV-C 254 nm) was determined using a logarithmically grown culture (Brock´s basal salt medium with 0.1% (w/v) tryptone and 0.2% (w/v) D-arabinose). For the UV-treatment, 10 ml of culture was transferred to a 110 mm plastic petri dish and treated with a defined UV-dose of 200 $J/m^2$, 150 $J/m^2$, 100 $J/m^2$, 75 $J/m^2$, 50 $J/m^2$ or 25 $J/m^2$, respectively (λ 254 nm, UV-Stratalinker 1800, Stratagene). Additionally, two independent treatments were performed with the UV lamp (W20, Min UVis, DEGESA) for 45 sec at 254 nm. All treatments were performed under red light. For the control culture, exactly the same procedure was followed under red light (incubation 45 sec) without UV-treatment. The treated cultures were stored in the dark at RT for 15 min and were subsequently plated (dilutions of $10^{-4}$, $10^{-5}$ and $10^{-6}$). Plates were incubated for five days at 78°C.

Numbers of colony forming units per ml were determined and the surviving rate in % was calculated.

*Microscopy and analysis of cell aggregate formation*

Cell aggregates were analysed in a phase-contrast microscope (Axioskop, Zeiss) 1000-fold magnified. To fix the cells, microscopic slides were coated with solid media. One ml of 2-fold concentrated Brock salt solution (Grogan, 1989) with 2% $MgCl_2$ and pH 3, without carbon sources was preheated to 78°C and mixed with 1 ml of melted 1.3% gelrite (Merck and Co., Kelco Division) in distilled 78°C water. 500 µl of the solution was immediately poured on a microscopic slide and a cover slip was added. After about 1 min, the cover slip was removed. Then 5 µl pure *S. solfataricus* culture or 1:2 diluted with 1 x Brock salt solution, pH 3, 78°C was added and a cover slip was placed on top before microscopy was performed. To quantify the formation of aggregates, the amount of cells in aggregates were counted until 1000 or 500 single cells were observed, respectively. For the statistic analysis the percentage of cells in aggregates against the single cells was calculated from three independent experiments. The images were digitalized with a microscope-coupled device camera (Power Shot G6, Canon) connected to a computerized image analysis system (Remote Capture, Canon). To analyse the cell vitality the LIVE DEAD Baclight ™ (Invitrogen) assay was used.

*FACS analysis*

Cells were fixed by addition to ice-cold 80% ethanol (70% ethanol final concentration). For staining, cells were washed two times with 10 mM TrisHCl (pH 7.4), 10 mM $MgCl_2$, and then resuspended in this buffer containing an additional 20 $\mu g.ml^{-1}$ propidium iodide and 100 $\mu g \cdot ml^{-1}$ RNAseA. Flow Cytometric analysis was carried out using a MoFlo high-speed cell sorter (Dakto-Cytomation) as described previously (Robinson *et al.*, 2007).

*Analysis of chromosomal DNA in pulse field gel electrophoresis (PFGE)*

To identify double-strand breaks (DSB) in DNA, cells were embedded in agarose plugs before cells lysis was performed. 15 ml of the cell culture was harvested for approximately three plugs. Cells were washed twice with 10 ml TEN solution (50 mM Tris-HCl, pH 8, 50 mM EDTA, pH 8, 100 mM NaCl) and centrifuged at 4000 rpm for 10 min at 4°C, between

each washing step. Washed cells were finally resuspended in 100 µl TEN-solution. The cell-solution was briefly warmed to 37°C, and then immediately mixed with 100 µl of 0.8% low melting point agarose (pre-cooled to 37°C). 85 µl of the agarose cell solution was dispensed into wells of a disposable plug mold (Bio-Rad, cat. No. 1703706) and incubated for 15 min at 4°C. Congealed plugs were transferred to a 2 ml tube containing 2 ml of NDS solution (0.5 M EDTA, 0.12% Tris-HCL, 0.55 M NaOH), pH 9.0 and 1 mg/ml proteinase K and were incubated over night at 37°C. The liquid was replaced by 2 ml NDS, pH 8.0 with 1 mg/ml proteinase K added and incubated again over night at 37°C. Plugs were then washed with 2 ml NDS, pH 8.0 and stored in 2 ml of NDS, pH 8.0 at 4°C. For PFGE we used a CHEF-DR® III Pulse Field Electrophoresis System (Bio-Rad) with a permanent circulation of 0.5 x TBE buffer (Tris-Borat-EDTA). Electrophoresis was run for 20 h at 6 V/cm (and increasing pulse time from 5 to 50 sec) at 14°C. The PFGE gel was stained for 10 min with Sybr Green (1:100000 in 1 x TBE, Fluka) and the image was created with Phospho-Imager (FLA-5000, Fujifilm).

*RNA preparation and analysis*

Total RNA was extracted using a standard procedure (Chirgwin *et al.*, 1979). RNA quality was determined by agarose gel electrophoresis and by determination of the ratio of absorption at 260 nm and 280 nm. Only RNA samples with a ratio between 2.1 and 1.9 were used for further experiments.

*Microarray Design and Fabrication*

Each microarray consisted of 3456 70-mer oligonucleotides spotted onto glass slides. The design and fabrication methodologies for the microarrays were the same as those described in detail in another recent study (Fröls *et al.*, 2007).

*cDNA labeling and microarray hybridization*

Labeling of cDNA and microarray hybridizations with cyanine-3 or cyanine-5 (Cy-3/Cy-5) fluorescent molecules were performed as described recently (Fröls *et al.*, 2007). Each slide hybridization experiment was repeated as a dye-swap, and each time point was analysed by combining results of hybridizations from 4 independent UV-experiments. This resulted in a total of 6 to 12 data points for each gene at each time point, as the basis for

the quantitative and statistical analysis. In total, 62 successful hybridizations were performed in order to obtain the 8.5 h time-series.

*Microarray data analysis*

Qualitative and statistical analyses of the data were performed as recently described (Fröls *et al.*, 2007). A K-means clustering analysis was performed using the TIGR Multi Expression Viewer (Saeed *et al.*, 2003) integrated into the program Bluejay (Turinsky *et al.*, 2005) (see Fig. S.5.2 in the suppl. material). The K-Means were calculated from the genomic microarray data set of the virus-infected PH1(SSV1) and the non-infected strain PH1. The genes were sorted into 40 clusters with a maximum of 50 iterations. To identify the main genomic UV-answer we focused our analysis on genes, which showed a significant change in the expression rate over the time course. The KMC-analysis identified the same genes, as we identified by manual analysis for both tested strains. Beside the KMC analysis we surveyed the genomic microarray data for specific gene groups of main cellular events, such as cell cycle, replication, transcription, translation, and repair based on COG categories (Clusters of Orthologous Groups of proteins, http://www.ncbi.nlm.nih.gov). The PH1 and PH1(SSV1) dataset will be downloadable from http://bluejay.ucalgary.ca/sulfolobus.

## 5.6 Supplementary data

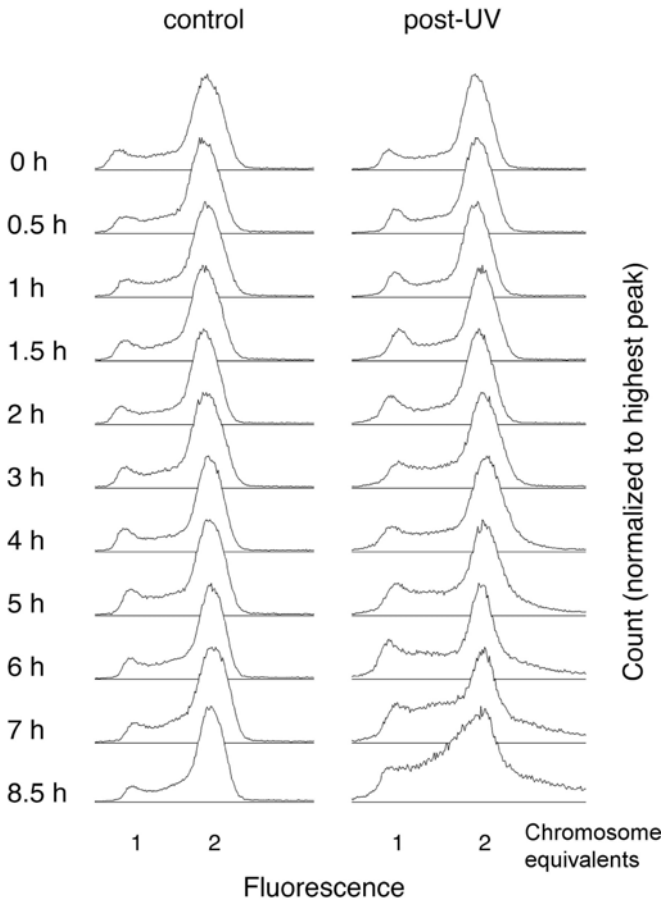



**Figure S.5.1:** Flow cytometry analysis of *S. solfataricus* cells after UV-treatment, strain PH1(SSV1). Cells were fixed in 80% ice-cold ethanol and the DNA contents were measured by fluorescence (same analysis as shown in Fig. 4 for strain PH1).

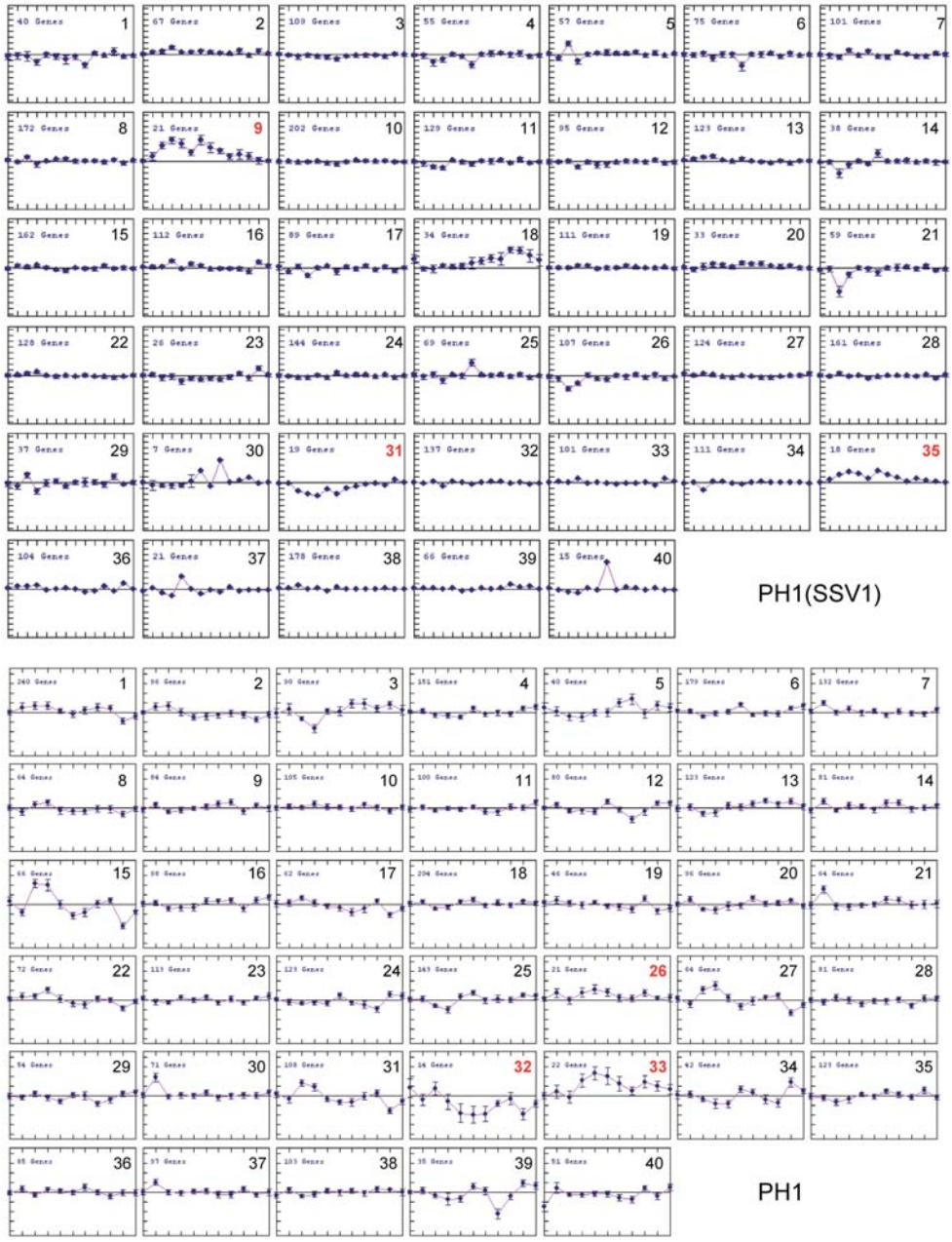



**Figure S.5.2:** Graphical output of the KMC-cluster analysis, using the TIGR Multi Expression Viewer with the program Bluejay (Turinsky *et al.*, 2005). The K-Means were calculated from the genomic microarray data set of strains PH1(SSV1) and PH1. The genes were sorted into 40 clusters with a maximum iteration of 50. To identify the main genomic UV-answer we focused our analysis on genes exhibiting significant changes in the expression rate over the time course. The clusters representing the UV-dependent gene groups as shown in Table 5.2 and Fig. S.5.3, are highlighted red. In the case of PH1(SSV1) we found 21 genes (group Ia, cluster 9) with a strong significant up-regulation from 0.5 h to 7 h after UV-treatment. A second cluster of 18 significantly up-regulated genes (group Ib, cluster 35) from 0.5 h to 6 h was identified and a third cluster of 19 significantly down regulated genes (group II, cluster 31) from 0.5 h to 5 h. In the case of PH1, one cluster with 22 highly up-regulated genes (group Ia, cluster 33) from 1 h to 5 h was identified, a second cluster of 21 moderately induced genes (group Ib, cluster 26), following the same expression profile, and a third cluster of 14 genes (group II, cluster 32) with an average of 4-fold down regulation (from 1 h to 6 h).

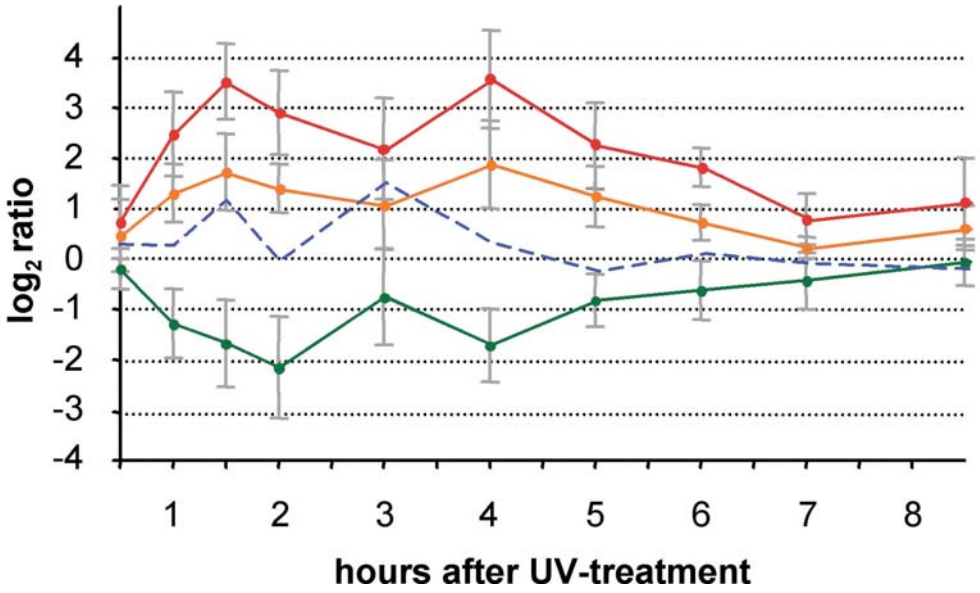



**Figure S.5.3:** Expression profiles of the general transcriptional response after UV-treatment of the SSV1-infected strain PH1(SSV1). The curves display the mean average of the three identified UV-dependent regulated gene groups as displayed in Table 2, the highly induced group of 19 genes (red curve), induced gene group of 14 genes (orange) and the down-regulated group of 22 genes (green). The blue dashed curve was generated from the averages of 11 genes, but represents qualitatively the pattern of approximately 400 genes that are mostly involved in the transcription and translation processes. Errors bars represent the standard deviation of the displayed means average.

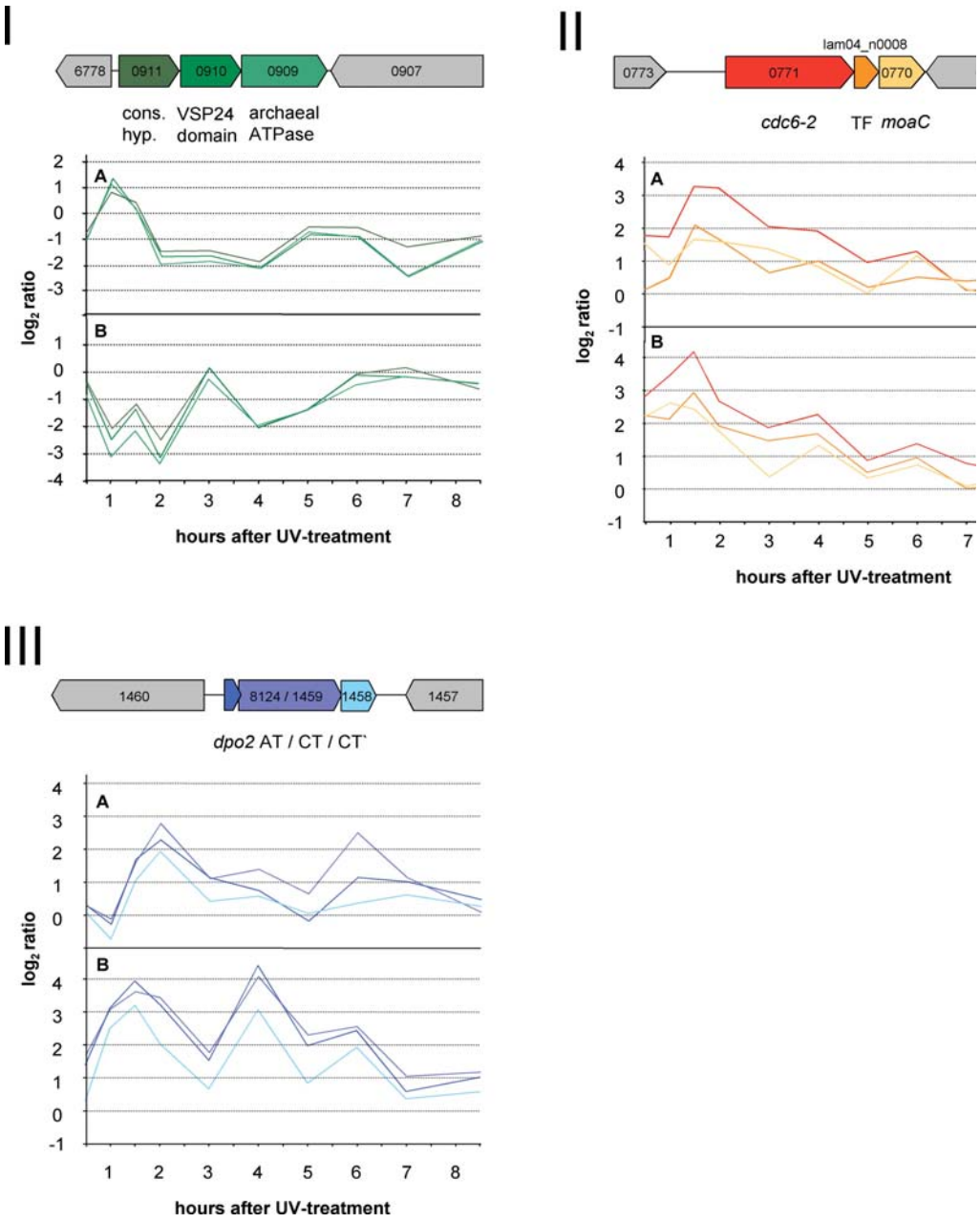



**Figure S.5.4:** Expression profiles of three strongly regulated UV-dependent operons, in strains PH1 (**A**) and PH1(SSV1) (**B**).The predicted gene functions were identified by bioinformatics analysis. (**I**) Operon with three most strongly and quickly down-regulated genes in both tested strains. *Sso0910* exhibits a VSP24 domain and *sso0909* a putative ATPase with a conserved archaeal domain. The same operon was found in *S. tokodaii* and *S. acidocaldarius.* (**II**) The first gene encodes the putative replication repressor *cdc6-2*. The second gene, *lam04_n0008*, encodes a small protein of 77 amino acids, with similarities to transcriptional factors. The third gene (*moaC*) encodes a putative cofactor for molybdenum biosynthesis. The same gene order was only found in *S.acidocaldarius*. (**III**) The operon of DNA-polymerase 2 (*dpo2*) of *S. solfataricus. sso1458* shows highest similarities to proteins of unknown function or proteins involved in transcriptional processes (RNA helicase, DNA dep. RNA-polymerase) The small gene *sso8124* encodes a protein homologous to the amino terminal, *sso1459* to the carboxy terminal of Dpo2. The same gene order, including the frame shift between *sso8124* and *sso1459*, was also found in *S. tokodaii* and *S. acidocaldarius.*

**Table S.5.1:** Corresponding ratios of the UV-specific regulated genes of *S. solfataricus* PH1(SSV1)

| **UV-specific regulated genes of PH1(SSV1) after UV-treatment *** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| gene ID | | $log_2$ ratios at time in hours after UV-treatment | | | | | | | | | |
| NCBI | magpie | 0.5 | 1 | 1.5 | 2 | 3 | 4 | 5 | 6 | 7 | 8.5 |
| **up regulated genes** | | | | | | | | | | | |
| SSO0691 | c10_039 | 1.23 | 1.46 | 4.42 | 4.19 | 3.75 | 5.17 | 3.25 | 1.88 | 1.11 | 2.69 |
| SSO0280 | bac13_072 | 1.41 | 3.46 | 3.83 | 3.12 | 1.15 | 3.83 | 1.76 | 1.45 | 0.49 | 0.47 |
| SSO3146 | c38_012 | 0.40 | 2.42 | 2.98 | 2.91 | 1.71 | 2.91 | 2.05 | 1.62 | 0.75 | 0.52 |
| SSO1501 | bac15_040 | 1.18 | 2.21 | 3.00 | 2.49 | 3.24 | 3.25 | 1.63 | 1.70 | 1.35 | 1.05 |
| SSO0771 | lam04_020 | 2.82 | 3.46 | 4.17 | 2.68 | 1.88 | 2.38 | 0.90 | 1.39 | 0.80 | 0.35 |
| | c54_n0003_1 | 0.08 | 2.23 | 2.77 | 2.15 | 1.82 | 4.23 | 2.50 | 1.84 | 0.68 | 1.25 |
| SSO0120 | c04_002 | 1.35 | 4.06 | 4.29 | 4.10 | 4.83 | 4.10 | 4.06 | 2.69 | 2.04 | 3.18 |
| SSO0152 | bac25_060 | 0.37 | 2.81 | 3.86 | 2.27 | 1.95 | 4.36 | 1.83 | 1.33 | 0.49 | 0.58 |
| SSO2338 | bac19_049 | 0.48 | 1.83 | 4.02 | 3.56 | 3.04 | 3.57 | 3.06 | 2.06 | 0.42 | 1.52 |
| SSO8124 | bac15_n0021_1 | 1.70 | 3.10 | 3.62 | 3.44 | 1.77 | 4.08 | 2.30 | 2.56 | 1.05 | 1.17 |
| SSO0283 | bac13_074 | 0.46 | 2.06 | 3.56 | 1.90 | 0.78 | 3.43 | 1.77 | 1.27 | 0.01 | 0.30 |
| SSO0117 | c04_005 | 0.50 | 1.85 | 3.82 | 3.01 | 2.71 | 4.13 | 2.73 | 1.39 | 0.83 | 1.43 |
| SSO0118 | c04_004 | 1.35 | 4.06 | 4.29 | 4.1 | 4.83 | 4.1 | 4.06 | 2.69 | 2.04 | 3.18 |
| SSO1053 | bac16_043 | 0.77 | 2.85 | 3.79 | 2.38 | 1.74 | 4.01 | 1.86 | 1.73 | 0.15 | 0.22 |
| SSO0037 | bac10_041 | 0.21 | 2.92 | 4.59 | 3.29 | 2.73 | 4.03 | 2.61 | 1.58 | 1.08 | 1.39 |
| SSO2395 | c32_011 | 0.30 | 2.58 | 3.45 | 4.41 | 1.99 | 1.14 | 3.06 | 1.98 | 0.63 | 1.45 |
| SSO1458 | bac15_083 | 0.32 | 2.54 | 3.22 | 2.07 | 0.69 | 3.08 | 0.86 | 1.95 | 0.38 | 0.59 |
| SSO1459 | bac15_082 | 1.39 | 3.13 | 3.94 | 3.25 | 1.53 | 4.41 | 1.98 | 2.44 | 0.58 | 1.02 |
| SSO1436<br>SSO1654<br>SSO2320<br>SSO2950 | bac30_007_1 | 0.44 | 2.17 | 2.34 | 2.42 | 2.49 | 3.49 | 1.76 | 1.90 | 0.64 | 0.45 |
| SSV1_027 | SSV1_027 | 0.71 | 2.68 | 3.50 | 4.21 | 3.54 | 4.79 | 3.73 | 2.08 | 2.13 | 2.96 |
| SSO0001 | bac10_077 | 0.00 | 1.33 | 2.15 | 1.11 | 0.01 | 1.50 | 0.74 | 0.42 | 0.20 | 0.15 |
| SSO1460 | bac15_081 | 0.28 | 1.59 | 1.01 | 1.31 | 0.01 | 1.22 | 0.7 | 0.76 | 0.11 | 0.22 |
| SSO3177 | bac09_016 | 0.45 | 1.35 | 1.53 | 1.30 | 1.74 | 1.67 | 1.48 | 0.57 | 0.45 | 1.02 |
| SSO2251 | bac26_054 | 0.31 | 1.35 | 1.82 | 1.60 | 0.86 | 1.22 | 1.00 | 1.18 | 0.27 | 0.81 |
| | c43_n0005 | 0.76 | 1.11 | 0.70 | 1.04 | -0.23 | 1.49 | 0.89 | 0.52 | 0.53 | 0.90 |
| SSO2121 | c02_016 | 0.85 | 1.54 | 1.64 | 0.93 | 1.68 | 1.54 | 1.14 | 0.15 | -0.23 | 1.54 |
| | bac18_n0002 | 0.42 | 0.19 | 1.31 | 1.72 | 0.53 | 2.00 | 1.88 | 0.84 | 0.14 | 0.61 |
| SSO0823 | c39_033 | -0.31 | 0.99 | 1.22 | 1.54 | 1.56 | 1.73 | 1.83 | 0.83 | 0.59 | 0.69 |
| SSO0770 | lam04_019 | 2.20 | 2.60 | 2.41 | 1.73 | 0.83 | 1.32 | 0.34 | 0.72 | 0.08 | 0.60 |
| SSO1823 | bac07_019 | 1.14 | 1.80 | 2.42 | 1.75 | 2.10 | 1.62 | 1.53 | 0.07 | 0.34 | 1.00 |
| | c44_n0017 | 0.32 | 1.79 | 2.05 | 1.31 | 0.54 | 2.75 | 1.47 | 1.04 | 0.15 | 0.36 |
| | bac03_n0001 | 0.18 | 0.41 | 2.45 | 1.43 | 2.68 | 2.53 | 2.14 | 1.06 | 0.56 | 1.03 |
| | lam04_n0008 | 2.24 | 2.14 | 2.94 | 1.92 | 1.48 | 1.69 | 0.58 | 0.97 | 0.05 | 0.11 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | bac18_n0006 | 0.15 | 1.04 | 1.32 | 1.73 | 1.85 | 3.69 | 1.93 | 1.08 | 0.18 | 0.54 |
| | bac18_n0001 | -0.24 | 1.61 | 1.63 | 1.47 | 0.43 | 3.67 | 1.72 | 1.15 | 0.23 | 0.90 |
| SSO0121 | bac02_046 | -0.14 | 1.22 | 2.68 | 2.29 | 0.96 | 2.96 | 1.93 | 1.02 | 0.32 | 0.84 |
| SSO0119 | c04_003 | 0.07 | 1.25 | 2.83 | 2.08 | 1.17 | 3.10 | 1.65 | 1.29 | 0.26 | 0.73 |
| **down regulated genes** | | | | | | | | | | | |
| SSO0909 | bac21_079 | -0.35 | -2.04 | -1.13 | -2.45 | 0.20 | -2.01 | -1.35 | -0.02 | 0.21 | -0.57 |
| SSO2288 | bac29_048 | -0.13 | -0.92 | -1.79 | -1.45 | 0.17 | -1.52 | -0.47 | -0.24 | -0.37 | -0.01 |
| SSO3242 | bac28_014 | 0.08 | -0.98 | -2.49 | -2.91 | -0.98 | -2.03 | -0.91 | -1.51 | -1.19 | -0.20 |
| SSO2750 | c97_038 | -0.43 | -1.76 | -1.86 | -1.74 | -0.27 | -1.80 | 0.02 | 0.22 | 0.06 | -0.88 |
| SSO2751 | c97_037 | -0.25 | -2.07 | -1.02 | -2.04 | -0.53 | -2.21 | -0.52 | 0.06 | 0.15 | -0.68 |
| SSO0257 | bac13_048 | -0.01 | -0.90 | -2.38 | -1.94 | -0.30 | -2.01 | -0.61 | -0.40 | -0.19 | 0.31 |
| SSO0034 | bac10_044 | -0.41 | -0.84 | -1.79 | -1.48 | -1.16 | -1.76 | -0.61 | -0.24 | -0.12 | -0.66 |
| | c50_020 | -1.33 | -1.33 | -1.38 | -0.60 | -2.68 | -3.73 | -0.95 | 0.00 | -1.18 | 0.34 |
| SSO0910 | bac21_081 | -0.52 | -2.57 | -1.45 | -3.25 | 0.08 | -2.11 | -1.47 | -0.18 | -0.25 | -0.49 |
| SSO0881 | bac21_052 | -0.30 | -2.56 | -3.88 | -5.19 | -3.08 | -2.00 | -2.01 | -1.74 | -1.37 | -0.01 |
| SSO0858 | bac21_029 | 0.15 | -1.00 | -1.36 | -1.83 | -0.86 | -2.04 | -0.92 | -0.69 | -0.08 | -0.13 |
| | c41_n0006 | 0.28 | -1.32 | -2.88 | -2.99 | -0.36 | -1.69 | -1.14 | -1.57 | -0.94 | 0.22 |
| | bac21_n0010 | 0.12 | -0.85 | -1.65 | -1.76 | -0.89 | -2.00 | -0.76 | -0.73 | -0.14 | -0.37 |
| | bac08_n0006 | 0.64 | -0.75 | -1.75 | -2.69 | -2.10 | -2.16 | -1.40 | -1.15 | -0.66 | 0.23 |
| SSO3207 | bac09_046 | -0.33 | -1.35 | -2.17 | -1.28 | -0.91 | -1.60 | -0.22 | -0.28 | -0.34 | 0.12 |
| SSO0451 | c43_014 | 0.29 | -0.95 | -3.07 | -4.05 | -2.52 | -2.80 | -1.98 | -1.96 | -1.39 | 0.17 |
| SSO2200 | bac26_003 | -0.84 | -1.53 | -1.17 | -1.41 | -0.47 | -0.99 | -0.45 | -0.06 | 0.49 | -0.23 |
| SSO0911 | bac21_082 | -0.76 | -3.05 | -2.09 | -3.30 | -0.17 | -1.88 | -1.32 | -0.38 | -0.08 | -0.35 |
| SSO3066 | bac23_051 | -0.26 | -0.34 | 0.19 | -1.30 | 0.05 | -0.58 | -0.25 | -0.33 | 0.40 | 0.37 |
| SSO9180 | c44_n0016 | 0.00 | -0.66 | -0.81 | -1.69 | -0.91 | -0.39 | -0.29 | -0.85 | -0.60 | 0.91 |
| SSO10610 | c58_n0004_1 | 0.05 | -0.66 | -0.68 | -1.90 | -0.79 | -0.45 | -0.38 | -0.89 | -0.41 | 0.39 |
| SSO9535 | c54_n0002_1 | -0.10 | -0.76 | -0.99 | -1.84 | -0.66 | -0.69 | -0.42 | -1.14 | -1.48 | 1.01 |
| SSO0271 | bac13_062 | -0.28 | -0.71 | -0.59 | -1.50 | 0.22 | -1.30 | -0.64 | -0.47 | -0.70 | -0.39 |

*estimate $\log_2$ ratios of the UV-specific regulated genes in PH1(SSV1) after UV-treatment (correspond to table 5.2)
Figure S.5.3 display the mean average of the identified UV-specific regulated gene groups.

**Table S.5.2:** Corresponding ratios of the UV-specific regulated genes of *S. solfataricus* PH1

| **UV-specific regulated genes of PH1 after UV-treatment *** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| gene ID | | $log_2$ ratios at time in hours after UV-treatment | | | | | | | | | |
| NCBI | magpie | 0.5 | 1 | 1.5 | 2 | 3 | 4 | 5 | 6 | 7 | 8.5 |
| **up regulated genes** | | | | | | | | | | | |
| SSO0691 | c10_039 | 2.54 | -0.20 | 1.91 | 3.70 | 2.87 | 3.11 | 0.14 | 2.89 | 1.91 | 0.77 |
| SSO0280 | bac13_072 | 0.19 | -0.02 | 2.03 | 2.19 | 1.85 | 0.38 | 0.28 | 1.30 | 0.95 | 0.11 |
| SSO3146 | c38_012 | 0.30 | -0.57 | 1.08 | 2.34 | 0.19 | 1.54 | 0.70 | 0.80 | 0.83 | 0.02 |
| SSO1501 | bac15_040 | 0.08 | 1.27 | 2.67 | 2.27 | 1.46 | 1.15 | 0.77 | 2.25 | 0.00 | 0.47 |
| SSO0771 | lam04_020 | 1.77 | 1.73 | 3.25 | 3.20 | 2.04 | 1.91 | 0.96 | 1.29 | 0.12 | 0.04 |
| | c54_n0003_1 | 0.08 | -0.57 | 0.43 | 2.24 | 1.62 | 0.77 | 0.09 | 0.41 | 0.90 | 0.20 |
| SSO0120 | c04_002 | 0.44 | -0.84 | 1.00 | 1.48 | 2.29 | 1.19 | 0.67 | 1.16 | 1.05 | 0.95 |
| SSO0152 | bac25_060 | 0.51 | -0.37 | 1.64 | 2.10 | 1.90 | 1.65 | 0.14 | 1.80 | 0.90 | 0.78 |
| SSO2338 | bac19_049 | 0.12 | -0.26 | 0.93 | 2.75 | 2.69 | 0.81 | -0.06 | 1.56 | 1.11 | 1.49 |
| SSO8124 | bac15_n0021_1 | 0.30 | -0.11 | 1.62 | 2.79 | 1.12 | 1.40 | 0.66 | 2.51 | 1.15 | 0.10 |
| SSO0283 | bac13_074 | -0.07 | -0.17 | 1.20 | 1.80 | 1.94 | 0.65 | 0.80 | 1.21 | 0.61 | 0.81 |
| SSO0117 | c04_005 | 0.13 | -0.42 | 1.71 | 1.84 | 2.88 | 1.22 | 0.32 | 1.29 | 1.48 | 1.35 |
| SSO0118 | c04_004 | 0.86 | 0.33 | 2.49 | 3.78 | 2.83 | 3.33 | 1.53 | 2.60 | 2.05 | 0.83 |
| SSO1053 | bac16_043 | 0.15 | -0.59 | 1.90 | 2.62 | 3.63 | 0.77 | 0.64 | 1.19 | 0.58 | 0.00 |
| SSO0037 | bac10_041 | 1.06 | -0.36 | 2.30 | 2.31 | 3.02 | 2.04 | 0.42 | 1.73 | 1.27 | 0.85 |
| SSO2395 | c32_011 | 0.41 | -0.42 | 0.98 | 2.58 | 2.40 | 1.33 | 0.80 | 0.96 | 1.24 | 1.42 |
| SSO1458 | bac15_083 | 0.07 | -0.74 | 1.05 | 1.92 | 0.41 | 0.57 | 0.05 | 0.35 | 0.61 | 0.26 |
| SSO1459 | bac15_082 | 0.31 | -0.26 | 1.72 | 2.29 | 1.14 | 0.75 | -0.16 | 1.15 | 1.02 | 0.48 |
| SSO1436<br>SSO1654<br>SSO2320<br>SSO2950 | bac30_007_1 | -0.17 | -0.51 | 0.85 | 1.95 | 0.94 | 0.80 | 0.33 | 1.47 | 1.24 | -0.06 |
| SSV1_027 | SSV1_027 | 1.69 | -0.51 | -0.10 | -0.10 | 0.13 | 0.41 | 0.89 | 0.33 | 0.94 | 0.96 |
| SSO0001 | bac10_077 | 0.58 | -0.15 | 1.11 | 0.51 | 1.11 | 0.50 | -0.22 | 0.69 | 0.36 | 0.50 |
| SSO1460 | bac15_081 | -0.62 | -0.64 | 0.33 | 0.56 | -0.17 | 0.43 | -0.21 | 0.07 | 0.94 | -0.36 |
| SSO3177 | bac09_016 | 0.98 | 0.33 | -0.10 | 0.56 | 1.17 | 0.68 | 0.24 | 1.27 | 0.37 | -0.70 |
| SSO2251 | bac26_054 | 1.03 | -0.02 | 0.93 | 1.30 | 0.54 | 0.41 | 0.87 | 0.69 | 0.03 | 0.43 |
| | c43_n0005 | 1.25 | -0.28 | 0.50 | 0.81 | 1.22 | -0.24 | -0.20 | 1.00 | 0.02 | 0.59 |
| SSO2121 | c02_016 | 0.54 | 0.77 | 0.50 | 0.73 | 0.54 | 0.75 | 1.05 | 1.22 | -0.56 | -0.68 |
| | bac18_n0002 | 0.17 | -0.57 | 0.24 | 0.47 | 0.88 | 0.51 | 0.00 | 0.05 | 0.99 | 0.48 |
| SSO0823 | c39_033 | 0.20 | -0.53 | 0.54 | 0.60 | 1.21 | -0.01 | -0.28 | 0.54 | 0.58 | 1.08 |
| SSO0770 | lam04_019 | 1.52 | 0.87 | 1.65 | 1.59 | 1.36 | 0.84 | 0.02 | 1.15 | 0.15 | -0.42 |
| SSO1823 | bac07_019 | 1.10 | 0.54 | 0.74 | 1.77 | 1.58 | 0.38 | -0.17 | 1.37 | -0.04 | 0.28 |
| | c44_n0017 | 0.71 | -0.61 | 0.61 | 1.30 | 0.91 | 0.15 | 0.23 | -0.02 | 0.28 | 0.01 |
| | bac03_n0001 | 0.09 | -0.28 | 0.82 | 1.34 | 1.08 | -0.28 | -0.40 | 0.69 | 0.24 | 0.59 |
| | lam04_n0008 | 0.14 | 0.50 | 2.08 | 1.64 | 0.64 | 1.00 | 0.21 | 0.51 | 0.39 | 0.59 |
| | bac18_n0006 | 0.58 | -0.58 | 0.09 | 1.36 | 0.85 | -0.17 | 0.01 | 0.79 | 0.55 | -0.04 |
| | bac18_n0001 | 0.03 | -0.67 | 0.07 | 1.24 | 0.74 | 0.75 | 0.11 | 0.32 | 1.06 | -0.08 |
| SSO0121 | bac02_046 | -0.08 | -0.41 | 1.10 | 0.95 | 2.10 | 0.71 | 0.17 | 0.82 | 0.73 | 1.38 |
| SSO0119 | c04_003 | -0.14 | -0.49 | 1.09 | 1.38 | 2.13 | 0.45 | 0.41 | 0.98 | 0.80 | 1.30 |
| **down regulated genes** | | | | | | | | | | | |
| SSO0909 | bac21_079 | -0.79 | 0.86 | 0.45 | -1.56 | -1.53 | -1.98 | -0.56 | -0.58 | -1.37 | -0.92 |
| SSO2288 | bac29_048 | 0.04 | 0.33 | -0.94 | -1.03 | -0.02 | -0.32 | -0.10 | -0.80 | 0.19 | 0.72 |
| SSO3242 | bac28_014 | -0.61 | 0.55 | -0.80 | -1.74 | -2.19 | -2.42 | -0.96 | -0.19 | -1.95 | -0.56 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| SSO2750 | c97_038 | 0.70 | 0.14 | -0.09 | -0.67 | -0.18 | 0.09 | -0.06 | -0.92 | -0.77 | 0.04 |
| SSO2751 | c97_037 | 0.21 | 0.34 | 0.06 | -0.42 | -0.29 | -0.34 | -0.04 | -0.27 | -1.47 | -0.26 |
| SSO0257 | bac13_048 | -0.84 | -0.03 | -0.23 | -1.07 | -1.05 | -0.92 | -0.18 | -0.63 | -0.41 | -0.01 |
| SSO0034 | bac10_044 | 0.79 | -0.37 | -0.04 | -0.83 | -0.78 | -0.38 | -0.45 | -0.99 | -0.33 | -0.08 |
| | c50_020 | 0.80 | 0.18 | -0.55 | 0.00 | 0.00 | 0.94 | 0.36 | -0.57 | 1.20 | 0.23 |
| SSO0910 | bac21_081 | -0.98 | 1.28 | 0.30 | -2.00 | -1.88 | -2.15 | -0.68 | -0.88 | -2.47 | -1.01 |
| SSO0881 | bac21_052 | -1.04 | 1.73 | -2.16 | -4.55 | -4.31 | -3.79 | -1.49 | -1.15 | -3.38 | -1.40 |
| SSO0858 | bac21_029 | 0.31 | 0.53 | -0.57 | -1.27 | -0.83 | -0.74 | -0.09 | -0.45 | -0.94 | -0.40 |
| | c41_n0006 | -0.02 | -0.16 | -1.64 | -2.05 | -1.33 | -1.18 | -1.15 | -0.13 | -1.79 | -0.28 |
| | bac21_n0010 | 0.04 | 0.15 | -0.50 | -1.40 | -0.73 | -1.11 | 0.02 | -0.32 | -1.01 | -0.37 |
| | bac08_n0006 | 1.20 | 0.77 | -0.90 | -1.92 | -2.04 | -1.59 | -0.90 | -0.45 | -1.46 | -0.40 |
| SSO3207 | bac09_046 | 0.05 | -0.04 | -0.57 | -1.21 | -0.08 | -0.23 | -1.47 | -0.69 | -0.13 | 0.55 |
| SSO0451 | c43_014 | -0.66 | 1.22 | -1.23 | -2.56 | -3.24 | -2.37 | -0.74 | -0.66 | -2.39 | -1.33 |
| SSO2200 | bac26_003 | -0.49 | 0.10 | 0.30 | -0.73 | -0.96 | -1.05 | 0.00 | -0.49 | -1.28 | -0.16 |
| SSO0911 | bac21_082 | -0.94 | 1.46 | 0.13 | -1.73 | -1.70 | -2.19 | -0.82 | -0.90 | -2.53 | -1.09 |
| SSO3066 | bac23_051 | -0.60 | 1.57 | -0.84 | -1.34 | -1.23 | -0.45 | -0.78 | -0.61 | -1.49 | -0.54 |
| SSO9180 | c44_n0016 | -0.86 | -0.10 | -0.38 | -0.89 | -1.50 | -1.55 | -0.65 | 0.65 | -1.41 | -0.86 |
| SSO10610 | c58_n0004_1 | -0.85 | 0.03 | -0.34 | -1.02 | -1.65 | -1.87 | -0.70 | 0.59 | -1.89 | -0.67 |
| SSO9535 | c54_n0002_1 | -0.50 | -0.20 | -0.71 | -1.01 | -1.47 | -1.23 | -0.49 | 0.71 | -1.43 | -0.72 |
| SSO0271 | bac13_062 | -0.62 | 0.44 | 0.09 | -1.01 | -1.58 | -1.94 | -0.81 | -0.24 | -1.67 | -0.65 |

*estimate $\log_2$ ratios of the UV-specific regulated genes in PH1 after UV-treatment (correspond to table 5.2)
Figure 5.5 display the mean average of the identified UV-specific regulated gene groups.

**Table S.5.3:** Genes potentially involved in DNA-repair systems of *S. solfataricus*

**Genes of *S. solfataricus* of putative DNA-repair systems**

| DNA-Repair System | Gene ID | | COG | Predicted function | Comments | Predominant found in | Reference |
|---|---|---|---|---|---|---|---|
| **PR - photoreversal** | *phrB* | SSO2472 | COG0415L | DNA photolyase | | *bacteria eukarya* | Grogan, 2000; Wood *et al.*,1997 |
| **DR - damage reversal** | *ogt* | SSO2487 | COG0350L | Methylated DNA protein, cysteine methyltransferase | paralog of *ada* | *bacteria* | Aravind *et al.*,1999 |
| | *dcD* | SSO0190 | COG0717F | Deoxycytidine triphosphate deaminase | paralog to *dut*, universal in archaea | *bacteria* | Aravind *et al.*, 1999 |
| | *dcD-2* | SSO2954 | COG0717F | Deoxycytidine triphosphate deaminase | | *bacteria* | Aravind *et al.*, 1999 |
| **BER - base excision** | *ogG* | SSO0904 | COG1059L | N-glycosylase/ DNA lyase | putative mutY familiy of a glycosidases/ endonucleases (Nth and AlkA are paralogs) | *bacteria* | Aravind *et al.*, 1999 |
| | *ntH-1* | SSO0116 | COG0177L | DNA endonuclease III | | *bacteria* | Aravind *et al.*, 1999 |
| | *ntH-2* | SSO2484 | COG2231L | DNA endonuclease III | | *bacteria* | Aravind *et al.*, 1999 |
| | *nfi* | SSO2454 | COG1515L | DNA endonuclease V | | *bacteria* | Aravind *et al.*, 1999 |
| **NER – nucleotide excision** | *ssXPF* | SSO0729 | COG1948L | DNA endonucleas | orthologues to XPF/RAD1; in *Sulfolobus* only C-terminal nuclease domian | *eukarya* | Roberts *et al.*, 2003 |
| | *ssXPG* | SSO0179 | COG0258L | DNA endonuclease | orthologues to XPG/RAD2 | *eukarya* | Salerno *et al.*, 2003 |
| | *ssXpb-I* | SSO0959 | COG1061KL | DNA helicase | orthologues to XPB/RAD25 | *eukarya* | Salerno *et al.*, 2003 |
| | *ssXpb-II* | SSO0473 | | | | | |
| | *ssXPD* | SSO0313 | COG1199KL | DNA helicase | orthologues to XPD/RAD3 | *eukarya* | Salerno *et al.*, 2003 |
| **TR - trans lesion** | *dinP* (*dpo4*) | SSO2448 | COG0389L | DNA polymerase IV | DinB-like DNA polymerase with lesion-bypass properties | *bacteria* | Wang *et al.*, 2006 |
| **RER - recombinational** | *radA* | SSO0250 | COG0468L | RecA/RadA recombinase | archaea and eucarya contain in addition an N-terminal HhH-like domain | *bacteria eukarya* | Sandler *et al.*, 1996 Seitz *et al.*, 1998; |
| | *nurA* | SSO2248 | COG1630S | DNA exonuclease | | *archaea* | Constantinesco *et al.*, 2002 |
| | *rad50* | SSO2249 | COG0419L | Chromatin modifying ATPase | | *eukarya* | Constantinesco *et al.*, 2002 |
| | *mre11* | SSO2250 | COG0420L | DNA nuclease | 3´- 5´ exo- and endonucelase; as a complex with Rad50, involved non-homologous joining of DNA ends | *eukarya* | Constantinesco *et al.*, 2002 |
| | *herA* | SSO2251 | COG0433R | bipolar DNA helicase | | *archaea* | Constantinesco *et al.*, 2002 |

## 5.7 References


Albers, S.V. and Driessen, A.J. (2005) Analysis of ATPases of putative secretion operons in the thermoacidophilic archaeon Sulfolobus solfataricus. *Microbiology*, 151, 763-773.

Aravind, L., Walker, D.R. and Koonin, E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res*, 27, 1223-1242.

Baliga, N.S., Bjork, S.J., Bonneau, R., Pan, M., Iloanusi, C., Kottemann, M.C., Hood, L. and DiRuggiero, J. (2004) Systems level insights into the stress response to UV radiation in the halophilic archaeon Halobacterium NRC-1. *Genome Res*, 14, 1025-1035.

Bonura, T. and Smith, K.C. (1975) Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of Escherichia coli K-12. *J Bacteriol*, 121, 511-517.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, 18, 5294-5299.

Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L. and Elie, C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Res*, 32, 1439-1447.

Crowley, D.J., Boubriak, I., Berquist, B.R., Clark, M., Richard, E., Sullivan, L., Dassarma, S. and McCready, S. (2006) The uvrA, uvrB and uvrC genes are required for repair of ultraviolet light induced DNA photoproducts in Halobacterium sp. NRC-1. *Saline Systems*, 2, 11.

Dorazi, R., Gotz, D., Munro, S., Bernander, R. and White, M.F. (2007) Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus. *Mol Microbiol*, 63, 521-529.

Fröls, S., Gordon, P.M., Panlilio, M.A., Schleper, C. and Sensen, C.W. (2007) Elucidating the transcription cycle of the UV-inducible hyperthermophilic archaeal virus SSV1 by DNA microarrays. *Virology* 365:48-59.

Garinis, G.A., Mitchell, J.R., Moorhouse, M.J., Hanada, K., de Waard, H., Vandeputte, D., Jans, J., Brand, K., Smid, M., van der Spek, P.J., Hoeijmakers, J.H., Kanaar, R. and van der Horst, G.T. (2005) Transcriptome analysis reveals cyclobutane pyrimidine dimers as a major source of UV-induced DNA breaks. *Embo J*, 24, 3952-3962.

Grogan, D.W. (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. *J Bacteriol*, 171, 6710-6719.

Grogan, D.W. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol*, 8, 180-185.

Grogan, D.W. (2004) Stability and repair of DNA in hyperthermophilic Archaea. *Curr Issues Mol Biol*, 6, 137-144.

Grogan, D.W., Carver, G.T. and Drake, J.W. (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius. *Proc Natl Acad Sci U S A*, 98, 7928-7933.

Haber, J.E. (2006) Chromosome breakage and repair. *Genetics*, 173, 1181-1185.

Hans-Peter Klenk, R.A.G. (2006) *Archaea: Evolution, Physiology and Molecular Biology*. Blackwell Publishing Professional.

Hansen, J.E., Dill, A.C. and Grogan, D.W. (2005) Conjugational genetic exchange in the hyperthermophilic archaeon Sulfolobus acidocaldarius: intragenic recombination with minimal dependence on marker separation. *J Bacteriol*, 187, 805-809.

Hanzelmann, P. and Schindelin, H. (2006) Binding of 5'-GTP to the C-terminal FeS cluster of the radical S-adenosylmethionine enzyme MoaA provides insights into its mechanism. *Proc Natl Acad Sci U S A*, 103, 6829-6834.

Hayashi, M., Fukuzawa, T., Sorimachi, H. and Maeda, T. (2005) Constitutive activation of the pH-responsive Rim101 pathway in yeast mutants defective in late steps of the MVB/ESCRT pathway. *Mol Cell Biol*, 25, 9478-9490.

Hjort, K. and Bernander, R. (1999) Changes in cell size and DNA content in Sulfolobus cultures during dilution and temperature shift experiments. *J Bacteriol*, 181, 5669-5675.

Hopfner, K.P., Putnam, C.D. and Tainer, J.A. (2002) DNA double-strand break repair from head to tail. *Curr Opin Struct Biol*, 12, 115-122.

Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G.C., Lukas, J. and Jackson, S.P. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol*, 8, 37-45.

Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, 8, 669-676.

Kulaeva, O.I., Koonin, E.V., McDonald, J.P., Randall, S.K., Rabinovich, N., Connaughton, J.F., Levine, A.S. and Woodgate, R. (1996) Identification of a DinB/UmuC homolog in the archeon Sulfolobus solfataricus. *Mutat Res*, 357, 245-253.

Lundgren, M. and Bernander, R. (2007) Genome-wide transcription map of an archaeal cell cycle. *Proc Natl Acad Sci U S A*, 104, 2939-2944.

Makarova, K.S., Aravind, L., Grishin, N.V., Rogozin, I.B. and Koonin, E.V. (2002) A DNA repair system specific for thermophilic Archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Res*, 30, 482-496.

Martin, A., Yeats, S., Janekovic, D., Reiter, W.D., Aicher, W. and Zillig, W. (1984) SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium Sulfolobus acidocaldarius isolate B12. *Embo J*, 3, 2165-2168.

Martusewitsch, E., Sensen, C.W. and Schleper, C. (2000) High spontaneous mutation rate in the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by transposable elements. *J Bacteriol*, 182, 2574-2581.

McCready, S. and Marcello, L. (2003) Repair of UV damage in Halobacterium salinarum. *Biochem Soc Trans*, 31, 694-698.

McCready, S., Muller, J.A., Boubriak, I., Berquist, B.R., Ng, W.L. and Dassarma, S. (2005) UV irradiation induces homologous recombination genes in the model archaeon, Halobacterium sp. NRC-1. *Saline Systems*, 1, 3.

Papke, R.T., Koenig, J.E., Rodriguez-Valera, F. and Doolittle, W.F. (2004) Frequent recombination in a saltern population of Halorubrum. *Science*, 306, 1928-1929.

Raychaudhuri, S., Karmakar, P., Choudhary, D., Sarma, A. and Thakur, A.R. (2003) Effect of heavy ion irradiation on DNA DSB repair in Methanosarcina barkeri. *Anaerobe*, 9, 15-21.

Reilly, M.S. and Grogan, D.W. (2001) Characterization of intragenic recombination in a hyperthermophilic archaeon via conjugational DNA exchange. *J Bacteriol*, 183, 2943-2946.

Reilly, M.S. and Grogan, D.W. (2002) Biological effects of DNA damage in the hyperthermophilic archaeon Sulfolobus acidocaldarius. *FEMS Microbiol Lett*, 208, 29-34.

Reiter, W.D., Palm, P. and Zillig, W. (1988) Analysis of transcription in the archaebacterium Sulfolobus indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res*, 16, 1-19.

Robb, F.T., Maeder, D.L., Brown, J.R., DiRuggiero, J., Stump, M.D., Yeh, R.K., Weiss, R.B. and Dunn, D.M. (2001) Genomic sequence of hyperthermophile, Pyrococcus furiosus: implications for physiology and enzymology. *Methods Enzymol*, 330, 134-157.

Robinson, N.P., Blood, K.A., McCallum, S.A., Edwards, P.A. and Bell, S.D. (2007) Sister chromatid junctions in the hyperthermophilic archaeon Sulfolobus solfataricus. *Embo J*, 26, 816-824.

Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004) Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. *Cell*, 116, 25-38.

Romano, V., Napoli, A., Salerno, V., Valenti, A., Rossi, M. and Ciaramella, M. (2006) Lack of Strand-specific Repair of UV-induced DNA Lesions in Three Genes of the Archaeon Sulfolobus solfataricus. *J Mol Biol*.

Rosenshine, I., Tchelet, R. and Mevarech, M. (1989) The mechanism of DNA transfer in the mating system of an archaebacterium. *Science*, 245, 1387-1389.

Saeed, A.I., Sharov, V., White, J., Li, J., Liang, W., Bhagabati, N., Braisted, J., Klapa, M., Currier, T., Thiagarajan, M., Sturn, A., Snuffin, M., Rezantsev, A., Popov, D., Ryltsov, A., Kostukovich, E., Borisovsky, I., Liu, Z., Vinsavich, A., Trush, V. and Quackenbush, J. (2003) TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*, 34, 374-378.

Salerno, V., Napoli, A., White, M.F., Rossi, M. and Ciaramella, M. (2003) Transcriptional response to DNA damage in the archaeon Sulfolobus solfataricus. *Nucleic Acids Res*, 31, 6127-6138.

Sandler, S.J., Satin, L.H., Samra, H.S. and Clark, A.J. (1996) recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast Saccharomyces cerevisiae. *Nucleic Acids Res*, 24, 2125-2132.

Schleper, C., Holz, I., Janekovic, D., Murphy, J. and Zillig, W. (1995) A multicopy plasmid of the extremely thermophilic archaeon Sulfolobus effects its transfer to recipients by mating. *J Bacteriol*, 177, 4417-4426.

Schleper, C., Kubo, K. and Zillig, W. (1992) The particle SSV1 from the extremely thermophilic archaeon Sulfolobus is a virus: demonstration of infectivity and of transfection with viral DNA. *Proc Natl Acad Sci U S A*, 89, 7645-7649.

Schleper, C., Roder, R., Singer, T. and Zillig, W. (1994) An insertion element of the extremely thermophilic archaeon Sulfolobus solfataricus transposes into the endogenous beta-galactosidase gene. *Mol Gen Genet*, 243, 91-96.

Schmidt, K.J., Beck, K.E. and Grogan, D.W. (1999) UV stimulation of chromosomal marker exchange in Sulfolobus acidocaldarius: implications for DNA repair, conjugation and homologous recombination at extremely high temperatures. *Genetics*, 152, 1407-1415.

Stetter, K. (2006) History of discovery of the first hyperthermophiles. *Extremophiles*, 10, 357-362.

Stohl, E.A. and Seifert, H.S. (2006) Neisseria gonorrhoeae DNA recombination and repair enzymes protect against oxidative damage caused by hydrogen peroxide. *J Bacteriol*, 188, 7645-7651.

Turinsky, A.L., Ah-Seng, A.C., Gordon, P.M., Stromer, J.N., Taschuk, M.L., Xu, E.W. and Sensen, C.W. (2005) Bioinformatics visualization and integration with open standards: the Bluejay genomic browser. *In Silico Biol*, 5, 187-198.

Williams, E., Lowe, T.M., Savas, J. and Diruggiero, J. (2006) Microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus exposed to gamma irradiation. *Extremophiles*.

Wood, E.R., Ghane, F. and Grogan, D.W. (1997) Genetic responses of the thermophilic archaeon Sulfolobus acidocaldarius to short-wavelength UV light. *J Bacteriol*, 179, 5693-5698.

**Acknowledgement**

We thank Christiane Elie and Achim Quaiser for helpful comments on the function of the Mre11 operon and other information processing genes as well as Sonja Albers for sharing information on the SSO0121 operon.

This project was sponsored through the German ministery, BMBF, Metagenomics cluster, grant 4.1 of the Göttingen GenoMics network and through funding of the Centre of excellence for GeoBiology of the University of Bergen. The creation of the microarrays was funded through the Canada Foundation for Innovation (CFI) and the Alberta Science and Research Authority (ASRA). The creation of the software for the analysis of the gene chip experiments was in part funded through contributions from Genome Canada and Genome Alberta. SDB and IGD are funded by the Medical Research Council, UK.

All software created by the team is available from the authors upon request.

Chapter 6

# UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation

Manuscript submitted

April 2008

# UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation

**Sabrina Fröls** , Behnam Zolghadr , Michaela Wagner , Magorzata Ajon , Mihaela Folea , Egbert J. Boekema , Arnold J.M. Driessen , Christa Schleper and Sonja-Verena Albers

## 6.1 Abstract

The hyperthermophilic archaeon *Sulfolobus* solfataricus exhibits a complex transcriptional response to UV-irradiation involving 55 genes (Fröls *et al.*, 2007). Among the most strongly UV-induced genes a putative pili biogenesis operon had been observed. Here we provide the first detailed description of pili in the domain Archaea encoded by an operon of six genes, including a potential secretion ATPase, two prepilins, a putative transmembrane protein and a protein of unknown function. The pili formation was inducible by UV-light (254 nm). Electron microscopy and image reconstruction showed that the pili were straight, variable in length, they were not bundled or polarised and they were composed of three evenly spaced helices with 100 Å in diameter, thereby clearly being distinguishable from the archaeal flagella. Both prepilin proteins possess a class III signal sequence and the cleavage of the SSO0118 prepilin by an archaeal typeIV prepilin peptidase was demonstrated *in vitro*. A deletion mutant, replacing the central typeII/IV secretion ATPase (SSO0120), verified that the pili were encoded by the UV-dependently induced operon (*sso0117* through *sso0121*) named *ups*-operon (**U**V-inducible **p**ili operon of **S**ulfolobus) and that they mediated the cellular aggregation. We showed further that the cellular aggregation was a UV-dose dependent, dynamic process, not inducible by other stressors like temperature or pH, but stimulated by chemically induced double strand breaks in DNA.

We also showed that UV-irradiation strongly increases the conjugative activity of *S. solfataricus*. In the context of further transcriptional data we conclude that pili-mediated cellular aggregation probably mediates enhanced conjugation, which eventually leads to an enhanced repair of UV-damaged DNA in *S. solfataricus via* homologous recombination.

## 6.2 Introduction

The ability of Bacteria and Archaea to form multicellular structures is observed in a variety of biological systems. This fascinating phenomenon of a collective behaviour can be manifested in the formation of biofilms from mixed microbial mats, cellular aggregates or microcolonies. Multicellular structures represent an essential strategy for adaptation to changing environmental conditions or even survival (Shapiro, 1998; Davey & O´Toole, 2000; Battin *et al.*, 2007). Cells organised in biofilm-like structures show a higher resistance to toxic compounds, as for example antimicrobials (Patel, 2005) or to physical stress, like shifts in temperature or pH, or exposure to UV-light (Ojanen *et al.*, 1997; Roine *et al.*, 1998; Elasri & Miller, 1999; Martinez & Casadevall, 2007). In addition, microorganisms benefit from the attachment on substrates like e.g. suspended particles, which provides a higher nutrient availability (Davey & O´Toole, 2000). Also genetic transfer, i.e. DNA exchanges *via* conjugation plays an important role in biofilms to disseminate niche genes of metabolic pathways (Gasson & Davies, 1980; Molin & Tolker-Nielsen, 2003). The rate of conjugative DNA-exchange in biofilm structures is enhanced and conjugative pili stabilise the biofilm structure (Gasson & Davies, 1980; Ghigo, 2001; Molin & Tolker-Nielsen, 2003; Reisner *et al.*, 2006)

Cellular aggregation is mainly reported for organisms of the domain Bacteria, while comparably few but quite diverse examples have been found in the domain of the Archaea. A complex biofilm-structure of a marine hydrothermal vent system was formed by methanogenic archaea of the order *Methanococcales* and by *Thermococcales* and *Archaeaoglobales* (Schrenk *et al.*, 2003). Species of the order *Desulfosarcina* and *Desulfococcus* generate synergistic communities with sulphate-reducing bacteria in microcolonies (Boetius *et al.*, 2000). An unusual microbial community organised in string-of-pearls was found in cold sulphurous water. It is formed by the euryarchaeon SM1 that grows in close association with the bacterium *Thiotrix sp.* and forms complex and unusual cellular appendages (hami) (Moissl *et al.*, 2003, 2005). Single strain cultures of the hyperthermophilic euryarchaeote *Archaeoglobus fulgidus*form a protein-, metal- and polysaccharide- containing heterogeneous biofilm, which is inducible by environmental stressors like UV-light (LaPaliga & Hartzell, 1997). *Pyrococcus furiosus* can form surface attached microcolony structures mediated by multifunctional flagella, which can also form cable-like structures to mediate cell-cell contacts (Näther *et al.*, 2006). Beside adherent multicellular structures that are found attached to diverse surfaces, non-adherent floating

multicellular structures are also described. *Methanosacrina mazei* e.g. forms aggregates during exponential growth (Mayerhofer *et al.*, 1992) and halophilic archaea do so in the presence of divalent cations by (Kawakami *et al.*, 2005, 2007). For the halophilic euryarchaeote *Halobacterium volcanii* and the hyperthermophilic crenarchaeote *Sulfolobus ssp.* cellular aggregation was observed in the context with conjugative DNA-transfer (Rosenshine *et al.*, 1989; Schleper *et al.*, 1995; Prangishvili *et al.*, 1998).

Characteristic for all types of cellular aggregation is the attachment between single cells, mostly mediated or stabilised by exoploysaccharides (EPS) and/or proteins (Davey & O´Toole, 2000; Klemm *et al.*, 2004; Kawakami *et al.*, 2007). Some microorganisms like *Xanthomonas* and *Pseudomonas*, use type IV pili to initiate or mediate the cellular aggregation (Ojanen-Reuhs *et al.*, 1996; Bhattacharjee *et al.*, 2001). A mutant defective in the typeIV pilus biogenesis of *Pseudomonas aeruginosa* was unable to attach on surfaces and form microcolonies (O´Toole & Kolter, 1998, O´Toole *et al.*, 2000). Type IV pili are also required for the twitching mobility mechanism (Mattik, 2002), like the light regulated mobility of *Synechocystis* PCC6803 (Bhaya *et al.*, 2001), or the coordinated cell movement in fruitbody development of *Myxococcus xanthus* (Wall & Kaiser, 1999). Type IV pili mediate as well the DNA-uptake in natural transformation systems of mesophilic and thermophilic bacteria (Graupner *et al.*, 2001; Friedrich *et* al. 2003; Averhoff & Friedrich, 2003; Averhoff, 2004). In addition they act as receptors for bacteriophages found in *Pseudomonas* (Roine *et al.,* 1998). The type IV pili biogenesis pathways and the the type II protein secretion systems are very closely related to each other, as has been demonstrated with the *E. coli* type IV pilus biogenesis secretory machinery that was able to assemble PulG, of *Klebsiella oxytoca* into pilus-like structures (Sauvonnet *et al.*, 2000; Köhler *et al.*, 2004).

The bacterial type IV pili machinery is also closely related to the archaeal flagella systems, as was shown by bioinformatic, biochemical and structural analyses (Faguy *et al.*, 1994; Bardy & Jarrell, 2002; Peabody *et al.*, 2003; Cohne-Krausz & Trachtenberg, 2002, 2008). The core components of the bacterial and archaeal systems are (I) a type II/IV secretion system ATPase, (II) a multispanning transmembrane protein and (III) the structure giving prepilin with a characteristic N-terminal signal sequence termed class III signal peptides (Peabody *et al.*, 2003). In addition, it has been shown that the flagella of *Halobacterium salinarium* and *Sulfolobus shibatae* are in symmetry and structure more closely related to the bacterial type IV pili than to bacterial flagella. The archaeal flagellum is arranged in a

three-start helical symmetry and lacks the central core (Cohen-Krausz & Trachtenberg, 2002, 2008). All archaeal flagellins exhibit class III signal peptides and are related to bacterial pili (Faguy *et al.*, 1994; Jarrell *et al.*, 1996; Bardy *et al.*, 2004).

In the genome of the crenarchaeot *Sulfolobus solfataricus* three putative type IV pili loci, containing a typeII/IV secretion ATPase, a predicted integral membrane protein and at least one ORFs containing a class III signal peptide were identified *(*Albers & Driessen, 2005). The operon SSO2316, named after the central ATPase, codes for the flagellum of *S. solfataricus.* A pilus-like bundled structure of 14 nm in diameter is responsible for mobility on surfaces *(*Szabó *et al.*, 2007b). The operon SSO2680 encodes a recently described bindosome assembly system (Bas) that is needed for the functional surface localisation of sugar binding proteins (Zolghadr *et al.*, 2007). The biological function of the third operon SSO0120, spanning ORFs *sso0117* through *sso0121* was unclear. Using whole genome microarray studies to analyse the UV-response of *S. solfataricus* we observed that the genes *sso0117* to *sso0121* were among the most highly induced genes, using UV-light of 75 J/m$^2$ at 254 nm (Frols *et al.*, 2007). A strong up-regulation of the operon was also observed by an independent study of M. White and co-workers using a higher UV-dose of 200 J/m$^2$, with *S. solfataricus* and *S. acidocaldarius* (Götz *et al.*, 2007). In the same time span as the transcriptional response we observed a massive aggregation of the cells, which disappeared after the cellular regeneration (Frols *et al.*, 2007).

In this study we demonstrate that extracellular pili-like structures are formed upon UV-light treatment and by using targeted deletion mutants we demonstrate that they are encoded by the UV-inducible pili operon SSO0120. Furthermore, we show that these pili-like structures are essential for the UV-dependent auto-aggregation of *S. solfataricus* cells and that this phenomenon is driven by double-strand breaks in the DNA, but not by many other stressors. In addition, UV-dependent increased conjugation events further suggest, that pili formation and aggregate formation in *Sulfolobus* mediate a repair mechanism *via* homologous recombination among chromosomes of sister cells.

## 6.3 Results

### 6.3.1 UV-inducible Induction of operon SSO0120

Induction of the genes *sso0118* and *sso0117* belonged to the strongest and fastest transcriptional reactions that were detected earlier in a genome-wide microarray study upon exposure of *S. solfataricus* cells to UV-light (Frölς *et al.*, 2007). These genes belonged to a cluster, and possibly operon of five genes (*sso0117* through *sso0121*) all of which were very strongly induced with a maximal induction of 14-fold for *sso0118* (Fig. 6.1). The transcriptional increases were observed at 1.5 h to 5 hours after UV-treatment. A similar transcriptional reaction pattern for (*sso0121*, *0120*, *0119* and *0117*) of these genes was observed over the time course of 8.5 h hours, but not for the genes flanking upstream or downstream (*sso0116* and *sso0115*, *sso0122*) indicating transcription from a common promoter, as suggested earlier under non-inducing growth conditions (Albers & Driessen, 2005). Only gene *sso0118* deviated from the UV-dependent pattern and appeared up to 3.5-fold higher induced, which may indicate an additional promoter in front of the gene or alternatively, a higher stability of the transcript (Fig. 6.1 B).

Bioinformatic analysis indicated a putative type IV pili biogenesis operon, represented by a type II/IV secretion system ATPase (SSO0120) and an integral trans-membrane protein (SSO0119) (Fig. 6.1 A). The deduced protein sequence of the ATPase contains Walker A/B sites and the conserved domain (virB11-related ATPase COG630N) was found from position 85 to 369 aa and clustered by sequence analysis into the TadA subfamiliy of the type IV ATPases (Planet *et al.*, 2001). Nine transmembrane helices can be predicted for SSO0119 and a conserved domain was found from position 109 to 456 (TadC, COG2064N). Both proteins were found to be homologous to the Tad system (TadA and TadB/TadC), which conveys non-specific tight adherence of *Actionobacillus* on surfaces (Kachlany *et al.*, 2001). The last two proteins, SSO0118 and SSO0117, exhibited an N-terminal signal sequence as found in type IV pilin precursors which belong to the class III secretory signal peptides. No functional predictions could be made for the first gene, *sso0121*, which encodes a highly hydrophilic protein and is exclusive found in the genomes of Sulfolobales.

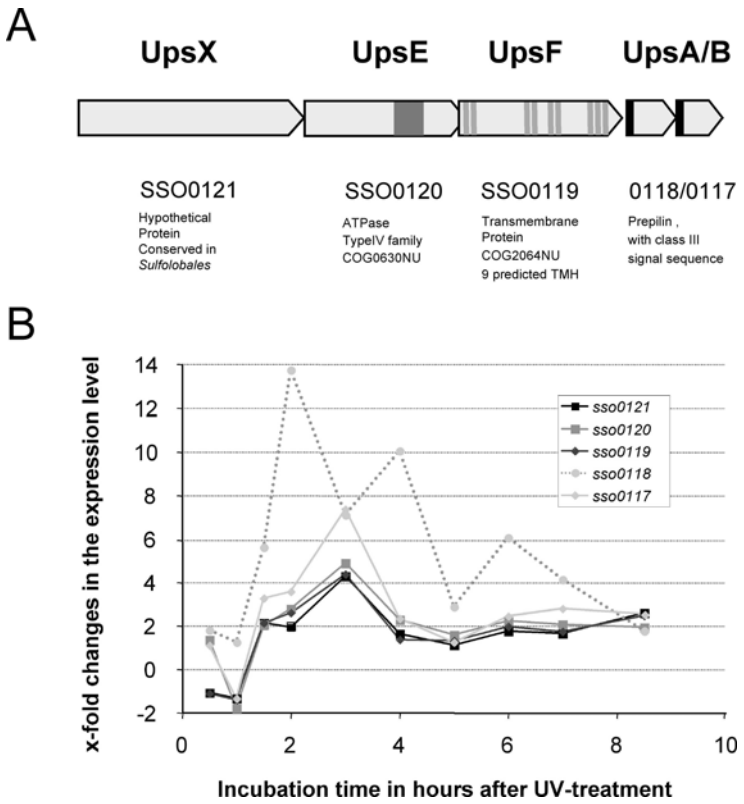



**Fig. 6.1:** (**A**) Composition and characteristics of the *ups* operon (**U**V-inducible **p**ili operon of **S**ulfolobus) (top) and its expression profile upon UV-treatment (bottom), extracted from whole genome microarray data (Fröls *et al.*, 2007). Putative functional regions, (on protein level) were marked: UpsE, cons. ATPase domain (dark grey); UpsF, predicted TMD (grey); UspA and B, class III signal sequence. (**B**) The curves display mean values of 2 to 4 biological replicates for each time point. A UV-dose of 75 J/m$^2$ at 254 nm was applied to exponentially grown cells of a *S. solfataricus* culture.

Bioinformatic analysis indicated a putative type IV pili biogenesis operon, represented by a type II/IV secretion system ATPase (SSO0120) and an integral trans-membrane protein (SSO0119) (Fig. 6.1 A). The deduced protein sequence of the ATPase contains Walker A/B sites and the conserved domain (virB11-related ATPase COG630N) was found from position 85 to 369 aa and clustered by sequence analysis into the TadA subfamiliy of the type IV ATPases (Planet *et al.*, 2001). Nine transmembrane helices can be predicted for SSO0119 and a conserved domain was found from position 109 to 456 (TadC, COG2064N). Both proteins were found to be homologous to the Tad system (TadA and TadB/TadC), which conveys non-specific tight adherence of *Actionobacillus* on surfaces (Kachlany *et al.*, 2001). The last two proteins, SSO0118 and SSO0117, exhibited an N-terminal signal sequence as found in type IV pilin precursors which belong to the class III secretory signal peptides. No functional predictions could be made for the first gene, *sso0121*, which encodes a highly hydrophilic protein and is exclusive found in the genomes of Sulfolobales. The putative pili operon is highly conserved in the order

Sulfolobales, with the same gene arrangement in the strains *S. tokodaii* and *S. acidocaldarius* (Table S.6.1 and Szabó *et al.*, 2007a*)*. Further similarities were only found to genes of the hyperthermophilic crenarchaeon *Metallosphera sedula*, belonging to a closely related order. In *M. sedula*, homologues of *sso0120*, *sso0119* and *sso0117* formed an operon structure whereas *sso0118* was found in a different genomic region.

### 6.3.2 Maturation of prepilins

Both SS0117 and SSO0118 contained the predicted cleavage site for the type IV prepilin peptidase PibD, an aspartyl type IV prepilin signal peptidase (Albers *et al.*, 2003; Szabó *et al.*, 2007a). In SSO0117 only 6 amino acids and in SSO0118, 16 amino acids would be cleaved by PibD (Fig. 6.2 A). The ORFs of SSO0117 and SSO0118 were cloned into an *E. coli* expression vector already containing PibD (Szabó *et al.*, 2007a).


**Fig. 6.2:** Analysis of the pili proteins UpsA/B (SSO0117/118). (**A**) The upper panel shows the N-termini of both proteins. The processing site of PibD is indicated by an arrow. (**B**) The lower panel shows the result of an in vivo cleavage assay of SSO0118-HA by PibD in *E. coli*. Expression of SSO0118-HA was detected by by Western Blot analysis using HA-tag antibodies. Lane 1: expression of SSO118-HA in the absence of PibD; lane 2: expression of SSO0118-HA two hours after arabinose induction; lane 3: 2 h after induction of PibD by IPTG

Using the *in vivo* assay the expression of the prepilin proteins was induced for two hours before the expression of the peptidase was induced. Western blot analysis of crude membrane extracts of the recombinant *E. coli* cells showed that SSO0118 was processed by PibD resulting in a faster running species when compared to the full length protein (Fig. 6.2 B, lane 2 and 3). Cleavage of the signal peptide of SSO0117 could not be observed, most probably because the difference between the pre-protein and the

processed form do not differ enough to be separated on SDS-PAGE. Experiments to separate these two forms in IEF gel electrophoresis failed.

### 6.3.3 UV-induced pili formation

We used electron microscopy to analyse cellular surfaces in order to search for pili formation upon UV-treatment. To exclude that any extracellular structures were not artefacts of flagella we used the *S. solfataricus* knock-out strain ΔFlaJ that does not produce flagella (Szabó *et al.*, 2007b). Only on the surface of the UV-treated cells, we observed pili-like structures (Fig. 6.3 A), none were observed on untreated cells. These pili structures were spread over the whole surface and were not polarised at one cell side. Most of the cells of a UV-treated culture contained many pili, some had less or very few (only 2 to 3 pili), only few cells did not expose pili on their surfaces at all. A time series experiment showed that first pili structures were observed at 1 h after UV-treatment.

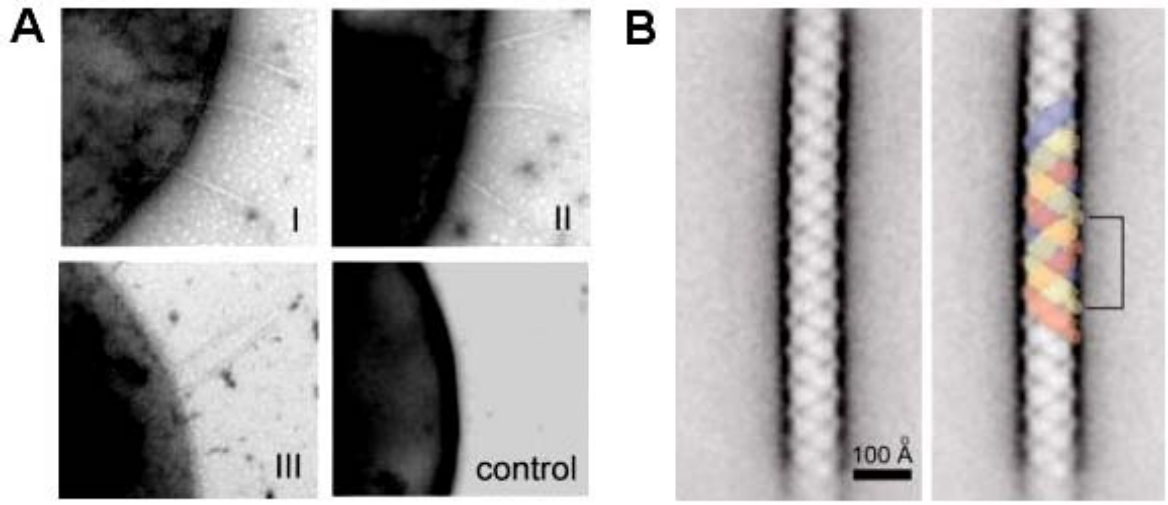



**Fig. 6.3:** Electron microscopically analysis of UV-inducible pili in *S. solfataricus*. (**A**) ΔFlaJ cells were analyzed by electron microscopy 3 h after UV light treatment (A I, II and III) and mock treatment for the control. (**B**) Image processing of pili. Left: projection map obtained after processing 700 non-overlapping fragments of straight pili. Right: scheme of 3-stranded helical arrangement of the pili overlayed. The horizontal lines indicate the pitch of the structure, which is 155 Å

In comparison to the flexible flagella, the pili showed a more straight and rigid structure. Pili of up to 16 µm in length or even longer were observed. However, such long filaments where only found detached from the cells, which indicates that they are more fragile than flagella. Because the pili appeared straight for most of their parts, it was possible to

process them by single particle analysis selecting straight segments of almost up to 1000 Å. About 700 hundred segments were extracted from long pili, aligned and summed. The final average projection map is shown in Fig 6.3 B. The structure appeared to be build up from three evenly spaced helices. The pitch (repeating unit) of the pili was 155 Å and the maximal diameter was about 100 Å. In the negatively stained samples the single helices appeared almost uniformly stained and there were no clear density differences that could give clues about the handedness (left- or right handed) of the helices.

### 6.3.4 Cellular aggregation after UV-treatment

The appearance of pili upon UV-treatment that could mediate cell-to-cell contacts inspired us to analyse microscopically the formation of cell aggregates (Fröls *et al.*, 2007). We have shown earlier that aggregation occurred with high frequency independent of the *S. solfataricus* genotypes, because experiments with four different strains (P1, PH1, PH1-M16, PH1(SSV1)) showed the same phenotypic reaction. With increasing time after UV-treatment, an increasing number of cells were found in aggregates with the highest amount of aggregation found at 6 to 8 hours after UV-treatment (Fröls *et al.*, 2007 and Fig. 6.7). The aggregates increased also in size. While three to five cells were found in the early aggregates, bigger complexes formed at later time points. The shape of the early aggregates seemed to be random, as variations of pyramids, circle shapes, straight and branching chains were observed (not shown). In the later stages (6 h) the cells accumulated to big clusters of > 100 cells. As it was impossible to count the number of cells in such aggregates, our quantitative data (% cells in aggregates of total cell count) generally represent an underestimate.

Attempts to destroy the cell-cell connections by shear force experiments resulted in destroyed cells at all stages but not in disaggregation, indicating a high stability of the aggregates. The induction of cellular aggregation was UV-dose dependent (see Fig. 6.4 A). We treated the cells with seven different UV-doses ranging from 5 J/m$^2$ to 1000 J/m$^2$. Growth retardation of the respective cultures was directly proportional to the applied UV-dose (data not shown). Highest cellular aggregation was observed 6 h after UV–treatment, i.e. at the expected maximum. The highest amount of cellular aggregation was found with 75 J/m$^2$ (at least 50 to 70% of cells were found in aggregates) whereas with a UV-dose of 50 J/m$^2$ ca. 40 to 45% of the cells were found in aggregates of ≥ 3 cells. Even the lowest dose of UV-light of 5 J/m$^2$ induced the cellular aggregation,

whereas the high UV-doses of 200 J/m$^2$ and 1000 J/m$^2$ showed a very low, respectively, no significant aggregation reaction.

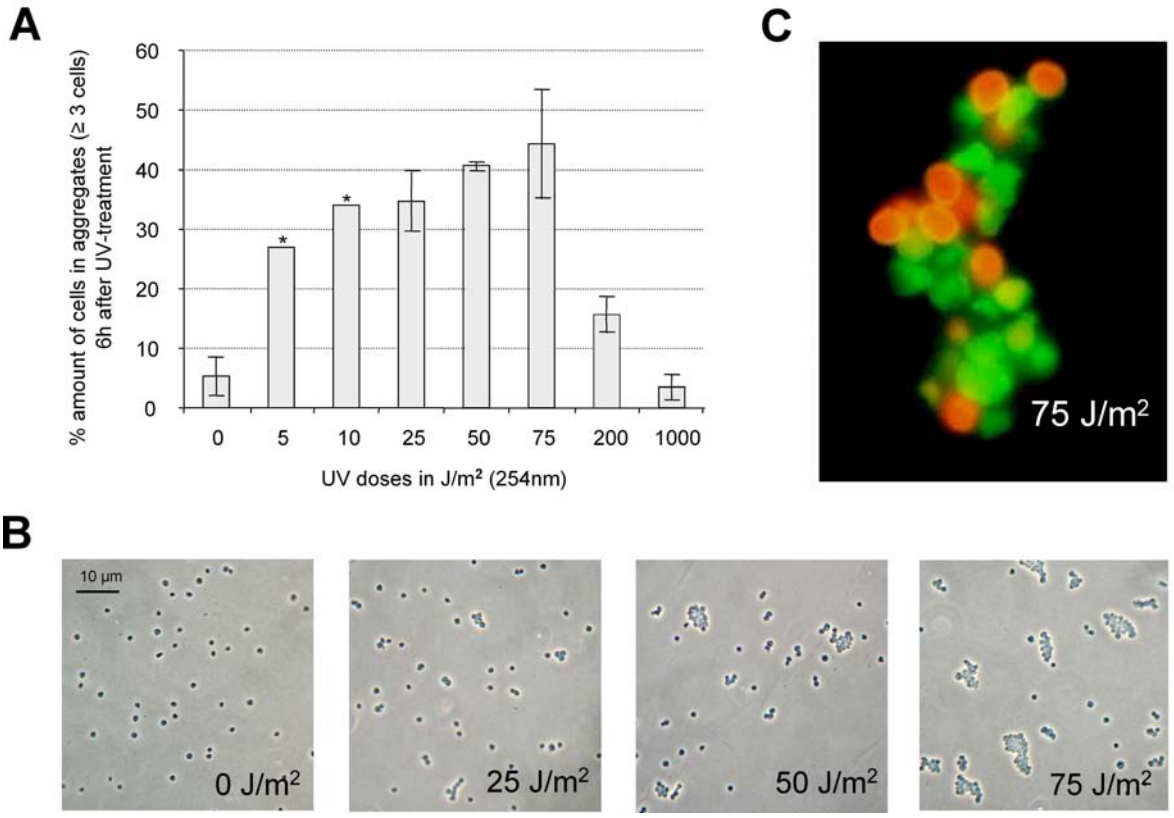



**Fig. 6.4:** Aggregation of *S. solfataricus* cells after treatment with different UV-doses (**A**) Quantitative analysis of cellular aggregation at 6 h after UV-treatment. Exponential cultures were treated with 0, 5, 10, 25, 50, 75, 200 and 1000 J/m$^2$. The % amount of cells in aggregates (≥ 3 cells) is given in relation to the total cells. For each UV-dose the amount of cells in and outside aggregates were counted until 500 single cells were found. The bars display the mean of three independent experiments, except for 5 and 10 J/m$^2$ (see asterisk), where only one experiment was performed. (**B**) Light micrograph of *S. solfataricus* cell aggregates at 6 h after UV-treatment with different UV-doses. The size of the aggregates increased with the UV-dose, the biggest aggregates were found after treatment with 50 J/m$^2$ and 75 J/m$^2$. (**C**) Fluorescence micrograph of a *S. solfataricus* cell aggregate at 6 h after UV treatment at 75 J/m$^2$. Cells were stained with the LIVE DEAD Baclight $^{TM}$ (Invitrogen) assay. Living cells are labelled in green and dead cells in red. Big aggregates of > 20 cells were mostly found at 3 h after treatment. For quantitative analysis of the cell vitality at different UV-doses see table 6.1.

We also observed a strong correlation between the size and amount of cellular aggregates (Fig. 6.4 B). Low UV-doses of 5 J/m$^2$ and 10 J/m$^2$ resulted in cellular aggregates of < 7 cells. Only upon a UV-dose of 50 J/m$^2$ to 75 J/m$^2$, big aggregates of 10 to 20 cells or

more were generated frequently (average size 12-15 cells). In the case of the high UV-doses of 200 J/m$^2$ and 1000 J/m$^2$ no aggregates > 4 cells were observed.

By using different life and death staining techniques (see methods and material), we investigated if the cellular aggregation could represent an accumulation of dead cells (see Tab. 6.1). In the case of the lowest UV-dose of 5 J/m$^2$ only 8% of the total cells in aggregates (≥ 3 cells) were dead. The amount of dead cells increased proportionally with the UV-dose but was far lower than the number of living cells. 64% of the total cells in the infrequent aggregates at 200 J/m$^2$ were not alive. But at lower UV-dose, like 75 J/m$^2$ even big aggregates of > 20 cells were almost uniformly constituted of living cells (Fig. 6.4 C).

**Table 6.1:** Cell vitality of *S. solfataricus* cells in aggregates after treatment with UV-light

| UV-dose (J/m$^2$) | Amount of dead cells in % [a] |
|---|---|
| 5 | 8 |
| 10 | 12 |
| 25 | 17 |
| 50 | 34 |
| 75 | 44 |
| 200 | 64 |

[a] at 6 h after UV-treatment a live and death stain was performed (see methods and material and also Fig. 6.4 C). A minimum of 50 aggregates ≥ 3 cells, were counted per each UV-dose and the fraction of dead cells is given in relation to the total cells found in aggregates.

### 6.3.5 The gene products of the UV-inducible pili operon are responsible for pili formation and mediate cellular aggregation

To proof that pili were indeed assembled from components expressed from the putative pili operon, a deletion mutant was constructed in which the ATPase (SSO0120) was replaced by insertion of the *lacS* gene *via* a double cross over. The successful knock-out of the *sso0120* gene, was confirmed by Southern-analysis and RT-PCR (see Fig. 6.5 A and B). The RT-PCR showed that under inducing conditions the *sso0120* mRNA was absent, while the downstream genes of the operon were still expressed.

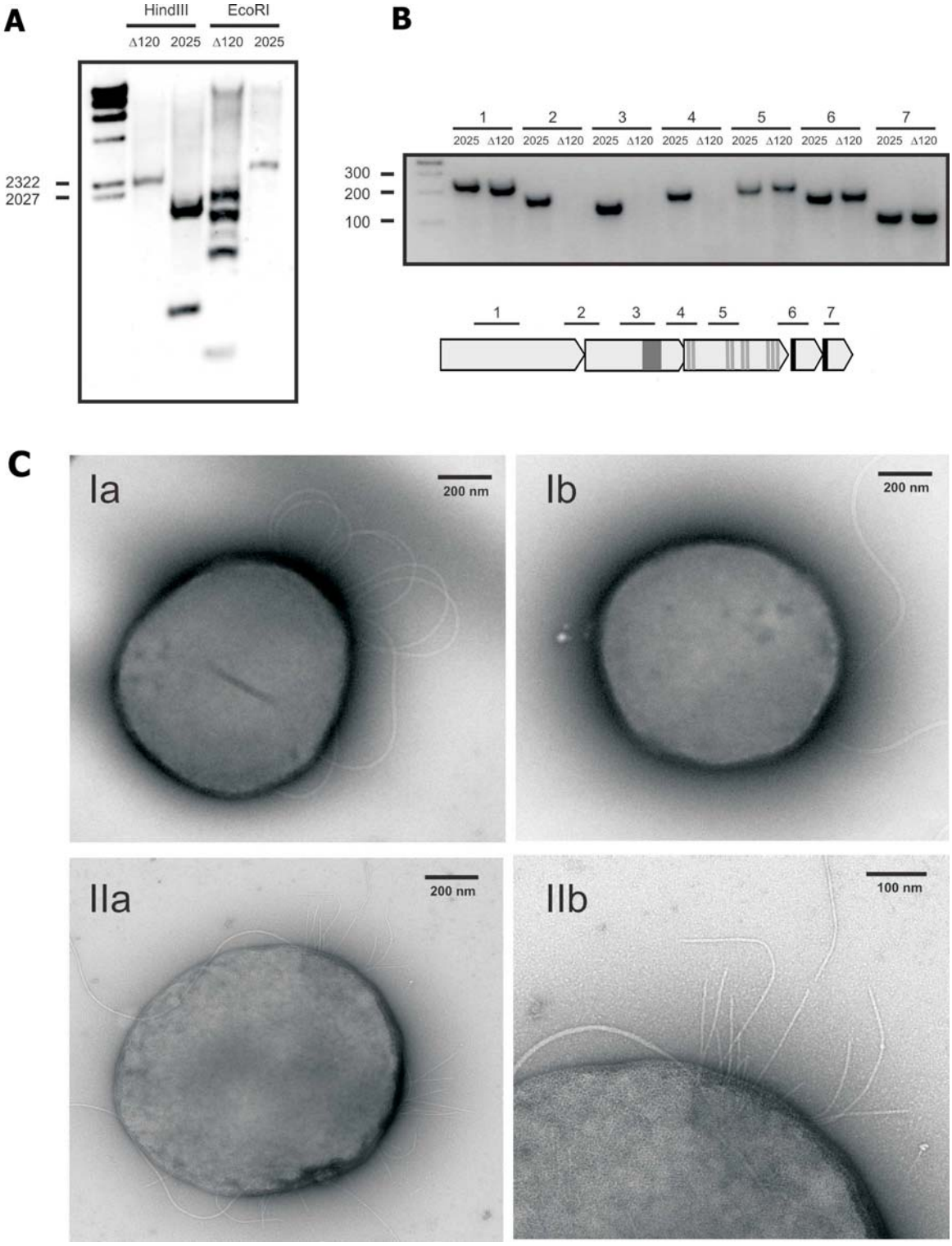



**Fig. 6.5:** Analysis of the SSO0120 knock-out strain. (**A**) Southern Blot analysis of wildtype PBL2025 and the ΔSSO0120 strain. Genomic DNA of both strains was digested either with *Hind*III or *EcoR*I. (**B**) RT-PCR analysis of PBL2025 and ΔSSO0120 strain after UV-stress. The position of the primers used for the PCR reactions are indicated by the same number above the gel and the map of the operon. (**C**) Electron micrographs of PBL2025 (IIa and b) and ΔSSO0120 (Ia and b) 3 h after UV-treatment.

After UV-treatment of the mutant ΔSSO0120 we could not observe any pili-like structures on the cellular surfaces by electron microscopy (Fig. 6.5 C Ia and Ib). As control we used the parent strain PBL2025 (Schelert *et al.*, 2004), which clearly showed pili-like structures beside the flagella on its cellular surface upon exposure to UV-light (Fig. 6.5 C IIa and IIb). Overexpression of both pilin genes, *sso0117* and *sso0118*, in the ΔFlaJ strain using the virus based vector construct pSVA96 resulted in the assembly of fewer, but extremely long and irregular pili (Fig. 6.6). This demonstrated that the two prepilin genes indeed form the UV-inducible pili.

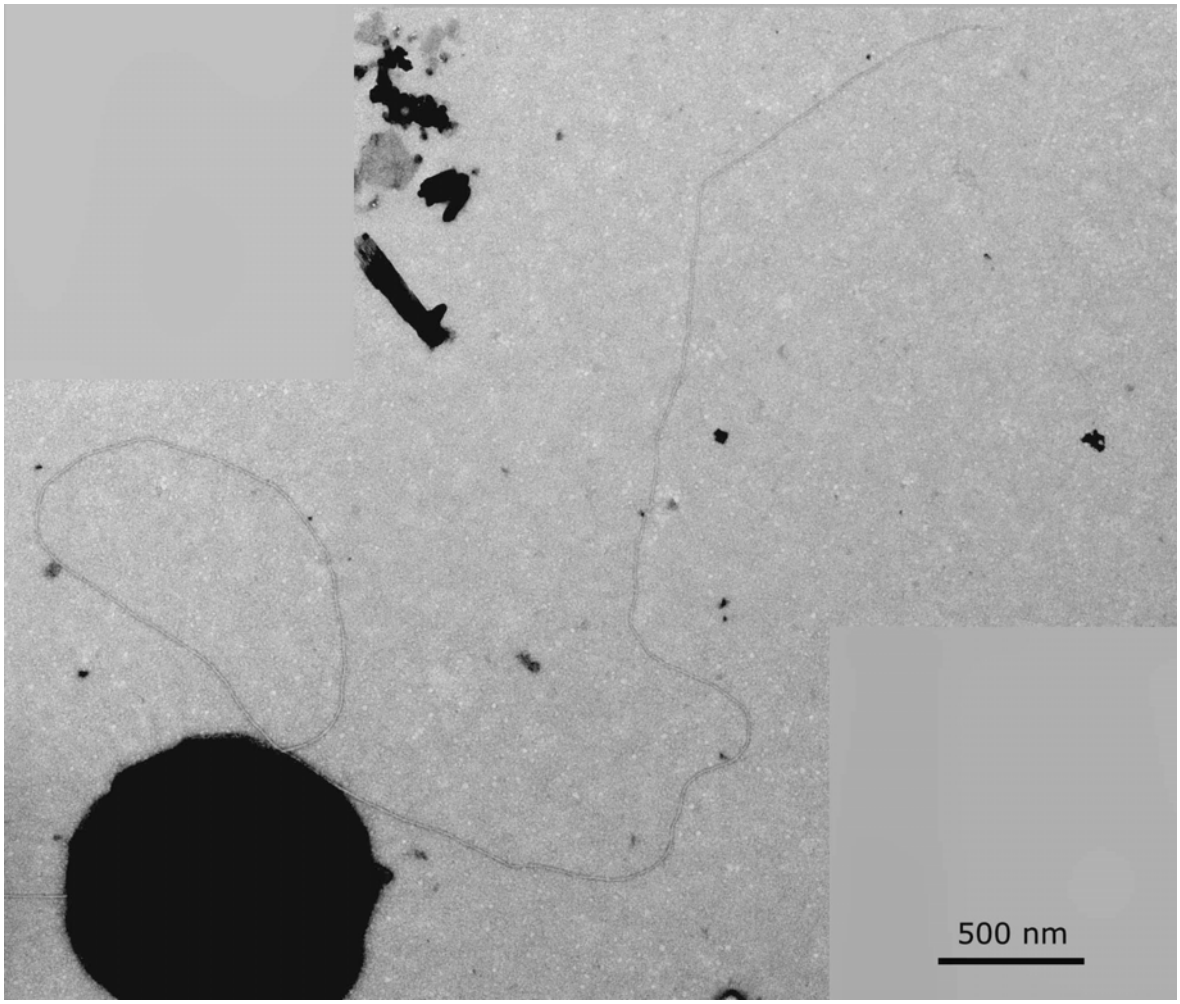



**Fig. 6.6:** Electron micrograph of a ΔFlaJ cell overexpressing the pilin SSO117/118. Two pictures were assembled to show the length of the pilus.

The ΔSSO0120 strain was also tested for its ability to form cellular aggregates upon UV-exposure. After a treatment with a UV–dose of 50 J/m$^2$ no significant cellular aggregation of more than 4 cells was observed (Fig. 6.7). The amount of cells in aggregates

accounted for less than 10% in the UV-treated culture and the control (mock-treated) culture, similar to the % amount of cells in aggregates observed for the mock-treated cultures of the other four tested *S. solfataricus* strains.

The *S. solfataricus* strain P1 and the PH1-M16 (P1 ΔlacS) showed a maximum aggregation at 6 h to 8 h after treatment, with an average of 45-50% cells in aggregates. In the same experiment, the PBL2025 and the ΔFlaJ strains exhibited a shifted maximum at 8 h to 10 h and a lower amount of aggregation with an average of 20%. The weaker reaction is most probably due to the different genotypes of these strains, which stem from PBL2025, an isolate from Yellowstone Nationalpark *S. solfataricus* 98/2s (Schelert *et al.*, 2004). Comparable results were observed when using a lower UV-dose of 25 J/m$^2$ (see Fig. S.6.1). Again, no significant cellular aggregation was observed for strain ΔSSO0120. The P1 and PH1-M16 strains showed a lower amount of aggregation with 30-40%, as expected in relation to the lower UV-dose.

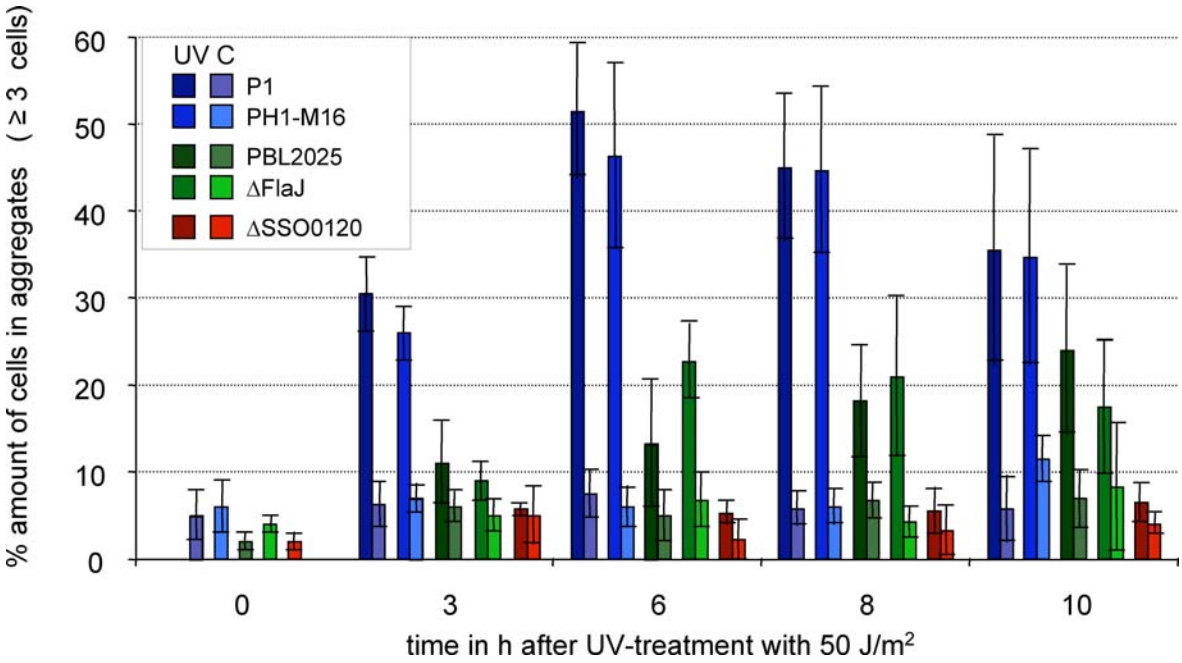



**Fig. 6.7:** Quantitative analysis of the UV-induced cellular aggregation of different *S. solfataricus* strains at 0 to 10 hours after treatment. The graph is based on four independent UV-experiments for each strain. Cellular aggregation was observed at 3 h, 6 h, 8 h, and 10 h after UV-treatment with 50 J/m$^2$ (254 nm). The bars display the % amount of cells in aggregates (≥ 3 cells) in relation to the total amount of evaluated cells (500 to 1000 single cells were counted). No UV-induced cellular aggregation was observed in the knock-out strain ΔSSO0120. Similar results were observed by using a UV-dose of 25 J/m$^2$ (254 nm) displayed in supplementary material Fig. S.6.1.

The amount of cells in aggregates in the case of the PBL2025 strain stayed the same, whereas in the case of the ΔFlaJ strain the amount of cells in aggregates increased to > 30% and the maximum shifted to 6 h.

Based on these results we concluded that the UV-induction of the putative pilin operon, the inducible pili production and the cellular aggregation were functionally linked to each other. Based on the observed properties we named the newly identified operon **UPS** for **U**V-inducible **p**ili operon of **S**ulfolobus, represented by the genes *upsX*, *upsE* (ATPase), *upsF* (TM protein), *upsA* and upsB (prepilin).

### 6.3.6 Cellular aggregation is not inducible by other environmental stressors or in late growth phases

To analyse if cellular aggregation can be induced by other conditions than UV-exposure, four strains that harbour the wild-type of the *ups*-operon were used: *S. solfataricus* strains P1, PH1-M16, PBL2025 and ΔFlaJ. We monitored and quantified the extent of cellular aggregation after a temperature shift from 78°C to 88°C (heat shock) and down to 65°C (Kagawa *et al.*, 2003), which corresponded to non-lethal heat- and cold-shock conditions that might be often encountered in hot springs. PH shifts from pH 3 to pH 4 and down to pH 2.5 were similarly investigated. No significant cellular aggregation was observed under the tested conditions in none of the four tested strains. The % amount of cells in aggregates (≥ 3 cells) were all below 10% (see Fig. S.6.2).

We also monitored the extent of cellular aggregation in the late growth phases of the cultures, from stationary to death phase (three time points for each phase). Only at the beginning of the late stationary phase, i.e. at the beginning of growth retardation, a slightly increased cellular aggregation was noted. For strain P1 up to 24% of the cells were found in aggregates of 4 to 7 cells at most, while the % amount of cells in aggregates (≥ 3 cells) were lower than 10% in all other growth phases (see Fig. S.6.2).

### 6.3.7 Cellular aggregation is induced through treatment with DSB-inducing argents

As a response of *S. solfataricus* to UV-light we observed earlier the formation of double-strand breaks in the genomic DNA (DSB). Whereas cis-syn-cyclobutane pyrimidine dimers

(CPD) represent direct DNA damages caused by UV-light effect, DSB are probably formed as a result of collapsing replication forks at unrepaired sites in the genomic DNA (Frölś *et al.*, 2007). It has been speculated earlier, that DSB might represent an intracellular signal for further cellular reactions (Garinis *et al.*, 2005). We were therefore inspired to test whether the formation of DSB is connected to the formation of cellular aggregates. We tested the induction of cellular aggregation, of the different *S. solfataricus* strains P1, PH1-M16, PBL2025 and ΔSSO0120 in response to the DSB inducing argents bleomycin (Fig. 6.8) and mitomycin C (Tab. 6.2).

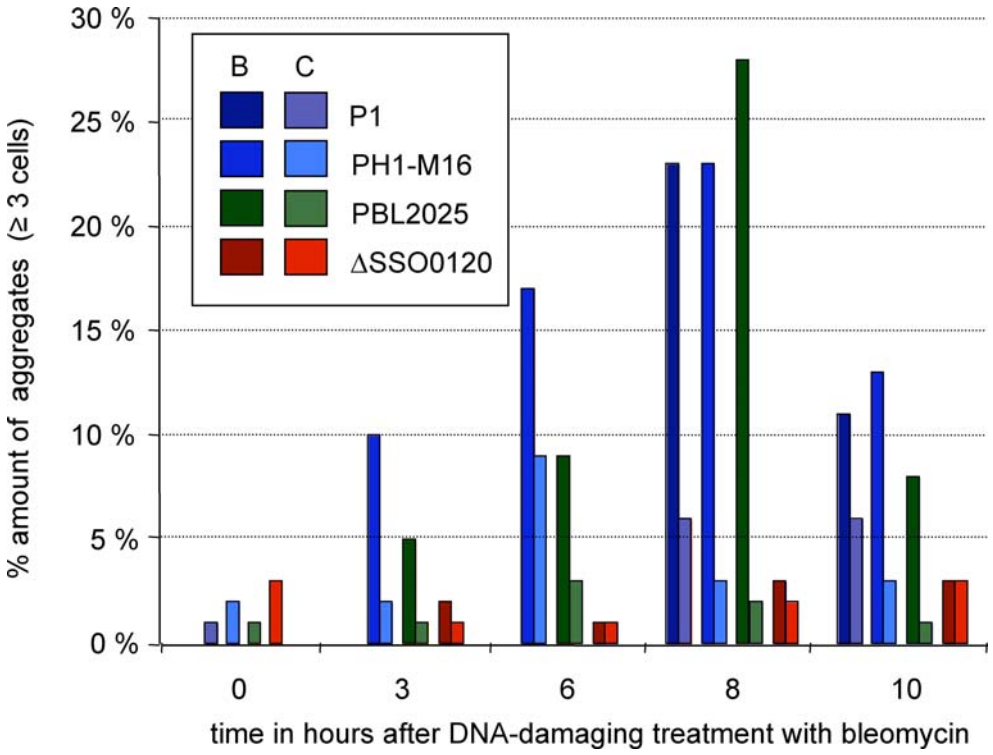



**Fig. 6.8:** Aggregate formation of different *S. solfataricus* strains after treatment with bleomycin (3 µg/ml). No significant cell aggregation was observed with the knock-out strain ΔSSO0120. The bars display the % amount of cells in aggregates (≥ 3 cells) in relation to the total amount of evaluated cells (500 single cells were counted).

Cells were treated with different doses of bleomycin (3 ug/ml) and mitomycin C (5, 10 and 15ug/ml) (Cannio *et al.*, 1998; Grogan *et al.*, 2001; Kosa *et al.*, 2004). These concentrations were non-lethal to the cells as investigated by plating efficiencies and growth behaviours in liquid cultures (data not shown). Cellular aggregation was monitored at 3 h, 6 h and 8 h after the treatment, in the treatment with bleomycin additionally at 1 h and 10 h. All tested strains, except the ΔSSO0120 strain showed a significant cellular aggregation in response to the agents. Eight hours after the treatment with bleomycin

strains P1, PH1-M16 and PBL2025 exhibited 25-35% amount of cells in aggregates (Fig. 8), while aggregation in the mock treated cultures and the bleomycin-treated strain ΔSSO0120 remained below 10%. Similarly, although less strongly, mitomycin C induced aggregate formation in the *ups*-operon containing wildtype strains (P1, PBL2025), but not in the knock-out strain (Tab. 6.2). These observations indicate that DNA-damage and in particular DSBs might be a direct or indirect signal for inducing aggregate formation.

**Table 6.2:** Cellular aggregation in % after treatment with the DNA DBS-inducing agent mitomycin C [a]

| Strains | Sample | time in hours after treatment | | |
|---|---|---|---|---|
| | | 3 h | 6 h | 8 h |
| P1 (wild type strain) | 5 ug /ml | 5 | 10 | 10 |
| | 10 ug /ml | 8 | 10 | 16 |
| | control | 1 | 2 | 2 |
| PBL2025 (parent strain) | 5 ug /ml | 1 | 10 | 9 |
| | 15 ug /ml | 6 | 12 | 10 |
| | control | 1 | 1 | 3 |
| ΔSSO0120 (knock-out strain) | 5 ug /ml | 1 | 1 | 2 |
| | 15 ug /ml | 0 | 0 | 0 |
| | control | 0 | 1 | 0 |

[a] 20 ml of an exponential cultures of the *S. solfataricus* strains were treated with 5, 10 and 15 µg/ml mitomycin C (Sigma) and re-incubated at 78°C and 150 rpm. The % amount of cells in aggregates (≥ 3 cells) in relation to the total amount of evaluated cells is given (500 single cells were counted).

### 6.3.8 UV-light induced conjugation in *S. solfataricus*

In order to study, if conjugation frequencies in *S. solfataricus* are enhanced upon UV-light exposure, we have developed a conjugation assay based on the exchange of the reporter gene ß-galactosidase (*lacS*); between the *lacS* wild type strain (P1; *lacS* $^{+}$/ *pyrEF* $^{+}$) and a *lacS*-deficient strain (PH1-M16; *lacS* $^{-}$/ *pyrEF* $^{-}$). To distinguish the strains on plates, a 5-FOA negative selection was used, similar to that described by Grogan, (1996). Only uracil auxotrophic cells were able to grow under these conditions and positive conjugants formed additionally blue colonies when exposed to the chromophore X-Gal. To calculate the conjugation frequency, we separately determined the recombination, reversion and mutation frequencies of the mixed and single cultures, respectively (for more details see

material and methods). No reversion events of the mutated *pyrEF* or *lacS* genes of the PH1-M16 (*lacS* $^{-}$/ *pyrEF* $^{-}$) strain were observed with or without UV-irradiation. The mutation frequency of strain P1 to *pyrEF* auxotrophy was on average $10^{-5}$/cell and experiment (see Tab. 6.3) (similar to the determined mutation frequency for *S. solfatarcius* in Martusewitsch *et al.*, (2000). Using a UV-dose of 75 J/m$^2$ a recombination frequency up to 1.11 x $10^{-2}$ was observed with 1:1 mixtures of strain P1 and PH1-M16, indicating that about 1 out of 100 cells exchanged the marker upon conjugation per one generation. At 50 J/m$^2$ the conjugation reactions decreased by one order of magnitude to an observed recombination frequency of 1.09 x $10^{-3}$ (see also Tab. 6.3). No significant conjugational activity, i.e. enhanced recombinational events were observed without UV-treatment.

**Table 6.3:** Recombination and mutation frequencies in *S. solfataricus* upon UV-light

| Experiment | | Recombination mix cfu / ml (*lacS* $^{-}$ /*pyrEF* $^{-}$) x (*lacS* $^{+}$ /*pyrEF* $^{+}$) | Recombination frequencies[a] (*lacS* $^{-}$ /*pyrEF* $^{+}$) | Mutation frequencies[b] (*lacS* $^{-}$ /*pyrEF* $^{+}$) |
|---|---|---|---|---|
| A | 75 J/m$^2$ | 3.80 x $10^{7}$ | 1.11 x $10^{-2}$ | 1.90 x $10^{-5}$ |
| | Control | 1.87 x $10^{8}$ | - | 1.59 x $10^{-5}$ |
| B | 75 J/m$^2$ | 2.77 x $10^{7}$ | 4.85 x $10^{-3}$ | 3.21 x $10^{-6}$ |
| | Control | 2.10 x $10^{8}$ | - | 1.00 x $10^{-4}$ |
| A | 50 J/m$^2$ | 3.67 x $10^{7}$ | 1.09 x $10^{-3}$ | n. d. |
| | Control | 8.80 x $10^{7}$ | - | n. d. |
| B | 50 J/m$^2$ | 1.42 x $10^{8}$ | 3.11 x $10^{-3}$ | 2.14 x $10^{-5}$ |
| | Control | 2.39 x $10^{8}$ | - | 2.15 x $10^{-5}$ |

Abberrations: -, no conjugation events observed; n. d., not decetd

[a] recombination frequencies were determined as event/cell; the median of the positive conjugates (*lacS* $^{-}$ /*pyrEF* $^{+}$) cfu/ml with selection (5-FOA) were determined and divided by the median of the cfu/ml from all observed colonies under non-selective conditions.

[b] mutation frequencies were given for strain P1 as event/cell

## 6.4 Discussion

The special living conditions of Archaea in extreme environments make them interesting objects to study adaptations and stress responses. In particular hyperthermophilic and acidophilic archaea like *S. solfataricus* have to deal with a constant stress and DNA-damage in their harsh environments.

Here we present the identification and characterization of an archaeal pili system that mediates cellular aggregation of *S. solfataricus* in response to UV-damage. The genes encoding the now called *ups*-operon for ***U****V-inducible* ***p****ili operon of* ***S****ulfolobus* had earlier been identified to be UV-dependently induced in a genome-wide DNA-microarray analysis (Fröls *et al.*, 2007).

To our knowledge this is the first reported study on a UV-inducible pili-mediated auto-aggregation system. As discussed below, its induction seems to be coupled to the DNA DSBs caused by UV-irradiation. We suspect that cellular aggregation mediates DNA-repair *via* conjugation, as we find increase in conjugation activity upon UV-irradiation and in transcripts of genes involved in homologous recombination.

### 6.4.1 UV-inducible pili mediate cellular aggregation

By electron-microscopic analysis we found a strong correlation between the formation of extracellular pili on the cellular surface after UV-treatment and the expression of the *ups*-operon, both of which appeared first at 1 h after UV-treatment and reached a maximum within 5 to 6 hours. To proof our hypothesis that the gene products of the *ups*-operon are responsible for the production of the pili we used the recently developed genetic system (Albers & Driessen, 2007) to produce a specific knock-out of the putative secretion ATPase SSO0120. No pili structures were observed on the cellular surface of the ΔSSO0120 strain.

By testing the ΔSSO0120 *strain* in a quantitative cellular aggregation analysis, we proved that the pili are necessary for the cellular aggregation of *S. solfataricus* after UV-treatment. Cellular aggregates were as infrequent (i.e. lower 10% of all cells) as in mock-treated controls of four different *S. solfataricus* strains. Image analysis of isolated pili structures showed that the pili are much thinner in diameter and clearly distinguishable from the flagella of *S. solfataricus* (Szabó *et al.*, 2007b).

These pili like structures are spread over the whole cellular surface. They are not bundled or polarized like the cable-like flagellar-bundles of *Pyrococcus furiosus*, which mediate cell attachment (Näther *et al.*, 2006) or the typeIV pili of the Tad system from *Actinobacillus* species that mediate non-specific adherence (Kachalany *et al.*, 2000, 2001). Experiments to disconnect cellular aggregates by shearing forces failed, indicating that the cell-cell contacts were highly stable once formed and showing that the cellular aggregates were not a result of an unspecific accumulation. The latter was also ruled out by a live and death stain analysis (Fig. 6.4).

The detailed mechanisms of auto-aggregation is, however, still unknown. It has been reported that the bacterial typeIV pili are bound with their tip on surface structures or other cells (Mattick, 2002). We did not observe any attachment to surfaces. However, our experiments were performed under moderate shaking in glass flasks, such that one cannot rule out the possibility of surface attachment under different conditions.

### 6.4.2 UV-inducible cellular aggregation is highly dynamic

A quantitative assay was developed in this study to analyze the dynamics of cellular aggregation in more detail. We showed that the aggregation is a fast process induced by the UV-dependent reaction of *S. solfataricus* and seems to occur in two phases. First, small aggregates of 3-5 cells accumulate, which later aggregate to bigger forms. The maximum of aggregation was reached at 6 h to 8 h after UV-treatment, followed by a clear disappearance, interpreted as an active disaggregation. One has to note that the absolute amount of cellular aggregation is by far underestimated because cell aggregates of more than 20 cells were uncountable. Maxi-aggregates with even up to 100 and more cells were found frequently at 6 h after UV-treatment. Furthermore, cell aggregates of two were not incorporated in the calculations in order to exclude dividing cells.

In correlation to the cell cycle length of *S. solfataricus,* which is around 7 h, the dynamics of this process are relatively fast. For example, the cellular packets of *M. mazei* need 2 to 6 days to form the lamina structures, and then remain stable over 6-11 days until the culture reaches stationary growth phase and the lamina disaggregate (Mayerhofer *et al.*, 1992). The stress induced biofilm formation of *Archaeoglobus fulgidus* occurs in 2-12 hours. But in this case no disaggregation was reported (LaPaglia & Hartzell, 1997).

### 6.4.3 UV-light is the only identified stressor to induce auto-aggregation

It is reported that cells organized in multicellular structures show a higher resistance to different environmental stressors, like temperature, pH and also UV-light (Ojanen-Reuhs *et al.*, 1996; Roine *et al.*, 1998; Martinez & Casadevall, 2006). Treatment of the hyperthermophilic archaeon *Archaeoglobus fulgidus*, with a high dose of UV-light and other physical or toxic stressors results in a biofilm production (LaPaliga & Harzell, 1997). Mutants defect in the auto-aggregation of the plant pathogen bacteria *Pseudomonas syringae* and *Xantomonas campestris* showed a higher sensibility to UV-irradiation, than the wildtype capable of forming multicellular structures (Ojanen-Reuhs *et al.*, 1996; Roine *et al.*, 1998). A decreased sensitivity to UV-light and other environmental stressors was also reported for biofilms of the yeast-like fungi *Cryptococcus neoformans* (Martinez & Casadevall, 2007). None of the stressors that we used for *S. solfataricus* induced cellular aggregation, nor did late growth phase stages. This stands in contrast to all given examples of multicellular structures, which are typically interpreted as an advantageous life form under harsh or specialized environmental conditions. Thus *S. solfataricus* shows a unique multicellular formation, which is not a general effect of a stress response.

Interestingly, the extent of cellular aggregation (aggregate sizes and % amount cells involved) was dependent on the UV-dose. Relatively high doses of UV-light, like 200 J/m$^2$ or 1000 J/m$^2$ resulted in an insignificant amount of small aggregates (≤ 4 cells) and killed most of the cells. In contrast a relatively low dose of UV-light, like 5 J/m$^2$ induced cellular aggregation. In nature sunlight with ca. 96% UV-A and 4% UV-B reaches the ground and is the most DNA-damaging factor. The daily dose of DNA-damaging UV-B light on a sunny day in the northern and southern world hemispheres is measured between 1-3 kJ/m$^2$ over 24 h (depending on the season). The experimentally used UV-C (254 nm) is about 100 fold ($10^2$) more effective than UV-B in inducing CPDs (Kuluncsics *et al.*, 1999). With reference to the observation that even low dose of UV-light significantly induce the cellular aggregation of *S. solfataricus* we conclude that this phenotypic effect reflects the behaviour of the organism to the sunlight in the natural environment.

### 6.4.4 Cellular aggregation is induced by double strand breaks and might mediate a recombinational repair system *via* conjugation

Between 2 and 8 hours after UV-treatment we observed the formation of DNA double-strand breaks (DSB), probably resulting from replication fork collapse at CPD damaged DNA-sites (Fröls *et al.*, 2007). These observations inspired us to investigate in this study if the cellular aggregation is causally linked to the presence of DSBs in the genome. Indeed, the DSB inducing argents bleomycin and mitomycin C caused the same phenotype of cellular aggregation as UV-light. Similarly, the proliferation of the *Sulfolobus shibatae* virus 1 (SSV1) can be induced by mitomycin C as well as UV-light (Reiter *et al.*, 1988) indicating that the same internal signal cascades are involved.

However, it is still unclear how DSB DNA might be sensed in the cells and how the signal is further transferred to induce the cellular aggregation and DNA-repair reactions. A phototaxis mechanism is reported for *Halobacterium salinarum* that regulates the motor switch of the flagella. The UV-light is sensed by the sensory rhodopsin (Htr) and activates a Che-like two-component system (Nutsch *et al.*, 2003). However, both such components are not known in *Sulfolobales*. In *Synechocystis* PCC6803 Che-like histidine kinases control the cell orientation to the light and type IV pilus biosynthesis (Bhaya *et al.*, 2001).

During our experiment with the DSB inducing argents we observed that the ΔSSO0120 strain reacted more sensitively to the treatment than the *ups*-wildtype strains. This leads to the hypothesis that the cellular aggregation is needed for an efficient DNA-repair. But it could as well be possible that aggregation is acting as efficiently as a physical protection mechanism simply providing shade within the tight cellular consortium.

However, several lines of evidence indicate, that cellular aggregation might play an important role for mediating conjugation dependent DNA-repair *via* homologous recombinations:

(I) Cellular aggregation of *Sulfolobus* has previously been reported in the context of conjugation mediated by plasmids (Schleper *et al.*, 1995). The structure of the cellular aggregates was highly similar to the here described UV-caused phenotype.
(II) We have observed UV-light enhanced conjugative activity in this study at the same UV-dose ranges as used for inducing cellular aggregation and with a frequency that

correlated with the applied UV-dose.
(III) UV-inducible conjugation has been described earlier for *S. acidocaldarius*. After UV-treatment with 70 J/$m^2$ the highest recombination frequencies were reached (between $10^{-4}$ to $10^{-1}$) (Wood *et al.*, 1997).
(IV) We found a slight, but significant up-regulation of the *rad50*/*mre11* operon upon UV-treatment in *S. solfataricus* using whole-genome microarrays (Fröls *et al.*, 2007). This operon encodes homologues proteins of the eukaryotic system involved in the DSB repair *via* homologous recombination (Hopfner *et al.*, 2002; Constantinesco et *al.*, 2004).

By integrating our observations we think that recombinational repair *via* homologous recombination and DNA-exchange via cell-cell contacts might be an important strategy to overcome DNA-damage in *Sulfolobus* caused by UV-light. Future studies will aim at investigating in more detail the nature of the cell-cell-contacts, e.g. if exo-polysaccharides are produced to stabilize the complexes and if aggregation can occur among closely related species. It will also be interesting to elucidate the transcriptional regulation of the UV-induced genes, with the perspective to clarify the signal transduction pathways that sense UV-irradiation or DNA-damage in crenarchaeota.

Although the specific UV-inducible operon involved in pili formation in *Sulfolobus* is not conserved outside the order Sulfolobales, we consider it likely that cellular aggregation and enhanced conjugation might be a more general principle employed by microorganisms that are regularly exposed to sunlight.

## 6.5 Material and Methods

*Strains and growth*

*S. solfataricus* PH1 (Schleper *et al*., 1994), PH1-M16 (Martusewitch *et al.*, 2000) and *S. solfataricus* PBL2025 (Schelert *et al.*, 2004) and the several deletion mutants were grown aerobically at 80°C in Brock´s medium (Grogan, 1989) adjusted to pH 3 with sulfuric acid and supplemented with 0.1% (w/v) of trypton and 0.2% (w/v) of arabinose under moderate agitation (ca. 150 rpm in NewBrunswick shaker). Growth of cells was monitored by measuring the optical density at 600 nm. Solid media were prepared by adding gelrite to a final concentration of 0.6% and $Mg^{2+}$ and $Ca^{2+}$ to 0.3 and 0.1 M, respectively. Plates were incubated for five days at 78°C. For the propagation of plasmids *E. coli* strain DH5$\alpha$ was used. For the virus containing plasmids ElectroMAX™ *E. coli* Stbl4™ cells (Invitrogen, Germany) were used.

*Treatment with UV–light*

For the UV-treatment all preparations were performed under red dimmed light. Aliquots of 10 ml *S. solfataricus* culture ($OD_{600\ nm}$ 0.3 - 0.5) were transferred to a 110 mm plastic petri dish and treated with a defined UV-dose ($\lambda$ 254 nm, UV-Stratalinker 1800, Stratagene). The treated cultures were combined to a final volume of 20 ml. The mock treated cultures were set for 5 sec under red dimmed light. The treated cultures were stored in the dark at RT for 15 min and were re-incubated at 78°C and 150 rpm. Samples at different time points were used for microscopy, cell vitality and electron microscopy analysis.

*Electron microscopy and single-particle analysis*

For image processing, cells with attached pili were negatively stained with 2% uranyl acetate on glow-discharged carbon-coated copper grids. Electron microscopy was performed on a Philips CM120 electron microscope operating at 120 kV with a LaB6 filament. Images were recorded with a 4000 SP 4K slow-scan CCD camera at 60.000 x magnification at a pixel size of 5.0 Å at the specimen level with “GRACE” software (Oostergetel *et al.*, 1996). Single particle analysis was performed with the Groningen Image Processing (“GRIP”) software package on a PC cluster. Non-overlapping pili

segments were extracted from the micrographs and aligned with correlation methods. The aligned projections were treated with multivariate statistical analysis in combination with hierarchical classification before final averaging (Penczek *et al.*, 1992; van Heel *et al.*, 2000).

*Plasmid construction for expression in S. solfataricus and E. coli*

The genes *sso0117* and *sso0118* were cloned in the same arrangement as found in the genomic context into the virus based expression vector pMJ05 (Jonuscheit *et al.*, 2003) *via* the entry vector pMZ1 (Zolghadr *et al.*, 2007). The expression of both genes from the resulting plasmid pSVA96 could then be induced by the addition of 0.4% arabinose to the medium of transformed cells. To construct the expression plasmid for the signal peptide cleavage assay, SSO0118 was amplified using 118-forward-*NcoI* and 118-reverse-*BamHI* and cloned into pZA7, which added an HA-tag to the protein and resulted in pSVA133. By *Nco*I-*Hind*III restriction the sso0118-HA part was then transferred into pBAD/Myc-HisA yielding pSVA134. To achieve co-expression with the peptidase a fragment containing *pibD* under the control of the T7 promoter was cloned from pUC18-pibD into pSVA134 by *Sph*I restriction resulting in pSVA135. All plasmid vectors used in this study are listed in table S.6.2.

*Expression of SSO0117/118 in ΔFlaJ*

For the expression of SSO117/118 *Δ*FLaJ was grown to an $OD_{600\ nm}$ of 0.2. Cells were then transformed as described by Jonscheit *et al.*, 2003 with pSVA96. After two days the cultures were transferred to medium containing 0.4% arabinose to induce the expression of SSO0117/118. At an $OD_{600\ nm}$ of 0.5 cells were analyzed by electron microscopy.

*Construction of plasmids for the directed deletion of genes*

The up- and downstream flanking regions of *sso0120* were amplified using primer pairs KO-UP forward/KO-UP reverse and KO-DOWN forward/KO-DOWN reverse, respectively (Tab. S.6.3). The resulting fragments were cloned using *Kpn*I/*Nco*I for the upstream flanking region (1099 bp) or *BamH*I/ *Not*I for the downstream flanking region (924 bp) in pET2268, a vector containing the *lacS* cassette for selection, yielding pSVA37. Electroporation of the knockout plasmids and selection for correct deletion mutants were performed as described in Albers and Driessen, 2007.

*Southern blot*

Genomic DNA (8 ηg) was digested with the appropriate enzymes and separated on 0.8% agarose gel. The gel was equilibrated in 20x SSC and the DNA was transferred overnight to a positively charged nylon membrane (BIO-RAD, the Netherlands). DNA hybridization was performed in standard hybridization buffer. PCR products of both *lacS* and *sso120* gene were digoxigenin-labelled with the HighPrime Kit (Roche, the Netherlands). Detection was performed as recommended by the manufacturer using a LumiImager (Roche, the Netherlands).

*Expression analysis in knockout strain*

Total RNA isolation and cDNA synthesis were performed as described previously (Zolghadr *et al.*, 2007). Gene specific primer sets (1-7, Tab. S.6.3) were used to detect the presence of the genes in the ups-operon. PCR products were analyzed on 0.8 % agarose gels.

*Growth conditions and preparation of E. coli crude membranes*

BL21(DE3) (pLysS) *E. coli* cells were transformed with plasmids pSW017, pSW018, pSW019 and pSW020. The cleavage assay was performed as described before in (Szabó *et al.*, 2007a). At an $OD_{600\ nm}$ of 0.6 the expression of the precursor genes, *sso0117/0118*, was induced by addition of 0.2% L-arabinose for 2 h. Subsequently, the expression of PibD was induced with 0.1 mM IPTG (isopropyl-b-D-thiogalactopyranoside) for 2 h. The cultures were harvested the cell pellets were resuspended in 2 ml of buffer (50 mM Tris-HCl, pH 7.5, 1 mM EDTA). Crude membranes were isolated as described previously (Szabó *et al.*, 2007a) and resuspended in 50 mM Tris-HCl, pH 7.5. Cleavage of substrates was determined by SDS-PAGE and Western blot analysis of 5 µg of crude membranes. Substrate proteins were detected using monoclonal anti-hemagglutinin antibodies (Sigma).

*Treatment with UV–light*

For the UV-treatment all preparations were performed under red dimmed light. Aliquots of 10 ml *S. solfataricus* culture ($OD_{600\ nm}$ 0.3-0.5) were transferred to a 110 mm plastic petri dish and treated with a defined UV-dose (λ 254 nm, UV-Stratalinker 1800, Stratagene).

The treated cultures were combined to a final volume of 20 ml. The mock treated cultures were set for 5 sec under red dimmed light. The treated cultures were stored in the dark at RT for 15 min and were re-incubated at 78°C and 150 rpm. Samples at different time points were used for microscopy, cell vitality and electron microscopy analysis.

*Microscopy and quantitative analysis of cell aggregate formation*

Cell aggregate microscopy was performed as described in Fröls *et al.*, 2007. To quantify the formation of aggregates, the frequency of cells in aggregates and the amount of aggregates were counted until 1000 or 500 single cells were observed and at least seven fields of views were analysed for each time point. To exclude that the cellular aggregates were not artefacts of the microscopic slide preparation only fields of views were analysed where the cells showed an even spreading. For the statistic analysis the percentage of cells in aggregates ≥ 3 cells (to exclude the dividing pairs of cells), against the total amount of cells was calculated. Additionally the percentage of each aggregates size (from 3 to 15 cells) against the total amount of cells were analysed to observe the time and dose dependent formation of cellular aggregates in a higher resolution.

*Analysis of the cell vitality*

To analyse the cell vitality the LIVE DEAD Baclight $^{TM}$ (Invitrogen) assay was used under manufacture instructions. Alternatively, a combined DAPI propidium iodide stain was used. At 6 h after UV-treatment the 20 µl liquid cultures were mixed with 2 µl propidium iodide (1:30 dilution in 10 mM Tris HCL pH 7.5) and incubated for 15 min in the dark at room temperature. Microscopic slides were coated with 1% agar (10 mM Tris HCL pH 7.5) containing 0.2 µg/ml DAPI. 5 µl of the propidium iodide were spread on the coated slide and immediately examined. To analyse the amount of dead cells in aggregates in relation to the living cells in aggregates, a minimum of 50 cellular aggregates of ≥ 3 cells were counted for each UV-dose.

*Testing of various stress factors*

For the temperature shift, 20 ml aliquots of an exponentially grown *S. solfataricus* culture ($OD_{600\ nm}$ 0.3 - 0.5) were transferred from 78°C to 88°C or 65°C, additionally control

cultures were transferred to 78°C and 150 rpm. Samples for the quantitative analysis of the cellular aggregation were taken at 0 h, 1 h, 3 h, 6 h, 8 h, and 10 h after transfer.
For the pH shift, 20 ml of an exponential *S. solfataricus* culture were harvested for 10 min at 4000 rpm and 4°C. The supernatants were removed and cell pellets were resuspended in 20 ml of Brock's basal salt medium supplemented with D-arabinose (0.2%) and tryptone (0.1%) at pH 4, pH 2.5 and pH 3, respectively. Freshly inoculated cultures were incubated at 78°C and 150 rpm. Samples for the quantitative analysis of the cellular aggregation were taken at 0 h, 3 h, 6 h, and 8 h after pH-shift. For the treatment with the DSB inducing antibiotics bleomycin and mitomycin C, 20 ml of an exponential *S. solfataricus* culture were treated with 3 ug/ml bleomycin (Bleocin $^{TM}$, Calbiochem) or 5, 10 and 15 µg/ml mitomycin C (Sigma) and re-incubated at 78°C and 150 rpm. The treated cultures and the control cultures were plated on Brock's solid media at 1.5 h after re-incubation. Survival rates confirmed the use of a non-lethal drug concentration for both antibiotics as descript in the literature (Cannio *et al.*, 1998; Grogan *et al.*, 2001; Kosa *et al.*, 2004). Samples for the quantitative analysis of the cellular aggregation were taken at 0 h, 3 h, 6 h, 8 h and 10 h after drug treatment.

*Conjugation assay for S. solfataricus*

The assay was based on the transfer of the ß-galactosidase encoding reporter gene *lacS* from the wildtype strain P1 to the *lacS* deficient strain PH1-M16. To distinguish both strains on plates we selected for *pyrEF* auxotrophic cells with 5-FOA. Only derivate conjugates and PH1-M16 are able to grow under these conditions and only positive conjugates would additionally form blue colonies. For each conjugation test a minimum volume of 60 ml of an exponentially grown *S. solfataricus* culture ($OD_{600\ nm}$ 0.2 – 0.4) was used. The treatment was performed as described above under red dimmed light. Aliquots of 10 ml culture were poured into petri dishes and evenly spread by moderate shaking. Cultures were treated with UV-light in the UV-crosslinker (254 nm, 75 or 50 J/m$^2$). Control cultures were treated under red light only. All experimental cultures had a final volume of 20 ml, for the recombination mix the *S. solfataricus* strains P1 (*lacS* $^+$/ *pyrEF* $^+$) and PH1-M16 (*lacS* $^-$/ *pyrEF* $^-$) were mixed in a ratio of 1:1. Flasks were stored in the dark for 15 min at room temperature, 40 µl of uracil (12.5 µg /ml final concentration) were added and re-incubated at 78°C for 6 h and 150 rpm. Samples for plating were taken at 6 h after UV-treatment and diluted in Brock's basal salt pH 3, without carbon sources.
To determine the colony forming units (cfu) without selection, cells were plated on

Brock's basal salt solid media with D-arabinose (0.2%) and tryptone (0.1%) with 10 μg /ml uracil.

To identify the positive conjugative recombinants (*lacS* $^{-}$/ *pyrEF* $^{-}$), *pyrEF* auxotrophic mutations of strain P1 (*pyrEF* $^{+}$ to *pyrEF* $^{-}$) and *pyrEF* revertants of strain PH1-M16 (*pyrEF* $^{-}$ to *pyrEF* $^{+}$), 50 µg/ml 5-fluor-orotic-acid (5-FOA, Sigma; Grogan & Gunsalus, 1993) was added to the solid media. As negative control basic plates supplemented with 5-FOA but without uracil were used. For each single stain culture (P1 UV, PH1-M16 UV, P1 C, PH1-M16 C) three solutions were plated on selective (5-FOA) and non-selective plates. For the recombination mix (P1xPH1-M16 UV, P1xPH1-M16 C), three solutions were plated on non-selective plates and five solutions on selective plates (with 2-3 solutions in replicates for the P1xPH1-M16 UV cultures). Plates were incubated for 5-6 days at 78°C under wet atmosphere.

To identify the *lacS* (ß-galactosidase) genotype, of the positive conjugative recombinants (*lacS* $^{+}$/*pyrEF* $^{-}$), *lacS* $^{-}$ mutations of strain P1 (*lacS* $^{+}$ to *lacS* $^{-}$) and revertants of strain PH1-M16 (*lacS* $^{-}$ to *lacS* $^{+}$), the plates were sprayed with x-gal (10 mg/ml in DSMF) and incubated for 30 min at 78°C, only *lacS* $^{+}$ cells resulted in blue colonies.

To analyse the revertation and mutation frequencies event/cell, the median of the cfu/ml with selection (5-FOA) were determined and divided by the median of the cfu/ml under non-selective conditions. To identify *pyrEF* auxotrophic mutations of strain P1 (*pyrEF* $^{+}$ to *pyrEF* $^{-}$) 200 to 800 cells, for the *pyrEF* revertants of strain PH1-PH1-M16 (*pyrEF* $^{-}$ to *pyrEF* $^{+}$) 500 to 1500, cells were observed for each experiment. To identify *lacS* $^{-}$ mutations of strain P1 (*lacS* $^{+}$ to *lacS* $^{-}$) 700 to 2500 cells and revertations of strain PH1-M16 (*lacS* $^{-}$ to *lacS* $^{+}$) 1000 to 3000 cells were observed by microscopy for each experiment. To determine the recombination frequency event/cell the median of the positive conjugates (*lacS* $^{+}$/ *pyrEF* $^{-}$) cfu/ml with selection (5-FOA) were determined and divided by the median of the cfu/ml from all observed colonies under non-selective conditions. To identify positive ex-conjugates (*lacS* $^{+}$/*pyrEF* $^{-}$) cfu/ml with selection (5-FOA), in total of all UV-experiments >13000 cells were counted and 788 positive ex-conjugates were identified. In the case of the control experiments in total >11000 cells were counted and 5 conjugation events were identified, which correspond to the determined mutation frequency of P1 of $10^{-5}$ events /cell.

## 6.6 Supplementary data

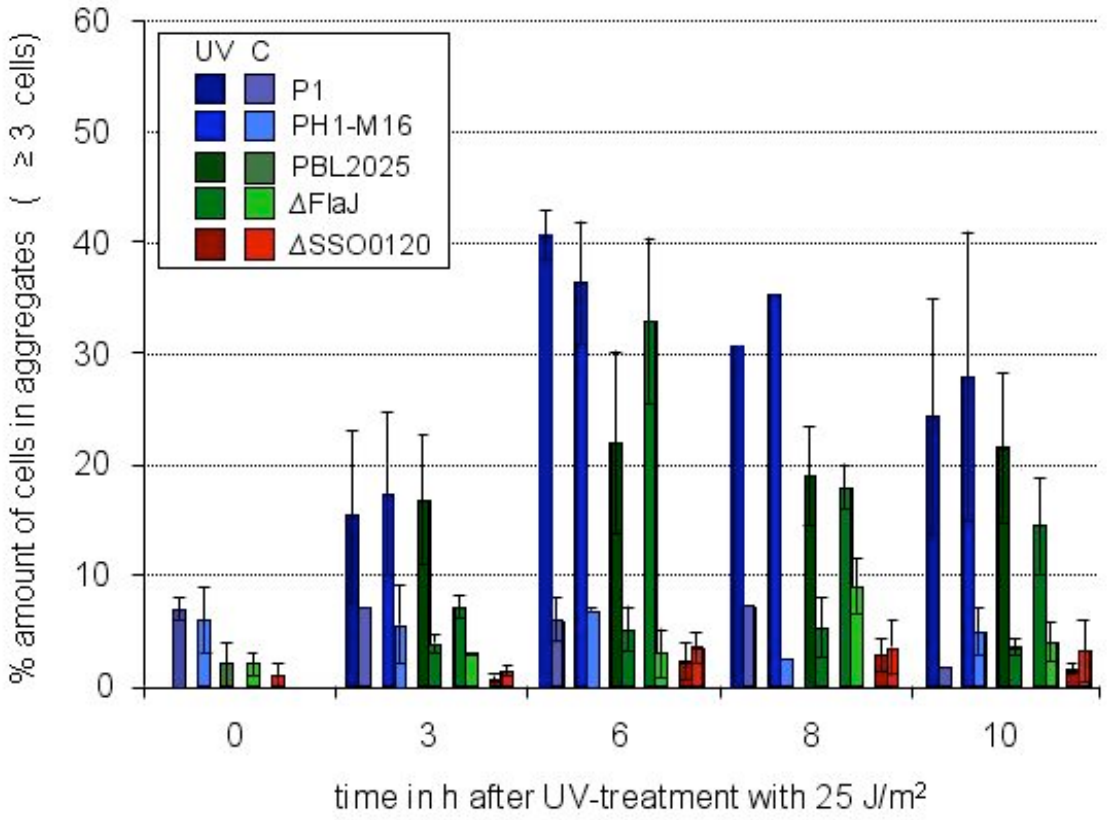



**Fig. S.6.1:** Quantitative analysis of the UV-induced cellular aggregation of different *S. solfataricus* strains. Cellular aggregation was observed at 3 h, 6 h, 8 h, and 10 h after UV-treatment with 25 J/m$^2$ (254 nm). The graph is based on three independent UV-experiments; in the case of the strains PBL2025, ΔFlaJ and ΔSSO0120 and 2 independent UV-experiments for the strains P1 and PH1-M16 (only one experiment at 8 h, respectively). The bars display the % amount of cells in aggregates (≥ 3 cells) in relation to the total amount of evaluated cells (1000 single cells, minimum of 500 single cells were counted, respectively).

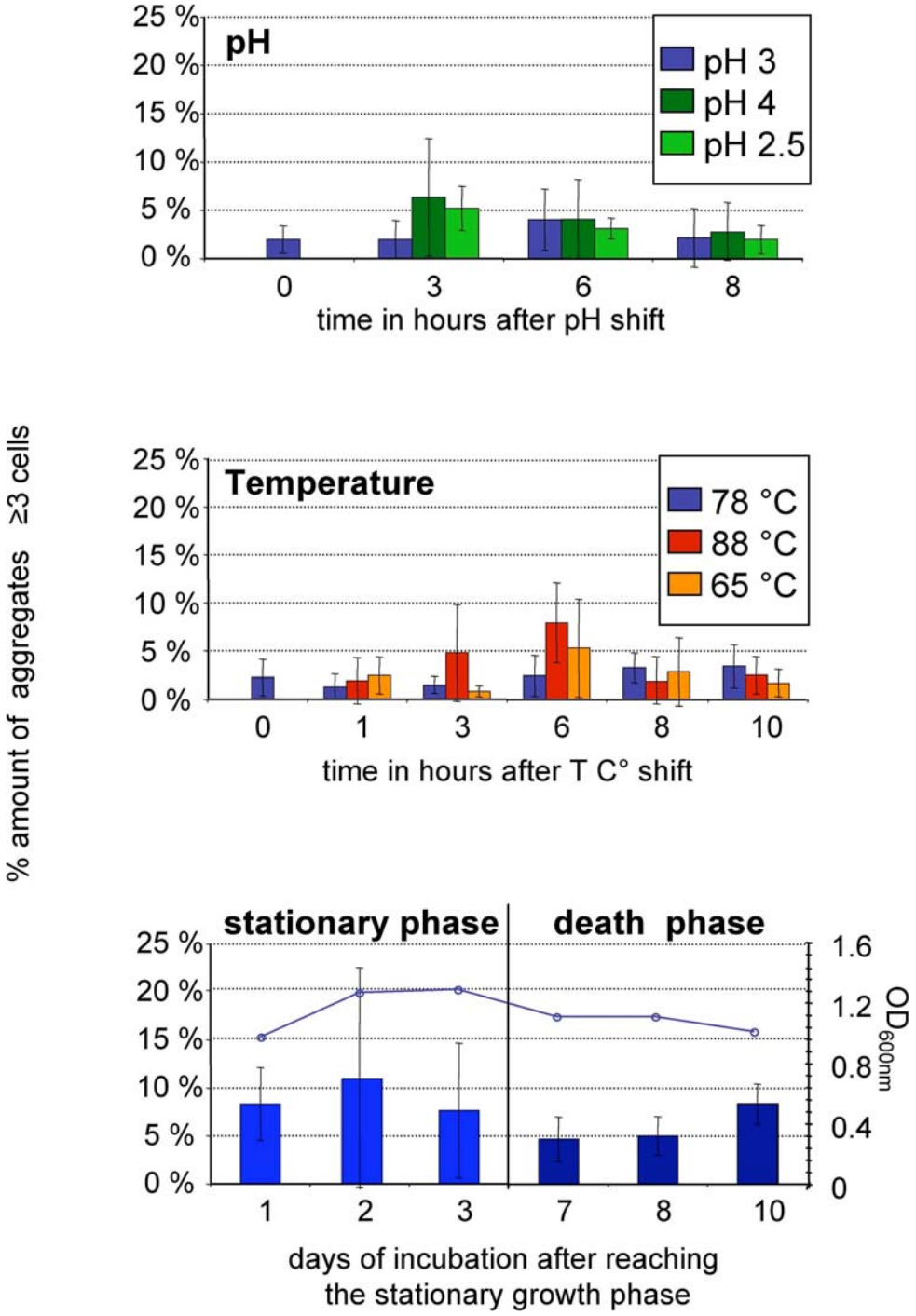



**Fig. S.6.2:** Quantitative analysis of the cellular aggregation under different environmental stressors and cell growth: non-lethal pH shift, from pH 3 to pH 4 and down to pH 2.5 and a non-lethal temperature shift from 78°C up to 88°C and down to 65°C, early stationary growth phase until death phase. The bars represent the mean of the results obtained for the four *sso0120* wildtype strains; P1, PH1-M16, PBL2025 and ΔFlaJ. The % amount of cells in aggregates (≥ 3 cells) in relation to the total amount of evaluated cells (500 single cells, or a minimum of 250 single cells were counted).

**Table S.6.1:** Conservation of the *ups*-operon in the order Sulfolobales

| **S. solfataricus** | Length in aa | ***S.tokodaii*** | ***S. acidocaldarius*** | ***Metalosphera sedula*** |
|---|---|---|---|---|
| SSO0121 | 695 | ST1396 | Saci_1493 | Msed_2103 |
| SSO0120 | 483 | ST1397 | Saci_1494 | Msed_2104 |
| SSO0119 | 481 | ST1398 | Saci_1495 | Msed_2105 |
| SSO0118 | 154 | ST1399 | Saci_1496 | Msed_1193 |
| SSO0117 | 137 | ST1400 | * | Msed_2107 |

* not official annotated, found by tblastn search on position 1275628-1275993 in the genome of *S. acidocaldarius* (aa identity 30%, aa similarity %60, score $2e^{-09}$, pos. 1-125)

**Table S.6.2:** List of Plasmids

| **Plasmid name** | **Description** | **Reference** |
|---|---|---|
| pET2268 | Vector containing *lacS* cassette | Szabó *et al.*, 2007b |
| pZA7 | Transfer vector to add HA-tag to proteins | Szabó *et al.*, 2006 |
| puC18-*pibD* | pUC18 containing pibD under the control of a T7 promoter | Szabó *et al.*, 2007a |
| pMJ05 | Virus based vector for S. solfataricus | Jonuscheit *et al.*, 2003 |
| pSVA37 | pET2268 containing up- and downstream flanking regions of *sso0120* | This study |
| pSVA96 | pMJ05 containing *sso117/118* under control of *araS* promoter | This study |
| pSVA133 | pZA7 containing *sso0118* | This study |
| pSVA134 | pBAD/*Myc*-HisA containing *sso0118*-HA | This study |
| pSVA135 | pSVA134 containing *pibD* | This study |

**Table S.6.3:** List of primers

| **Primer name** | **Sequenz (5'-3')** |
|---|---|
| 1 RT forward | ATAGGTCAAGTGATGGGTTA |
| 1 RT reverse | CATCTGCTGCAAGTATCTTT |
| 2 RT forward | GCCTATACGCATGGTTTCAC |
| 2 RT reverse | AAGGGTCAGCTAAGGGTACA |
| 3 RT forward | AGCAAGAAGATCACGTACTA |
| 3 RT reverse | CTGGAGTATCCTCTATGGTAAT |
| 4 RT forward | GATCTAGAAGAGTTCAGTGTT |
| 4 RT reverse | AGACCTTGGCTCTGCTTTCC |
| 5 RT forward | ACACAAGTGGTGAGTCAATA |
| 5 RT reverse | TTTGCAGCGAGTTCTCCTAA |
| 6 RT forward | AGGGCAGTTGGCAACTTAGA |
| 6 RT reverse | ATATCTGTGTGCTGCCGGTA |
| 7 RT forward | GCTGGGTGGTCTACTTTATG |
| 7 RT reverse | AGTACTGCCCAGCAGTTA |
| 118 forward-NcoI | CCCCCCCATGGTACAACTAATGATGAAAGGAGG |
| 118 reverse-BamHI | CCCCCGGATCCCGCTATTGAAGCCAGCA |
| KO-UP forward | CCCGGTACCGTGCGTATTATCTACGTTA |
| KO-UP reverse | CCCCCATGGCAGTGTTTATTTAAAGAA |
| KO-DOWN forward | CCCGGATCCGGAGAATATTCATGATAC |
| KO-DOWN reverse | CCCCCCCCCGCGGCCGCCGAGTGCAAAGATACTTG |

## 5.7 References

Albers, S.V. and Driessen, A.J. (2005) Analysis of ATPases of putative secretion operons in the thermoacidophilic archaeon Sulfolobus solfataricus. *Microbiology*, 151, 763-773.

Albers, S.V. and Driessen, A.J. (2007) Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome. *Archaea*, 2, 145-149.

Albers, S.V., Szabo, Z. and Driessen, A.J. (2003) Archaeal homolog of bacterial type IV prepilin signal peptidases with broad substrate specificity. *J Bacteriol*, 185, 3918-3925.

Averhoff, B. (2004) DNA transport and natural transformation in mesophilic and thermophilic bacteria. *J Bioenerg Biomembr*, 36, 25-33.

Averhoff, B. and Friedrich, A. (2003) Type IV pili-related natural transformation systems: DNA transport in mesophilic and thermophilic bacteria. *Arch Microbiol*, 180, 385-393.

Bardy, S.L. and Jarrell, K.F. (2002) FlaK of the archaeon Methanococcus maripaludis possesses preflagellin peptidase activity. *FEMS Microbiol Lett*, 208, 53-59.

Bardy, S.L., Ng, S.Y. and Jarrell, K.F. (2004) Recent advances in the structure and assembly of the archaeal flagellum. *J Mol Microbiol Biotechnol*, 7, 41-51.

Battin, T.J., Sloan, W.T., Kjelleberg, S., Daims, H., Head, I.M., Curtis, T.P. and Eberl, L. (2007) Microbial landscapes: new paths to biofilm research. *Nat Rev Microbiol*, 5, 76-81.

Bhattacharjee, M.K., Kachlany, S.C., Fine, D.H. and Figurski, D.H. (2001) Nonspecific adherence and fibril biogenesis by Actinobacillus actinomycetemcomitans: TadA protein is an ATPase. *J Bacteriol*, 183, 5927-5936.

Bhaya, D., Takahashi, A. and Grossman, A.R. (2001) Light regulation of type IV pilus-dependent motility by chemosensor-like elements in Synechocystis PCC6803. *Proc Natl Acad Sci U S A*, 98, 7540-7545.

Boetius, A., Ravenschlag, K., Schubert, C.J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jorgensen, B.B., Witte, U. and Pfannkuche, O. (2000) A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature*, 407, 623-626.

Cannio, R., Contursi, P., Rossi, M. and Bartolucci, S. (1998) An autonomously replicating transforming vector for Sulfolobus solfataricus. *J Bacteriol*, 180, 3237-3240.

Cohen-Krausz, S. and Trachtenberg, S. (2002) The structure of the archeabacterial flagellar filament of the extreme halophile Halobacterium salinarum R1M1 and its relation to eubacterial flagellar filaments and type IV pili. *J Mol Biol*, 321, 383-395.

Cohen-Krausz, S. and Trachtenberg, S. (2008) The flagellar filament structure of the extreme acidothermophile Sulfolobus shibatae B12 suggests that archaeabacterial flagella have a unique and common symmetry and design. *J Mol Biol*, 375, 1113-1124.

Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L. and Elie, C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Res*, 32, 1439-1447.

Davey, M.E. and O'Toole G, A. (2000) Microbial biofilms: from ecology to molecular genetics. *Microbiol Mol Biol Rev*, 64, 847-867.

Elasri, M.O. and Miller, R.V. (1999) Study of the response of a biofilm bacterial community to UV radiation. *Appl Environ Microbiol*, 65, 2025-2031.

Faguy, D.M., Jarrell, K.F., Kuzio, J. and Kalmokoff, M.L. (1994) Molecular analysis of archael flagellins: similarity to the type IV pilin-transport superfamily widespread in bacteria. *Can J Microbiol*, 40, 67-71.

Friedrich, A., Rumszauer, J., Henne, A. and Averhoff, B. (2003) Pilin-like proteins in the extremely thermophilic bacterium Thermus thermophilus HB27: implication in competence for natural

transformation and links to type IV pilus biogenesis. *Appl Environ Microbiol*, 69, 3695-3700.

Fröls, S., Gordon, P.M., Panlilio, M.A., Duggin, I.G., Bell, S.D., Sensen, C.W. and Schleper, C. (2007) Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage. *J Bacteriol*, 189, 8708-8718.

Garinis, G.A., Mitchell, J.R., Moorhouse, M.J., Hanada, K., de Waard, H., Vandeputte, D., Jans, J., Brand, K., Smid, M., van der Spek, P.J., Hoeijmakers, J.H., Kanaar, R. and van der Horst, G.T. (2005) Transcriptome analysis reveals cyclobutane pyrimidine dimers as a major source of UV-induced DNA breaks. *Embo J*, 24, 3952-3962.

Gasson, M.J. and Davies, F.L. (1980) High-frequency conjugation associated with Streptococcus lactis donor cell aggregation. *J Bacteriol*, 143, 1260-1264.

Ghigo, J.M. (2001) Natural conjugative plasmids induce bacterial biofilm development. *Nature*, 412, 442-445.

Graupner, S., Weger, N., Sohni, M. and Wackernagel, W. (2001) Requirement of novel competence genes pilT and pilU of Pseudomonas stutzeri for natural transformation and suppression of pilT deficiency by a hexahistidine tag on the type IV pilus protein PilAI. *J Bacteriol*, 183, 4694-4701.

Grogan, D.W. (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. *J Bacteriol*, 171, 6710-6719.

Grogan, D.W. (1996) Exchange of genetic markers at extremely high temperatures in the archaeon Sulfolobus acidocaldarius. *J Bacteriol*, 178, 3207-3211.

Grogan, D.W., Carver, G.T. and Drake, J.W. (2001) Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon Sulfolobus acidocaldarius. *Proc Natl Acad Sci U S A*, 98, 7928-7933.

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

Hopfner, K.P., Putnam, C.D. and Tainer, J.A. (2002) DNA double-strand break repair from head to tail. *Curr Opin Struct Biol*, 12, 115-122.

Jarrell, K.F., Bayley, D.P. and Kostyukova, A.S. (1996) The archaeal flagellum: a unique motility structure. *J Bacteriol*, 178, 5057-5064.

Jonuscheit, M., Martusewitsch, E., Stedman, K.M. and Schleper, C. (2003) A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector. *Mol Microbiol*, 48, 1241-1252.

Kachlany, S.C., Planet, P.J., Bhattacharjee, M.K., Kollia, E., DeSalle, R., Fine, D.H. and Figurski, D.H. (2000) Nonspecific adherence by Actinobacillus actinomycetemcomitans requires genes widespread in bacteria and archaea. *J Bacteriol*, 182, 6169-6176.

Kachlany, S.C., Planet, P.J., DeSalle, R., Fine, D.H. and Figurski, D.H. (2001) Genes for tight adherence of Actinobacillus actinomycetemcomitans: from plaque to plague to pond scum. *Trends Microbiol*, 9, 429-437.

Kagawa, H.K., Yaoi, T., Brocchieri, L., McMillan, R.A., Alton, T. and Trent, J.D. (2003) The composition, structure and stability of a group II chaperonin are temperature regulated in a hyperthermophilic archaeon. *Mol Microbiol*, 48, 143-156.

Kawakami, Y., Hayashi, N., Ema, M. and Nakayama, M. (2007) Effects of divalent cations on Halobacterium salinarum cell aggregation. *J Biosci Bioeng*, 104, 42-46.

Kawakami, Y., Ito, T., Kamekura, M. and Nakayama, M. (2005) Ca(2+)-dependent cell aggregation of halophilic archaeon, Halobacterium salinarum. *J Biosci Bioeng*, 100, 681-684.

Klemm, P., Hjerrild, L., Gjermansen, M. and Schembri, M.A. (2004) Structure-function analysis of the self-recognizing Antigen 43 autotransporter protein from Escherichia coli. *Mol Microbiol*, 51, 283-296.

Köhler, R., Schafer, K., Muller, S., Vignon, G., Diederichs, K., Philippsen, A., Ringler, P., Pugsley, A.P., Engel, A. and Welte, W. (2004) Structure and assembly of the pseudopilin PulG. *Mol Microbiol*, 54, 647-664.

Kosa, J.L., Zdraveski, Z.Z., Currier, S., Marinus, M.G. and Essigmann, J.M. (2004) RecN and RecG are required for Escherichia coli survival of Bleomycin-induced damage. *Mutat Res*, 554, 149-157.

Kuluncsics, Z., Perdiz, D., Brulay, E., Muel, B. and Sage, E. (1999) Wavelength dependence of ultraviolet-induced DNA damage distribution: involvement of direct or indirect mechanisms and possible artefacts. *J Photochem Photobiol B*, 49, 71-80.

LaPaglia, C. and Hartzell, P.L. (1997) Stress-Induced Production of Biofilm in the Hyperthermophile Archaeoglobus fulgidus. *Appl Environ Microbiol*, 63, 3158-3163.

Martinez, L.R. and Casadevall, A. (2007) Cryptococcus neoformans biofilm formation depends on surface support and carbon source and reduces fungal cell susceptibility to heat, cold, and UV light. *Appl Environ Microbiol*, 73, 4592-4601.

Martusewitsch, E., Sensen, C.W. and Schleper, C. (2000) High spontaneous mutation rate in the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by transposable elements. *J Bacteriol*, 182, 2574-2581.

Mattick, J.S. (2002) Type IV pili and twitching motility. *Annu Rev Microbiol*, 56, 289-314.

Mayerhofer, L.E., Macario, A.J. and de Macario, E.C. (1992) Lamina, a novel multicellular form of Methanosarcina mazei S-6. *J Bacteriol*, 174, 309-314.

Moissl, C., Rachel, R., Briegel, A., Engelhardt, H. and Huber, R. (2005) The unique structure of archaeal 'hami', highly complex cell appendages with nano-grappling hooks. *Mol Microbiol*, 56, 361-370.

Moissl, C., Rudolph, C., Rachel, R., Koch, M. and Huber, R. (2003) In situ growth of the novel SM1 euryarchaeon from a string-of-pearls-like microbial community in its cold biotope, its physical separation and insights into its structure and physiology. *Arch Microbiol*, 180, 211-217.

Molin, S. and Tolker-Nielsen, T. (2003) Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Curr Opin Biotechnol*, 14, 255-261.

Näther, D.J., Rachel, R., Wanner, G. and Wirth, R. (2006) Flagella of Pyrococcus furiosus: multifunctional organelles, made for swimming, adhesion to various surfaces, and cell-cell contacts. *J Bacteriol*, 188, 6915-6923.

Nutsch, T., Marwan, W., Oesterhelt, D. and Gilles, E.D. (2003) Signal processing and flagellar motor switching during phototaxis of Halobacterium salinarum. *Genome Res*, 13, 2406-2412.

O'Toole, G.A., Gibbs, K.A., Hager, P.W., Phibbs, P.V., Jr. and Kolter, R. (2000) The global carbon metabolism regulator Crc is a component of a signal transduction pathway required for biofilm development by Pseudomonas aeruginosa. *J Bacteriol*, 182, 425-431.

O'Toole, G.A. and Kolter, R. (1998) Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Mol Microbiol*, 30, 295-304.

Ojanen-Reuhs, T., Kalkkinen, N., Westerlund-Wikstrom, B., van Doorn, J., Haahtela, K., Nurmiaho-Lassila, E.L., Wengelnik, K., Bonas, U. and Korhonen, T.K. (1997) Characterization of the fimA gene encoding bundle-forming fimbriae of the plant pathogen Xanthomonas campestris pv. vesicatoria. *J Bacteriol*, 179, 1280-1290.

Oostergetel, G.T., Keegstra, W. and Brisson, A. (1998) Automation of specimen selection and data acquisition for protein electron crystallography. *Ultramicroscopy*, 74, 47-59.

Patel, R. (2005) Biofilms and antimicrobial resistance. *Clin Orthop Relat Res*, 41-47.

Peabody, C.R., Chung, Y.J., Yen, M.R., Vidal-Ingigliardi, D., Pugsley, A.P. and Saier, M.H., Jr. (2003) Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology*, 149, 3051-3072.

Penczek, P., Radermacher, M. and Frank, J. (1992) Three-dimensional reconstruction of single particles embedded in ice. *Ultramicroscopy*, 40, 33-53.

Planet, P.J., Kachlany, S.C., DeSalle, R. and Figurski, D.H. (2001) Phylogeny of genes for secretion NTPases: identification of the widespread tadA subfamily and development of a diagnostic key for gene classification. *Proc Natl Acad Sci U S A*, 98, 2503-2508.

Prangishvili, D., Albers, S.V., Holz, I., Arnold, H.P., Stedman, K., Klein, T., Singh, H., Hiort, J., Schweier, A., Kristjansson, J.K. and Zillig, W. (1998) Conjugation in archaea: frequent occurrence of conjugative plasmids in Sulfolobus. *Plasmid*, 40, 190-202.

Reisner, A., Holler, B.M., Molin, S. and Zechner, E.L. (2006) Synergistic effects in mixed Escherichia coli biofilms: conjugative plasmid transfer drives biofilm expansion. *J Bacteriol*, 188, 3582-3588.

Reiter, W.D., Zillig, W. and Palm, P. (1988) Archaebacterial viruses. *Adv Virus Res*, 34, 143-188.

Roine, E., Raineri, D.M., Romantschuk, M., Wilson, M. and Nunn, D.N. (1998) Characterization of type IV pilus genes in Pseudomonas syringae pv. tomato DC3000. *Mol Plant Microbe Interact*, 11, 1048-1056.

Rosenshine, I., Tchelet, R. and Mevarech, M. (1989) The mechanism of DNA transfer in the mating system of an archaebacterium. *Science*, 245, 1387-1389.

Sauvonnet, N., Vignon, G., Pugsley, A.P. and Gounon, P. (2000) Pilus formation and protein secretion by the same machinery in Escherichia coli. *Embo J*, 19, 2221-2228.

Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M. and Blum, P. (2004) Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon Sulfolobus solfataricus by use of gene disruption. *J Bacteriol*, 186, 427-437.

Schleper, C., Holz, I., Janekovic, D., Murphy, J. and Zillig, W. (1995) A multicopy plasmid of the extremely thermophilic archaeon Sulfolobus effects its transfer to recipients by mating. *J Bacteriol*, 177, 4417-4426.

Schleper, C., Roder, R., Singer, T. and Zillig, W. (1994) An insertion element of the extremely thermophilic archaeon Sulfolobus solfataricus transposes into the endogenous beta-galactosidase gene. *Mol Gen Genet*, 243, 91-96.

Schrenk, M.O., Kelley, D.S., Delaney, J.R. and Baross, J.A. (2003) Incidence and diversity of microorganisms within the walls of an active deep-sea sulfide chimney. *Appl Environ Microbiol*, 69, 3580-3592.

Shapiro, J.A. (1998) Thinking about bacterial populations as multicellular organisms. *Annu Rev Microbiol*, 52, 81-104.

Szabó, Z., Albers, S.V. and Driessen, A.J. (2006) Active-site residues in the type IV prepilin peptidase homologue PibD from the archaeon Sulfolobus solfataricus. *J Bacteriol*, 188, 1437-1443.

Szabó, Z., Sani, M., Groeneveld, M., Zolghadr, B., Schelert, J., Albers, S.V., Blum, P., Boekema, E.J. and Driessen, A.J. (2007b) Flagellar motility and structure in the hyperthermoacidophilic archaeon Sulfolobus solfataricus. *J Bacteriol*, 189, 4305-4309.

Szabó, Z., Stahl, A.O., Albers, S.V., Kissinger, J.C., Driessen, A.J. and Pohlschroder, M. (2007a) Identification of diverse archaeal proteins with class III signal peptides cleaved by distinct archaeal prepilin peptidases. *J Bacteriol*, 189, 772-778.

van Heel, M., Schatz, M. and Orlova, E. (1992) Correlation functions revisited. *Ultramicroscopy*, 46, 307-316.

Wall, D. and Kaiser, D. (1999) Type IV pili and cell motility. *Mol Microbiol*, 32, 1-10.

Wood, E.R., Ghane, F. and Grogan, D.W. (1997) Genetic responses of the thermophilic archaeon Sulfolobus acidocaldarius to short-wavelength UV light. *J Bacteriol*, 179, 5693-5698.

Zolghadr, B., Weber, S., Szabo, Z., Driessen, A.J. and Albers, S.V. (2007) Identification of a system required for the functional surface localization of sugar binding proteins with class III signal peptides in Sulfolobus solfataricus. *Mol Microbiol*, 64, 795-806.

# 7. General Discussion

This work addresses the transcriptional and cellular reaction of the archaeon *Sulfolobus solfataricus* to UV-light. All organisms exposed to sunlight have to deal with a constant damage of their DNA by the natural UV-irradiation. They therefore need efficient DNA-damage repair mechanisms. By comparative genome analysis, potential archaeal genes that are related to Bacteria and Eukarya involved in DNA-repair had been identified (Aravind *et al.,* 1999; White, 2003). But their function in DNA-repair and the role in Archaea are comparably little understood (Grogan, 2000; Kelman & White, 2005*)*.

Two different *S. solfataricus* strains, a non-infected and a SSV1 infected strain were used to analyse the global transcriptional reaction upon UV-exposure applying a DNA-microarray approach. The insights from the global transcriptional reaction were used further to develop new hypotheses, especially about the UV-damage DNA-repair mechanism in *S. solfataricus*. These hypotheses were subsequently aimed to verify by additional experiments. The analysis of the virus-infected strain made it also possible to study the transcriptional reaction of SSV1, which was known to be inducible by UV-light (Martin *et al.,* 1984). Furthermore the comparison of the infected strain with the non-infected strain was used to shed light on some specific virus-host interactions. In addition, it was attempted to study putative protein-protein interactions of SSV1 in a yeast two-hybrid screen.

This general discussion will briefly summarise the central results of this study and discuss them in the view of the response of SSV1 and *S. solfataricus* to UV-light. The obtained data are discussed in the context of the newest literature and a comparative analysis with relevant DNA-microarray studies is performed. Finally all relevant data of this study and of other groups are integrated to present a general model about the UV-response in *S. solfataricus*.

## 7.1 The UV-induced proliferation of the *Sulfolobus shibatae* virus 1

By investigation of the transcriptional reaction of the virus to UV-light, a complete and detailed scenario of the viral cycle was obtained. SSV1 exhibited a chronological transcription cycle after UV-induction, starting 1 h after UV-exposure of infected *S. solfataricus* cells (see summary Fig. 3.9).

This UV-induced transcription cycle is rather complex compared to the other up to now described crenarchaeal viruses or conjugative plasmids. The crenarchaeal viruses SIRV1 and SIRV2, for example, showed a comparably simple cycle with the simultaneous transcription of nearly all ORFs shortly after infection (Kessler *et al.*, 2004). Also the plasmid pSSVx, which propagates like a satellite virus together with SSV2, showed only slight changes in the transcriptional activity after induction of replication in the late exponential growth phase (Contursi *et al.,* 2007).

### 7.1.1 The UV-induced proliferation of SSV1 is correlated with the host response

The data of this study give indications that SSV1, and more precisely, the first transcript, T-ind, is activated *via* a DNA-damage-responsive host factor. We infer this, because the transcriptome data from the infected strain *S. solfataricus* showed, that T-ind has the same transcriptional pattern as the highly induced UV-dependent genes (Fig. 3.7). Furthermore, the replication inhibitor mitomycin C is the only known SSV1 inducing factor so far beside UV-light (Reiter *et al.,* 1988c). Other stressors like temperature or pH shifts could not induce the virus proliferation (C. Schleper, personal communication). Both treatments cause double-strand breaks in DNA, indicating that T-ind is specifically activated by an interaction with a cellular DNA-damage response.

Comparative genome analysis with other SSV types showed that T-ind is only found in SSV1 (Wiedenheft *et al.*, 2004) and not in other homologous SSV type viruses. The latter are also not inducible by UV-light. This makes SSV1 a unique model system for the study of DNA-damage dependent virus proliferation in Archaea.

### 7.1.2 T-ind may be involved in the UV-induced transcription and replication

Earlier studies of W. D. Reiter proposed, that, T-ind acts on the RNA level and does probably not translate into a protein, due to the small predicted size of only 49 aminoacids (Reiter *et al.,* 1987). However, data based on the yeast two-hybrid screen performed in this study suggested that the product of T-ind forms homo-dimers (or possibly also homo-multimers, Fig. 4.1.), a suggestion that needs experimental biochemical verification. In principle, the results from the yeast-two-hybrid screen all remain preliminary without further biochemical studies. But one has to note that the observed homo-dimerization of D-335 (the viral integrase) and F-93 (a putative transcription factor) confirm known results from earlier biochemical and crystallisation studies (Muskhelishvill *et al.*, 1994; Kraft *et al.*, 2004).

According to the observed transcriptional cycle, the protein product of T-ind could act in the transcriptional activation of the early transcripts T5 and T6 (and probably also T9) (Fig. 3.4) or in the replication of SSV1. Similar to what is shown for the large T-antigen from the simian virus 40, which activates the viral transcription and replication (Borowiec *et al.,* 1990; Coulombe *et al.*, 1992; Fradet-Turcotte *et al.*, 2007). In case of the replication the T-antigen binds to the viral origin and mediates the initiation of replication by interacting with replication protein A (RPA) (Han *et al.,* 1999).

### 7.1.3 SSV1 probably encodes its own transcriptional repressor

Protein B-115 of SSV1 shows similarities on the sequence level to an ArsR-like transcriptional repressor and is encoded by the last induced transcript of the transcriptional cycle, as a late extended version of the transcript T7/8.

A recently published study of Qureshi *et al.* (2007) gives also hints for the presence of a SSV1-encoded repressor. The identified protein showed a specific binding *in vitro* on both 11nt long direct repeats in the promoter regions of the transcripts T5 and T6. It was therefore termed STRIP for **S**SV1-**tr**anscript-**i**nteracting-**p**rotein. The binding sites masked the BRE and the initiator region. In addition STRIP caused bending of the DNA, which suggests a negative influence on the transcriptional initiation. However, the direct effect of STRIP on DNA binding is not clear, as repression in an *in vitro* transcription assay was apparently only observed at high concentrations of the protein (Qureshi *et al.*, 2007). It

would be very interesting for further studies to investigate if STRIP is encoded by ORF B-115 and if it is responsible for the transcriptional down-regulation or early termination of the early transcripts T5/T6 at the end of the UV-induced cycle of SSV1.

### 7.1.4. First insights into the virus-host interaction after UV-induction

To characterize specific virus-host interaction during the response to UV-light, the obtained data of the infected and non-infected *S. solfataricus* strains were compared and 22 genes were identified with a significantly different strength in the transcriptional reaction (Tab. 3.1). Only six genes showed an exclusive transcriptional reaction in the infected strain after UV-treatment (Fig. 3.8). Among these, only two genes encoding the subunits of topoisomerase VI could be assigned a clear function. Its potential role in DNA topology for SSV1 has been discussed in chapter 3.

During the production of SSV1 particles the cells showed an extended growth retardation in comparison to the non-infected cells (Fig. 3.1). This extended cell cycle arrest may be associated with the genes *sso2750* (a putative ATPase) and *sso2751* (a putative kinase). Both genes were identified as the strongest repressed genes in the infected strain and further analysis of the genomic region suggested an operon-like structure. A cell cycle dependent cyclic up-regulation was reported of the homologous genes in *S. acidocaldarius* (*saci*1228 *and saci*1229) and it was suggested that these genes are probably involved in mitosis, i.e. in division events during the G2 phase (Lundgren *et al.*, 2007). Because the regulation of the cell cycle is unknown in Archaea, it might be very interesting to study these two genes in more detail.

### 7.1.5 Conclusions and outlook

Because this was the first microarray analysis of a hyperthermophilic crenarchaeote so far, it was not clear in the beginning of the study if the presence of e.g. RNA binding proteins allowed us to generate representative data to study transcriptional changes. The analysis of the transcriptional regulation of SSV1 after UV-induction supplied data with a high resolution. We were able to describe the tightly regulated transcriptional cycle of SSV1, as it has not been seen before. The data corresponded nicely with the earlier studies of W.D. Reiter (Reiter *et al.,* 1987; 1988a; 1988b; 1989; 1990). In addition, based on the microarray data a new SSV1 transcript called Tx was identified (Fig. 3.4, Fig. 3.5),

the time point of replication was verified and first insights into the virus-host interaction were observed.

The results of this study represent a good basis for further studies aimed at elucidating the UV-induced regulation of the virus and at understanding viral gene functions.

## 7.2 The transcriptional and cellular reactions of *S. solfataricus* to UV-light

In this study it was observed that *S. solfataricus* exhibits a complex and specific transcriptional reaction to UV-light with 55 genes involved. The genes were quite diverse in function, involved in e.g. transcription, replication, repair, cell cycle and secretion processes. They clustered into three groups of different strength and transcriptional reaction types: 19 were highly induced genes, 14 were moderately induced and 22 genes were repressed in reaction to UV-damage (Tab. 5.2).

### 7.2.1 Genes involved in replication and transcription

Of genes involved in replication the most dominant reaction was observed for the ORC1/cdc6 encoding genes (Fig. 5.6). While Cdc6-1 and Cdc6-3 were proposed to promote replication, Cdc6-2 may act as a negative regulator for replication (Robinson *et al.*, 2004). The strong transcriptional up-regulation of *cdc6-2* and the simultaneous down-regulation of *cdc6-1* indicated a repression of the replication initiation during the UV-response of *S. solfataricus*.

With exception of the DNA-Polymerase 2 (Dpo2), no further transcriptional reaction of genes involved in replication were observed, like e.g. those encoding Dpo1, Dpo3, the sliding clamp (PCNA), the clamp loader RFC and the MCM helicase. This observation was very interesting, because until now the exact role of the three different polymerases and in particular of Dpo2 is not clear, in *S. solfataricus.* Dpo1 and Dpo3 are found in both kingdoms of the *Archaea*, whereas Dpo2 is unique to the Crenarchaeota (Edgell *et al.*, 1997; Cann & Ishino 1999). For Dpo1 there exists evidence for an involvement in Okazaki-fragment processing (Dionne *et al.*, 2003) and the Dpo3 probably mediates leading strand synthesis (Duggin & Bell, 2006), but the role of the Dpo2 was still unclear.

The result of this study indicated for the first time a possible function of Dpo2 in DNA-repair or recombination, respectively.

The minimally required factors for the *in vitro* initiation of transcription for *Sulfolobus* are the TBP (TATA binding protein), the RNA-polymerase and TFB (transcription factor 2B homolog) (Bell *et al.*, 1998). Whereas most of the *Archaea* have only one copy of TFB, *S. solfataricus* possesses three copies of TFB proteins, which probably regulate the transcription of different gene sets (Baliga *et al.*, 2000; Shockley *et al.*, 2003). Surprisingly *tfb-3* (*sso0280*) showed a strong UV-dependent up-regulation comparable with the reaction of *cdc6-2*. This exclusive reaction might indicate a regulatory role in the UV-dependent transcriptional regulation. However in comparison to TFB-1 and TFB-2, TFB-3 is truncated and lacks the B finger domain for transcriptional initiation and for binding to the BRE (Götz *et al.*, 2007). Therefore, it is unclear if TFB-3 is able to act as a transcriptional regulator. A current model suggests, that TFB-3 may acts as a competitive inhibitor of transcription initiation. It was shown *in vitro* that TFB-3 interacts with the RNA-polymerase and that it competes with TFB-1 at increasing concentrations (Götz *et al.*, 2007).

### 7.2.2 Transcriptional reactions of the predicted UV-damage repair genes in *S. solfataricus*

Quite little is still known about the UV-damage repair mechanisms in the domain Archaea. Hints for putative repair enzymes were obtained from comparative genomic analyses with Bacteria and Eukarya (Aravind *et al.*, 1999) and some of the predicted enzymes were biochemically analysed (reviewed in White, 2007). This present study attempted to analyse the UV-response based on the transcriptional reaction of the genes. We did therefore not expect that our study would identify all genes involved in DNA repair mechanisms, but rather only those that are not constitutively expressed.

After UV irradiation the first mechanism likely to be active is the photoreactivation; a mechanism that mediates the light-dependent direct reversion of the UV products in DNA (cis-syn-cyclobutane pyrimidine dimers, i.e. CPD) by the action of the photolyase (Phr), (White, 2007). The recently analysed crystal structure of the first archaeal photolyase from *S. tokodaii*, showed that the structure is highly similar to those of *E. coli*, *Anacystis nidulans* and *Thermus thermophilus* (Fujihashi *et al.*, 2007). A recent study demonstrated that *S. solfataricus* is able to remove 50% of the CPD within the first 30 min after UV-

irradiation, whereas the removing rate was 3 to 5-fold decreased without light (Dorazi *et al.*, 2007). However, no transcriptional induction of the corresponding *phrB* gene (*sso2472*) was observed. These data suggest that the PhrB protein is constantly present in the cell and that its transcriptional regulation is not coupled to UV-damage. Similar results were observed for the trans lesion DNA-polymerase (Dpo4) of *S. solfataricus*. The Dpo4 is able to bypass DNA-lesions like CPDs and 6-4 photoproducts (Boudsocq *et al.*, 2001). No specific transcriptional induction was observed after UV-treatment. This might confirm the previously postulated accessorial function during the replication of *S. solfataricus* (Gruz *et al.*, 2003) and not a UV-specific regulation.

One of the biggest open questions in Archaea is the existence and character of a nucleotide excision repair system (NER) system, which removes photoproducts from DNA (Kelman & White, 2005). Most archaea have homologues to the eukaryotic NER nucleases XPF (Rad1) / XPG (Rad2) and helicases XPB (Rad25) / XPD (Rad3), whereas homologous genes to the bacterial-like UvrABC and mismatch repair pathway are only present in mesophilic archaea (Grogan, 2000).

For *S. solfataricus*, a UV-caused transcriptional induction of the potentially NER-associated genes XPF, XPG and XPBI was reported after a treatment with 200 J/m$^2$ (Salerno *et al.*, 2003). This is in contrast to the present study where no UV-dependent induction of these genes was observed. They were also not found to be induced in the global transcriptional study of Götz *et al.* 2007, which used 200 J/m$^2$. However, in our study the corresponding genes of the NER system did show a transcriptional induction (XPF and XPG) before and after the UV-dependent response (at 1.5 h and 5 h). Thus, they followed the pattern of the constitutive, highly transcribed "house-keeping" genes, like e.g. those involved in transcription and translation (see Fig. 5.5). We therefore assume, that in the (non-genome-wide) study of Salerno *et al.* (2003) it was not possible to distinguish between transcriptional induction based on UV-light as opposed to transcriptional enhancement of house-keeping genes that was based on some synchronization of the cells after UV irradiation.

Although no UV-dependent transcriptional activity of the potential NER system was observed in this study, two recently published studies demonstrated a functional NER mechanism in *S. solfataricus* (Romano *et al.,* 2006; Dorazi *et al.,* 2007). The removal of

photoproducts, during permanent dark incubation was demonstrated *in vivo* for *S. solfataricus.* Interestingly, no evidence for a transcription coupled repair (TCR) system was observed. This is in contrast to analyses in *E. coli, S. cerevisiae* and human cells, where the TCR repair is at least two-fold more efficient than the global genomic repair system (Dorazi *et al.,* 2007).

Of all genes potentially involved in DNA-repair systems only the genes belonging to the archaeal *rad50*/*mre11* operon showed a UV-dependent transcriptional induction (Fig. 5.8). In Eukarya and Bacteria the related proteins Rad50/Mre11 and SbcC/SbcD are involved in double strand break (DSB) repair *via* homologous recombination (HR) or non-homologous end joining (NHEJ), DNA-damage detection and cell cycle checkpoint signalling (D´Amours & Jackson, 2002). All archaeal genomes possess homologues of the proteins Rad50 and Mre11 (Shin *et al.,* 2004). The core nuclease Mre11 exhibits strand dissociation and strand annealing properties, with an intrinsic DNA binding activity. The nuclease activity is regulated by sequence homology of the DNA substrates and by the interaction with Rad50, which probably binds DNA ends and brings them together (D´Amours & Jackson, 2002). In most thermophilic archaea *rad50* and *mre11* are clustered with the two genes *nurA* and *herA*.: NurA is a 5´ to 3´ exonuclease and HerA is a DNA helicase able to utilize both 3´ or 5´ single-stranded DNA extensions (Constantinesco *et al*., 2004). In contrast to the genes of the *rad50*/*mre11* operon, the *radA* gene (*sso0250*) characterised to catalyze the process of homologous DNA-paring and strand exchange (Seitz *et al*., 1998) did not show a UV-dependent pattern (Fig. 5.8). The transcriptional reaction followed, similar to the genes of the NER system, the pattern of the constitutive, highly transcribed genes. These results were supported by observation of Sandler *et al*., (1996), who did also not observe increased levels of *radA* mRNA after UV-irradiation (10 J/$m^2$) in *S. solfataricus*. The significance of the various repair mechanisms and in particular the UV-inducible transcription of the mre11 operon are discussed in a broader context in the last chapter of the discussion.

### 7.2.3 Comparison of DNA-microarray analyses in *S. solfataricus*

A critical question regarding DNA-microarray approaches is, how comparable are these kind of studies, in particular when they are performed by different laboratories or on different scientific issues. Reproducibility of quantitative results becomes particularly an issue for systems biology approaches and their data management, because different

experiments from various laboratories are usually included in the global datasets. Since two microarray studies of *S. solfataricus* related to ours were recently published, we had the opportunity to compare and analyse the datasets and obtain an impression about the robustness of the different experimental approaches.

The first comparison was performed to the study of M. White and co-workers on the transcriptional reactions of *S. solfataricus* and *S. acidocaldarius* to a higher UV-dose of 200 J/m$^2$ (Götz *et al.,* 2007). While the experimental setup and the used DNA-microarrays were different; the identified UV-dependent regulated genes of *S. solfataricus* and *S. acidocaldarius* were highly overlapping with those identified in this work. This is even more remarkable, as we have used a UV-dose of only 75 J/m$^2$. The comparison of the UV-dependent regulated genes after UV-treatment with 75 J/m$^2$ (Fröls *et al.*, 2007) versus 200 J/m$^2$ (Götz *et al.*, 2007) showed > 90% overlap in the genes (45 out of 49, ORFs) (Fig 7.1). This means, for example, that 17 out of 19 most strongly induced genes of our study are found among the top induced genes of the Götz et al. study (6 more genes of our top list were not represented on the microarray study of Götz *et al.*, as they were < 300bp in length). And *vice versa*: 14 out of 20 most induced genes found in Götz *et al.*, are also found in our study. In consensus with our study, Götz *et al.*, observed a UV-dependent reaction of genes like e.g. *cdc6-2* (*sso0771*), *dpo2* (*sso8124/1459/1458*), *tfb-3* (*ss*o0280), the putative transmembrane proteins *sso0691* and *sso03146* or the genes of the *ups*-operon. From genes involved in DNA-repair, only genes of the *rad50*/*mre11* operon showed a slightly UV-dependent transcriptional up-regulation.

Due to the higher UV-dose of 200 J/m$^2$ (Götz *et al.*, 2007) additional transcriptional reactions were observed in the second study. The genes *sso2078-2080* probably involved in the detoxification of DNA-damage by reactive oxygen species (ROS) showed an up-regulation in both strains. None of these genes showed a transcriptional reaction in our study after the irradiation with 75 J/m$^2$. We had also performed a single experiment with 200 J/m$^2$ (Fröls *et al.*, unpublished) and a subtractive analysis by using the present data set (75 J/m$^2$) was done to investigate UV-dose dependent reactions. Interestingly, the *sso2078-2080* genes were identified as the most pronounced differences between both UV-doses.

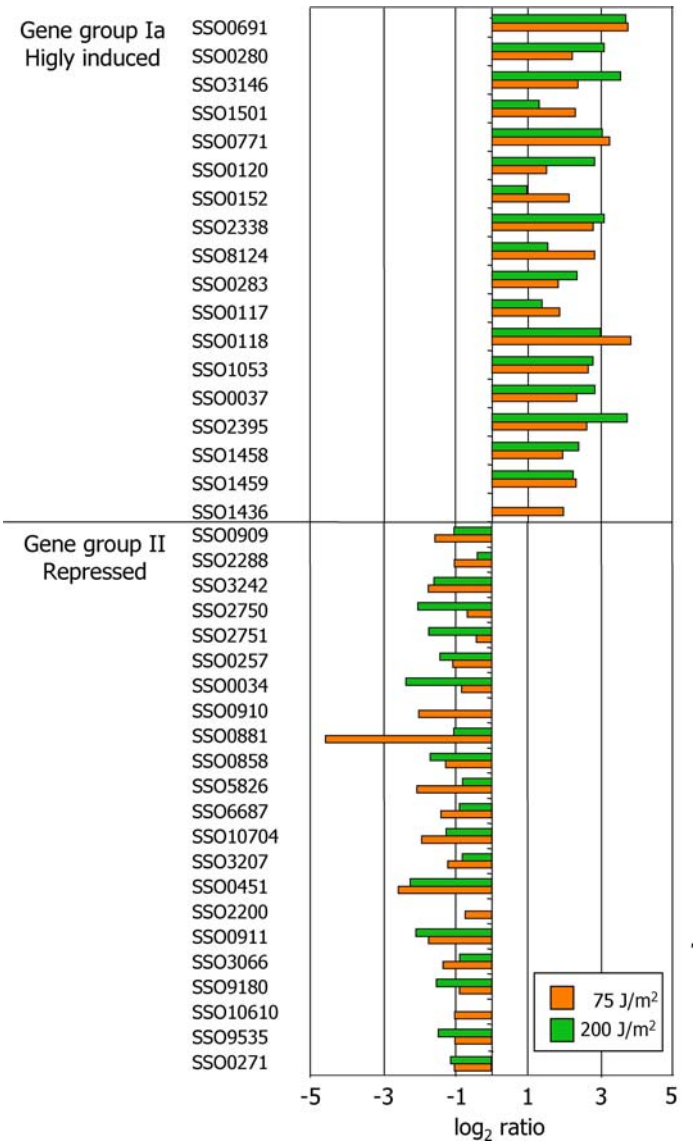



Fig. 7.1: Data comparison of the UV-dependent regulated genes of *S. solfataricus* from two independent DNA-micorarray studies (this study versus Götz *et al.,* 2007). The reaction of the highly induced and repressed UV-dependent regulated genes of this study were used for the comparison of both datasets. The orange bars display the observed transcriptional reaction after a UV-irradiation with 75 J/m$^2$ (this study) and the green bars with 200 J/m$^2$ (Götz *et al.,* 2007).

Another difference caused by higher UV dose might be the up-regulation of the genes for the beta–carotene biosynthesis (*sso2905* and *sso2906*) that were observed by Götz *et al*., (2007). The production of pigments probably represents an additional protection mechanism to UV-light. In the present study no transcriptional reaction was observed, while an up-regulation at 0.5 h was also seen in our single 200 J/m$^2$ dataset.

The second comparison was performed with the study of M. Lundgren *et al*. (2007), in which the cell cycle-dependent regulated genes of *S. acidocaldarius* were investigated. The comparison showed that only 4 of the UV-dependent up-regulated genes were also cell cycle-dependent regulated. In comparison nearly 80% (18 out of 22) of the UV-dependent repressed genes were shown to be also cell cycle dependent regulated. (Lundgren *et al.,* 2007).

The comparative analysis with Lundgren *et al*. (2007) indicated that, the UV-dependent cellular reaction induced a transcriptionally regulated cell cycle arrest. This corresponds to the observed growth retardation during the UV-dependent response from 2 h to 5 h after treatment (Fig. 3.1 and Fig. 5.1).

The comparison of the DNA-microarray data of this study with two other approaches with *S. solfataricus* and *S. acidocaldarius* (Götz *et al*., 2007; Lundgren *et al*., 2007) showed a high comparability and revealed the possible integration of datasets from different experimental approaches.

### 7.2.4. UV-irradiation causes double-strand breaks in DNA

By analysing the reaction of the putative repair systems it was observed that only the genes of the *rad50*/*mre11* operon showed a UV-dependent transcriptional pattern. Since these genes are involved in processing double-strand breaks (DSB) in DNA (see below), this finding suggested the significant presence of DSB in the chromosomal DNA upon UV-exposure. Indeed, a fragmentation of the DNA was observed in this study. It did not seem to directly cause by the UV-treatment but was most probably due to cellular reactions (Fig. 5.4). Our experiments showed for the first time for Archaea that UV-treatment causes DSB in DNA, as was found earlier in Eukarya and Bacteria (Bonura *et al.,* 1975; Garinis *et al*., 2005; Haber, 2006). The formation of DSB is probably a result of stalled or collapsed replication forks caused by unremoved photoproducts (Kuzminov, 2001). This observation may imply that the NER system of *S. solfataricus* is not efficient enough to repair the DNA-damage directly caused by the UV-irradiation. Thie effect might be more apparent in *S. solfataricus* rather than in Bacteria or Eukarya, due to the lack of a transcription coupled repair NER system. However these observations and the studies mentioned above, suggest that the formation of DSB in DNA is a general, secondary DNA-damage effect of UV-irradiation in all three domains of live.

*S. solfataricus* contains several proteins known to be involved in recombination processes like Mre11, Rad50, RadA and Hj. However the general mechanism is still unclear. In Eukarya the mechanism of non-homologous end joining (NHEJ) is the major pathway to remove DSB DNA (Hefferin & Tomkinson, 2005). The repair of DSB *via* NHEJ occurs during the G0/G1 phase of the cells and is catalysed by the DNA-dependent kinase (DNA-

PK) and DNA ligase IV/XRCC4 (Shin *et al*., 2004). No homologues to the DNA-PK are known for Archaea so far. In the S and G2 phase of the eukaryotic cell cycle the homologous recombination pathway predominates, which requires homologous chromosomal DNA as a template. The major proteins involved in homolgous recombination are Mre11/Rad50, in complex with NBS1 (in human) / XRS2 (in yeast), Rad51/RecA (RadA) and RPA (Shin *et al*., 2004). In yeast and in Bacteria the homologous recombination is predominant (Hefferin & Tomkinson, 2005), while the situation in Archaea is unclear. The study of Quaiser *et al*. (2008) showed recently that the proteins Rad50, HerA, Mre11 and RadA are constantly present in exponentially grown *S. solfataricus* cells and 50% of Rad50 was found as DNA-bound molecules. It was also shown, that Mre11 interacted with Rad50 and HerA. In addition both gamma irradiation and a stop of replication recruited Mre11 (and RadA) to the DNA (Quaiser *et al*., 2008). These results strongly suggest that in *S. solfataricus* the Rad50/Mre11/HerA protein complex is involved in DSB repair *via* homologous recombination.

### 7.2.5. How is UV-damage sensed in *Sulfolobus*?

The sensing of DNA-damage is crucial for the survival of the cell, because non-repaired DNA-damage can lead to mutations and genomic rearrangements (Jackson, 2002). The existence of a UV-dependent transcriptional response in *S. solfataricus* strongly suggests a DNA-damage-specific recognition mechanism. In *Bacteria* the SOS system is responsible for the DNA-damage recognition. RecA, bound to ssDNA gaps and DSB ends, replaces the single strand binding protein (SSB) and stimulates the auto-proteolysis of LexA, the repressor of the damage inducible genes (Meyer & Laine, 1990). In eukaryotic cells the two DNA-damage checkpoint kinases, ataxia-telangiectasia (ATM) and ATM- and Rad3-realted (ATR) are able to sense DNA-damage and recruit further response regulators, which in turn affect the cell cycle arrest, DNA-repair and apoptosis (D´Amours & Jackson, 2002; Bradbury & Jackson, 2003). In response to DSB DNA or stop of replication the ATM and ATR are recruited to DSB DNA by the Mre11-Rad50-Nbs1 complex. It was observed that, Mre11 processes ssDNA ends, which were coated by RPA. In the case of stalled replication forks, RPA accumulates on the ssDNA (Jazayeri *et al.,* 2006).

All archaea possess SSB proteins, which are more similar to the eukaryotic RPA than to the bacterial SSB (Wadsworth & White, 2001). In response to UV-light the overall transcriptional reaction of the *ssb* gene followed the pattern of the constitutive, highly

transcribed genes. However, it is notable that the transcription of *ssb* was up to 8-fold induced, perhaps indicating a very early reaction to UV-damage. Like in Bacteria and Eukarya the SSB has been implicated in DNA-damage recognition. In presence of DNA lesions the SSB of *S. solfataricus* is able to recognise and melt DNA duplexes containing CPD. In addition it was observed *in vitro* that the SSB-coated DNA is able to interact with putative repair proteins including Rad50 and NurA, (involved in homologous recombination) and XPBI (involved in NER) (Cubeddu & White, 2005; Wei *et al.,* 2007).

### 7.2.6 The UV-induced cellular aggregation and conjugation of *S. solfataricus*

Of special interest was the observed UV-induced cellular aggregation of *S. solfataricus* (Chapter 6). It was verified that the cellular aggregation is dependent of the pili, encoded by the UV-dependent induced genes of the *ups*-operon (**U**V-inducible typeIV **p**ili operon of **S**ulfolobus). Furthermore it was shown that cellular aggregation is also induced by treatment with DSB-inducing argents (mitomycin C and bleomycin) (Fig. 6.8, Tab. 6.2), but not by other conditions that simulate environmental stress (temperature and pH shifts) (S. 6.2). This observation is compatible with the UV and mitomycin C-inducible proliferation of the SSV1 replication (Reiter *et al.*, 1988c). The biological function of the UV-induced cellular aggregation of *S. solfataricus* is still not clear. The observation that even very low doses of UV-light down to 5 J/m$^2$ (Fig. 6.4) induced the reaction, implies that cellular aggregation may be a natural response of the cells and probably represents a protection mechanism like the stress-induced biofilm formation of *Archaeoglobus fulgidus* (LaPaglia et. al. 1997). Alternatively (or additionally) it might mediate a mechanism to overcome DNA-damage, such as conjugation.

In addition to the cellular aggregation a UV-dependent conjugative activity of *S. solfataricus* was identified (Tab. 6.3). Frequent conjugation events in *S. solfataricus* were only observed after UV-irradiation. Using a UV-dose of 75 J/m$^2$ a recombination frequency of up to 1.11 x $10^{-2}$ events/cell was determined. Perhaps the most interesting findings leading to novel hypotheses in the present study are the observed DNA-fragmentation, cellular aggregation and mediated conjugation in response to UV-irradiation and that these processes might be connected and represent a novel mechanism to overcome DNA-damage by the exchange of DNA. To estimate theoretically if the determined recombination frequency could link the DNA-fragmentation and the cellular aggregation, all available data were integrated into a simplified calculation. The observed DNA

fragmentation after UV-irradiation with 75 J/m$^2$ resulted in fragments of 100 to 600 kb (Fig. 5.4). The genome of *S. solfataricus* is 2.9 Mbp in size; this would result in 30 fragments of 100 kb or 5 fragments of 600 kb, respectively. The exchange of one fragment, encoding the marker gene, between two cells would result in a theoretical frequency of 3 x $10^{-2}$ events/cell (in the case of a 100 kb fragmentation). At the observed maximum of cellular aggregation at 6 h at least 50-70% of the cells (probably more) were found to be connected to at last one neighbouring cell (Fig. 5.2). This would result in a theoretical frequency of at last 1.5 x $10^{-2}$ events/cell, which is quite close to the determined recombination frequency of 1.11 x $10^{-2}$ events/cell. Although this calculation is very simplified; it shows that the UV-induced DNA-fragmentation, cellular aggregation and conjugation could be linked to each other.

Two cases of conjugative exchange of DNA in combination with an observed cellular aggregation have been described so far in archaeal species, i.e. in *S. solfataricus* and *Halobacterium volcanii*. In both examples intercellular bridges were shown in addition to close cellular connections (Rosenshine *et al.*, 1989; Schleper *et al.*, 1995). While the observed pili probably mediated the initial cellular attachment, it is still unclear which mechanisms or factors are involved in the cellular connection. Interestingly, besides the genes of the *ups*-operon, the genes of the putative transmembrane proteins SSO0691 and SSO03146 showed a strong transcriptional up-regulation after the UV-treatment. The gene *sso0691* represented the highest induced transcriptional reaction in accordance to the study of Götz *et al.* (2007, see 7.2.3). It is possible that these putative membrane proteins are involved in forming the intercellular connections.

### 7.2.7 Conclusions

This present study showed that *S. solfataricus* exhibits a complex transcriptional regulation and also cellular reactions to UV-light. Based on the results of this study and by comparison to studies from other groups, the cellular response to UV-damage appears to be a complex reaction by an interaction of non UV-inducible (PhrB, Dpo4, NER, RadA) probably permanently present and in addition UV-inducible (Rad50/Mer11) DNA-repair mechanisms. Beside the reaction of the DNA-repair mechanisms, the UV-dependent transcriptional reactive genes involved many different cellular networks like e.g. the cell cycle or hierarchic levels like the observed cellular aggregation, mediated by a UV-regulated pili secretion system.

Based on the results of the present study and their interpretation the following scenario of the UV-caused reaction of *S. solfataricus* is created.

(a) Firstly, 0.5 to 1 h after UV-treatment the cells start to re-grow and recover from the growth arrest, caused by the treatment under room temperature. At that stage, the induced photoproducts are probably removed by the constitutively present photolyase and the NER system and genes involved in the central information processing, e.g. transcription, translation, replication and DNA-repair (NER, RadA, SSB) were up regulated.
(b) Not all photoproducts were removed until the start of replication. The non-repaired DNA-lesions block the replication; this results in a collapsing of the replication forks and in double-strand breaks (DSB) in DNA.
(c) This secondary effect of DNA-lesions induces the UV-dependent transcriptional reaction. The presence of the DSB DNA-damage might probably be sensed by the SSB and might be mediated by the Mre11 and Rad50 proteins. The UV-dependent response causes (i) a cell cycle arrest, (ii) a repression of initiation of replication (reaction of the *cdc6* genes), (iii) induction of the transcription factor TFB-3 (*sso0280*), (iii) induction of genes, putatively involved in DNA-repair mechanism (*mre11*/*rad50* operon, *dpo2*, *sso0001* encode a putative RecB nucelase) and (iv) an induction of a pili system (*ups*-operon) and further genes involved in secretion.
(d) The pili production, mediates aggregation of *S. solfataricus* cells, thereby initiating enhanced conjugative exchange of DNA to increase repair of the genome *via* homologous recombination.

### 7.2.8 Outlook

The investigation of the transcriptional and cellular reactions of *S. solfataricus* to UV-light raised a lot of open questions that should be addressed in further studies. Many of the identified UV-dependent genes are still of unknown function, like e.g. putative transcription regulators. For some genes, like the Dpo2, a transcriptional activity linked with a biological context was shown for the first time. Even the observed reactions of DNA-repair mechanisms, cellular aggregation and conjugation to UV-light are not understood in detail yet.

In addition the highly similar transcriptional pattern of the identified UV-dependent regulated gene groups provide a high potential to identify UV-specific regulatory promoter motifs. Beside comparative analysis of the promoter regions within the UV-dependent regulated genes of *S. solfataricus*, it is also possible to include the related genes from the genomes of *S. acidocaldarius*, *S. tokodaii* strain 7 or the more distantly related species *Metallosphaera sedula.* First studies to characterise the UV-specific regulatory motifs were performed (Daniela Teichmann, paper in preparation). It was shown by *in vivo* experiments, that the UV-dependent regulatory sequence are located at the proximal-promoter region within the first 50-60 nt upstream of the transcriptional start site. In addition the UV-specific regulatory promoter sequences are a useful tool to identify the regulatory proteins, e.g. by magnetic DNA affinity purification experiments. A successful application of this technique in Archaea, was shown for the identification of the host encoded viral transcription factor Sta1 (Kessler *et al.*, 2006).

Beside *in vivo* transcriptional studies with reporter constructs, the targetted knock-out of genes is now being used to dissect functions of the most strongly UV-induced genes that encode unknown proteins. In connection with detailled biochemical studies this will lead to a deeper understanding of the UV-response in *Sulfolobus* and other archaea.

# 9. Literature

Regarding to General Introduction and General Discussion:


Albers, S.V. and Driessen, A.J. (2007) Conditions for gene disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome. *Archaea*, 145-149.

Albers, S.V., Jonuscheit, M., Dinkelaker, S., Urich, T., Kletzin, A., Tampe, R., Driessen, A.J. and Schleper, C. (2006) Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus. *Appl Environ Microbiol*, 72, 102-111.

Aravind, L., Walker, D.R. and Koonin, E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res*, 27, 1223-1242.

Auchtung, T.A., Takacs-Vesbach, C.D. and Cavanaugh, C.M. (2006) 16S rRNA phylogenetic investigation of the candidate division "Korarchaeota". *Appl Environ Microbiol*, 72, 5077-5082.

Baliga, N.S., Goo, Y.A., Ng, W.V., Hood, L., Daniels, C.J. and DasSarma, S. (2000) Is gene expression in Halobacterium NRC-1 regulated by multiple TBP and TFB transcription factors? *Mol Microbiol*, 36, 1184-1185.

Barry, E.R. and Bell, S.D. (2006) DNA replication in the archaea. *Microbiol Mol Biol Rev*, 70, 876-887.

Bell, S.D. and Jackson, S.P. (1998) Transcription in Archaea. *Cold Spring Harb Symp Quant Biol*, 63, 41-51.

Bell, S.D., Jaxel, C., Nadal, M., Kosa, P.F. and Jackson, S.P. (1998) Temperature, template topology, and factor requirements of archaeal transcription. *Proc Natl Acad Sci U S A*, 95, 15218-15222.

Bernander, R. (1998) Archaea and the cell cycle. *Mol Microbiol*, 29, 955-961.

Bernander, R. (2000) Chromosome replication, nucleoid segregation and cell division in archaea. *Trends Microbiol*, 8, 278-283.

Bernander, R. and Poplawski, A. (1997) Cell cycle characteristics of thermophilic archaea. *J Bacteriol*, 179, 4963-4969.

Bonura, T. and Smith, K.C. (1975) Enzymatic production of deoxyribonucleic acid double-strand breaks after ultraviolet irradiation of Escherichia coli K-12. *J Bacteriol*, 121, 511-517.

Borowiec, J.A., Dean, F.B., Bullock, P.A. and Hurwitz, J. (1990) Binding and unwinding--how T antigen engages the SV40 origin of DNA replication. *Cell*, 60, 181-184.

Boudsocq, F., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) Sulfolobus solfataricus P2 DNA polymerase IV (Dpo4): an archaeal DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic poleta. *Nucleic Acids Res*, 29, 4607-4616.

Bradbury, J.M. and Jackson, S.P. (2003) The complex matter of DNA double-strand break detection. *Biochem Soc Trans*, 31, 40-44.

Brochier-Armanet, C., Boussau, B., Gribaldo, S. and Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol*, 6, 245-252.

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

Cann, I.K. and Ishino, Y. (1999) Archaeal DNA replication: identifying the pieces to solve a puzzle. *Genetics*, 152, 1249-1267.

Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H.P. and Garrett, R.A. (2005) The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota. *J Bacteriol*, 187, 4992-4999.

Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L. and Elie, C. (2004) A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nurA genes in thermophilic archaea. *Nucleic Acids Res*, 32, 1439-1447.

Contursi, P., Cannio, R., Prato, S., She, Q., Rossi, M. and Bartolucci, S. (2007) Transcriptional analysis of the genetic element pSSVx: differential and temporal regulation of gene expression reveals correlation between transcription and replication. *J Bacteriol*, 189, 6339-6350.

Coulombe, J., Berger, L., Smith, D.B., Hehl, R.K. and Wildeman, A.G. (1992) Activation of simian virus 40 transcription in vitro by T antigen. *J Virol*, 66, 4591-4596.

Cubeddu, L. and White, M.F. (2005) DNA damage detection by an archaeal single-stranded DNA-binding protein. *J Mol Biol*, 353, 507-516.

D'Amours, D. and Jackson, S.P. (2002) The Mre11 complex: at the crossroads of dna repair and checkpoint signalling. *Nat Rev Mol Cell Biol*, 3, 317-327.

de la Torre, J.R., Walker, C.B., Ingalls, A.E., Konneke, M. and Stahl, D.A. (2008) Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol. *Environ Microbiol*, 10, 810-818.

DeLong, E.F. and Pace, N.R. (2001) Environmental diversity of bacteria and archaea. *Syst Biol*, 50, 470-478.

Di Giulio, M. (2007) The tree of life might be rooted in the branch leading to Nanoarchaeota. *Gene*, 401, 108-113.

Dionne, I., Robinson, N.P., McGeoch, A.T., Marsh, V.L., Reddish, A. and Bell, S.D. (2003) DNA replication in the hyperthermophilic archaeon Sulfolobus solfataricus. *Biochem Soc Trans*, 31, 674-676.

Dorazi, R., Gotz, D., Munro, S., Bernander, R. and White, M.F. (2007) Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in Sulfolobus solfataricus. *Mol Microbiol*, 63, 521-529.

Duggin, I.G. and Bell, S.D. (2006) The chromosome replication machinery of the archaeon Sulfolobus solfataricus. *J Biol Chem*, 281, 15029-15032.

Edgell, D.R., Klenk, H.P. and Doolittle, W.F. (1997) Gene duplications in evolution of archaeal family B DNA polymerases. *J Bacteriol*, 179, 2632-2640.

Elie, C., Baucher, M.F., Fondrat, C. and Forterre, P. (1997) A protein related to eucaryal and bacterial DNA-motor proteins in the hyperthermophilic archaeon Sulfolobus acidocaldarius. *J Mol Evol*, 45, 107-114.

Fradet-Turcotte, A., Vincent, C., Joubert, S., Bullock, P.A. and Archambault, J. (2007) Quantitative analysis of the binding of simian virus 40 large T antigen to DNA. *J Virol*, 81, 9162-9174.

Fröls, S., Gordon, P.M., Panlilio, M.A., Duggin, I.G., Bell, S.D., Sensen, C.W. and Schleper, C. (2007) Response of the hyperthermophilic archaeon Sulfolobus solfataricus to UV damage. *J Bacteriol*, 189, 8708-8718.

Fuchs, T., Huber, H., Burggraf, S. and Stetter, K.O. (1996) 16S rDNA-based phylogenyof the archaeal order Sulfolobales and reclassification of Desulfurolobus ambivalens as Acidianus ambivalens comb. nov.. *System. Appl. Microbiol.*, 56-60.

Fujihashi, M., Numoto, N., Kobayashi, Y., Mizushima, A., Tsujimura, M., Nakamura, A., Kawarabayasi, Y. and Miki, K. (2007) Crystal structure of archaeal photolyase from Sulfolobus tokodaii with two FAD molecules: implication of a novel light-harvesting cofactor. *J Mol Biol*, 365, 903-910.

Garinis, G.A., Mitchell, J.R., Moorhouse, M.J., Hanada, K., de Waard, H., Vandeputte, D., Jans, J., Brand, K., Smid, M., van der Spek, P.J., Hoeijmakers, J.H., Kanaar, R. and van der Horst, G.T. (2005) Transcriptome analysis reveals cyclobutane pyrimidine dimers as a major source of UV-induced DNA breaks. *Embo J*, 24, 3952-3962.

Geiduschek, E.P. and Ouhammouch, M. (2005) Archaeal transcription and its regulators. *Mol Microbiol*, 56, 1397-1407.

Gohl, H.P., Grondahl, B. and Thomm, M. (1995) Promoter recognition in archaea is mediated by transcription factors: identification of transcription factor aTFB from Methanococcus thermolithotrophicus as archaeal TATA-binding protein. *Nucleic Acids Res*, 23, 3837-3841.

Götz, D., Paytubi, S., Munro, S., Lundgren, M., Bernander, R. and White, M.F. (2007) Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol*, 8, R220.

Grogan, D.W. (1989) Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains. *J Bacteriol*, 171, 6710-6719.

Grogan, D.W. (2000) The question of DNA repair in hyperthermophilic archaea. *Trends Microbiol*, 8, 180-185.

Gruz, P., Shimizu, M., Pisani, F.M., De Felice, M., Kanke, Y. and Nohmi, T. (2003) Processing of DNA lesions by archaeal DNA polymerases from Sulfolobus solfataricus. *Nucleic Acids Res*, 31, 4024-4030.

Haber, J.E. (2006) Chromosome breakage and repair. *Genetics*, 173, 1181-1185.

Han, Y., Loo, Y.M., Militello, K.T. and Melendy, T. (1999) Interactions of the papovavirus DNA replication initiator proteins, bovine papillomavirus type 1 E1 and simian virus 40 large T antigen, with human replication protein A. *J Virol*, 73, 4899-4907.

Hausner, W., Wettach, J., Hethke, C. and Thomm, M. (1996) Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J Biol Chem*, 271, 30144-30148.

Hefferin, M.L. and Tomkinson, A.E. (2005) Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair (Amst)*, 4, 639-648.

Hirata, A., Klein, B.J. and Murakami, K.S. (2008) The X-ray crystal structure of RNA polymerase from Archaea. *Nature*.

Huet, J., Schnabel, R., Sentenac, A. and Zillig, W. (1983) Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. *Embo J*, 2, 1291-1294.

Jackson, S.P. (2002) Sensing and repairing DNA double-strand breaks. *Carcinogenesis*, 23, 687-696.

Jazayeri, A., Falck, J., Lukas, C., Bartek, J., Smith, G.C., Lukas, J. and Jackson, S.P. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol*, 8, 37-45.

Jiang, P.X., Wang, J., Feng, Y. and He, Z.G. (2007) Divergent functions of multiple eukaryote-like Orc1/Cdc6 proteins on modulating the loading of the MCM helicase onto the origins of the hyperthermophilic archaeon Sulfolobus solfataricus P2. *Biochem Biophys Res Commun*, 361, 651-658.

Jonuscheit, M., Martusewitsch, E., Stedman, K.M. and Schleper, C. (2003) A reporter gene system for the hyperthermophilic archaeon Sulfolobus solfataricus based on a selectable and integrative shuttle vector. *Mol Microbiol*, 48, 1241-1252.

Kawarabayasi, Y., Hino, Y., Horikawa, H., Jin-no, K., Takahashi, M., Sekine, M., Baba, S., Ankai, A., Kosugi, H., Hosoyama, A., Fukui, S., Nagai, Y., Nishijima, K., Otsuka, R., Nakazawa, H., Takamiya, M., Kato, Y., Yoshizawa, T., Tanaka, T., Kudoh, Y., Yamazaki, J., Kushida, N., Oguchi, A., Aoki, K., Masuda, S., Yanagii, M., Nishimura, M., Yamagishi, A., Oshima, T. and Kikuchi, H. (2001) Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7. *DNA Res*, 8, 123-140.

Kelman, Z. and White, M.F. (2005) Archaeal DNA replication and repair. *Curr Opin Microbiol*, 8, 669-676.

Kessler, A., Brinkman, A.B., van der Oost, J. and Prangishvili, D. (2004) Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon sulfolobus. *J Bacteriol*, 186, 7745-7753.

Kessler, A., Sezonov, G., Guijarro, J.I., Desnoues, N., Rose, T., Delepierre, M., Bell, S.D. and Prangishvili, D. (2006) A novel archaeal regulatory protein, Sta1, activates transcription from viral promoters. *Nucleic Acids Res*, 34, 4837-4845.

Kitano, H. (2002) Systems biology: a brief overview. *Science*, 295, 1662-1664.

Konneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B. and Stahl, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature*, 437, 543-546.

Kraft, P., Oeckinghaus, A., Kummel, D., Gauss, G.H., Gilmore, J., Wiedenheft, B., Young, M. and Lawrence, C.M. (2004) Crystal structure of F-93 from Sulfolobus spindle-shaped virus 1, a winged-helix DNA binding protein. *J Virol*, 78, 11544-11550.

Kuluncsics, Z., Perdiz, D., Brulay, E., Muel, B. and Sage, E. (1999) Wavelength dependence of ultraviolet-induced DNA damage distribution: involvement of direct or indirect mechanisms and possible artefacts. *J Photochem Photobiol B*, 49, 71-80.

Kuzminov, A. (2001) Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A*, 98, 8241-8246.

Lao-Sirieix, S., Marsh, V.L. and Bell, S.D. (2007) *Archaea Molecular and Cellular Biology*;*Chapter 3 DNA Replication and Cell Cycle* American Society for Microbiology.

Lapaglia, C. and Hartzell, P.L. (1997) Stress-Induced Production of Biofilm in the Hyperthermophile Archaeoglobus fulgidus. *Appl Environ Microbiol*, 63, 3158-3163.

Lecompte, O., Ripp, R., Thierry, J.C., Moras, D. and Poch, O. (2002) Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. *Nucleic Acids Res*, 30, 5382-5390.

Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., Prosser, J.I., Schuster, S.C. and Schleper, C. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature*, 442, 806-809.

Londei, P. (2005) Evolution of translational initiation: new insights from the archaea. *FEMS Microbiol Rev*, 29, 185-200.

Londei, P. (2007) *Archaea Molecular and Cellular Biology; Chapter 8 Translation*. American Society for Microbiology.

Londei, P., Teixido, J., Acca, M., Cammarano, P. and Amils, R. (1986) Total reconstitution of active large ribosomal subunits of the thermoacidophilic archaebacterium Sulfolobus solfataricus. *Nucleic Acids Res*, 14, 2269-2285.

Lundgren, M., Andersson, A., Chen, L., Nilsson, P. and Bernander, R. (2004) Three replication origins in Sulfolobus species: synchronous initiation of chromosome replication and asynchronous termination. *Proc Natl Acad Sci U S A*, 101, 7046-7051.

Lundgren, M. and Bernander, R. (2007) Genome-wide transcription map of an archaeal cell cycle. *Proc Natl Acad Sci U S A*, 104, 2939-2944.

Martin, A., Yeats, S., Janekovic, D., Reiter, W.D., Aicher, W. and Zillig, W. (1984) SAV 1, a temperate u.v.-inducible DNA virus-like particle from the archaebacterium Sulfolobus acidocaldarius isolate B12. *Embo J*, 3, 2165-2168.

Martusewitsch, E., Sensen, C.W. and Schleper, C. (2000) High spontaneous mutation rate in the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by transposable elements. *J Bacteriol*, 182, 2574-2581.

Meyer, R.R. and Laine, P.S. (1990) The single-stranded DNA-binding protein of Escherichia coli. *Microbiol Rev*, 54, 342-380.

Muskhelishvili, G., Palm, P. and Zillig, W. (1993) SSV1-encoded site-specific recombination system in Sulfolobus shibatae. *Mol Gen Genet*, 237, 334-342.

Nikolov, D.B. and Burley, S.K. (1997) RNA polymerase II transcription initiation: a structural view. *Proc Natl Acad Sci U S A*, 94, 15-22.

Palmer, J.R. and Daniels, C.J. (1995) In vivo definition of an archaeal promoter. *J Bacteriol*, 177, 1844-1849.

Poplawski, A. and Bernander, R. (1997) Nucleoid structure and distribution in thermophilic Archaea. *J Bacteriol*, 179, 7625-7630.

Prangishvili, D. and Garrett, R.A. (2004) Exceptionally diverse morphotypes and genomes of crenarchaeal hyperthermophilic viruses. *Biochem Soc Trans*, 32, 204-208.

Prangishvili, D. and Garrett, R.A. (2005) Viruses of hyperthermophilic Crenarchaea. *Trends Microbiol*, 13, 535-542.

Quaiser, A., Constantinesco, F., White, M.F., Forterre, P. and Elie, C. (2008) The Mre11 protein interacts with both Rad50 and the HerA bipolar helicase and is recruited to DNA following gamma irradiation in the archaeon Sulfolobus acidocaldarius. *BMC Mol Biol*, 9, 25.

Qureshi, S.A. (2007) Protein-DNA interactions at the Sulfolobus spindle-shaped virus-1 (SSV1) T5 and T6 gene promoters. *Can J Microbiol*, 53, 1076-1083.

Qureshi, S.A., Bell, S.D. and Jackson, S.P. (1997) Factor requirements for transcription in the Archaeon Sulfolobus shibatae. *Embo J*, 16, 2927-2936.

Qureshi, S.A. and Jackson, S.P. (1998) Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength. *Mol Cell*, 1, 389-400.

Reiter, W.D. and Palm, P. (1990) Identification and characterization of a defective SSV1 genome integrated into a tRNA gene in the archaebacterium Sulfolobus sp. B12. *Mol Gen Genet*, 221, 65-71.

Reiter, W.D., Palm, P. and Yeats, S. (1989) Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. *Nucleic Acids Res*, 17, 1907-1914.

Reiter, W.D., Palm, P., Yeats, S. and Zillig, W. (1987) Gene expression in archaebacteria: physical mapping of constitutive and UV-inducible transcripts from the Sulfolobus virus-like particle SSV1. *Mol Gen Genet*, 209, 270-275.

Reiter, W.D., Palm, P. and Zillig, W. (1988a) Analysis of transcription in the archaebacterium Sulfolobus indicates that archaebacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res*, 16, 1-19.

Reiter, W.D., Palm, P. and Zillig, W. (1988b) Transcription termination in the archaebacterium Sulfolobus: signal structures and linkage to transcription initiation. *Nucleic Acids Res*, 16, 2445-2459.

Reiter, W.D., Zillig, W. and Palm, P. (1988c) Archaebacterial viruses. *Adv Virus Res*, 34, 143-188.

Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S., McDermott, T. and Young, M.J. (2001) Viruses from extreme thermal environments. *Proc Natl Acad Sci U S A*, 98, 13341-13345.

Robinson, N.P., Blood, K.A., McCallum, S.A., Edwards, P.A. and Bell, S.D. (2007) Sister chromatid junctions in the hyperthermophilic archaeon Sulfolobus solfataricus. *Embo J*, 26, 816-824.

Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004) Identification of two origins of replication in the single chromosome of the archaeon Sulfolobus solfataricus. *Cell*, 116, 25-38.

Romano, V., Napoli, A., Salerno, V., Valenti, A., Rossi, M. and Ciaramella, M. (2006) Lack of Strand-specific Repair of UV-induced DNA Lesions in Three Genes of the Archaeon Sulfolobus solfataricus. *J Mol Biol*.

Rosenshine, I., Tchelet, R. and Mevarech, M. (1989) The mechanism of DNA transfer in the mating system of an archaebacterium. *Science*, 245, 1387-1389.

Salerno, V., Napoli, A., White, M.F., Rossi, M. and Ciaramella, M. (2003) Transcriptional response to DNA damage in the archaeon Sulfolobus solfataricus. *Nucleic Acids Res*, 31, 6127-6138.

Sandler, S.J., Satin, L.H., Samra, H.S. and Clark, A.J. (1996) recA-like genes from three archaean species with putative protein products similar to Rad51 and Dmc1 proteins of the yeast Saccharomyces cerevisiae. *Nucleic Acids Res*, 24, 2125-2132.

Schelert, J., Dixit, V., Hoang, V., Simbahan, J., Drozda, M. and Blum, P. (2004) Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon Sulfolobus solfataricus by use of gene disruption. *J Bacteriol*, 186, 427-437.

Schleper, C., Holz, I., Janekovic, D., Murphy, J. and Zillig, W. (1995) A multicopy plasmid of the extremely thermophilic archaeon Sulfolobus effects its transfer to recipients by mating. *J Bacteriol*, 177, 4417-4426.

Seitz, E.M., Brockman, J.P., Sandler, S.J., Clark, A.J. and Kowalczykowski, S.C. (1998) RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev*, 12, 1248-1253.

She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De Moors, A., Erauso, G., Fletcher, C., Gordon, P.M., Heikamp-de Jong, I., Jeffries, A.C., Kozera, C.J., Medina, N., Peng, X., Thi-Ngoc, H.P., Redder, P., Schenk, M.E., Theriault, C., Tolstrup, N., Charlebois, R.L., Doolittle, W.F., Duguet, M., Gaasterland, T., Garrett, R.A., Ragan, M.A., Sensen, C.W. and Van der Oost, J. (2001) The complete genome of the crenarchaeon Sulfolobus solfataricus P2. *Proc Natl Acad Sci U S A*, 98, 7835-7840.

Shin, D.S., Chahwan, C., Huffman, J.L. and Tainer, J.A. (2004) Structure and function of the double-strand break repair machinery. *DNA Repair (Amst)*, 3, 863-873.

Shockley, K.R., Ward, D.E., Chhabra, S.R., Conners, S.B., Montero, C.I. and Kelly, R.M. (2003) Heat shock response by the hyperthermophilic archaeon Pyrococcus furiosus. *Appl Environ Microbiol*, 69, 2365-2371.

Shuck, S.C., Short, E.A. and Turchi, J.J. (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology. *Cell Res*, 18, 64-72.

Stetter, K.O. (1989) *Order III. Sulfolobales ord. nov. ; Archaeobacteria, Cyanobacteria, and remaining Gram-negative Bacteria ; Bergey's Manual of Systematic Bacteriology*.

Stetter, K.O. (2006) Hyperthermophiles in the history of life. *Philos Trans R Soc Lond B Biol Sci*, 361, 1837-1842; discussion 1842-1833.

Stetter, K.O., Fiala, G., Huber, G., Huber, R. and Segerer, A. (1990) Hyperthermophilic microorganisms. *FEMS Microbiol. Rev.*, 75, 117-124.

Thomm, M. (1996) Archaeal transcription factors and their role in transcription initiation. *FEMS Microbiol Rev*, 18, 159-171.

Thomm, M. (2007) *Archaea Molecular and Cellular Biology*; *Chapter 6 Transcription Mechanism and Regulation.* American Society for Microbiology.

Wadsworth, R.I. and White, M.F. (2001) Identification and properties of the crenarchaeal single-stranded DNA binding protein from Sulfolobus solfataricus. *Nucleic Acids Res*, 29, 914-920.

Wei, T., Zhang, S., Zhu, S., Sheng, D., Ni, J. and Shen, Y. (2008) Physical and functional interaction between archaeal single-stranded DNA-binding protein and the 5'-3' nuclease NurA. *Biochem Biophys Res Commun*, 367, 523-529.

Wettach, J., Gohl, H.P., Tschochner, H. and Thomm, M. (1995) Functional interaction of yeast and human TATA-binding proteins with an archaeal RNA polymerase and promoter. *Proc Natl Acad Sci U S A*, 92, 472-476.

White, M.F. (2003) Archaeal DNA repair: paradigms and puzzles. *Biochem Soc Trans*, 31, 690-693.

White, M.F. (2007) *Archaea Evolution, Physiology, and Molecular Biology*; *Chapter 15 DNA Repair* Blackwell Publishing.

Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A.K., Zoeller, L., Snyder, J., Douglas, T. and Young, M. (2004) Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. *J Virol*, 78, 1954-1961.

Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A*, 74, 5088-5090.

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

Woese, C.R., Magrum, L.J. and Fox, G.E. (1978) Archaebacteria. *J Mol Evol*, 11, 245-251.

Wuchter, C., Abbas, B., Coolen, M.J., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., Schouten, S. and Sinninghe Damste, J.S. (2006) Archaeal nitrification in the ocean. *Proc Natl Acad Sci U S A*, 103, 12317-12322.

Zillig, W., Arnold, H.P., Holz, I., Prangishvili, D., Schweier, A., Stedman, K., She, Q., Phan, H., Garrett, R. and Kristjansson, J.K. (1998) Genetic elements in the extremely thermophilic archaeon Sulfolobus. *Extremophiles*, 2, 131-140.

Zillig, W., Holz, I., Janekovic, D., Klenk, H.P., Imsel, E., Trent, J., Wunderl, S., Forjaz, V.H., Coutinho, R. and Ferreira, T. (1990) Hyperthermus butylicus, a hyperthermophilic sulfur-reducing archaebacterium that ferments peptides. *J Bacteriol*, 172, 3959-3965.

Zillig, W., Palm, P., Langer, D., Klenk, H.P., Lanzendorfer, M., Hudepohl, U. and Hain, J. (1992) RNA polymerases and transcription in archaebacteria. *Biochem Soc Symp*, 58, 79-88.

Zillig, W., Prangishvilli, D., Schleper, C., Elferink, M., Holz, I., Albers, S., Janekovic, D. and Gotz, D. (1996) Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea. *FEMS Microbiol Rev*, 18, 225-236.

Zillig, W., Stetter, K.O. and Janekovic, D. (1979) DNA-dependent RNA polymerase from the archaebacterium Sulfolobus acidocaldarius. *Eur J Biochem*, 96, 597-604.

Zillig, W., Stetter, K.O., Wunderl, S., Schulz, W., Priess, H. and Scholz, I. (1980) The Sulfolobus "Caldariella" group: taxonomy on the basis of the structure of DNA-dependet RNA polymerases. *Arch. Microbiol.*, 125, 259-269.

# 9. Abbreviations

| | |
|---|---|
| Å | Ångström |
| bp | base pairs |
| Mbp | mega base pairs |
| °C | degrees Celsius |
| CPD | cis-syn-cyclobutane pyrimidine dimers |
| DSB | double-strand break |
| DNA | Desoxyribonucleic acid |
| e. g. | for example |
| et al. | and others |
| h | hour |
| i. e. | that is |
| kb | kilo bases |
| l | liter |
| M | molar |
| Mbp | mega base pairs |
| min | minute |
| nt | nucleotide |
| OD | optical density |
| PCR | Polymerase chain reaction |
| J | Joule |
| rpm | rounds per minute |
| sec | second |
| UV | Ultra-violet |
| w/v | weight per volume |

# 10. Eidesstattliche Erklärung

Die vorliegende Arbeit wurde als Promotionsarbeit am Institut für Mikrobiologie und Genetik des Fachbereichs Biologie der Technischen Universität Darmstadt unter der Leitung von **Frau Prof. Christa Schleper** in der Abteilung von Frau Prof. Felicitas Pfeifer im Zeitraum von Dezember 2003 bis Juni 2007 und von Juli 2007 bis September 2007 am Institut für Biologie der Universität Bergen in Norwegen angefertigt und abschließend von Oktober 2007 bis April 2008 an der Universität Wien vollendet.

Ich erkläre hiermit an Eides Statt, dass ich den materiellen Inhalt der vorliegenden Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel und Quellen angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken und Fakten, sofern sie nicht wissenschaftliches Allgemeingut sind, wurden als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Darmstadt, den 09. April 2008

Sabrina Fröls

# 11. Curriculum Vitae

## Personal information

| | |
|---|---|
| Name | **Sabrina Fröls** |
| Date of birth | 30th of June, 1977 |
| Place of birth | Frankfurt am Main, Germany |
| Private address | Weiherstrasse. 11, 63303 Dreieich, Germany |

## University education

| | |
|---|---|
| 1996 – 2003 | **Studies of biology** at Darmstadt University of Technology<br>Major fields: Microbiology, Genetics, Cell and Developmental Biology |
| 2003 | **Diploma Thesis** in Microbiology<br>Title: "Construction of a shuttle vector system for the hyperthermophilic archaeon *Sulfolobus solfataricus* – based on the multisite gateway technology"<br><br>Average mark: very good<br>Advisors: Prof. Dr. Christa Schleper, Prof. Dr. Felicitas Pfeifer<br>Degree earned: Diplom-Biologe |
| December 2003 to present | **PhD student** at the Institute of Microbiology and Genetics, Darmstadt University of Technology |

## Research stays abroad

| | |
|---|---|
| 2001 | Research project at the University of Manchester, England (February - April), Group of Prof. D. Denning |
| 2005 - 2007 | Temporary visits at the Department of Biology, University of Bergen, Norway<br>In 2005 (Oct - Dec), 2006 (Sept - Oct) and 2007 (March - June) |
| 2007 - 2008 | Since October 2007 at the Department of Genetics in Ecology, University of Vienna, Austria<br><br>Group of Prof. Dr. Christa Schleper |

## School education

| | |
|---|---|
| 1983 - 1987 | Primary school, Ludwig-Erk-Schule, Dreieich, Germany |
| 1987 - 1993 | High school, Weibelfeldschule, Gesamtschule, Dreieich, Germany |
| 1993 - 1996 | High school, Berufliches Gymnasium Fachrichtung Technik, Germany Max-Eyth-Schule, Dreieich |
| | Abitur, average mark: good |

## Oral presentations and poster

| | |
|---|---|
| 2007 | Thermophiles Conference<br>24th to 27th of September 2007 Bergen, Norway<br>**Talk**: "Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV-irradiation" |
| 2007 | FEBS Advanced Lecture Course on Systems Biology<br>10th to 16th March, 2007 Gosau, Austria<br>**Poster**: "Genome-wide transcriptional response of the hyperthermophilic archaeon *S. solfataricus* and its virus SSV1 to UV-treatment" |
| 2006 | Extremophiles Conference<br>17th to 21th September 2006 Brest, France<br>**Talk**: "UV-dependent transcriptional regulation of the archaeal virus SSV1 and of its hyperthermophilic host as seen through genome-wide microarrays" |
| 2005 | VAAM (Association for General and Applied Microbiology)<br>25th to 28th September 2005 Göttingen, Germany<br>**Talk**: "Genome-wide microarray analyses to study the UV-dependent transcriptional regulation of the virus SSV1 and of its host, the hyperthermophilic archaeon *Sulfolobus solfataricus"* |

# 12. Acknowledgments

At first, I would like to thank my remarkable advisor Prof. Christa Schleper, for everything that I was able to learn from her during the last years. For the support and development of my project in the field of the Archaea. For great support in all situations of science and live, the trust in my work as well as the possibility to work independently.

My co-advisor Prof. Felicitas Pfeifer I would like to thank with warmest regards for allowing me to create a bridge between Darmstadt and Bergen. I also would like to thank her for the interesting conversations, the support of my work and the knowledge, that scientists are also only humans.

Dr. habil. Arnulf Kletzin I would like to thank for the door that was always open, his never-ending patience and kindness in his answers to my never-ending questions.

I also would like to thank my collaborators and co-authors, without them this work would not have been possible. First Christoph Sensen and Paul Gordan from the University of Calgary, who developed the DNA-microarrays that I was working with and took care of the statistical analysis. Peter Uetz and Seesandra Rajagopala from Forschungszentrum Karlsruhe, who gave us the opportunity to perform the 2-hybrid screen with SSV1. Also, Sonja-Verena Albers and Behnam Zolghadr from the University of Groningen, for the discussion regarding the pili and for providing the SSO0120 knock-out strain. I will never forget these exiting moments, when we saw the first pili in the electron microscope. In addition I would like to thank all the people of the Archaea community who were always open for help and discussion - it was a pleasure to be one of you.

I would like to thank my old and new colleagues from the AG Schleper, spread out over Darmstadt, Bergen and Vienna for the best, relaxed and casual working climate that I can dream of. Dear all, I really enjoyed the time with you during the lab work and outside of the lab at Bergen´s most beautiful places.

I would also like to thank my colleagues at Darmstadt: all the members of the AG Pfeifer, AG Heider, AG Göringer and the people on the third floor for the very good and helpful working climate, the help with theoretical and practical problems as well as several coffee breaks on the balcony.

Special thanks goes to Daniela Teichmann, Tamara Hess and Andreas Veith for the personal care and a great collaboration during the practical course, diploma thesis and other research projects.

Also to Tim Urich, Cordula Böhm, Moritz Niemann and Annette Adler; who never got tired to: discuss with me my newest results, listen to my crazy home-town stories and to wait for me – thank you.

Also to all my friends from home, who after all still have no clue, what I am really doing but were always listening to me.

The great support of my family over all the years was so fantastic – without you, I would not have been able to go through all this.

Finally, I am deeply grateful for having Maik on my side. I am pretty sure it is not easy to live with a biologist; you are doing a great job!