„Functional analysis of multiple general transcription factors in *Sulfolobus acidocaldarius*"

Dissertation

Zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

- Dr. rer. nat. -

vorgelegt von

Bernadette Rauch

geboren am 13.07.1984, in Dernbach

Biofilm Centre, Molekulare Enzymtechnologie und Biochemie

Fachbereich Chemie

der Universität Duisburg-Essen

2013

Die vorliegende Arbeit wurde im Zeitraum von November 2009 bis Juli 2013 in der Arbeitsgruppe von Prof. Dr. Bettina Siebers in der Fakultät Chemie, Biofilm Centre, Molekulare Enzymtechnologie und Biochemie der Universität Duisburg-Essen durchgeführt.

Tag der Disputation: 04.12.2013

Gutachter: Prof. Dr. Bettina Siebers

Prof. Dr. Perihan Nalbant

Vorsitzender: Prof. Dr. Ulbricht

Table of contents

List of Figures

List of Tables

1. Introduction

1.1. Archaea - an overview

On the basis of small subunit (16S) ribosomal RNA (rRNA) phylogenies, Carl Woese and colleagues discovered the third domain of life, the Archaea along with Bacteria and Eukaryotes. Archaea are divided into five phyla: (I) Crenarchaeota, (II) Euryarchaeota (most cultured and known organisms are included in these two phyla), (III) Korarchaeota which have been identified by DNA sequencing but remain uncultivable so far (IV) Nanoarchaeota represented by the smallest prokaryote *Nanoarchaeum equitans* and (V) Thaumarchaeota, the most recent discovered phylum including *Nitrososphaera gargensis*, *Nitrosopumilus maritimus* and *Cenarchaeum symbiosum* as revealed by comparative genomics (Pester et al., 2011). Unique to Archaea are their cell wall, cytoplasmic membrane and metabolic pathways especially the central carbohydrate metabolism. The composition of cell walls varies from that found in Bacteria. Some Archaea, like *Methanocaldococcus jannaschii* (for review see Eichler, 2003, Akca et al., 2002, Claus et al., 2002), *Halobacterium salinarum* (Lechner and Sumper, 1987) and *Sulfolobus acidocaldarius* (Grogan 1996a) possess an S-layer as cell envelope, some species, mostly methanogens, integrated pseudopeptidoglycan instead of peptidoglycan as cell wall component and some Archaea even lack cell walls (*Ferroplasma acidiphilum*). Pseudopeptidoglycan consists of N-acetylglucosamine and N-acetyltalosaminuronic acid linked by a β-1,3-glycosidic bond. N-acetyltalosaminuronic acid replaces N-acetylmuramic acid in bacterial peptidoglycan. The cytoplasmic membrane contains isoprenoid linked ether lipids compared to fatty-acid linked ester lipids found in Bacteria and Eukaryotes (van de Vossenberg et al., 1998). Some Archaea build tetraetherlipids thus providing a stable monolayer. Archaea exhibit a tremendous metabolic diversity and harbor unusual pathways and enzymes for their central carbohydrate metabolism which differ from the classical eukaryotic and bacterial pathways (e.g. modified Entner-Doudoroff pathway or Embden-Meyerhof-Parnas pathway) (Siebers and Schönheit, 2005, Ahmed et al., 2005, Sato and Atomi, 2011). Whereas the intermediates of these pathways show similarity to the classical routes, many enzymes are not homologous to bacterial or eukaryotic counterparts. The regulation of the pathways is not well understood, however, allosteric enzyme regulation as reported for Bacteria and Eukaryotes seems to be absent and in some species enzymes might rather be regulated on gene expression level (Siebers and Schönheit, 2005, Ahmed et al., 2005, Sato and Atomi, 2011). Archaeal cell morphology equals Bacteria as well as their genome size and structure. They lack a nucleus, possess one circular chromosome, genes are organized in operon structures and some species harbor plasmids. What they have in common with Eukaryotes is the presence of histones in Euryarchaeota for compaction of DNA and they share homologous proteins of mechanisms involved in information processing like transcription, translation, replication and DNA repair (Reeve, 2003, Zhang et al., 2012, Bell and Jackson, 1998a, b, Kyrpides and Ouzounis, 1999).

The basic dogma of biology states that genetic information encoded in the DNA is transcribed into mRNA and subsequently translated into a protein (Crick, 1958, Crick, 1970). Transcription and translation are processes that can be divided into three steps: initiation, elongation and termination. The genome organization is influenced by chromatin proteins and

transcription factors and transcription regulators control gene expression by formation of mRNA.

1.2. Chromatin proteins in Eukaryotes and Archaea

Modulation of gene expression on the level of transcription is an ubiquitous process found in all organisms and is facilitated by a large repertoire of different proteins. Chromatin proteins, general transcription factors and transcription regulators orchestrate transcription processes in the cell to maintain cell growth and adaptation to environmental conditions. Chromatin proteins can be found in every living organism on earth and facilitate compaction of DNA. Furthermore they play a role in replication and gene expression in general. Eukaryal histone octamers are composed of heterodimers formed by the histones H2A, H2B, H3 and H4 (for review see Ammar et al., 2012). A variety of DNA compaction proteins also evolved in the archaeal branch. Euryarchaeota and Nanoarchaeota possess histones that resemble eukaryotic histones H3 and H4 and form homo- or heterodimers as well as homotetramers (Ammar et al., 2012, Zhang et al., 2012, Sandmann and Reeve, 2005). A putative histone-encoding gene (*eag*3) from marine non-culturable microorganisms was synthesized and recombinant expressed in *Escherichia coli.* The EAG3 protein showed DNA-binding properties and formed complexes (Čuboňová et al., 2005) suggesting that histones evolved before the archaeal and eukaryal lineage diverged as it has been recently stated by Ammar et al., 2012. Whereas a number of histones have been investigated in Euryarchaeota the presence of histones in Crenarchaeota is still a matter of debate. In Crenarchaeota other proteins like Sul7d, Cren7, CC1 and Alba have evolved for chromosomal organization instead of histones (Zhang et al., 2012). In general, these proteins bind to DNA, increase the DNA melting temperature and introduce negative supercoiling (Reeve, 2003, Sandmann and Reeve, 2005, Luo et al., 2007, Zhang et al., 2012).

1.3. Transcription apparatus in Bacteria, Eukaryotes and Archaea

In Bacteria only one RNA polymerase (RNAP) facilitates transcription of DNA in the cell. The core enzyme of the RNAP consists of four catalytic subunits α_2, β, β' and ω. For transcription initiation a σ-factor recognizes promoter elements (-35 and -10 region) and binds to the core enzyme to form the holoenzyme (α_2, β, β', ω, σ) that directs transcription. The σ-factor σ^{70} is used at standard growth conditions and regulates transcription of house-keeping genes. Specialized σ-factors bind to particular promoters to change transcriptional response according to environmental conditions (e.g. heat shock: σ^{32}) (Grossman et al., 1987, Raina et al., 1995). Antagonists of σ-factors are anti-sigma factors which regulate cellular processes like heat-shock response (DnaK inhibits σ^{32}), sporulation (SpoIIAB inhibits σ^{F} and σ^{G}) or flagellar biosynthesis (FlgM inhibits σ^{28}) (for review see Hughes and Mathee, 1998). After transition from the closed to the open promoter complex the σ-subunit is released and the core enzyme migrates along the DNA strand, leaves the promoter region (promoter clearance) and mRNA is produced (Roberts et al., 2008). Termination of transcription is mediated by two different mechanisms in Bacteria. The first mechanism is performed by formation of a stem loop structure (intrinsic termination), which changes the conformation of the RNAP and leads

to destabilization and dissociation of the enzyme. Terminator sequences are composed of a GC-rich hairpin and an adjacent U-rich segment (Henkin, 2000, Roberts et al., 2008, Peters et al., 2011). The second mechanism is facilitated by the ATP-dependent RNA translocase Rho (Lowery-Goldhammer and Richardson, 1974). The homohexamer binds to C-rich sites in the RNA and with its ATPase and helicase activity it unwinds RNA-DNA hybrids and releases RNA (Richardson, 2002, Banerjee et al., 2006, Ciampi, 2006, Roberts et al., 2008). The proteins NusA and NusG are involved in the elongation step of transcription and Rho-dependent termination. NusA decreases the elongation rate and the termination efficiency and NusG acts opposite by increasing these activities. Interestingly, both factors do not compete with each other (Burns et al., 1998).

Transcription in Eukaryotes takes place in the nucleus. A plethora of transcription factors is involved in all three steps of transcription. Depending on the RNAPs, different transcription factors are involved. For transcription initiation at RNAPII promoters the TATA-binding protein (TBP) recognizes and binds to an AT-rich promoter element, the TATA-box, and forms the TFIID complex with fourteen other proteins, i. e. TBP-associated factors (TAFs). The transcription initiation factor IIB (TFIIB) subsequently binds to TBP and is stabilized by TFIIA. RNAPII is bound by TFIIF and this complex is recruited by TFIIB. TFIIE recruits and stimulates TFIIH, which unwinds the DNA at the promoter to form the open promoter complex and phosphorylates the carboxyterminal heptapeptide (CTD) repeat of the RNAPII (Orphanides et al., 1996, Roeder 1996, Hahn, 2004, Luse, 2013). Transcription elongation in Eukaryotes is orchestrated by a large set of transcription elongation factors with a general elongation and regulator function. The evolutionary conserved mediator complex is exclusively present in Eukaryotes and directly interacts with RNAPII and transcription factors. It recruits transcription elongation factors and pre-mRNA processing factors as well as controls the phosphorylation state of the CTD of the RNAPII (Conaway and Conaway, 2013). Transcription elongation factors DRB-sensitivity-inducing factor (DSIF, heterodimer of Spt4 and Spt5), negative elongation factor (NELF) along with positive transcription elongation factor b (pTEF-b) control the step of nascent RNA formation (Yamaguchi et al., 2013). Transcription elongation factors TFIIS (SII), transcription factor TFIIF, Elongin (SIII) and eleven-nineteen lysine-rich leukemia (ELL) are also involved in elongation. The elongation factors pTEF-b and SII prevent arrest of RNAPII and TFIIF, Elongin and ELL increase the elongation rate by suppression of polymerase pausing (Reines et al., 1989, Reines et al., 1999, Conaway and Conaway, 1993). TFIIF is unique in Eukaryotes and regulates transcription in initiation and elongation steps (Reines et al. 1989, Reines et al., 1999, Conaway and Conaway, 1993). The termination of coding-RNA genes is facilitated by recognition of a termination signal (PAS consensus sequence in human: AAUAAA, yeast: TAYRTA or AAWAAA with R: A/G, W: A/T and Y: C/T). Upon recognition of this signal the transcript is processed by addition of a polyA-tail and endonucleolytic cleavage. For that reaction the cleavage and polyadenylation factor (CPF) and cleavage factor (CF) bind to the pre-mRNA (Mischo and Proudfoot, 2013). Termination takes place 100 bp downstream of the 3'-end processing site. The termination of non-coding RNA genes, like snoRNA and snRNA, depends on CPF and CF accompanied by the NRD complex (composed of Nrd1-Nab3-Sen1) which recognizes specific RNA sequences. SnoRNA termination signals are defined by two motifs (Consensus sequence in yeast UCUUG and (U/A)GUA(A/G)) (Carroll et al., 2004,

Mischo and Proudfoot, 2013). After transcription the pre-mRNA in Eukaryotes undergoes several modifications to enlarge genetic variety by splicing and stabilization of mRNA is facilitated by addition of a 5'-cap and a 3'-polyA tail. The modified mRNA is exported from the nucleus and translated into a protein in the cytoplasm.

The archaeal transcription apparatus closely resembles the eukaryotic-RNAPII machinery but uses a minimal set of transcription factors. Homologs of general transcription factors (GTFs) TFIIB (TFB), TATA-binding protein (TBP) and TFIIE-α (TFE) were found in Archaea (Marsh et al., 1994, Langer et al., 1995, Kyrpides and Ouzounis 1999, Bell and Jackson, 2001, Hickey et al., 2002). For a long time it was thought that Archaea lack TFIIE-β, however sequence analysis revealed a possible candidate (RPC34 homolog, Saci_1342) (Blombach et al., 2009). Homologs of TFIIA, TFIIF and TFIIH are absent from Archaea. Archaea possess one multi-subunit RNAP which is similar to eukaryotic RNAPII in subunit composition and complexity (Langer and Zillig, 1993, Bell and Jackson 1998a, b). The model of pre-initiation complex formation includes binding of TBP to the TATA-box thereby kinking the DNA. Under certain conditions TFE facilitates binding of TBP to the TATA-box, but this transcription factor is discussed to be dispensable for sufficient initiation *in vitro* (Hanzelka et al., 2001, Bell et al., 2001, Werner and Weinzierl, 2005). The C-terminal part of TFB binds to the DNA-TBP complex and builds sequence-specific contacts with the TFB-responsive element (BRE). The RNAP is recruited by the N-terminal part of TFB. Then the closed promoter complex transits to the open promoter complex (Bell and Jackson, 1998a, b, Bell et al., 1999a, Bell and Jackson, 2001, Geiduscheck and Ouhammouch, 2005). Transcriptional elongation in Archaea is interplayed between archaeal, bacterial and eukaryal components. Archaeal proteins (Spt4, Spt5) that are involved in transcription elongation are homologs of bacterial proteins NusA and NusG, which are transcription factors that control elongation in Bacteria. In Archaea and Eukaryotes NusG is referred to as Spt5. In Archaea the heterodimeric Spt4/Spt5 complex binds to RNAP and prevents dissociation of nucleic acids and increases stability of the elongation complex (Martinez-Rucobo et al., 2011, Werner, 2013). TFE was shown to compete with Spt4/5 for RNAP binding which could facilitate promoter escape (Grohmann et al., 2011) and stimulates transcription processivity (Hirtreiter et al., 2010, Grohmann and Werner, 2011b). NusA binds to CU-rich regions near termination signals in the Crenarchaeote *Aeropyrum pernix* (Shibata et al., 2007). The subunits E and F of RNAP, also called Rpb4 and Rpb7 in yeast, respectively, form a complex which stimulates processivity of RNAP (Grohman and Werner 2011a, b, Hirtreiter et al., 2010). TFS, a homolog of the eukaryotic elongation factor TFIIS has been found in *Sulfolobus* and has been identified as part of the RNAP (Langer and Zillig, 1993, Bell and Jackson 1998a, b). TFS promotes transcript cleavage during elongation (Hausner et al, 2000, Grohmann and Werner 2011a). In contrast to Bacteria transcription termination in Archaea is factor independent and lacks secondary structures. Instead poly-dA sequences serve as termination signals comparable to the mechanism of RNAPIII (Brown et al., 1989, Santangelo and Reeve 2006, Grohmann and Werner, 2011a).

As outlined above, in Eukaryotes TBP interacts with TAFs and forms the TFIID complex that binds to the TATA-box. No homologs of TAFs have been found in Archaea (Paytubi and White 2009), but some archaeal genomes possess homologs of TBP-interacting proteins (TIP)

like *Archaeoglobus fulgidus*, *Pyrococcus horikoshii* (TIP49) and *Thermococcus kodakaraensis* (TIP26) (Soppa, 1999, Matsuda et al., 1999, Kyrpides and Ouzounis, 1999). *Tk*-TIP26 was shown to be a dimer (Yamamoto et al., 2006) and has an inhibitory effect of TBP binding to the DNA (Matsuda et al., 2001). TIP26 possesses a single Zn-finger motif but binding to DNA is weaker compared to binding of TBP to DNA. Sequence alignment with TIP49 from *A. fulgidus* (AF_1813) as input revealed that TIP proteins are not restricted to Euryarchaeota and they might also occur in Crenarchaeota like *Thermoproteus tenax* Kra1 (TTX_1511, E-value: 9e-148, identity: 53%), *S. acidocaldarius* DSM 639 (Saci_0552, E-value: 8e-144, identity: 51%) and *S. solfataricus* 98/2 (SsoI_0253, E-value: 5e-140, identity: 53%).

1.4. RNAPs in the three domains of life

All forms of life use DNA-dependent RNAPs for transcription. It has been suggested that the RNAPs are evolutionary related and several hypotheses address RNAPs evolution (Werner 2007, Werner, 2008, Werner and Grohmann, 2011, Iyer et al., 2003, Koonin et al., 2007). In all three domains of life RNAPs share a common structure and act by related mechanisms (Werner and Grohmann, 2011).

Bacteria possess one RNAP consisting of four subunits as core enzyme (α_2 β β' ω) and five as holoenzyme (core enzyme plus σ-factor) (Darst, 2001). The largest subunits β and β' are necessary for the catalytic function of RNAP. The α_2 subunits are identical in sequence but differ in position in the RNAP and are involved in assembly of the polymerase and transcriptional regulation. Polymerase assembly is also facilitated by the ω subunit (Minakhin et al., 2001).

Eukaryotes encompass up to five multi-subunit RNAPs (RNAPI-V). The three nuclear RNAPs (RNAP I-III) transcribe DNA into ribosomal RNA (rRNA), messenger RNA (mRNA) or transfer RNA (tRNA) and the smallest ribosomal subunit 5S rRNA, respectively. The other RNAPs are known to be plant-specific and generate small-interfering RNA (siRNA) (Haag and Pikaard, 2011, Ream, et al., 2009, Pontier et al., 2005). The RNAPII of *Saccharomyces cerevisiae* consists of twelve subunits designated Rpb1 – Rpb12 (Cramer et al., 2000), whereas RNAPI and RNAPIII are composed of 14 and 17 subunits, respectively.

Interestingly, although the transcription machinery of Archaea is quite similar to that of Eukaryotes, only one RNAP is present in Archaea (Langer et al. 1995, Best and Olsen, 2001). The archaeal RNAP closely resembles the eukaryotic RNAPII in complexity, subunit composition and function (Zillig et al., 1979, Langer, 1995, Werner, 2007). In 1979, Zillig and coworkers investigated the first archaeal RNAP from *S. acidocaldarius* and found that it possesses greater similarities to the eukaryal RNAP than to the bacterial one (Zillig et al., 1979). RNAP structures from *S. solfataricus* (Hirata et al., 2008), *S. shibatae* (Korkhin et al., 2009) and *Pyrococcus furiosus* (Kusser et al., 2008) have been solved. The archaeal RNAP exhibits nine (e. g. *Metallosphaera sedula*, Auernik et al. 2008) to fourteen subunits (e. g. *Nanoarchaeaum equitans*, Waters et al., 2003) depending on the organisms. The Crenarchaeotes *S. solfataricus* and *S. acidocaldarius* comprise twelve subunits (She et al.,

2001, Hirata et al., 2008, Zillig et al., 1979, Lanzendörfer et al., 1994, Chen et al., 2005) and the enzymes of *S. shibatae* and *Thermoproteus tenax* thirteen subunits (Korkhin et al., 2009, Siebers et al., 2011). One of the main differences between RNAPs from Eukaryotes or Archaea to that of Bacteria is the presence of a protruding stalk built by RpoE'/F (Rpb7/4 in *S. cerevisiae*) (Hirata et al., 2008, Werner and Grohmann, 2011). It promotes the open-complex formation and facilitates the action of TFE (Werner and Grohmann, 2011). All subunits of RNAPs can be assigned to different functions. These are assembly of the platform (D, L, N, P), catalytic function (A'/A", B'/B") as well as auxiliary function (E, F, H, K) in Archaea (Werner, 2007, Werner und Grohmann, 2011).

The four bacterial RNAP core subunits can be also found in Eukaryotes (Rpb3, Rpb2, Rpb1, Rpb6) and Archaea (RpoD, RpoB, RpoA, RpoK). The archaeal subunits E, F, G, H, L, N and P are homologous to eukaryal subunits Rpb 7, 4, 8, 5, 11, 10 and 12. The archaeal subunit RpoM has no eukaryotic counterpart and the *rpoM* gene possesses orthologs in the orders Sulfolobales and Desulfurococcales but is absent from Euryarchaeota (Wojtas et al., 2011). RpoG is conserved among hyperthermophilic Crenarchaeota and one Korarchaeote (*Korarchaeum* sp.) and resembles Rpb8 in the yeast strain *S. cerevisiae* (Koonin et al., 2007). Fig. 1 illustrates RNAP subunits and their homology from the three domains of life.

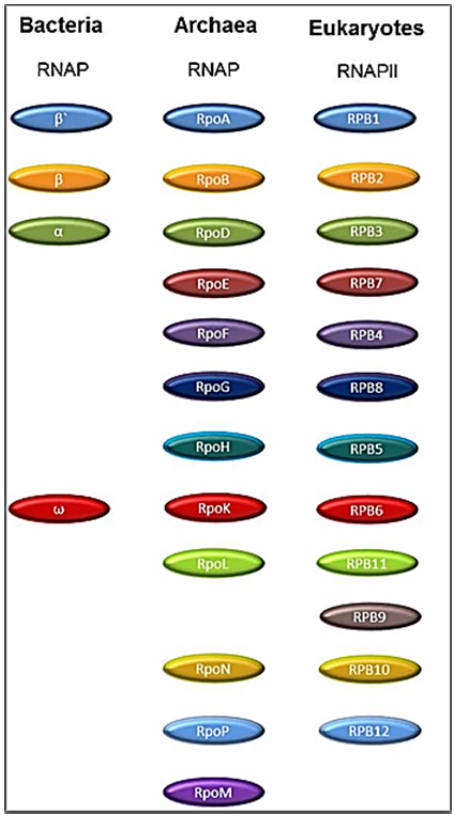


Fig. 1: Comparison of RNAP subunits of Bacteria, Archaea and RNAPII of Eukaryotes.
Homologous subunits are indicated in the same color (mod. from Goede et al., 2006).

A remarkable difference between eukaryal and archaeal RNAP is that the archaeal enzyme lacks the carboxyterminal heptapeptide (CTD) which is phosphorylated. Therefore Archaea do not need the transcription factors TFIIH and TFIIF, which are involved in phosphorylation. The CTD heptapeptide (Tyr1-Ser2-Pro3-Thr-4-Ser5-Pro6-Ser7) is repeated in varying

numbers from 26 in yeast up to 52 in vertebrates (for review see Hsin and Manley, 2012). The current model of RNAP phosphorylation states that Ser5 and Ser7 are phosphorylated at the stage of transcription initiation and that Ser2 and Thr4 are phosphorylated during elongation. Ser2 is phosphorylated by CDK9 (cyclin-dependent kinase) and Ser5 by the TFIIH-associated kinase CDK7 (Hsin and Manley, 2012, Hirose and Okhuma, 2007, Egloff and Murphy, 2008, Spain and Govind, 2011). Phosphatases, like Fcp1, Ssu72 and small CTD phosphatases (SCPs) that are not found in lower eukaryotes, are crucial for dephosphorylation of CTD. This domain of the RNAPII is involved in all steps of transcription and also provides the ability of histone modification and chromatin remodeling by recruiting factors. The CTD plays an essential role for pre-mRNA processing (capping, 3'-end processing, splicing) and termination of transcription (Hsin and Manley, 2012). Other post-translational modifications like glycosylation, ubiquitylation and methylation of the CTD are pointing towards its essential function in information processing.

1.5. Comparison of promoter structures in the three domains of life

Bacterial promoters consist of an UP-element, a -30 and a -10 region that is called Pribnow-Schaller box (consensus sequence in *E. coli*: -35: TTGACA, -10: TATAAT) (Harley and Reynolds, 1987, Pribnow, 1975, Schaller, 1975). The bacterial RNAP core enzyme (α_2, β, β', ω) binds weakly to the DNA without any sequence specificity. The σ-factor specifically recognizes the -35 and the -10 region of the DNA and enhances interactions between DNA and RNAP. The α-subunit additionally contacts the AT-rich UP-element.

RNAPII promoters that are found in Eukaryotes are composed of a BRE-site, a TATA-box, an initiator element (Inr) and a downstream-promoter element (DPE). The BRE-site harbors a pyrimidine-rich consensus sequence SSRCGCC (S: G/C, R: A/G) and mediates transcription initiation factor IIB (TFIIB) binding (Smale and Kadonaga, 2003). The TATA-box (*Drosophila melanogaster*: TATAAA, Smale and Kadonaga, 2003) is bound by the transcription initiation factor TFIID and it is positioned 25 bp upstream of the transcription start site (+1). The DPE is mostly present in TATA-less sequences in eukaryotic promoters and is located +28 to +32 downstream of the Inr element (for review see Smale and Kadonaga, 2003). TFIID binds to both elements in DPE-dependent promoters. The DPE consensus determined for *Drosophila* is RGWYV (R: A/G, W: A/T, Y: C/T, V: G/A/C) (for review see Smale and Kadonaga, 2003).

Archaeal promoter structures are similar to eukaryotic RNAPII promoters and possess the typical promoter elements (BRE, TATA-box, Inr). The BRE-site is a purine-rich sequence (Consensus sequence in Crenarchaeota: RNWAAW; *Pyrococcus*: VRAAA with N: any, R: A/G, V: A/C/G, W: A/T; van de Werken et al., 2006) which is highly conserved. It is located two to six bp upstream of the TATA-box and determines the orientation of the pre-initiation complex assembly and contributes to promoter strength. The TATA-box is recognized by TBP and comprises an eight bp A/T-rich sequence (Consensus sequence in Crenarchaeota: YTTTAWA; Halophiles: TTTWWW; Methanogens: YTTATATA; *Pyrococcus*: TTWWWAW with Y: C/T and W: A/T; Soppa, 1999b) positioned approximately 24 bp upstream of the transcription start site (+1). The initiator element determines the transcription

start site and shows only low sequence conservation consisting of a purine/pyrimidine dinucleotide pair which is probably recognized by RNAP. Recent experiments suggest that the presence of the Inr element might contribute to gene regulation in *Sulfolobus* (Ao et al., 2013).

1.6. Structure and function of transcription factors in Archaea and Eukaryotes

1.6.1. The TATA-binding protein TBP

TBP in Eukaryotes comprises an N-terminal part with little sequence conservation and a highly conserved C-terminal core domain which is required for TATA-box recognition and binding (Fikes et al., 1990, Nikolov et al., 1996, Burley et al., 1996). The C-terminal domain consists of two imperfect direct repeats which are flanked by a basic repeat (Kim and Roeder, 1994, Nikolov et., al 1996). Several X-ray structures of TBP and its isoform TBP2 of *Arabidopsis thaliana* have been reported (for review see Burley et al., 1996). Crystal structures of yeast and human TBP and TBP-TFB-DNA complexes have been solved (Chasman et al., 1993, Nikolov et al., 1992, Nikolov et al., 1995, Tsai and Sigler, 2000). They revealed that TBP contacts the minor groove of the DNA, bends the DNA and that TFB binds to the opposite site of the DNA helix (Nikolov, 1995, Tsai and Sigler, 2000).

TBP in Archaea consists of two imperfect direct repeats and in some cases a highly acidic six to ten amino acid long C-terminus (Bell and Jackson, 1998a, b, Hickey et al., 2002). The role of the acidic tail of the protein is still unclear. Archaeal and eukaryotic TBPs have a saddle like shape, but the archaeal TBP has a greater symmetry in regard to its three dimensional structure (Chasman et al., 1993, Nikolov et al., 1996, DeDecker et al., 1996, Bell et al., 1999a). It has been assumed that TBP evolved from a homodimer by gene duplication and fusion events (Bell et al., 1999a, Ouhammouch et al., 2009). Three dimensional structures of *P. woesei* and *M. jannaschii* TBP (DeDecker et al., 1996, Adachi et al., 2004, Adachi et al., 2008) as well as of the pre-initiation complex of *P. woesei* have been solved (Kosa et al., 1997, Littlefield et al., 1999).

1.6.2. Eukaryotic transcription factor TFIIB and its archaeal counterpart TFB

TFIIB in Eukaryotes consists of three domains namely the N-terminal part containing a Zn-finger motif, the linker comprising a B-reader and B-finger followed by the C-terminal domain consisting of two direct repeats (Bushnell et al., 2004, Deng and Roberts, 2007, Kostrewa et al., 2009, Wang and Roberts, 2010, Wiesler and Weinzierl, 2011). The co-crystal structure of a RNAP-TFIIB complex from yeast revealed that the N-terminal Zn-ribbon fold contacts the polymerase docking domain, the B-finger domain is positioned in the active center and the C-terminal part interacts with the polymerase and the TBP-DNA complex (Bushnell et al., 2004). Another crystal structure of the RNAPII-TFIIB complex of *S. cerevisiae* showed that the B-reader assists the recognition of the initiator sequence by binding the DNA template strand (Sainsbury et al., 2013). Mutations in the N-terminal part of TFIIB from *S. cerevisiae* point to a function of this domain in stable polymerase binding and

start site selection (Pardee et al., 1998). The Zn-ribbon interacts with the polymerase and TFIIF (Deng and Roberts, 2007, Bushnell et al., 2004).

Archaeal TFB proteins are also composed of three domains. The N-terminal region comprises a Zn-ribbon motif and a B-finger and the C-terminal part an imperfect direct repeat with a helix-turn-helix motif (HTH motif) (Bell and Jackson, 1998a, b, Bell and Jackson, 2000a, Hickey et al., 2002). The N- and the C-terminal part are linked by a flexible linker region. The N-terminal domain is necessary for RNAP recruitment and the B-finger is known to be critical for start site selection in Eukaryotes like yeast and humans as well as in Archaea (Zhu et al., 1996, Chen and Hampsey 2004, Bell and Jackson, 2000a, Bushnell et al., 2004). The B-finger functions in stabilization of the RNAP-DNA template thus stimulating transcription (Werner and Weinzierl, 2005, Weinzierl, 2010). Many crenarchaeal TFB sequences lack two of the four conserved Cys residues of the Zn-ribbon motif and it is not clear if they bind Zn (Dixit et al., 2004). The C-terminal part interacts with TBP and contacts the BRE-site (Littlefield, 1999). The linker region is involved in RNAP recruitment, stabilization of the pre-initiation complex and stimulates transcription ("superstimulation") (Wiesler and Weinzierl, 2011). The three-dimensional structure solved by Littlefield and colleagues revealed that two conserved residues (R283, V280) are crucial for formation of hydrogen bonds and van der Waals forces between *P. woesei* TFB and the BRE-site (Littlefield et al., 1999).

1.6.3. Eukaryotic transcription factor TFIIE and its archaeal homolog TFE

Human TFIIE consists of one α and one β subunit forming a heterodimer. TFIIE-α comprises a leucine-repeat, a Zn-finger and a HTH-domain, an alanine-rich region and an acidic region (Ohkuma et al., 1990, Ohkuma, 1997). TFIIE-β has a serine-rich region, a leucine repeat, and a basic C-terminus (Ohkuma et al., 1990, Ohkuma, 1997).

The N-terminus of TFE is weakly conserved and comprises a leucine-rich region within a winged HTH motif. The C-terminal Zn-ribbon normally consists of four cysteine residues, but in some species like *S. solfataricus* and *Archaeoglobus fulgidus* the second cysteine is replaced by aspartate or methionine, respectively. *Pyrococci* and *Methanobacterium thermoautotrophicum* possess proline or glycine residues at that position (Bell et al., 2001). The X-ray structure of *S. solfataricus* TFE revealed an unusual negatively charged groove and long wing, which contrasts to the DNA-binding ability of TFE (Meinhart et al., 2003). Indeed in electrophoretic mobility shift assays binding of TFE to DNA was not observed (Meinhart et al., 2003). TFE is involved in initiation and early elongation processes during transcription unlike its eukaryotic counterpart which is only present in initiation. Therefore in Archaea, this transcription factor is discussed not to be essential but affects transcription at weak promoters (Hanzelka et al., 2001, Bell et al., 2001, Grünberg et al., 2007). TFE binds to the pre-initiation complex and maybe stabilizes the transcription bubble during elongation by probably enhancing DNA melting and DNA loading (Grünberg et al., 2007). TFE directly interacts with subunits E and F of the RNAP and its function depends on the presence of the E/F heterodimer (Werner, 2007, Grünberg et al., 2007, Naji et al., 2007, Ouhammouch et al, 2004).

1.6.4. Multiple general transcription factors in Archaea and Eukaryotes

Intriguingly, in higher eukaryotes vertebrate-specific TBP paralogs exist which are designated TBP2 (TBP-related factor TRF2) and probably evolved by gene duplication (Gazdag et al., 2007). Experiments with *Xenopus* and zebrafish embryos revealed that TBP2 is crucial for embryogenesis. Akhtar and Veenstra (2009) showed that TBP2 is the major TBP in *Xenopus* oocytes and the TBP2 protein level decreases in eggs and early embryos whereas in the same time the level of TBP starts accumulating (Müller and Tora, 2009). Also plants developed multiple TBP and TFIIB proteins. The plant *Arabidopsis thaliana* possesses one TBP-isoform (TBP2) and two TFIIB proteins. Another example is *Oryza sativa,* which encompasses three TFIIBs and five TFIIB-like proteins (Brf1, Rrn7, Brp1, Brp2, Brp5) (Knutson, 2013).

In contrast to Eukaryotes, Archaea only possess a subset of general transcription factors. Strikingly, most Archaea possess multiple homologs of these GTFs, the actual number depending on the species. The highest number of transcription factor TFB was identified in *Nitrososphaera gargensis* with eleven TFB-encoding genes and one TBP-encoding gene (Spang et al., 2012). *Halobacterium salinarum* NRC-1 possesses six *tbp* (*tbp*A-F) and seven *tfb* genes (*tfb*A-G) that are encoded on one chromosome and two mini-chromosomes (Ng et al., 2000). The methanogen *Methanosarcina mazei* harbors only one TFB and three TBP proteins and *P. furiosus* as well as *Thermococcus kodakaraensis* possess two TFBs and one TBP (Deppenmeier et al., 2002, Fukui et al., 2005). *S. acidocaldarius* and *S. solfataricus* possess three TFB and one TBP-encoding sequences and *Thermoproteus tenax* has four TFB and one TBP-encoding genes (Chen et al., 2005, She et al., 2001, Siebers et al., 2011).

Adaptation to changing environmental conditions requires the ability of the cell to vary and regulate intracellular processes in order to survive in extreme and quickly changing environments. In Archaea it is still a matter of debate if multiple TFB and TBP homologs, which are mostly supposed to be derived from gene duplication, might be involved in stress response and thus might act like bacterial sigma-factors (Baliga, 2000, Micorescu et al., 2008, Lu et al., 2008, Peng et al., 2009, Paytubi and White, 2009, Turkarslan et al., 2011).

There is first evidence that a lineage-specific expansion of transcription factors occurred as specialization for an adaptation to unique and extreme environments in the halophile *H. salinarum* NRC-1, meaning that only a subset of transcription factors is used at one certain growth condition (Turkarslan et al., 2011). Mapping of transcription factors binding-sites of *H. salinarum* NRC-1 by ChIP-chip revealed that over 34% of all genes might be regulated by more than one transcription factor (Facciotti et al., 2007). In the same study binding of TFBf to promoters of genes, which encode proteins that are involved in ribosomal biogenesis and TFBg to promoters of genes involved in phosphate transport and phototrophy related proteins was observed (Facciotti et al., 2007). In addition, an important role in oxidative stress response has been proposed for TFBc (Kaur et al., 2010), and for TFBa and TBPd in heat response by microarray studies and growth experiments with deletion mutants (Turkarslan et al., 2011, Coker and DasSarma, 2007). For TFBe several functions like response to low temperature, high pH and low salinity were suggested (Turkarslan et al., 2011). From this study in the extreme halophile *H. salinarum* NRC-1 it was concluded that transcription factors, which were achieved by horizontal gene transfer or gene duplication, do not seem to

be essential and lead to adaptation to environmental conditions (Turkarslan et al., 2011). TFB2 from the hyperthermophilic Euryarchaeote *P. furiosus* appears to be necessary for heat-shock response (Micorescu et al., 2008) but no difference in function of TFB1 and TFB2 was observed in *T. kodakaraensis* using a mutational approach, suggesting that these TFBs can compensate each other (Santangelo et al., 2007). Moreover TBP of the mesophilic Euryarchaeote *M. mazei* was shown to be involved in heat shock response (Thomsen et al., 2001, DeBiase et al., 2002).

In the thermoacidophilic Crenarchaeote *S. acidocaldarius* the function of TFB1 is well established and it was shown to be the classical TFIIB homolog (Bell and Jackson 1999a, Bell and Jackson, 2001, Geiduscheck and Ouhammouch, 2005). Microarray studies indicated that TFB2 might be cell-cycle dependent and only present at the transition from G_1/S-phase (Lundgren and Bernander, 2007), but so far no experiments on protein level have been performed for TFB2. In contrast to TFB1 and TFB2 that are full-length proteins, TFB3 is only half of the size in *S. acidocaldarius* and lacks the B-finger and the first cycline-fold. In case of TFB3 microarray studies revealed a significant up-regulation after UV-irradiation suggesting a role in UV-induced stress response and a competition with TFB1 for binding to RNAP has been observed (Götz et al., 2007, Paytubi and White, 2009). TFE of *S. solfataricus* was shown to be selectively depleted upon heat shock. TFE protein levels were undetectable after 1 hour upon heat shock but recovered after 48 hours (Iqbal and Qureshi, 2010).

1.7. Transcriptional regulation in Archaea

Notably, the expression from the eukaryotic-like archaeal transcription apparatus in Archaea is controlled mostly by bacterial-like regulators (Kyrpides and Ouzounis, 1999, Aravind and Koonin, 1999). Regulators that were identified and characterized so far act by binding to promoter regions and thereby blocking the access of transcription factors to basal promoter elements and inhibiting RNAP recruitment or promoting transcription by recruitment of transcription factors (Geiduscheck and Ouhammouch, 2005). TFB-recruitment factor (TFB-RF, PF1088) from *P. furiosus* is the first transcriptional regulator in Archaea which activates transcription by recruitment of transcription factor B (Ochs et al., 2012). Numerous transcriptional regulators have been characterized so far (Brinkman et al., 2002, Brinkman et al., 2003, Sedelnikova et al., 2001, Dahlke and Thomm, 2002, Leonard et al., 2001, Yokoyama et al., 2007, Peeters et al., 2004, Peeters and Charlier 2010, Napoli et al., 1999, Bell and Jackson, 2000b, Fiorentino et al., 2003, Fiorentino et al., 2007, Fiorentino et al., 2011, Cohen-Kupiec et al., 1997, Lie et al., 2005). For example the transcriptional activators Ptr2 from *M jannaschii* regulates genes encoding proteins involved in electron transfer (Ouhammouch and Geiduscheck, 2001, Ouhammouch et al., 2003, Ouhammouch et al., 2005) and CopR from *S. solfataricus* responds to copper levels (Villafane et al., 2011). Examples of negative regulators are the metal-dependent repressor MDR1 from *A. fulgidus* (Bell et al., 1999b) and GvpD (Scheuch et al., 2008, Hofacker et al., 2004, Pfeifer et al., 2001). Gas vesicle formation in Haloarchaea is regulated by the repressor GvpD and by the activator GvpE (Scheuch et al., 2008, Hofacker et al., 2004, Pfeifer et al., 2001). Interestingly, GvpE possesses an eukaryal-type basic leucine zipper (bZIP) region making it one of the few

examples for eukaryal-like regulators in Archaea (Krüger et al., 1998, Plößer and Pfeifer, 2002).

1.8. Aims of this study

Whereas multiple *tfb* and *tbp* homologs have been deeply investigated in the euryarchaeal branch, the role of multiple GTFs in Crenarchaeota remains poorly understood. It has been proposed that various *tfb* homologs might serve like σ-factors in Bacteria and fulfill specialized functions according to stress response (Micorescu et al., 2008, Peng et al., 2009, Paytubi and White, 2009). The aim of this study is to provide multiple GTF proteins required for experiments like EMSAs and *in vitro* transcription and to gain a deeper insight into their function in *S. acidocaldarius* by the generation of reporter gene constructs, protein-protein interaction studies and generation of knock-out mutants. The thermoacidophile *S. acidocaldarius* grows optimally at 80°C and pH 2-3 (Brock et al., 1972) and is an obligate aerobic organism; it grows heterotrophically and requires a limited range of polysaccharides or amino acids (e. g. dextrin, starch, tryptone and glutamate) for growth (Grogan et al., 1989). The major advantage working with *S. acidocaldarius*, compared to other Crenarchaeota like *T. tenax*, is that a genetic system is established (Wagner et al., 2009, Wagner et al., 2012, Sakofsky et al., 2011). An expression system is available and markerless knock-out mutants can be generated. The function of TFB1 and TFB3 has only been investigated in few studies in Crenarchaeota (Bell and Jackson, 2000a, Götz et al., 2007, Paytubi and White, 2009). A first hint for the function of TFB2 was given by microarray studies and a cell-cycle dependent role was proposed (Lundgren and Bernander, 2011). However, TFB2 from *S. acidocaldarius* has not been characterized on protein level so far. The role of GTFs will be investigated by expression of the proteins, reporter gene assays, protein-protein interaction studies with the RNAP and GTFs and knock-out mutant analysis for determination of *in vivo* function.

2. Material and Methods

2.1. Instruments

The instruments used in the present study are listed in Table 1.

Table 1: List of instruments used in this study

Instruments	Description
Fast Flow Liquid Chromatography (FPLC)	Äkta purifier GE Healthcare (Freiburg)
Chromatography columns	HiLoad 26/60 Superdex 200/75 prep grade, GE Healthcare (Freiburg), ResourceQ 6 ml GE Healthcare (Freiburg), Ni-TED, Macherey-Nagel (Mannheim), Strep-tag columns IBA (Göttingen)
Autoclaves	H+P Varioklav, 25T, Federgari autoclave, Integra Bioscience (IBS) (Fernwald) H+P Varioklav, 75S, Federgari autoclave, Integra Bioscience (IBS) (Fernwald)
Agarose gel electrophoresis system	Agagel Mini, Biometra (Göttingen) Power supply: Consort E143 (MS Laborgeräte)
Thermocycler	Mastercycler personal, Eppendorf (Hamburg) Thermocycler C1000, Bio-Rad Laboratories GmbH (Munich)
Vertical gel electrophoresis system	Mini-gel Twin, Biometra GmbH (Göttingen) Power supply: Consort E143 (MS Laborgeräte)
Ultraviolet light	TCP-20.C, Vilbert Lourmat (Marne La Valle, France)
Incubators	Infors HT Unitron (Bottmingen, Switzerland) Heraeus B6 Kendro (Langenselbold) Heraeus T20 Kendro (Langenselbold) Thermotron Infors
Sonicator	Ultraschallprozessor UP 200s, Oehmen Labortechnik (Hielscher)
French Press	French Pressure Cell, FA-078AE, Thermo Electron Corporation (Milford, USA)
Centrifuges	Sorvall centrifuge RC26, Kendro (Langenselbold) Rotor: Sorvall GSA, Kendro (Langenselbold), MiniSpin, centrifuge 5804R Eppendorf (Hamburg)
Heat Block	Thermoblock v4.6, Hardware & Service (Friedland)
Water bath	Thermomix UB, B. Braun (Melsungen)
Sterile work bench	Hera Safe KS12, 1,200x780x627 mm (Langenselbold)
Filtration system	KNF Lab, Laboport (Freiburg)
VersaDoc gel documentation	Versadoc Imaging System, 4000MP, Bio-Rad Laboratories GmbH (Munich) Tamron AF Aspherical 28 – 80 mm
Microwave	HP1612, Siemens (Munich)
Stratalinker®	StratalinkerTM 2400 Stratagene (La Jolla, USA)

2.2. Chemicals

The chemicals and enzymes used in this work were purchased from Bio-Rad Laboratories GmbH (Munich), Difco Laboratories (Augsburg), Fermentas (ThermoScientific, St. Leon Rot), Gerbu Biotechnik GmbH (Heidelberg), Qiagen (Hilden), Roche Diagnostics GmbH (Mannheim), Roth GmbH (Karlsruhe), Serva Electrophoresis GmbH (Heidelberg), Sigma-Aldrich (Taufkirchen), Takara Bio Europe/Clontech (Saint-Germain-en-Laye, France).

2.3. Plasmids

The plasmids from the pET vector series (pET11c, pET15b, pET30a) were purchased from Novagen and the directional cloning vector pET151/D-TOPO® from Invitrogen. The plasmids pPREX and pSVA1450 for homologous expression and reporter gene assays in *S. acidocaldarius* and pSVA406 for genetic manipulation (knock-in/knock-out) were kindly provided by Dr. Sonja-Verena Albers (MPI, Marburg). The vectors pGADT7, pGBKT7, pGADT7::*SsorpoK* and pGBKT7::*Ssotfb3* used in yeast two-hybrid experiments were provided by Prof. Dr. Malcolm White (University of St. Andrews, UK).

2.4. Strains and growth conditions

Escherichia coli strain DH5α (DSMZ 6897) (Hanahan, 1983) and *E. coli* One Shot® TOP10 cells (Invitrogen) were used for cloning and storage of the plasmids. The strains *E. coli* BL21 (DE3) (with and without pRIL), Rosetta (DE3), BLR (DE3) pLysS (Novagen), Lemo21 (DE3) (New England Biolabs) and BL21 Star™ (DE3) One Shot® (Invitrogen) (Studier and Moffat, 1986) were used for heterologous overexpression of *S. acidocaldarius* proteins. The strains were grown in 5 ml up to 2 L volumes of LB-medium (Sigma-Aldrich) in reaction tubes or Erlenmeyer-flasks which were incubated under aerobic conditions at 37°C in a shaker at 180 rpm. Agar plates were prepared containing 1.5% (w/v) agar.

Saccharomyces cerevisiae AH109 was grown aerobically at 30°C with shaking at 200 rpm in YPD medium (1% (w/v) Bacto yeast extract, 2% (w/v) Bacto peptone) in the presence of 2% (v/v) sterile glucose (40% (w/v) stock solution) or in transformation and selection medium (SD -Leucine/-Tryptophane or SD -Leucine, -Tryptophane, -Histidine, -Adenine, +X-α-galactose) in a volume of 5 ml up to 200 ml. For preparation of selective agar plates 2.5% (w/v) agar was added to the medium and autoclaved.

Sulfolobus acidocaldarius DSM 639 was incubated in Brock minimal salt medium (Brock et al., 1972) with 0.1% (v/v) tryptone (20% (w/v) stock solution) and 0.2% (v/v) dextrin (20% (w/v) stock solution) for growth under non-selective conditions. For first selection tryptone was replaced by 0.1% (v/v) N-Z-amine® (20% (w/v) stock solution) (Sigma-Aldrich). In case of second selection 2 ml uracil (final concentration: 10 µg/ml, 5 mg/ml stock) and 1 ml 5-fluoro-orotic acid (5-FOA, 100 mg/ml stock) were added per liter of Brock medium. For preparation of plates 6 g gellan gum (0.6% (w/v)) per one liter final volume were transferred into a sterile 1 L Erlenmeyer flask. One liter water was split into two halves and with one half the complete Brock medium (recipe for 1 L) was prepared and heated at 78°C for 30 min. Six ml of sterile $CaCl_2$ (0.5 M) and 10 ml $MgCl_2$ (1 M) were added to allow the polymerization of

the plates. The other half of water was added to the gellan gum and heated on a heat block with stirring until boiling. Both parts of the medium were combined and mixed on a magnetic stirrer before pouring the plates. For liquid cultures cells were grown under aerobic conditions at 78°C with shaking (180 rpm) in an incubator (Thermotron, Infors). After spreading cells from liquid cultures onto solid media, plates were wrapped in plastic for incubation to prevent desiccation. Plates were incubated at 78°C for five to seven days.

Brock I (500 ml):
476.14 mM $CaCl_2$ x 2 H_2O

Brock II (1 L):
98.38 mM $(NH_4)_2SO_4$
101.4 mM $MgSO_4$ x 7 H_2O
add 1.5 ml 50% (v/v) H_2SO_4

Brock III (500 ml):
411.52 mM KH_2PO_4
Volumes refer to 1% stock solutions of the following additives
18 ml $MnCl_2$ x 4 H_2O
45 µl $Na_2B_4O_7$ x 10 H_2O
2.2 ml $ZnSO_4$ x 7 H_2O
500 µl $CaCl_2$ x 2 H_2O
300 µl Na_2MoO_4 x 2 H_2O
300 µl $VoSO_4$ x 2 H_2O
100 µl $CoSO_4$ x 2 H_2O
750 µl of 50% (v/v) H_2SO_4

Autoclave Brock media at 121°C for 20 min

Final Brock`s medium (1 L):
1 ml Brock I
10 ml Brock II
5 ml Brock III
1 ml Fe_2Cl_3 x 6 H_2O solution (20 g/l, sterilize by filtration, store in the dark)
5 ml tryptone (20% (w/v))
10 ml dextrin (20% (w/v))
200 µl H_2SO_4 (50% (v/v))

Overexpression medium (1 L):
1 ml Brock I
10 ml Brock II
5 ml Brock III
1 ml Fe_2Cl_3 x 6 H_2O solution (20 g/l, sterilize by filtration, store in the dark)
5 ml N-Z-Amine® (20% (w/v))

10 ml sucrose (20% (w/v))
20 ml maltose (20% (w/v))
200 µl H_2SO_4 (50% (v/v))

2.5. Preparation of glycerol stocks of *S. acidocaldarius*

Cells were grown until OD_{600} reached 0.5 to 0.8 and 4 ml were collected twice in 2 ml centrifuge tubes in a bench top centrifuge (6320 x g, 2 min, RT). The pellet was resuspended in glycerol Brock medium (50 µl Brock I, 500 µl Brock II, 250 µl Brock III, 50 µl 20 g/l $Fe_2Cl_3^*$ 6 H_2O, 25 ml H_2O, 25 ml glycerol) and the cells were stored at -80°C.

2.6. Basic molecular biology methods

2.6.1. Preparation of chromosomal DNA from *S. acidocaldarius* and quantification of DNA

The Blood and Tissue DNA preparation kit from Qiagen was used for isolation of chromosomal DNA. *S. acidocaldarius* DSM 639 or the respective mutants were grown until OD_{600} reached 0.8. Cells of 4 ml culture were collected and the resulting pellet was treated according to manufacturer's instruction. For elution of the DNA from the column 90 µl DNase/RNase-free water was added and the column was incubated for 1 min at room temperature and centrifuged at 12,000 x g for 1 min. The concentration and quality of the DNA was controlled spectrophotometrically as well as by agarose gel electrophoresis. The DNA-concentration was measured at a wavelength of 260 nm using the photometer (Eppendorf). OD_{260} was assumed as 50 µg/ml of double stranded DNA (Sambrook et al., 1989). To control the purity of the DNA additional measurements at 230 nm and 280 nm were performed for determination of ethanol and protein contaminations (Marmur, 1961). A lower ratio of OD_{260}:OD_{280} than 1.8 indicates protein or phenol contamination. Contaminants like guanidine HCl used for DNA extraction and carbohydrates absorb at 230 nm and thereby lead to a decreased OD_{260}:OD_{230} ratio.

Pure DNA: OD_{260} : OD_{280} = 1.8 (Sambrook et al., 1989)
OD_{260} : OD_{230} = 2.0 - 2.2

2.6.2. Isolation of plasmid DNA from *E. coli*

For isolation of high-copy vectors (pGADT7/pGBKT7) cells from at least 5 ml overnight cultures and for low-copy vectors (pET vector series, pPREX, pSVA1450, pSVA406) 10 ml up to 50 ml were collected by centrifugation. Plasmid isolation was carried out by using the Plasmid Mini-preparation kit from Fermentas (St. Leon Rot, Germany).

2.6.3. Agarose gel electrophoresis

Depending on the size of the DNA samples 0.8 to 1.5% (w/v) agarose gels were used for separation of PCR products, restriction digests or genomic DNA. Agarose gels were prepared with 1x TAE buffer (40 mM Tris-Acetate, 1 mM EDTA, pH 8.2-8.4 at 25°C). For sample preparation and to monitor sample separation loading buffer containing the dye bromophenol blue was added (6x Loading buffer: 0.2% (w/v) bromophenol blue, 0.2% (w/v) xylencyanol FF, 60% (v/v) glycerol, 60 mM EDTA). The size marker GeneRuler^{TM} 1 kb DNA ladder

(Fermentas, St. Leon Rot) was loaded for determination of DNA fragment size. After electrophoresis at 100 V for 50 min to 1 h the gels were stained for 15 min in an ethidium bromide bath (2 mg/ml) and destained for 1 min in water. Detection of the DNA samples was performed under UV-light by the GelDoc-gel documentation system (Bio-Rad, Munich).
Purification of DNA samples from gels was achieved by the Wizard clean up PCR purification and Gel Extraction kit from Promega (Mannheim).

2.6.4. Amplification of DNA fragments via PCR

For amplification of DNA-fragments 1 pg to 100 ng plasmid DNA or 50 ng to 250 ng genomic DNA per 50 µl reaction were added as template. 5x Phusion HF buffer was diluted to a final 1x concentration, 200 µM of each dNTP was added as well as 0.5 µM forward and reverse primer, 3% (v/v) DMSO, and 0.02 U/µl Phusion polymerase (Biozym). PCR was performed using a thermocycler (Eppendorf, Bio-Rad). The two-step protocol encompassed an initial denaturation step at 98°C for 30 s, followed by a 98°C denaturation for 10 s, 55°C primer annealing for 30 s and elongation time lasted according to the calculated PCR product size at 72°C. The primer annealing temperature of 55°C was used for all amplification reactions in this study. In general, Phusion polymerase elongates 1000 bp per 15 s to 30 s. The cycle was repeated for 32 times and the final extension was programmed at 72°C for 10 min.

2.6.5. Restriction of DNA

PCR-products and plasmids were restricted with fast digest enzymes from Fermentas (St. Leon Rot) according to the manufacturer´s instruction. Double-stranded DNA was incubated with an appropriate amount of enzymes at the optimal temperature and time. The restriction enzymes used in this study are listed together with the respective primer sequences (Table 5, Table 7-15).

2.6.6. Ligation of vector and insert

For ligation 75 ng of restricted vector and insert in a molar ratio of 1:3 (vector: insert) were used. Water, restricted insert and vector were incubated at 45°C for 5 min in 8 µl reaction volume. After cooling on ice and spinning down the samples, 1 µl ligase-buffer (Fermentas) and 1 µl T4-DNA-ligase were added and incubated overnight at 16°C. The reaction was heat-inactivated for 10 min at 65°C.

2.6.7. Preparation of chemically competent *E. coli* cells

Chemical competent *E. coli* cells of the strains DH5α, BL21 (DE3) (with and without pRIL), Rosetta (DE3), BLR (DE3) pLysS and Lemo21 (DE3) were prepared according to a modified calcium chloride method (Hanahan, 1983). 200 ml LB-medium were inoculated with an overnight culture of *E. coli* cells to a final concentration between 1% (v/v) and 3% (v/v). The cells were incubated at 37°C with shaking (180 rpm) until an OD_{600} of 0.4 was reached. The cell suspension was chilled on ice for 10 min and then centrifuged at 8,450 x g for 7 min at 4°C. The cells were resuspended in ice-cold $CaCl_2$ solution (60 mM $CaCl_2$, 15% (v/v) glycerol, 10 mM PIPES, pH7.0) and centrifuged a second time at 8,450 x g for 5 min at 4°C.

The pellet was carefully resuspended in 2 ml ice-cold $CaCl_2$ solution. 100 µl aliquots were frozen in liquid nitrogen and stored at -80°C.

2.6.8. Transformation of competent cells by heat-shock

10 µl of the ligation mixture or 200 ng plasmid were added to chemical competent *E. coli* cells and incubated on ice for 30 min. Then the cells were heat-shocked for 90 s at 42°C and directly transferred to ice for 2 min. After addition of 500 µl S.O.C. medium (0.5% (w/v) yeast extract, 2 % (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose (0.5 M glucose stock solution) were added after autoclaving) the cells were incubated at 37°C with shaking (180 rpm) for 1 hour. The cells were harvested at 16,910 x g for 45 s and the resulting pellet was resuspended in 100 µl S.O.C medium and spread on selective LB-agar-plates containing the appropriate antibiotics.

2.6.9. DNA sequencing

100 ng plasmid or 10 to 40 ng PCR product were sequenced by LGC Genomics (Berlin) using the automated Sanger sequencing (Sanger et al., 1977). If required, 2 µl (10 µM) of primers were added to the reaction. The sequences were analyzed using the Clone Manager program (Scientific & Educational Software, USA).

2.6.10. Sequence analysis

For BLAST (Basic local alignment search tool, Altschul et al, 1990) searches on protein or DNA level the databases KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/), NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/) and IMG (Integrated Microbial Genomes, Markowitz et al., 2012, http://img.jgi.doe.gov/home.html) were used. The ClustalW program was used for sequence alignments (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The genomic context of *tfb*2 was determined by GCview available on the website of Bioinformatics Toolkit at Max-Planck-Institute for Developmental Biology in Tübingen (http://toolkit.tuebingen.mpg.de/gcview) and using the SyntTax program (http://archaea.u-psud.fr/synttax/).

2.6.11. Gene synthesis

Due to the fact that the codon usage is very specific for each organism the codon-usage of all three *tfb* genes of *S. acidocaldarius* (*tfb*1, *tfb*2, *tfb*3) was adapted to *E. coli*. The resulting sequences were synthesized by the company GenScript (Piscataway, USA) and flanked with *Nde*I and *BamH*I restriction sites and were introduced into the cloning vector pUC57. The plasmids were restricted with *Nde*I and *BamH*I and the genes were ligated into the expression vectors pET11c and pET15b (Novagen, USA).

2.7. Analysis of genome organization of *tfb*2

2.7.1. Isolation of total RNA from *S. acidocaldarius* cells

S. acidocaldarius wild-type and *S. acidocaldarius* MW001 cells were grown to early-log, mid-log and stationary phase. The cell pellets of 15 ml culture were resuspended in 750 µl

Trizol per g cells wet weight. After addition of 200 µl chloroform, samples were incubated for 3 min at room temperature (RT) and subsequent centrifugation at 20,800 x g for 15 min at 4°C resulted in phase separation. The supernatant was transferred into a fresh reaction tube and 350 µl RLT buffer (premixed with 10 µl β-mercaptoethanol (14.3 M) on 1 ml buffer) were added to the solution. The samples were mixed by vigorous shaking and centrifuged at 20,800 x g for 2 min at room temperature. The supernatant was removed and 250 µl 100% ethanol were added. 700 µl of this solution were loaded onto a Qiagen RNeasy MiniKit column and centrifuged at 20,800 x g for 1 min at room temperature. Afterwards, the columns were transferred into fresh 2 ml tubes and 500 µl RPE buffer were used for washing. Washing steps were carried out two times at 20,800 x g at room temperature for 1 min and 2 min. For removal of residual ethanol the tubes were centrifuged for another 5 min with the same settings as mentioned before. The tubes were placed into fresh 1.5 ml reaction tubes and 30-50 µl DNase/RNase-free water was added onto the center of the columns and incubated for 1 min. RNA was eluted by centrifugation at 20,800 x g for 1 min. An OD_{260} of 1.0 resembles 40 µg/ml of single-stranded RNA (Sambrook et al., 1989). RNA concentration and purity was measured photometrically at 260 nm in 10 mM Tris-HCl pH 7.5 buffer. A lower ratio of OD_{260}:OD_{280} than 1.7 indicates protein or phenol contamination. If the OD_{260}:OD_{230} ratio is lower than the expected value of 2.0 it indicates the presence of contaminants that absorb at 230 nm like EDTA, carbohydrates and phenol.

Pure RNA: $OD_{260} : OD_{280} = 1.7$ to 2.0 (Sambrook et al., 1989)

$OD_{260} : OD_{230} = 2.0 - 2.2$

2.7.2. Removal of genomic DNA by DNaseI treatment

DNaseI digestion was performed following manufacturer's instruction (ThermoScientific). Reactions with 1 µg RNA, 1 µl reaction buffer, 1 µl DNaseI (1 U, RNase-free) were scaled up to 10 µl with DEPC-treated water. Samples were incubated for 30 min at 37°C. To stop the reaction 1 µl 50 mM EDTA was added and the RNase-free tubes were incubated at 65°C for 10 min. For further purification the samples were cleaned using the RNeasy Mini kit (Qiagen).

2.7.3. First strand cDNA synthesis

Reactions were carried out according to the GoScript Reverse transcription system manual. Up to 5 µg RNA, 1 µl of random and $Oligo(dT)_{15}$ primer and nuclease-free water were combined to a final volume of 5 µl. The mixture was heated at 70°C for 5 min and immediately chilled on ice for at least 5 min. Each tube was centrifuged to obtain the original volume. The tubes were kept on ice until the reverse transcription mixture was added. 4 µl reaction buffer, 1.2 µl (1.5 mM) $MgCl_2$, 1 µl PCR nucleotide mix, 0.12 µl ribonuclease inhibitor and 1 µl reverse transcriptase were combined and filled up to a final volume of 15 µl with nuclease-free water. 15 µl of this reverse transcription mix were added to 5 µl RNA/primer mix. For the annealing step samples were incubated at 25°C for 5 min and extension was performed for one hour at 42°C. The reaction was heat-inactivated at 70°C for

15 min. Resulting cDNA was stored at -20°C or used directly for PCR. In Table 2 the primer sequences used for the operon prediction are summarized.

Table 2: Primer sequences used for operon prediction of Saci_1339, *gar*1, *tfb*2 and Saci_1342

Gene	Sequence (5'-3')
1476_Saci_1340_F *(gar1)*	GTGAAACAAATAAAGCTTGTTGAAGCTG
1477_Saci_1340_R *(gar1)*	TCATCTTCTTCTCCTCCCTTTC
1478_Saci_1341_F *(tfb2)*	ATGAAATGTAAAATCTGTGGCAGTGAG
1479_Saci_1341_R *(tfb2)*	TCAAAGCTTTACTTCTATGTCAAAAG
1480_Saci_1342_F (hypothetical protein)	ATGGGGAAAGAGAATATATTATTTGG
1481_Saci_1342_R (hypothetical protein)	TTATCCAGCAGAAACACTAGTTAAG

2.8. Basic biochemical methods

2.8.1. Protein quantification using Bradford reagent

Protein concentration was measured at 280 nm using the Bio-Rad Bradford protein assay according to the manufacturer`s instruction (based on Bradford, 1976). A standard curve was prepared using different concentrations of bovine serum albumin (0-10 µg/ml). Samples were incubated for 5 min at RT in the dark.

2.8.2. BC assay for protein quantification (Brown et al., 1989)

The colorimetric BC assay (Uptima) was performed according to the manufacturer´s instruction. A standard curve was prepared using bovine serum albumin (20 µg/ml – 2 mg/ml). 25 µl of standard or sample were pipetted into a 96-well plate. 200 µl BC assay reagent (mix A + B 50:1) were added. The plates were incubated at 37°C for 30 min. After cooling the plates to RT absorbance was determined at 562 nm in the microplate reader (Tecan).

.

2.8.3. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was prepared for protein separation according to Laemmli (Laemmli, 1970). Depending on the molecular size of the proteins acrylamide concentrations ranged between 10%, 12.5% and 15% (v/v) (Table 3). The resolving gel components were mixed and poured between two glass plates and covered with isopropanol. After polymerization isopropanol was fully removed and the stacking gel (Table 4) was poured and a comb was placed between the two glass plates. Protein samples were mixed with 4 x loading buffer (final concentration: 62.5 mM Tris-HCl pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 5% (v/v) β-mercaptoethanol, 0.005% (w/v) bromophenol blue) and boiled for 10 min at 95°C and centrifuged (6,000 x g, 30 s, RT) before application. The running buffer consisted of 25 mM Tris-HCl, pH 8.3,

190 mM glycine and 0.1% (w/v) SDS. After sample application the run was performed for one gel at 25 mA and for two gels at 50 mA for 50 min using a Bio-Rad chamber. Protein standards with various molecular weight ranges were used for size determination (PageRuler™ Unstained Protein Ladder, PageRuler™ Unstained Low Range Protein Ladder, Fermentas). Pre-stained molecular weight markers were used for western blotting (PageRuler™ Prestained Protein Ladder 10-170 kDa, Spectra™ Multicolor Broad Range Protein Ladder, Spectra™ Multicolor High Range Protein Ladder, Spectra™ Multicolor Low Range Protein Ladder, Fermentas). After electrophoresis the gels were stained for 30 min in Coomassie Brilliant Blue solution (40% (v/v) methanol, 10% (v/v) acetic acid, 0.25% (w/v) Coomassie Brilliant Blue R-250) and destained with water using a microwave. SDS-polyacrylamide (SDS-PAA) gels were analyzed with the Bio-Rad GelDoc system.

Table 3: Composition of SDS-PAA resolving gel with different polyacrylamide concentrations

	10% (v/v)	12.5% (v/v)	15% (v/v)
H_2O	4.2 ml	3.3 ml	2.5 ml
Resolving gel buffer (375 mM Tris-HCl, pH 8.8, 0.1% (w/v) SDS)	2.5 ml	2.5 ml	2.5 ml
Acrylamide solution (30% (w/v)) 37.5 : 1	3.3 ml	4.2 ml	5 ml
10% (w/v) Ammoniumpersulfate	70 µl		
TEMED	5 µl		

Table 4: Composition of 4% (v/v) stacking gel for SDS-PAA gels

H_2O	1.85 ml
Stacking gel buffer (125 mM Tris-HCl, pH 6.8, 0.1% (w/v) SDS)	750 µl
Acrylamide solution (30% (w/v)) 37.5 : 1	400 µl
10% (w/v) Ammoniumpersulfate	14 µl
TEMED	3 µl

2.9. Heterologous expression of transcription factors in *E. coli*

2.9.1. Cloning of GTFs from *S. acidocaldarius* into overexpression vectors

Primer sequences used for cloning of the transcription factors *tfb*1, *tfb*2, *tfb*3 and *tbp* into pET11c are listed in Table 5. Codon usage optimized sequences of *tfb*1, *tfb*2 and *tfb*3 were used for cloning into the expression vector pET15b and the corresponding optimized gene sequences are shown in the supplement (Fig. 37).

Table 5: Primer sequences for cloning of transcription factors *tfb*1, *tfb*2, *tfb*3 and *tbp* into pET15b and pET11c

Gene	Primer sequence (5'-3')	Vector
*tfb*1 (Saci_0866)	synthesized gene, *Nde*I/*BamH*I	pET15b (N-terminal His tag)
*tfb*2 (Saci_1341)	synthesized gene, *Nde*I/*BamH*I	pET15b (N-terminal His tag)
*tfb*3 (Saci_0665)	synthesized gene, *Nde*I/*BamH*I	pET15b (N-terminal His tag)
*tfb*1_*Nde*I_F (Saci_0866)	GCGCGGCATATGGTAGAACAAAGTAAGGTTCCT TCT	pET11c
*tfb*1_*Bam*HI_R (Saci_0866)	GGCGCCGGATCCATGGTAGAACAAAGTAAGGTT CCTTC	
*tfb*2_*Nde*I_F (Saci_1341)	GCGGCGCATATGAAATGTAAAATCTG TGGCAGTG (internal *Nde*I restriction site)	pET11c
*tfb*2_*Bam*HI_F (Saci_1341)	GGCGGGATCCTCAAAGCTTTACTTCT ATGTCAAAAGC	
*tfb*3_*Nde*I_F (Saci_0665)	CGCCGCGCATATGGAATGTCCCGAGTGTAAA	pET11c
*tfb*3_*Bam*HI_R (Saci_0665)	GGCGCCGGATCCTTACATCGTTTTTAACGTTTGG	
*tbp_NdeI*_F (Saci_1336)	GCCGCGCATATGGAAAGAAGCGTACCAAAC	pET11c
*tbp_BamHI*_R (Saci_1336)	GGCGCCGGATCCTTAAAATTCTAACTCTTCTTC	
T7 prom_F	TAATACGACTCACTATAGGG	sequencing
T7 term _R	CTAGTTATTGCTCAGCGGTG	sequencing

underlined: restriction site

The genes *tfb*1, *tfb*2, *tfb*3 and *tbp* were amplified via PCR using genomic DNA isolated from *S. acidocaldarius* as template and were cloned into pET11c. The primers used for PCR are listed in Table 5.

Directional TOPO® cloning was carried out according to the manufacturer´s instruction (Invitrogen). Directional cloning enables cloning of blunt-end PCR products into TOPO® vectors. The plasmid pET151/D-TOPO® (Invitrogen) contains a single-strand 5'-GTGG overhang and a 3' blunt-end. The overhang invades double-strand DNA of the generated PCR product and anneals to the 5'-CACC sequence, which was introduced by the forward primer. Primers used for TOPO® cloning are listed in Table 6. Topoisomerase I ligates the PCR in the correct orientation. The vector pET151/D-TOPO® contains an N-terminal His-tag and a TEV-protease cleavage site for removal of the tag after purification of the protein. A 0.5:1 to 2:1 molar ration of PCR product: vector was used for the following cloning reactions. 1 µl buffer, 1 µl vector, 0.5 to 4 µl PCR product and water (final volume of 5 µl) were mixed and incubated for 5 min at RT. The mixture was cooled on ice and 3 µl were added to one vial of chemical competent *E. coli* OneShot® TOP10 cells (Invitrogen). The sample was incubated

on ice for 30 min and cells were incubated at 42°C for 30 s. Immediately after heat shock the tube was placed on ice for 2 min. 250 µl pre-warmed S.O.C. medium were added and the cells were incubated at 37°C for 1 h with shaking (180 rpm). 100 µl of the cell suspension were spread on LB-plates containing ampicillin (100 µg/ml) for selection and plates were incubated over night at 37°C.

Table 6: Primer sequences for cloning of transcription factors into pET151/D-TOPO®

Gene	Primer sequence (5'-3')
*tfb*1_TOPO_F	CACCATGGTAGAACAAAGTAAGGTTCCTTCT
*tfb*1_TOPO_R	CTCGAGTTGACTTGGAATCTGTATC
*tfb*2_TOPO_F	CACCATGAAATGTAAAATCTGTGGCAGT
*tfb*2_TOPO_R	TCAAAGCTTTACTTCTATGTCAAAAGC
*tfb*3_TOPO_F	CACCATGGAATGTCCCGAGTGTAAAA
*tfb*3_TOPO_R	TTACATCGTTTTTAACGTTTGGC
*tbp*_TOPO_F	CACCATGGCCCAGATAAGAGAGATTAAG
*tbp*_TOPO_R	TCATTTCAATAATAGATTCTCATATTCC
017_seq_pET151/D-TOPO_F*	CCATGGTAAGCCTATCCCTAAC
T7 term _R*	CTAGTTATTGCTCAGCGGTG

underlined: directional cloning site; * sequencing primer

The genes *tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe* and *rpo*G were amplified via PCR mutagenesis from genomic DNA from *S. acidocaldarius*. The PCR products were restricted with *Nde*I and *Xho*I and introduced into pET30a (C-terminal His-tag) by ligation. Primers used for cloning are given in Table 7. For sequencing reactions primers T7 prom and T7 term (Table 5) were used.

Table 7: Primer sequences for cloning of transcription factors and the *rpo*G subunit of the RNAP into pET30a

Gene	Primer sequence (5'-3')
066_*tfb*1 pET30a *Nde*I f	GGGTTCCATATGGTAGAACAAAGTAAGG
067_*tfb*1 pET30a *Xho*I r	GAATTCCTCGAGTTGACTTGGAATCTG
068_*tfb*2 pET30a *Nde*I f	GGCGCGCATATGAAATGTAAAATCTGTG
069_*tfb*2 pET30a *Xho*I r	GGCGCGCTCGAGAAGCTTTACTTCTATG
070_*tfb*3 pET30a *Nde*I f	GGCACGCATATGGAATGTCCCGAGTGT
071_*tfb*3 pET30a *Xho*I r	GGCGATCTCGAGCATCGTTTTTAACG
*tbp_NdeI*_F	GCCGCGCATATGGAAAGAAGCGTACCAAAC
Saci_1336_Y2H_*Xho*I R*	CATCTGCAGCTCGAGAAATTCTAACTCTTCTTCTTC
*tfe_NdeI*_F (Saci_0652)	GCGCATATGGCTAAAGAGCTAGTTGG
*tfe_XhoI*_R (Saci_0652)	GCGCTCGAGATAACTTTTACTGGACTCACG
*rpoG_NdeI*_F* (Saci_0661)	GATTACGCTCATATGAAGGAAATGATGCAGGGAACCTG
*rpoG_XhoI*_R* (Saci_0661)	CATCTGCAGCTCGAGAGCCTTTTTCTTTACGCAATAAT

underlined: restriction sites, * same primers used for cloning of yeast two-hybrid construct

2.9.2. Expression and purification of SaTFB1

After gene synthesis of the codon adapted *tfb*1 sequence for expression in *E. coli* by GenScript the gene was cloned into pET15b using *Nde*I and *Bam*HI restriction sites. TFB1 was overexpressed using BL21 (DE3) cells and protein expression was induced by addition of 200 µM IPTG when an OD_{600} of 0.5 was reached. After induction the cultures were incubated at 22°C with shaking (180 rpm). After 18 hours cells were harvested by centrifugation (14,800 x g, 15 min, 4°C) and resuspended in 1x LEW buffer (50 mM NaH_2PO_4 pH 8, 300 mM NaCl, 1 mM $MgCl_2$). After disrupting the cells by French press (3 times, 12,000 psi) the sample was centrifuged for 30 min at 4°C and 15,500 x g. The crude extract was used for heat-precipitation (75°C, 20 min) and centrifuged (20,800 x g, 45 min, 4°C). The fraction after heat precipitation was applied onto a Ni-TED column (Macherey-Nagel), washed two times with LEW buffer and the protein was eluted by addition of LEW buffer containing 250 mM imidazole. The eluate fractions were analyzed for the presence of TFB1 by SDS-PAGE and the fractions containing TFB1 were combined. The samples were dialyzed against 50 mM Tris-HCl, 300 mM NaCl and 1 mM $MgCl_2$ over night at 4°C and were applied in gel filtration using a HiLoad Superdex 26/60 200 prep grade column (GE Healthcare). Pure TFB1 fractions were combined and 10% (v/v) glycerol was added for storage at -80°C. One mg of protein was yielded from 18 g (wet weight).

2.9.3. Expression and purification of SaTFB2 and SaTFB3

Codon adapted *tfb*2 and *tfb*3 sequences for expression in *E. coli* were cloned into pET15b via restriction with *Nde*I and *Bam*HI and subsequent ligation. TFB2 and TFB3 were expressed in BL21 (DE3) cells and at OD_{600} of 0.5 expression was induced by addition of 200 µM IPTG and incubation over night was performed at 22°C on a shaker at 180 rpm. The LB-medium for growth of BL21 (DE3) pET15b-*tfb*2 was supplemented with 100 µM $ZnSO_4$ and 1 mM $MgCl_2$. The induced cells were harvested by centrifugation (14,800 x g, 15 min, 4°C) and resuspended in LEW buffer supplemented with 1 mM DTT and 10% (v/v) glycerol in case of TFB2 and 20% (v/v) glycerol in case of TFB3. The cell suspensions were disrupted by French press (3 times, 12,000 psi) and centrifuged at 20,800 x g for 40 min at 4°C. Cells were resuspended in 10 ml LEW buffer per gram of wet weight and centrifuged at 20,800 x g for 30 min at 4°C. The supernatant was discarded and the cells were resuspended in LEW buffer with 1 mM DTT and 10% (v/v) glycerol (TFB2) or 20% (v/v) glycerol (TFB3) containing 6 M urea. The samples were dissolved by stirring at room temperature for 1 hour and separated for 30 min at 4°C and 20,800 x g to remove any insoluble material. The samples were filtered using a 0.45 µM filter before applying onto a Ni-TED column (Macherey-Nagel). The column was washed two times with LEW buffer (1 mM DTT, 10% (v/v) glycerol, 6 M urea) and the proteins were eluted by addition of 250 mM imidazole. First experiments to renature TFB3 were performed. The samples containing TFB3 were combined and diluted drop by drop using the Äkta FPLC system (GE Healthcare) in 30-fold volume LEW-buffer (1 mM DTT, 10% (v/v) glycerol) of the initial sample volume of TFB3 for refolding of the protein. Additionally, it was tried to refold denatured TFB3 bound onto a Ni-TED column (Macherey-Nagel) by washing it two times with 50 ml LEW buffer (1 mM DTT, 10% (v/v) glycerol) without urea. For verification of proper refolding the proteins need to be analyzed by CD-spectra.

2.9.4. Expression and purification of SaTBP

The encoding sequence of TBP (Saci_1336) was amplified from genomic DNA from *S. acidocaldarius* (Table 5). The PCR product was restricted with *Nde*I and *Bam*HI and ligated into pET11c. After Rosetta (DE3) cells reached an OD_{600} of 0.5 protein expression was induced with 1 mM IPTG and the cells were incubated for 3 hours at 37°C after induction. The cultures were harvested by centrifugation (14,800 x g, 4°C, 15 min) and the cells were resuspended in Tris-buffer (50 mM Tris-HCl, 50 mM NaCl pH 6.8) and disrupted by using a French press (3 times, 12,000 psi). After centrifugation (20,800 x g, 40 min, 4°C) to remove cell debris the supernatant was used for subsequent heat precipitation at 80°C for 20 min. The samples were centrifuged for 40 min at 4°C and 20,800 x g. The supernatant was dialyzed over night in buffer K (20 mM Tris-HCl, 50 mM NaCl, pH 8.8) at 4°C. The sample was applied onto an equilibrated ResourceQ column (GE Healthcare) for ion exchange chromatography. Afterwards, the combined TBP fractions were dialyzed in 50 mM Tris-HCl, 300 mM NaCl, pH 7.0 for gel filtration on a HiLoad Superdex 26/60 75 prep grade column (GE Healthcare). Elution fractions were analyzed by SDS-PAA gels and TBP-containing fractions were combined. 10% (v/v) glycerol and 10 µl β-mercaptoethanol (14.3 M) were added for storage at -80°C. Twelve g (wet weight) yielded in 0.35 mg protein.

2.9.5. Expression and purification of SaTFE

The TFE encoding gene (Saci_0652) was amplified from genomic DNA from *S. acidocaldarius* (Table 7). The PCR product was restricted with *Nde*I and *Xho*I and ligated into pET30a. LB medium was inoculated 2% with an overnight culture. The Rosetta (DE3) cells were incubated at 37°C on a shaker at 180 rpm. After reaching an OD_{600} of 0.5 expression of TFE was induced by addition of 200 µM IPTG and the cells were incubated at 22°C over night. Cells were harvested by centrifugation (14,800 x g, 4°C, 15 min) and resuspended in 3 ml LEW buffer (50 mM NaH_2PO_4, 300 mM NaCl, 2 mM DTT, pH8) per 1 g of cells (wet weight) and were disrupted by a three times passage through a French press (12,000 psi). The cell lysate was centrifuged for 30 min at 4°C and at 20,800 x g. As a first purification step the crude extract was incubated at 75°C for 15 min and then centrifuged for 30 min at 4°C and at 20,800 x g. The supernatant was dialyzed over night at 4°C in LEW buffer containing 2 mM DTT. The lysate was filtered (0.45 µM pore size) and applied onto a Ni-TED column (Macherey-Nagel) pre-packed with 3 g column material. The column was equilibrated with LEW buffer containing 2 mM DTT and washed twice with the same buffer after application. TFE was eluted in 1x LEW buffer containing 2 mM DTT and 180 mM imidazole. The fractions containing TFE were combined and dialyzed against 20 mM Tris-HCl pH 8.8 buffer with 2 mM DTT. 10% (v/v) glycerol was added to prevent aggregation of the protein for concentration. The protein was separated by anion exchange chromatography using a ResourceQ column (GE Healthcare) and eluted at 170 mM NaCl concentration in a linear salt gradient. The protein was dialyzed in 50 mM Tris-HCl pH 7.8 and 2 mM DTT and applied on a HiLoad Superdex 26/60 75 prep grade column (GE Healthcare) for size exclusion. Afterwards TFE fractions were collected, 10% (v/v) glycerol was added directly to the sample and the protein was concentrated by centrifugation (7,563 x g, 4 h, 4°C) using Vivaspin columns (Sartorius). Out of 18 g (wet weight) 2.1 mg protein was obtained and stored at -80°C.

2.9.6. Expression and purification of SaRpoG

The RpoG encoding gene (Saci_0661) was amplified from genomic DNA from *S. acidocaldarius* via PCR using the primers listed in Table 7. The resulting PCR product was restricted with *Nde*I and *Xho*I and ligated into pET30a. Growing and expression conditions were the same used for expression of TFE in section 2.9.5. Heat precipitation and Ni-TED column (Macherey-Nagel) purification were comparable to those described in part 2.9.5 for TFE, but the buffer contained KCl instead of NaCl. For size exclusion the protein was dialyzed in 50 mM Tris-HCl, 300 mM KCl, 0.5 mM DTT pH 7.8 and separated on a HiLoad 26/60 Superdex 75 prep grade column (GE Healthcare). RpoG containing fractions were combined and stored at -80°C after addition of 10% (v/v) glycerol.

2.10. Homologous expression using *S. acidocaldarius* MW001 as expression host

For homologous expression in *S. acidocaldarius* the plasmids pPREX and pSVA1450 were used for cloning and overexpression of the transcription factors (*tfb*1-*tfb*3 and *tbp*) in *S. acidocaldarius.* The plasmid pPREX was used for cloning and transformation into *E. coli* cells. This plasmid contains a C-terminal His- and Strep-tag for purification and detection. The genes encoding TFB1, TFB2, TFB3 and TBP were amplified from genomic DNA with the primers listed in Table 8. The inserts (*tfb*1-*tfb*3 and *tbp* genes) were cloned into pPREX after restriction with *Nco*I and *Bam*HI. For cloning into the *E. coli-Sulfolobus* shuttle vector pSVA1450 the genes were cut with *Nco*I and *Eag*I instead of *BamH*I to transfer the His-Strep-tag into the new vector. pSVA1450 is an *E. coli-Sulfolobus* shuttle vector containing a maltose inducible promoter and the *pyrEF* cassette for selection (- uracil). It also carries the *lacS* gene for blue-white selection. After ligation of the genes into pSVA1450 the plasmid was methylated by transformation into the *E. coli* strain ER1821 (harbors plasmid pM.EsaBC4I, New England Biolabs). Methylation prevents restriction of the plasmid by the restriction enzyme *Sua*I after transformation in *S. acidocaldarius* (Wagner et al., 2012). The methylated plasmid was isolated and electroporated into competent *S. acidocaldarius* MW001 cells. Overexpression was induced by addition of 0.4% (v/v) (from 20% (w/v) stock) maltose directly after inoculation. *S. acidocaldarius* MW001 cells harboring the plasmids pSVA1450-*tfb*1, pSVA1450-*tfb*2, pSVA1450-*tfb*3 or pSVA1450-*tbp* were grown at 78°C to an OD_{600} of 0.6 to 0.8 in Brock medium supplemented with 0.1% (v/v) N-Z-Amine®, 0.2% (v/v) sucrose, 0.4% (v/v) maltose. 8 L cultures were harvested at 14,800 x g for 15 min at 4°C and disrupted by sonication (0.6 cycle, 60 amplitude) or French press (12,000 psi).

Table 8: Primer sequences for cloning of *tfb*1, *tfb*2, *tfb*3 and *tbp* in the vector pPREX for subsequent cloning into pSVA1450 and expression in *S. acidocaldarius* MW001

Primer name	Sequence (5'-3')
024_Saci_*tfb*1_*Nco*I_F (Saci_0866)	GCGCCATGGTAGAACAAAGTAAGGTTCC
060_Sa*tfb*1_HE_ *Bam*H*I*_R (Saci_0866)	CGCGGATCCTTGACTTGGAATCTGTATC
061_Sa*tfb*2_HE_ *Nco*I_F (Saci_1341)	GGGCCATGGCCAAATGTAAAATCTGTGGCAGTGAG
062_Sa*tfb*2_HE_ *Bam*HI_R (Saci_1341)	CGCGGATCCAAGCTTTACTTCTATGTCAAAAG
026_Saci_*tfb*3_*Nco*I_F (Saci_0665)	GCGCCATGGAATGTCCCGAGTGTAAAAG
063_Sa*tfb*3_HE_ *Bam*HI_R (Saci_0665)	CGCGGATCCCATCGTTTTTAACGTTTGGCTAGGTAC
064_Sa*tbp*_ HE_*Nco*I_F (Saci_1336)	GCGCCATGGCCATACCTGATGAGATCCCGTATAAAG
065_Sa*tbp*_ HE_*Bam*HI_R (Saci_1336)	GGCGGATCCAAATTCTAACTCTTCTTCTTCAACGGGCTTTAC
127_pSVA1450_Seq F*	CGGAGGTGTCCTTAAGTTTAG
128_pSVA1450_Seq R*	CGGGCGTGATAAAGTCTGTCTC

underlined: restriction sites; * sequencing primer

2.10.1. Preparation of competent *S. acidocaldarius* MW001 cells

S. acidocaldarius cells were grown in 400 ml Brocks basal salt medium (see 2.4.) supplemented with uracil (10 µg/ml) to an OD_{600} of 0.2 to 0.3. The cells were cooled on ice and afterwards harvested by centrifugation for 20 min at 4°C and 4,000 x g. The cells were washed with an equal volume of ice-cold 20 mM sucrose and the cells were centrifuged for 20 min at 4°C and 4000 x g. Then the cells were resuspended in half of the initial volume of 20 mM sucrose solution as used before and centrifuged again (4,000 x g, 15 min, 4°C). The cells were washed with 1 ml of a 20 mM sucrose solution and harvested by centrifugation (4,000 x g, 15 min, 4°C). 20 mM sucrose was added to obtain a theoretical final OD_{600} of 10. 50 µl aliquots of competent cells were stored at -80°C or directly used for transformation (described in Wagner et al., 2012).

2.10.2. Electroporation of competent *S. acidocaldarius* MW001 cells

Competent MW001 cells were thawed on ice. 200 to 400 ng of methylated vector or PCR-product were added to the cells. The samples were transferred into cooled electroporation cuvettes (1mm). Immediately after pulsing the samples (600 µF, 2500 Ω, 1 mm) they were mixed with 50 µl recovery solution (1% (w/v) sucrose, 20 mM β-alanine/malate buffer pH 4.5, 10 mM $MgSO_4$) and incubated at 76°C for 30 min in a shaker (180 rpm). Afterwards 100 µl of the cell suspension were spread on first selection plates. The plates were incubated for five to seven days at 78°C and wrapped in plastic to avoid drying.

2.11. Purification of Strep-tag proteins

Proteins derived from homologous expression or tagged with an N- or C-terminal Strep-FLAG-tag were purified according to the manufacturer´s instruction (IBA Göttingen). Cells were resuspended in buffer W (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 10% (v/v) glycerol) and disrupted by sonication (0.6 cycle, 60 amplitude). The samples were centrifuged for 30 min at 4°C and 20,800 x g and the supernatant was applied to equilibrated Strep-tactin columns (IBA Göttingen). After addition of the crude extract, the columns were washed five times with two column volumes buffer W. The proteins were eluted from the column by addition of six times 0.5 column volumes of buffer E (buffer W including 2.5 mM desthiobiotin). For regeneration of the column buffer R (buffer W containing 1 mM HABA) was added three times in five column volumes.

2.12. Western blot and immunodetection

2.12.1. Antibody generation and immunodetection

Antibodies were generated by Eurogentec (Belgium) in a 28-day speedy program by immunizing two animals. The final bleed of two animals was used in western blots. Antibodies against *S. acidocaldarius* TFB1 and TFE were raised in rabbit. Antisera from two animals per protein were tested with recombinant proteins and *S. acidocaldarius* MW001 crude extract. After western blot over night at 4°C at 20 V the membranes were blocked with 5% (w/v) milk powder for one hour and the first antibody (α-TFB1, α-TFE) was incubated for 1.5 hours at room temperature (1:500, 3% (w/v) milk powder in TBST (20 mM Tris, 150 mM NaCl, pH 7.6, 0.3% (v/v) Tween-20)). The membranes were washed three times for 15 min. The second antibody (α-rabbit IgG linked with alkaline phosphatase, Sigma-Aldrich) was diluted 1:10,000 in TBST and incubated for another 1.5 hours at RT. Afterwards the membranes were washed three times with TBST for 15 min and washed two times with TBS for 10 min. Finally the membranes were incubated in 1 ml CDP-Star® (Disodium 2-chloro-5-(4-methoxyspiro{1,2-dioxetane-3,2′-(5′-chloro)tricyclo[3.3.1.1^{3,7}]decan}-4-yl)-1-phenyl phosphate, Sigma-Aldrich) and the proteins were detected using the VersaDoc system (Bio-Rad, Munich).

2.12.2. Antibody purification from antiserum using protein-A-agarose

All steps were carried out according to the user manual of protein A-agarose from Roche. 400 µl antiserum were thawed on ice and centrifuged for 10 min at 500 x g for removal of cells. The supernatant was removed and centrifuged for 10 min at 4°C and 10,000 x g. An equal volume of an ammonium sulfate solution (ammonium sulfate, saturated solution approx. 800 g/l, pH 7.0, filtered and pre-cooled) was added drop by drop while gently stirring and cooling the solution. After stirring the mixture in the refrigerator for 1 hour the sample was centrifuged for 10 min at 10,000 x g at 4°C. The supernatant was removed and the pellet was resuspended in starting buffer (100 mM Tris-HCl, pH 8.0, filtered). The sample was dialyzed against starting buffer (three changes) and centrifuged to pellet remaining debris.

100 µl protein-A-agarose column material was washed with three bed volumes of starting buffer. The pH of the crude antibody sample was adjusted to pH 7.5-8 by addition of a volume that equals 10% of the initial sample volume of antiserum. The antibody solution was

added to the column material, mixed by gentle pipetting and incubated for 1 min. The agarose material was collected by centrifugation at 4°C for 1 min and 8,000 x g. The supernatant was removed and the sample was washed with five bed volumes of wash buffer 1 (100 mM Tris-HCl, pH 8.0, filtered). The agarose was pelleted again with the same settings described above. Afterwards the sample was washed with five to ten bed volumes of wash buffer 2 (10 mM Tris-HCl, pH 8.0, filtered) until no absorbance could be detected at 280 nm. The immunoglobulin was removed by addition of elution buffer (100 mM glycine, pH 3). The eluate collection tubes contained neutralization buffer (1 M Tris-HCl, pH 8.0) equivalent to 20% of the fraction size. The eluates were mixed with neutralization buffer and stored on ice. For determination of the antibody concentration the absorbance was measured at 280 nm. The purified immunoglobulin fractions were dialyzed against PBS buffer (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, 1.46 mM KH_2PO_4, pH 7.4*)* and stored at 4°C.

2.12.3. Western blot and immunodetection (His-tag antibody)

The proteins were transferred to a PVDF membrane by wet blotting (25 V, over night, 4°C or 100 V, 1 hour, 4°C). The membrane was washed with TBST and stained with Ponceau red solution (0.5% (w/v) Ponceau-S-red, 1% (v/v) acetic acid) for 2 min to visualize proteins. Then the membrane was washed with 1x TBST to remove the staining solution and was blocked with TBST containing 5% (w/v) milk powder. The anti-His-tag antibody (abcam) was diluted 1:1,000 in 10 ml TBST with 3% (w/v) milk powder and incubated for one hour at room temperature. Wash steps to remove unbound antibody were performed in TBST for three times for 15 min and two times for 10 min in TBS. Then CDP-Star® was pipetted onto the membrane to visualize the protein bands.

2.12.4. Immunodetection using Strep-tag antibody

All steps were performed at room temperature. After western blot the membrane was washed three times for 5 min in TBS containing 0.3% (v/v) Tween-20. Blocking was performed for 2 h in TBST with 0.2% (w/v) I-Block (Tropix). After washing the membrane three times for 5 min with TBST including 0.1% (w/v) I-Block the anti-Strep-tag antibody linked to alkaline phosphatase (IBA) was diluted 1:4,000 in 10 ml TBST with 0.1% (w/v) I-Block and incubated for 1 h. Subsequent washing steps were performed three times for 5 min in TBST with 0.1% (w/v) blocking reagent and with TBS (2 x 10 min). For developing 1 ml CDP-Star® was added to the membrane and chemiluminescence was detected using the Bio-Rad VersaDoc system.

2.12.5. Immunodetection with FLAG-tag antibody

After western blot the PVDF membrane was blocked in 30 ml TBS-buffer containing 3% (v/v) Tween-20 and 5% (w/v) milk powder for one hour at room temperature. After washing with TBS containing 0.05% (v/v) Tween-20 (3x for 5 min) the monoclonal anti-Flag-tag antibody (Sigma-Aldrich, F2555) was diluted 1000-fold in TBS including 0.05% (v/v) Tween-20 and 1% (w/v) milk powder and incubated for 2 h at room temperature. Before incubation with the second antibody the membrane was washed three times for 15 min in TBST (0.3% (v/v) Tween-20). Anti-rabbit IgG antibody linked with alkaline phosphatase (Sigma-Aldrich) was diluted (1:10,000) in 10 ml TBST and incubated for 1 h. To remove

unbound antibody the membrane was washed three times for 10 min with TBST and this step was repeated by washing with TBS (3x 10 min). Afterwards 1 ml CDP-Star® was applied onto the membrane and the membrane was wrapped into a plastic foil and chemiluminescence signals were detected using the VersaDoc system (Bio-Rad).

2.13. Construction of knock-out strains by homologous recombination and insertion of *Sso-pyrEF* cassette as genetic marker in *S. acidocaldarius*

For knock-out mutant construction via homologous recombination the up- and downstream flanking regions of the gene of interest were amplified from chromosomal DNA. In the following overlap extension PCR the fragments were fused and the resulting purified DNA was restricted with *Bam*HI and *Nco*I. After ligation into pSVA406 the plasmid was methylated by transformation into the strain *E. coli* ER1821. After cloning, the methylated plasmid was transformed into the uracil auxotroph mutant *S. acidocaldarius* MW001. After electroporation the recovered cells were spread on first selection medium (Brock-medium, - uracil, +N-Z-Amine®). After an incubation time of five to seven days colonies were used for inoculation of liquid first selection medium. 75 µl to 100 µl of the medium were spread on second selection plates (+uracil, +5-FOA). Primer sequences used for generation of markerless in-frame deletion mutants are shown in Table 9.

In a subsequent colony-PCR cells were resuspended in 30 µl 0.2 M NaOH and streaked on replicate second selection plates. To neutralize the pH 70 µl of 0.26 M Tris-HCl pH 7.8 are added and 1 µl of this suspension is used in 30 µl PCR reaction.

Table 9: Primer sequences for markerless knock-out of the transcription factors *tfb*1, *tfb*2 and *tfb*3

Gene	Primer name	Sequence (5'-3')
*tfb*1 (Saci_0866)	*tfb*1_*Bam*HI F	GCCGGATCCCAAGGATTCAATCCTCAATAG
	upstream_R	GAATTATTTATTACGTTTAATGGTGATGGTGGTGATGGTGGTGTGCTCCTTGACTTGGAATCTGTATCTTTAATTC
	downstream_F	GATTCCAAGTCAAGGAGCACACCACCATCACCACCATCACCATTAAACGTAATAAATAATTCCCTTGAAC
	*tfb*1_*Nco*I_R	GCGCCATGGTCGCTTTCACTATCTTAACG
*tfb*2 (Saci_1341)	*tfb*2_*Bam*HI_F	GCCGGATCCCGAAATTAAACTTGCCTGAAC
	upstream_R	CTTTATTTCCGACTTCAATGGTGATGGTGGTGATGGTGGTGTGCTCCAAGCTTTACTTCTATGTCAAAAG
	downstream_F	GAAGTAAAGCTTGGAGCACACCACCATCACCACCATCACCATTGAAGTCGGAAATAAAGTAAATAAATGG
	*tfb*2_*Nco*I_R	GCGCCATGGAGTTAATGTTCCACGGAGAG
*tfb*3 (Saci_0665)	*tfb*2_*Bam*HI_F	GCCGGATCCAGTGCGGAAATTTAGTTTG
	upstream_R	CTATAAAAATCATTTTCAATGGTGATGGTGGTGATGGTGGTGTGCTCCCATCGTTTTTAACGTTTGGCTAGGTAC
	downstream_F	GTTAAAAACGATGGGAGCACACCACCATCACCACCATCACCATTGAAAATGATTTTTATAGATGGAAAAATTTTTATATAAC
	*tfb*2_*Nco*I_R	GCGCCATGGACTTGAGTTATGGCCTTTAC
sequencing	018_pSVA406_seq_F	CGGCCAGTGAATTGTAATACGACTC
sequencing	019_pSVA406_seq_R	CAGGGAAAGTAGTTTCATTAGC
sequencing	072_Sa*tfb*1_ko_PCR_F	GCGCCAATAAATACCAGACC
sequencing	073_Sa*tfb*1_ko_PCR_R	CAGGGAGATCATAGGGATAGAG
sequencing	074_Sa*tfb*2_ko_PCR_F	TGAAGCTGGTCGGTTCTAC
sequencing	075_Sa*tfb*2_ko_PCR_R	TATTTGCGACCCGTGCATC
sequencing	076_Sa*tfb*3_ko_PCR_F	GCCTCTACTCTTAAGGTCAG
sequencing	077_Sa*tfb*3_ko_PCR_R	TACCGCGGCTATGTTTATCC
sequencing	114_*tfb*1 ko seq F 2	GCCGAGGTATTTCTATTAGATC
sequencing	115_*tfb*2 ko seq F 2	GTTGATGTTTGAAGCCTCCATTC
sequencing	116_*tfb*3 ko seq F 2	TAACCTCGATGTCTGCAAAAG

underlined: restriction site

For generation of disruption strains 50 bp to 60 bp of the gene of interest were fused upstream or downstream of the *pyrEF* cassette amplified from the vector pSVA406. 200-300 ng of the PCR product were electroporated into *S. acidocaldarius* MW001. In contrast to the markerless knock-out by homologous recombination the cells were only selected on plates and medium containing N-Z-Amine® and lacking uracil. Primers used for disruption of the transcription factor encoding genes (*tfb*1, *tfb*2, *tfb*3) and the genes *gar*1 and Saci_1342 as well

as sequencing primers for verification of insertion of the marker cassette are given in Table 10.

Table 10: Primers for generation of disruption mutants of *tfb*1, *tfb*2, *tfb*3, *tfe*, *gar*1 and Saci_1342 and sequencing of genes *tfb*1, *tfb*2, *gar*1, Saci_1342 and *Sso-pyrEF*

Gene	Primer name	Sequence (5'-3')
*tfb*1 (Saci_0866)	132_Sa*tfb*1_Ins_pyrEF_F	AGGAGATTAGAAGTACTAAGATGGAGAAA ATGGCAGATAAGGACAAGGATTTTTGAGCA GTTCTAG
	133_Sa*tfb*1_Ins_pyrEF_R	TCTCTCAAGCTCGTTCATAGCTTGTGCTAGG TTTCTGTCTATTGAGGATTGGACCGGCTATT TTTTCAC
*tfb*2 (Saci_1341)	134_Sa*tfb*2_Ins_pyrEF_F	GAGAAAGAACAGGTTCTCCTCTAACAGCCA AAGTACATGACTTTGGTATAACTTTGAGC AGTTCTAG
	135_Sa*tfb*2_Ins_pyrEF_R	TAGCTTATGTACTTTTATTCGGTCTTTGATT TTAGTATATCCTATTTTAGTGACCGGCTATT TTTTCAC
*tfb*3 (Saci_0665)	176_Sa*tfb*3_Ins_pyrEF F	CACTTTTACTAACGTCACAATTTTAACTTAT AAAGATAAAATTGAGAATTTGAGCAGTTCT AG
	177_Sa*tfb*3_Ins_pyrEF_R	TTTTTGATAGTTTACTATTAAATTTTATTATT TTATCAATTTTGACCGGCTATTTTTTCAC
*gar*1 (Saci_1340)	159_Sa*gar*1 Ins pyrEF_F	GGTAAGGTAATACTAGATGAAAAAGGGAA TAGAGTAGCTAAAATTCTCGATTTTGAGCA GTTCTAG
	160_Sa*gar*1 Ins pyrEF_R	CTTTGTTAATGGATAAGCCAATATATATGG AGCTTTAACATTACCAATTACGACCGGCTA TTTTTTCAC
hypothetical protein (Saci_1342)	161_Saci_1342 Ins pyrEF_F	GGGTCTAGATGCACGGGTCGCAAATACAGT TATTAAGAAACTAATTGAACAGTTTGAGCA GTTCTAG
	162_Saci1342 Ins pyrEF_R	TTTTATAACATTCTTTCCATTTTCTTTAATTG ACTTTTTCTTAATCTTTTTGACCGGCTATTTT TTCAC
*tfb*1	153_*tfb*1 Ins_PCR_F*	CCTGAGTGGAGGGCATTTAC
*tfb*2	154_*tfb*2 Ins_PCR_F*	TCAAGGACCAGAATGGAGAG
tfb3	179_*tfb*3_Ins_Seq_F*	CATAACTATATCGATAATG
tfb3	180_*tfb*3_Ins_Seq_R*	CACTAGAGGATCTAATTATTG
*gar*1	167_*gar*1_Ins_Seq_F*	GATGAAAAAGGGAATAGAG
*gar*1	168_*gar*1_Ins_Seq_R*	CTTCTTCTCCTCCCTTTCTC
Sso-pyrEF	155_*pyrEF* Colony PCR_R*	GATATGAGAGAGGTTTATCC

underlined: *pyrEF* cassette; * primers used for sequencing reactions

2.13.1. Growth of the *tfb*3 disruption strain (*tfb3::pyrEF*) and *S. acidocaldarius* MW001 after UV-treatment

S. acidocaldarius MW001 *tfb3::pyrEF* and *S. acidocaldarius* MW001 cells were grown in Brock medium or Brock medium supplemented with uracil (10 µg/ml) in case of the parent strain MW001 to an OD_{600} of 0.6. The main cultures (50 ml) were inoculated to reach an OD_{600} of 0.02. All experiments were carried out in triplicates and the experiment was repeated twice. After reaching OD_{600} around 0.3 the whole culture (50 ml) was split into half and poured into two pre-warmed petri dishes. The UV-shock was performed according to Fröls et al., 2007. The plates were placed into the Stratalinker® (Stratagene) and UV-doses of 50 J/m^2, 75 J/m^2 and 200 J/m^2 were applied. The cultures were combined in an Erlenmeyer flask and incubated in the dark at room temperature for 15 min. Afterwards the flasks were incubated at 78°C in the dark. To make sure that the temperature drop does not interfere with the results obtained after UV-treatment samples were treated the same way without application of a dose. This control samples were placed into the Stratalinker® for 30s, a time that equals the duration to apply a dose of 200 J/m^2 without irradiation (Ctrl. 200). In addition untreated cells incubated at 78°C during the whole experiment served as controls (Ctrl.). Growth rate (µ) and doubling time (t_d) were calculated with the following equations. Time point zero (t_0) is the moment of UV-shock and $OD_{600}(0)$ the measured optical density at that point.

$$\mu\,[h^{-1}] = \frac{\log OD_{600} - \log OD_{600}(0)}{t - t_0}$$

$$t_d\,[h^{-1}] = \frac{\ln 2}{\mu}$$

2.14. Genomic knock-in of His-tag and Strep-FLAG tag

2.14.1. Genomic tagging of transcription factors and subunit RpoG

For knock-in of chromosomal tags the same procedure as described in part 2.13 was performed and tag sequences (His- and Strep-FLAG tag, Fig. 2) were introduced via PCR. The subunit RpoG of the RNAP was tagged with a C-terminal 8x-His-tag for an easier purification from *S. acidocaldarius* MW001 cells. Primers for knock-in generation and sequencing are given in Table 11. The transcription factors TFB1, TFB2, TFB3 and TFE were tagged at their N- and C-termini with a Strep-FLAG tag. Primers used to create these mutants as well as primers used for sequencing are listed in Table 12 and Table 13.

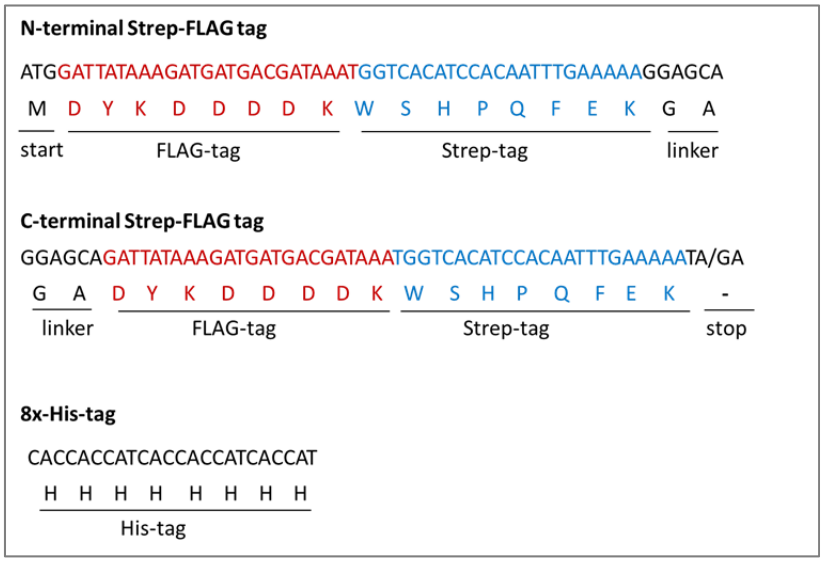


Fig. 2: Sequences of N- and C-terminal Strep-FLAG tag and 8x C-terminal His-tag used for generation of knock-in in *S. acidocaldarius.*

Table 11: Primers used for integration of a genomic C-terminal 8x-His-tag in RpoG

Gene	Primer name	Sequence (5'-3')
*rpo*G (Saci_0661)	upstream_*Eag*I	CGATCGGCCGCTCTCTGGGATCAAATAGG
	upstream_R	GGTG*ATGGTGGTGATGGTGGTG*AGCCTTTTTCTTTACGCAATAATATACGTGATCC
	downstream_F	GAAAAAGGCT*CACCACCATCACCACCATCACCAT*TAAATTTTTATGAAATCTTCATAGATATTGTTTTTGTTTATG
	downstream_*Avr*II	CCTACCTAGGGACCAAGGATGGTGTGGTATTAGC
sequencing	126_SarpoG_His_Seq_F	AGATGAAATCTCCCTGTATG
sequencing	158_SarpoG_ki_Seq_R	GAAGTCAGCAGTCGAAAGAG

underlined: restriction site, italic: 8x-His-tag

Table 12: Primers used for genomic integration of an N-terminal Strep-FLAG tag

*tfb*1	*BamH*I F	GCCGGATCCTCGAGACGGTAAGGAAAG
	overlap extension_R	CCTTACTTTGTTCTACCAT*TGCTCCTTTTTCAAA TTGTGGATGTGACCATTTATCGTCATCATCTTTATAATCCAT*TTCACC ACCCAAAATAC
	overlap extension_F	GTATTTTGGGTGGTGAA*ATGGATTATAAAGATG ATGACGATAAATGGTCACATCCACAATTTGAAAAAGGAGCA*ATGGT AGAACAAAGTAAGG
	*NcoI*_R	GCGCCATGGAATTGAAGCTGCAACGACAC
*tfb*2	*Bam*HI_F	GCCGGATCCCTTCTACAGCTTTAGAATACC
	overlap extension_R	GCCACAGATTTTACATTTCAT*TGCTCCTTTTTCAA ATTGTGGATGTGACCATTTATCGTCATCATCTTTATAATCCAT*CATCT TCTTCTCCTCCCTTTCTC
	overlap extension_F	GAGAAAGGGAGGAGAAGAAGATG*ATGGATTATAA AGATGATGACGATAAATGGTCACATCCACAATTTGAA*AAA*GGAGCA* ATGAAATGTAAAATCTGTGGC
	*Nco*I_R	GCGCCATGGGCAGGTAATCCCAATCTTTC
*tfb*3	*Bam*HI_F	GCCGGATCCATTGCGGTTCATCATATCTG
	overlap extension_R	CTACTTTTACACTCGGGACATTCCAT*TGCTCCTTTTTC AAATTGTGGATGTGACCATTTATCGTCATCATCTTTATAATCCAT*AGA ACAAAAAATCTCCTG
	overlap extension_F	CAGGAGATTTTTTGTTCT*ATGGATTATAAAGATGATG ACGATAAATGGTCACATCCACAATTTGAAAAAGGAGCA*ATGGAATG TCCCGAGTGTAAAAGTAG
	*Nco*I_R	GCGCCATGGTAACGTTTGGCTAGGTACTC
tfe	*Bam*HI_F	CGCGGATCCTTTAGTAGCGAGAGCCTTTG
	overlap extension_R	CAACTAGCTCTTTAGCCAA*TGCTCCTTTTTCAAATT GTGGATGTGACCATTTATCGTCATCATCTTTATAATCCAT*ATTAAGC ACTATATTTTCTATATC
	overlap extension_F	GATATAGAAAATATAGTGCTTAAT*ATGGATTATAAA GATGATGACGATAAATGGTCACATCCACAATTTGAAAAAGGAGCA*ATT GGCTAAAGAGCTAGTTG
	*Nco*I_R	GCGCCATGGAAAGGTTCACGGTTAAATC
sequencing	150_*tfb*1 ki N seq F	TTGAGAGGCCCGCATACTAC
	169_*tfb*1 ki N SF seq R	TCTTGTCAGGTGGACATAGG
	170_*tfb*2 ki SF N seq F	TTGAAGCTGGTCGGTTCTAC
	171_*tfb*2 ki SF N seq R	GTGGTTCATCGTCTACAG
	151_*tfb*3 ki N seq F	CCGGCTTGTCTAAATTCCCTG
	172_*tfb*3 ki SF N seq R	TTCCGCACTTGTAATCCC
	195_*tfe*_N-term_seq_F	GCCAAATTCTACATTATTCC
	196_*tfe*_N-term_intern_R	GTTCAATTCCTTAGCGATCTC

underlined: restriction site, italic: Strep-FLAG tag

Table 13: Primers used for genomic integration of a C-terminal Strep-FLAG tag

*tfb*1	*Bam*HI_F	GCCGGATCCCAAGGATTCAATCCTCAATAG
	overlap extension_R	CAAGGGAATTATTTATTACGT*TTATTTTTCAAATTGTGGATGTG ACCATTTATCGTCATCATCTTTATAATCTGCTCC*TTGACTTGGAA TCTGTATCTTTAATTC
	overlap extension_F	GATTCCAAGTCAA*GGAGCAGATTATAAAGATGATGACGATAAAT GGTCACATCCACAATTTGAAAAATAA*ACGTAATAAATAATTCCC TTGAAC
	*Nco*I_R	GCGCCATGGTCGCTTTCACTATCTTAACG
*tfb*2	*Bam*HI_F	GCCGGATCCCGAAATTAAACTTGCCTGAAC
	overlap extension_R	CTCTTTCCCCATTTATTTACTTTATTTCCGACT*TCATTTTTCAAA TTGTGGATGTGACCATTTATCGTCATCATCTTTATAATCTGCTCC*A AGCTTTACTTCTATGTCAAAAGCC
	overlap extension_F	GACATAGAAGTAAAGCTT*GGAGCAGATTATAAAGATGATGACG ATAAATGGTCACATCCACAATTTGAAAAATGA*AGTCGGAAATAA AGTAAATAAATGG
	*Nco*I_R	GCGCCATGGAGTTAATGTTCCACGGAGAG
*tfb*3	*Bam*HI_F	GCCGGATCCAGTGCGGAAATTTAGTTTG
	overlap extension_R	CCATCTATAAAAATCATTT*TTATTTTTCAAATTGTGGATGTGAC CATTTATCGTCATCATCTTTATAATCTGCTCC*CATCGTTTTTAAC GTTTGGCTAGG
	overlap extension_F	GCCAAACGTTAAAAACGATG*GGAGCAGATTATAAAGATGATGA CGATAAATGGTCACATCCACAATTTGAAAAATAA*AAATGATTTTT ATAGATGG
	*Nco*I_R	GCGCCATGGACTTGAGTTATGGCCTTTAC
tfe	*Bam*HI_F	CGCGGATCCATAGGAAATCAGCCACACTC
	overlap extension_R	CCTCCAGCCCCACTAAATAAA*TCATTTTTCAAATTGTGGATGTG ACCATTTATCGTCATCATCTTTATAATCTGCTCC*ATAACTTTTAC TGGACTCACGTCTAATCTC
	overlap extension_F	GAGATTAGACGTGAGTCCAGTAAAAGTTAT*GGAGCAGATTAT AAAGATGATGACGATAAATGGTCACATCCACAATTTGAAAAATGA* TTTATTTAGTGGGGCTGGAGG
	*Nco*I_R	GCGCCATGGGCCTAGCAAACTTTCTTATC
sequencing	107_Sa*tfb*1_knock-in_F	TGGGTGGTGAAATGGTAG
	173_*tfb*1 ki SF C seq R	GGGAGATCATAGGGATAGAG
	108_Sa*tfb*2 knock-in F	GTGGAGAAAGGGAGGAGAAG
	174_*tfb*2 ki SF C seq R	TGCGACCCGTGCATCTAGAC
	109_Sa*tfb*3 knock-in F	GCCTCTACTCTTAAGGTCAG
	175_*tfb*3 ki SF C seq R	CCCGTCTGTGGTCTCATAAG
	198_*tfe*_C-term_seq_F	GAAATAGATATGAGCAAGAGAAG
	200_*tfe*_C-term_seq_R	CGCAGCATGATTTATATCTAC
	152_SF tag seq_R	ATTGTGGATGTGACCATTTATCG

underlined: restriction site, italic: Strep-FLAG tag

2.14.2. Growing conditions of knock-in mutants and purification of transcription factors via genomic tags

The genetically modified *S. acidocaldarius* MW001 strains (MW001::*tfb*1C, MW001::*tfb*2N, MW001::*tfb*3N and MW001::*tfeC*) were incubated at 78°C in a shaker (180 rpm) until reaching OD_{600} 0.6-0.8 and collected by centrifugation at 14,800 x g for 15 min at 4°C. The cells were stored at -80°C or directly disrupted by sonication (0.6 cycle, 60 amplitude) or French press (12,000 psi, three times passage). For disruption the cells were resuspended in buffer W (100 mM Tris-HCl pH 8.0, 150 mM NaCl) supplemented with 10% (v/v) glycerol and protease inhibitor cocktail set (Roche). The samples were centrifuged for 40 min at 4°C and 20,800 x g. Crude extracts of MW001::*tfb*1C and MW001::*tfeC* were applied onto Strep-tag columns for affinity chromatography (IBA Göttingen) (refer to part 2.11.). For immunodetection of the proteins with anti-Strep-tag or anti-FLAG-tag antibody purified protein (TFB1 C-terminal tag, TFE C-terminal tag) or whole cells from 15 ml cultures (OD_{600} 0.6-0.8) of TFB2 (N-terminal tag) and TFB3 (N-terminal tag) were collected by centrifugation (6000 x g, 5 min, 4°C), resuspended in 50µl 4x SDS-PAGE loading buffer, boiled at 95°C for 10 min and applied on SDS-PAA gels.

2.14.3. Growth of *S. acidocaldarius* MW001::*rpo*G-His and purification of the endogenous RNAP from *S. acidocaldarius*

The strain *S. acidocaldarius* MW001::*rpo*G-His was grown in 18 L Brock medium supplemented with uracil (10 µg/ml) at 78°C and 180 rpm until reaching an OD_{600} of 0.8. Cells were harvested by centrifugation (14,800 x g, 15 min, 4°C) and stored at -80°C or disrupted by a three-times passage through French press (12,000 psi) immediately after centrifugation. The pellet was resuspended in 1x LEW buffer (1 ml buffer per mg wet weight of cells) containing 300 mM KCl instead of NaCl, 10% (v/v) glycerol, 0.5 mM DTT and protease inhibitor (Roche). The sample was centrifuged for 40 min at 4°C and at 20,800 x g and the cell-free extract was applied onto Ni-TED columns (Macherey-Nagel) for affinity chromatography. The protein complex was eluted with 250 mM imidazole. Before storage at -80°C the sample was dialyzed against Tris-buffer (50 mM Tris-HCl pH 7.8, 300 mM KCl, 0.5 mM DTT 10% (v/v) glycerol).

2.15. Reporter gene assays

2.15.1. Cloning of promoter regions of transcription factors for β-galactosidase reporter gene assays

The promoter regions of *tfb1*, *tfb2*, *tfb3*, *tbp*, *tfe* and *gar1* (-1 to -303 bp or -1 to -269 bp for *tfb*2) were cloned into pSVA1450. Primer sequences used for cloning and sequencing are listed in Table 14. The plasmid pSVA1450 originally contains the inducible *mal* promoter (Saci_1161) which was replaced by the mentioned promoter regions. By PCR 303 bp upstream of the start codons of the genes or in case of *tfb2* 269 bp were amplified from genomic *S. acidocaldarius* DNA. The PCR products were restricted with *Nco*I and *Sac*II and ligated into pSVA1450. The plasmids were transformed in competent *E. coli* DH5α cells. The plasmids were methylated by transformation in *E. coli* ER1821 and introduced into competent

S. acidocaldarius MW001 cells. The cells were selected on first selection medium (N-Z-Amine®). To verify the presence of the plasmid, cells were harvested by centrifugation (5000 x g, 2 min, RT) and plasmid and chromosomal DNA was isolated by use of the DNeasy Blood and Tissue kit (Qiagen). PCRs were performed with one promoter- and one plasmid specific primer. *S. acidocaldarius* MW001 DNA served as negative control for PCR.

Table 14: Primers used for generation of reporter gene constructs of *tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe* and *gar*1

Gene	Primer sequence (5'-3')
*tfb*1_prom_*Sac*II_F	GCGCCGCGGAAAGGTCATAAGTTGTGGAAATAAC
*tfb*1_prom_*Nco*I_R	GCGCCATGGTTCACCACCCAAAATAC
*tfb*2_prom_*Sac*II_F	GCGCCGCGGAAACAAATAAAGCTTGTTG
*tfb*2_prom_*Nco*I_R	GCGCCATGGCTTCTTCTCCTCCCTTTC
*tfb*3_prom_*Sac*II_F	GCGCCGCGGGTTACATCACCATATTTATC
*tfb*3_prom_*Nco*I_R	GCGCCATGGAGAACAAAAAATCTCCTG
*tfe*_prom_*Sac*II_F	GCGCCGCGGTGATAGATGATATGATAG
*tfe*_prom_*Nco*I_R	GCGCCATGGATTAAGCACTATATTTTC
*tbp*_prom_*Sac*II_F	GCGCCGCGGGTTACCTTATCTCACTCCACTTATATG
*tbp*_prom_*Nco*I_R	GCGCCATGGCGCATATAAATCCAATGTTTGATC
*gar*1_prom_*Sac*II_F	GCGCCGCGGTTCGTCTTCTGAAGATG
*gar*1_prom_*Nco*I_R	GCGCCATGGATTATTAAACTCGCATTTTG
1472_pSVA1450_seq_prom_F*	ACATGTTCTTTCCTGCGTTATCC
1473_pSVA1450_seq_prom_R*	GGATCATGAACCCATTTATACC
1482_pSVA1450_seq_prom_F2*	TTCGCCACCTCTGACTTGAGC
1483_pSVA1450_seq_prom_R2*	CTGGTAGATCTCCACTTACTAATC

underlined: restriction site; * sequencing primer

2.15.2. Treatment of cultures for β-galactosidase reporter gene assays

MW001-pSVA1450 *tfb*1-prom, MW001-PSVA1450-*gar*1-prom and MW001-*tfb*3-prom cells were grown in Brock medium supplemented with 0.1% (v/v) N-Z-Amine® (20% (w/v) stock solution) instead of tryptone to maintain selective pressure and grown to an OD_{600} of 0.6. The cells were directly tested for reporter gene activity (0 min) and for cold-shock experiments the temperature was shifted to 60°C while shaking in a Thermotron (Infors) incubator (180 rpm). 10 ml samples were removed after 30 min, 60 min, 90 min, 120 min and 240 min.

2.15.3. Sample preparation for β-galactosidase reporter gene assays

10 ml cultures from part 2.15.2 were harvested by centrifugation (9000 x g, 10 min, 4°C) and cells were resuspended in 1.5 ml 1x Z-buffer (40 mM NaH_2PO_4, 60 mM Na_2HPO_4, 10 mM KCl, 1 mM $MgSO_4*7H_2O$, 0.5% (v/v) Triton X-100, PMSF and protease inhibitor (Roche),

pH 7.0). The cells were disrupted by 30 s sonication (0.6 cycle, 60 amplitude) and centrifuged for 30 min at 4°C and 20,800 x g. The supernatant was used for β-galactosidase reporter gene assays. A solution with 3.62 mM ortho-nitrophenyl β-galactopyranosid (ONPG) was prepared containing 50 mM β-mercaptoethanol. The reaction mixture consisted of 437.5 µl 1x Z-buffer (without Triton X-100, with 50 mM β-mercaptoethanol) and 25 µl of the substrate. The reaction was started by addition of 37.5 µl cell suspension. Absorbance of triplicates was monitored at 410 nm and determined at 75°C in a spectrophotometer (Jasco®). For testing autohydrolysis buffer and substrate were measured alone. One unit is defined as the change of absorbance at 410 nm over time induced by a defined protein concentration. Calculation of modified Miller units for determination of β-galactosidase activity (Miller, 1972):

$$\text{mod. Miller Units [U/min]} = \frac{1000 * (Abs_{410}\ t_2\text{-}t_1 - Abs_{410}\ \text{autohydrolysis}\ t_2\text{-}t_1)}{t\ [\text{min}] * Vsample\ [\text{ml}] * c\ \text{protein}\ [\frac{\text{mg}}{\text{ml}}]}$$

2.16. Protein-protein interaction studies using the yeast two-hybrid system

2.16.1. Construction of yeast two-hybrid vectors

The vectors pGADT7 and pGBKT7 were restricted with *Nde*I/*Xho*I or *Nde*I/*Pst*I, respectively. The PCR products amplified from genomic DNA from *S. acidocaldarius* and the plasmids were used in the following In-Fusion reaction (Clontech). The primers used for PCR amplification are listed in Table 15. The reaction mixture was calculated by the help of the molar ratio calculator (http://bioinfo.clontech.com/infusion/molarRatio.do) from Clontech. The molar ratio of vector and purified PCR product of 2:1 was used. 2 µl 5x Fusion-buffer, the calculated amounts of vector and insert (100-400 ng vector, 50-200 ng insert) and 1 µl Fusion enzyme were mixed. The reaction volume was adjusted to 10 µl with water. If the reaction volume exceeded 7 µl (insert and vector) the reaction volume was doubled. The samples were incubated for 20 min at 37°C and the reaction was stopped at 60°C for 15 min. Afterwards the sample volume was adjusted to 100 µl with TE-buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA). 2.5 µl of the In-Fusion reaction were added to supercompetent *E. coli* DH5α cells (see section 2.16.2.) and transformed with heat shock (see 2.6.8.).

Table 15: Primer sequences used for cloning of yeast two-hybrid constructs

Gene	Primer name	Primer Sequence (5'-3')
*rpo*A1 (Saci_0692)	028_Saci_rpoA1_*Nde*I_F	GATTACGCTCATATGAGTGAGAAGATTATACGGGGCG
	029_Saci_rpoA1_*Xho*I_R	CATCTGCAGCTCGAGCCTCTTCCAACCAATAACTCTTT
*rpo*A2 (Saci_0691)	030_Saci_rpoA2_*Nde*I_F	GATTACGCTCATATGATAGATGAAAAGTTAAAGGGTTA
	031_Saci_rpoA2_*Xho*I_R	CATCTGCAGCTCGAGCCTCATATTAGGCTTCATTGTTA
*rpo*B (Saci_0693)	032_Saci_rpoB_*Nde*I_F	GATTACGCTCATATGTTAGATACAGAGTCTAGGTGGGC AATAG
	033_Saci_rpoB_*Xho*I_R	CATCTGCAGCTCGAGCTCATTTGAAGCACCTCCTAAG
*rpo*D (Saci_0083)	034_Saci_rpoD_*Nde*I_F	GATTACGCTCATATGCCTATTTCTTTAATCGAAAGAAAT
	035_Saci_rpoD_*Xho*I_R	CATCTGCAGCTCGAGCTTTCCAGCCTCCAACTTACTCT
*rpo*E (Saci_0834)	036_Saci_rpoE_*Nde*I_F	GATTACGCTCATATGTTTAAGCTAGTTAGAGCTAAAGG
	037_Saci_rpoE_*Xho*I_R	CATCTGCAGCTCGAGCTTACTCGCCTTTGCTATTTCTT
*rpo*F (Saci_0625)	038_Saci_rpoF_*Nde*I_F	GATTACGCTCATATGTCATATTCACTTAAAACTATTGA
	039_Saci_rpoF_*Xho*I_R	CATCTGCAGCTCGAGGCTCTCCACGTACTTTTTAATAA
*rpo*G (Saci_0661)	040_Saci_rpoG_*Nde*I_F	GATTACGCTCATATGAAGGAAATGATGCAGGGAACCTG
	041_Saci_rpoG_*Xho*I_R	CATCTGCAGCTCGAGAGCCTTTTTCTTTACGCAATAAT
*rpo*H (Saci_0694)	042_Saci_rpoH_*Nde*I_F	GATTACGCTCATATGCGTTCATCCTCAAAGAAAAAGAT
	043_Saci_rpoH_*Xho*I_R	CATCTGCAGCTCGAGCCCTGTTATAACATACCTATAAG
*rpo*K (Saci_1370)	044_Saci_rpoK_*Nde*I_F	GATTACGCTCATATGACAATAGACAAGATTAACGAAAT
	045_Saci_rpoK_*Xho*I_R	CATCTGCAGCTCGAGTCTGTTCTCTATTTTTCTTACAC
*rpo*L (Saci_0173)	046_Saci_rpoL_*Nde*I_F	GATTACGCTCATATGGAAATTAAAGTAATAAAAGAGGA
	047_Saci-_rpoL_*Xho*I_R	CATCTGCAGCTCGAGAATTTTTTTCACTTCATCTATGA
*rpo*M (Saci_0171)	048_Saci_rpoM_*Nde*I_F	GATTACGCTCATATGCAATTCTGTCCTAAATGTGGCGG
	049_Saci_rpoM_*Xho*I_R	CATCTGCAGCTCGAGCTCCTCTTCTTGTGATTCTTCTA
*rpo*P (Saci_0864)	050_Saci_rpoP_*Nde*I_F	GATTACGCTCATATGGCAAAGTACAGATGTGGAAAGTG
	051_Saci_rpoP_*Xho*I_R	CATCTGCAGCTCGAGAATCGCTTTCACTATCTTAACGG
*tfb*1 (Saci_0866)	052_Saci_tfb1_Y2H_*Nde*I_F	GAGGACCTGCATATGGTAGAACAAAGTAAGGTTCCTTC
	053_Saci_tfb1_Y2H_*Pst*I_R	ATGCGGCCGCTGCAGTTGACTTGGAATCTGTATCTTTA
*tfb*2	054_Saci_tfb2_Y2H_*Nde*I_F	GAGGACCTGCATATGAAATGTAAAATCTGTGGCAGTGA

(Saci_1341)	055_Saci_tfb2_Y2H_*Pst*I_R	ATGCGGCCGCTGCAGAAGCTTTACTTCTATGTCAAAAG
*tfb*3 (Saci_0665)	056_Saci_tfb3_Y2H_*Nde*I_F	GAGGACCTGCATATGGAATGTCCCGAGTGTAAAAGTAG
	057_Saci_tfb3_Y2H_*Pst*I_R	ATGCGGCCGCTGCAGCATCGTTTTTAACGTTTGGCTAG
tbp (Saci_1336)	118 Saci_1336_Y2H_*Nde*I F	GATTACGCTCATATGGAAAGAAGCGTACCAAAC
	119 Saci_1336_Y2H_*Xho*I R	CATCTGCAGCTCGAGAAATTCTAACTCTTCTTCTTC
sequencing	T7 prom_F	TAATACGACTCACTATAGGG
sequencing	058_ pGBKT7_ seq_R	GAAATTCGCCCGGAATTAGC
sequencing	059_pGADT7_seq_R	CATGGCCAAGATTGAAAC

underlined: restriction site

2.16.2. Preparation of supercompetent *E. coli* DH5α cells (Inoue et al., 1990)

250 ml SOB medium (0.5% (w/v) yeast extract, 2% (w/v) tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$) were inoculated with an overnight culture of *E. coli* DH5α cells and incubated at 22°C with moderate shaking. When the OD_{600} reached 0.5 the culture was transferred to an ice-bath for 10 min. Then the cells were harvested for 10 min at 4°C and 2,500 x g. The cells were resuspended by swirling in 80 ml ice-cold Inoue transformation buffer (55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl, 10 mM PIPES (stock solution: 0.5 M, pH 6.7), sterilized by filtration). After subsequent centrifugation of the cells for 10 min at 4°C and 2,500 x g they were resuspended in 20 ml ice-cold Inoue transformation buffer. A volume of 1.5 ml DMSO was added to the suspension which was incubated for 10 min on ice. 50 µl aliquots were frozen in liquid nitrogen and stored at -80°C until the cells were needed.

2.16.3. Preparation of competent yeast cells

100 ml YPD medium (1% (w/v) Bacto yeast extract, 2% (w/v) Bacto peptone, 2.5% (w/v) agar, after autoclaving add sterile filtered 40% (w/v) glucose (stock solution) to 2% (v/v) final concentration were inoculated with 5 ml of an overnight pre-culture of *Saccharomyces cerevisiae* AH109 and cultures were incubated in a shaker (200 rpm) at 30°C until an OD_{600} of 0.6 was reached. After centrifugation of the cells for 5 min at 4°C and 4,000 x g the cells were resuspended in 25 ml sorbitol buffer (1 M sorbitol, 10 mM bicine, pH 8.35, 3% (v/v) ethylenglycol). The cells were centrifuged again for 10 min at 4°C and 400 x g and resuspended in 1 ml sorbitol buffer and 55 µl DMSO. 100 µl aliquots of the competent cells were stored at -80°C.

2.16.4. Transformation of yeast cells

5 µl pre-heated salmon sperm DNA (Invitrogen, 10 mg/ml, 20 min, 70°C) and 1 µg of respective plasmids (pGBKT7/pGADT7 with inserts) were added to frozen competent yeast cells. The cells were incubated at 37°C for 5 min with shaking (180 rpm). Then 700 µl of PEG-bicine solution (40% (w/v) PEG1000, 200 mM bicine, pH 8.35) were added to the cells and incubated for 1 hour at 30°C without shaking. The suspension was centrifuged for 10 s at

RT and 13,000 x g to collect the cells. This step was repeated six times. After the final centrifugation 1 ml of a NaCl-bicine solution (150 mM NaCl, 10 mM bicine) was added to the cell pellet without suspending the cells. 900 µl of the supernatant was removed and the pellet dissolved in the residual 100 µl. The cells were spread onto SD -Leu/-Trp plates (SD -Leu/-Trp pouch with 2.5% (w/v) bacto agar) which were incubated for three days at 30°C. For testing the expression of the proteins only one vector was transformed. 125 µl of a 200 x Leu or Trp stock was spread on the plates with glass beads.

2.16.5. Protein-protein interaction studies

After transformation of the yeast cells with a pGADT7 and pGBKT7 plasmid one single colony was picked and transferred to liquid SD -Leu/-Trp medium. The cultures were grown over night until OD_{600} 1.5 was reached. Cultures were diluted to an OD_{600} of 0.1 and 5 µl were dropped onto SD -Leu/-Trp/-Ade/-His/+Xα-gal plates. The plates were incubated at 30°C for three days. Cells grown with empty pGADT7-pGBKT7 vectors served as negative control and pGADT7::Sso*rpo*K and pGBKT7::Sso*tfb*3 were positive controls kindly offered by Prof. M. White (University of St. Andrews, UK).

2.16.6. Testing false positive reactions

To analyze whether the interactions are true positives the pGADT7 vector with insert was co-transformed with empty pGBKT7 vector. In cases of a false positive reaction the sample fused to the activating domain could bind to DNA and lead to expression of the reporter genes without a partner protein fused to a DNA-binding domain. Colonies obtained after transformation were used for inoculation of liquid SD -Leu/-Trp medium. After over night incubation at 30°C in a shaker (200 rpm) 100 µl of 1:10 and 1:100 diluted cells were plated onto SD -Leu/-Trp/+X-α-gal and SD -Leu/-Trp/-Ade/-His/+X-α-gal plates. Colonies growing on the lower stringency plate (SD -Leu/-Trp/+X-α-gal) should stay white if the interaction is true positive whereas they would turn blue in case of a false positive interaction. No growth should be observed on higher stringency plates (SD -Leu/-Trp/-Ade/-His/+X-α-gal) if the interactions are true positive.

2.16.7. Preparation of yeast cell extracts

All steps were carried out according to the yeast protocols handbook (Clontech). For preparation of a pre-culture 5 ml SD-medium supplemented with Leu or Trp were inoculated with one single colony (1-2 mm in diameter) not older than 4 days. An untransformed yeast colony grown in YPD medium served as negative control. After an over night incubation the cultures were mixed for dispersal of cell clumps. 50 ml YPD medium was inoculated with the entire pre-culture and cells were grown to an OD_{600} of 0.4 to 0.6 at 30°C in a shaker (200 rpm). The cells were quickly cooled by pouring them into a centrifuge tube halfway filled with ice. The samples were centrifuged for 5 min at 4°C and 1,000 x g and the resulting pellet was resuspended in 50 ml ice-cold water. After repetition of the centrifugation step samples were frozen in liquid nitrogen and stored at -80°C.

2.16.8. Urea/SDS-method

Cracking buffer (8 M urea, 5% (w/v) SDS, 40 mM Tris-HCl pH 6.8, 0.1 mM EDTA, 0.4 mg/ml bromophenol blue, for preparing 1.13 ml complete cracking buffer add 1 ml

cracking buffer, 10 µl β-mercaptoethanol, protease inhibitor and 50 µl of a 100 x stock of PMSF) was prepared and pre-warmed to 60°C. 100 µl cracking buffer was used for 7.5 OD_{600} units (OD_{600} unit = culture volume * OD_{600} measured in a 1 ml sample). Frozen cell pellets were resuspended in the pre-warmed cracking buffer. Because the PMSF concentration quickly decreases over time, 1 µl of a 100x stock solution per 100 µl cracking buffer were added after initial 15 min and after 7 min until the samples are boiled at 100°C. For disruption of the cells the suspension was transferred into 1.5 ml reaction tubes containing 80 µl glass beads per 7.5 OD_{600} units of cells. The samples were heated at 70°C for 10 min and mixed by vortex for 1 min. To pellet cell debris samples were centrifuged for 5 min at 4°C and 15,500 x g. The first supernatant was removed and stored on ice. The resulting pellets were incubated at 100°C for 3-5 min, then mixed by vortex again for 1 min and centrifuged with the same settings used before. Both supernatants were combined, boiled for 5 min and applied on a SDS-PAA gel or stored at -80°C.

2.16.9. TCA-method

Cell pellets were thawed on ice and resuspended in 100 µl ice-cold TCA-buffer (20 mM Tris-HCl (pH 8.0), 50 mM ammonium acetate, 2 mM EDTA, protease inhibitor, 100 µl of 100x PMSF per 10 ml) per 7.5 OD_{600} unit of cells. The samples were transferred to microcentrifuge tubes containing glass beads and ice-cold 20% (w/v) TCA (each 100 µl per 7.5 OD_{600} of yeast cells). The solutions were mixed by vortexing four times for 1 min with 30 s breaks (incubation on ice). The supernatant was removed to a fresh microcentrifuge tube. The glass beads were washed with 500 µl of an ice-cold, pre-mixed 20% (w/v) TCA and TCA-buffer (1:1). The solutions were mixed by vortexing two times for 1 min. The supernatants were combined and centrifuged (15,500 x g, 10 min, 4°C). The supernatant was removed again and the pellet was mixed in TCA-Laemmli-buffer (10 µl per OD_{600} unit of cells). The TCA-loading buffer is composed of a 480 µl SDS/glycerol stock (7.3% (w/v) SDS, 29.1% (v/v) glycerol, 83.3 mM Tris (pH not adjusted) and spatula-tip of bromophenol blue), a Tris/EDTA solution (200 mM Tris (pH not adjusted), 20 mM EDTA), β-mercaptoethanol, protease-inhibitor (Roche) and PMSF. After boiling the samples at 100°C for 10 min they were centrifuged for 10 min at 4°C and 15,500 x g. The supernatant was transferred to a fresh tube and loaded onto a SDS-gel or stored at -80°C.

2.16.10. Immunodetection for verification of protein expression in the yeast two-hybrid system

After western blot of the proteins the PVDF membrane was equilibrated in TBST for 5 min. For blocking the membrane was incubated in 30 ml TBST with 5% (w/v) milk powder for 1 h. The anti-c-myc tag or HA-tag antibodies linked with alkaline-phosphatase (abcam, 1:1,000 in 10 ml TBST, 3% (w/v) milk powder) were incubated over night at 4°C with gentile rotation. The next day unbound antibody was removed by washing three times with TBST for 15 min. Final washing steps were carried out twice with TBS for 15 min. 1 ml CDP-Star® was added onto the membrane and detection of proteins was performed with the VersaDoc system (Bio-Rad). Immunodetection was additionally carried out with HA-tag antibody (Sigma-Aldrich) and secondary anti-mouse antibody linked to horseradish peroxidase (HRP) (Sigma-Aldrich). The PVDF membrane (0.45 µm pore size) was blocked for 1 hour in TBST

with 5% (w/v) milk powder. The HA-tag antibody was diluted 1:1,000 in TBST with 3% (w/v) milk powder and incubated for 1 hour at RT. The membrane was washed with TBST and incubated with the secondary anti-mouse antibody linked with HRP for 90 min. Before detection by enhanced chemiluminescent (ECL) substrate (Pierce**)** the membrane was washed with TBST (3x 15 min) and TBS (3x 10 min).

3. Results

Compared to Eukaryotes, Archaea possess only a minimal set of GTFs, however, depending on the species the number of GTF paralogs differs in Archaea. The multiplicity of TFBs and TBPs has been proposed to reflect an adaptation to quickly changing and extreme environments and the presence of multiple general transcription factors in Euryarchaeota could be assigned to different functions (Thomsen et al., 2001, De Biase et al., 2002, Facciotti et al., 2007, Coker and DasSarma, 2007, Micorescu et al., 2008, Kaur et al., 2010, Turkarslan et al., 2011). Specialized GTFs could exhibit an advantage for cell survival and genome propagation. However, the function of multiple TFBs in Crenarchaeota still needs to be elucidated. *S. acidocaldarius* possesses three TFB proteins and one TBP (Chen et al., 2005). The function of these factors is poorly understood so far but it is most likely that TFB3 plays a role in stress response (e.g. UV-irradiation) (Götz et al., 2007, Paytubi and White, 2009). TFB2 in *Sulfolobus* spp. has not been investigated on protein level so far and its role is still unclear. In this work the expression and purification of the proteins by heterologous, homologous expression and tag-in was accomplished in order to further investigate the function of these proteins. The genomic context of *tfb*2 was analyzed for a functional prediction. Moreover, transcriptional analyses were performed to find out whether *tfb*2 is embedded in an operon structure. In addition, promoters of GTFs were cloned for reporter gene assays (β-galactosidase activity) in order to follow expression under different stress conditions. Since it is known that TFBs bind to the RNAP during transcription initiation (Bell and Jackson, 1998a, b, Bell et al., 1999a, Bell and Jackson, 2001, Geiduscheck and Ouhammouch, 2005), protein interaction studies with all twelve RNAP subunits of *S. acidocaldarius* and the three TFB paralogs were performed to investigate which subunit is bound by the different TFBs. In general TFB binds to TBP-DNA complex for initiation of transcription (Bell and Jackson, 1998a, b, Bell et al., 1999a, Bell and Jackson, 2001, Geiduscheck and Ouhammouch, 2005). In addition yeast two-hybrid experiments were used to analyze TFB and TBP interactions. To gain further insight into *in vivo* functions of the GTFs from *S. acidocaldarius* it was aimed to generate knock-out mutants.

3.1. Expression and purification of GTFs from *S. acidocaldarius*

Not much is known about the role of multiple GTFs in Crenarchaeota. Only few studies investigate transcription factors in *S. acidocaldarius* on protein level (Bell and Jackson, 2000a, Paytubi and White, 2009). To elucidate the function of the transcription factors of *S. acidocaldarius* in further experiments, like chromatin immunoprecipitation (ChIP), EMSA and footprinting, expression and purification of the proteins was a pre-requisite. The transcription factor TFB2 is of special interest, because it has not been characterized on protein level so far. Three approaches, namely heterologous expression in *E. coli*, homologous expression in *S. acidocaldarius* and insertion of a Strep-FLAG tag into the chromosome, have been undertaken.

3.1.1. Heterologous expression of GTFs in *E. coli*

TFB1 (Saci_0866), TFB2 (Saci_1341), TFB3 (Saci_0665) and TBP (Saci_1336) encoding sequences were amplified from genomic DNA and cloned into pET11c (without tag) using *Nde*I and *Bam*HI restriction sites, ligated into pET30a (C-terminal His-tag, via *Nde*I/*Xho*I) or were introduced into the directional cloning vector pET151/D-TOPO® (N-terminal His-tag, TEV cleavage site). The respective primers are given in Table 5 to Table 7 section 2.9.1. Overexpression of TFB1, TFB2, TFB3 and TBP was tested under standard conditions. 50 ml LB-medium were inoculated with an over night pre-culture and expression was induced after reaching an OD_{600} of 0.5 by addition of IPTG (1 mM final IPTG concentration). The cultures were incubated at 37°C and 180 rpm for 3 hours. Expression of the proteins encoded on pET11c was investigated with two different expression strains (Rosetta (DE3), BL21 (DE3) pRIL) which harbor an additional plasmid encoding rare tRNAs Arg, Ile, Leu. Expression strain One Shot® BL21 Star™(DE3) transformed with pET151/D-TOPO® constructs were analyzed for overexpression using SDS-PAGE (data not shown). Although cloning was successful and the correct sequence was confirmed by sequencing no overexpression of TFB1, TFB2 and TFB3 could be detected using SDS-PAA gels (data not shown). Therefore it was decided to clone the respective sequences into pET30a (C-terminal His-tag) (Table 7 in 2.9.1.). The plasmids were transformed with competent *E. coli* Rosetta (DE3) and BLR (DE3) pLysS cells (according to Bell and Jackson, 2000). The expression strain BLR (DE3) pLysS is a derivative of BL21 (DE3) and harbors the pLysS plasmid expressing T7 lysozyme to prevent T7 RNA polymerase expression prior to induction via IPTG. After expression in *E. coli* Rosetta (DE3) and BLR (DE3) pLysS under standard conditions (LB-medium, 1 mM IPTG, 3 h induction at 37°C) no difference between uninduced and induced samples, crude extract and membrane fraction was visible using SDS-PAGE and immunodetection (data not shown).

The encoding TFB sequences were amplified from genomic DNA and overexpression was performed in *E. coli* (Rosetta (DE3), BL21 (DE3) pRIL, BLR (DE3) pLysS); therefore the differences in codon usage between *S. acidocaldarius* and *E. coli* might account for missing expression. Consequently, codon usage was adapted to the expression host *E. coli* via gene synthesis using the service of the company GenScript. The plasmids derived from GenScript (supplement Fig. 37) were restricted with *Nde*I and *Bam*HI and the codon usage optimized gene sequences were ligated into pET15b (N-terminal His-tag) to express a His-tag containing protein for easy purification and detection. TFB1, TFB2 and TFB3 were solely expressed in BL21 (DE3) after overnight expression at 22°C with 200 µM IPTG (Fig. 3, Fig. 4, Fig. 5). Cells were grown at 37°C with shaking at 180 rpm until an OD_{600} of 0.5 was reached. Then they were induced with 200 µM IPTG and incubated at 22°C and 180 rpm for 18 hours. For overexpression of TFB2 the LB-medium was supplemented with 100 µM $ZnSO_4$ and 1 mM $MgCl_2$ because TFB proteins comprise an N-terminal Zn-finger region (Chen et al., 2000, Hahn and Roberts, 2000, Hickey et al., 2002, Geiduschek and Ouhammouch, 2005, Grohmann and Werner, 2011a).

TFB1 (37 kDa) was expressed as soluble protein and could be purified to homogeneity by heat-precipitation (75°C, 15 min), subsequent Ni-TED column (Macherey-Nagel) purification and gel filtration (Fig. 3). All buffers used for purification of TFB1 were supplemented with

1 mM $MgCl_2$. For a detailed purification protocol see section 2.9.2. Purification from 18 g (wet weight) of cells yielded 1 mg protein after gel filtration. A double band was visible after Ni-TED purification and gel filtration (Fig. 3 A). However, only the upper band corresponds to TFB1 (theoretical size of 37 kDa) which was confirmed by immunodetection (Fig. 3 B).

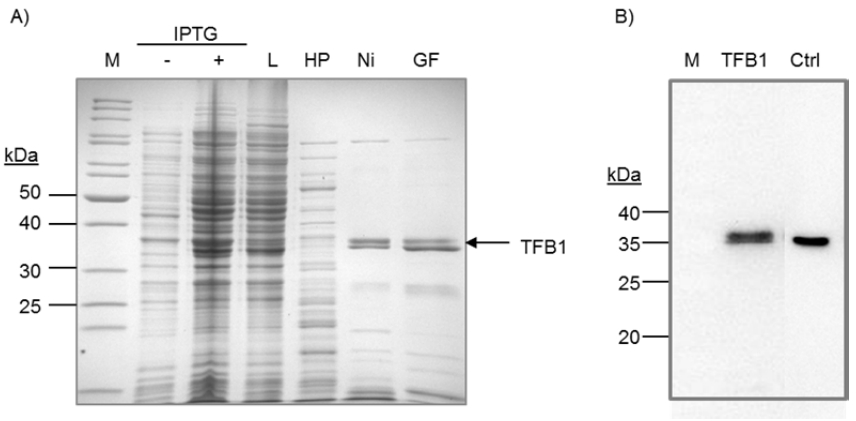


Fig. 3: Purification of recombinant TFB1 from *S. acidocaldarius* analyzed by SDS-PAGE and immunodetection.
Competent *E. coli* BL21 (DE3) cells were transformed with pET15b containing the codon optimized synthesized DNA sequence of *tfb*1 for expression in *E. coli*. A) The recombinant protein was purified via heat precipitation (75°C, 15 min) followed by Ni-affinity chromatography and gel filtration. Cell samples before (-) and after (+) induction, cell lysate (L; 20 µg) as well as after heat precipitation (HP; 15 µg), Ni-TED purification (Ni; 2 µg) and gel filtration (GF; 2 µg) were separated via a 12.5% SDS-PAA gel alongside the PageRuler™ Unstained Protein Ladder (M, Fermentas) and proteins were stained with Coomassie Brilliant Blue. B) Immunodetection of purified, recombinant TFB1-N-terminal 6x-His-tag (2 µg, 37 kDa) and control multi-tag protein (1 µg, 32 kDa, Roth) with alkaline-phosphatase-linked anti His-tag antibody (1:1,000).

TFB2 (35 kDa) and TFB3 (23 kDa) were expressed insoluble and required to be purified under denaturing conditions on Ni-TED columns (Fig. 4 and Fig. 5). Cell pellets of TFB2 and TFB3 were resuspended in 1x LEW buffer containing 1 mM DTT and 10% (v/v) glycerol (TFB2) or 20% (v/v) glycerol (TFB3). After disruption of the cells by French press and centrifugation the collected cells were resuspended in 1x LEW buffer supplemented with 1 mM DTT, 10% (v/v) or 20 % (v/v) glycerol and 6 M urea. The mixture was incubated for 1 hour at 22°C on a magnetic stirrer to denature proteins. The samples were centrifuged (20,800 x g, 30 min, 4°C) and the supernatants were filtered before application onto Ni-TED columns (section 2.9.3). The proteins bound to the column, were washed two times with LEW buffer containing 6 M urea and eluted with 250 mM imidazole. SDS-PAGE analysis of the purification revealed that TFB2 is present in the supernatant after urea treatment (Fig. 4 A). In addition, the flow through and wash fractions contained TFB2 maybe because too much protein solution was applied to the column and unbound protein was eluted. However, the main part of TFB2 eluted after addition of imidazole (Fig. 4 A). Before urea treatment TFB2 was only detectable in the membrane fraction and not in the crude extract as it is shown by immunodetection with anti-His-tag antibody (Fig. 4 B). Elution fractions after affinity chromatography under denaturing conditions contained a band at 35 kDa that corresponds to the theoretical size of TFB2 (Fig. 4). Since this is the first purification step the protein is not pure and additional protein bands were visible in the SDS-PAA gel and immunodetection (Fig. 4).

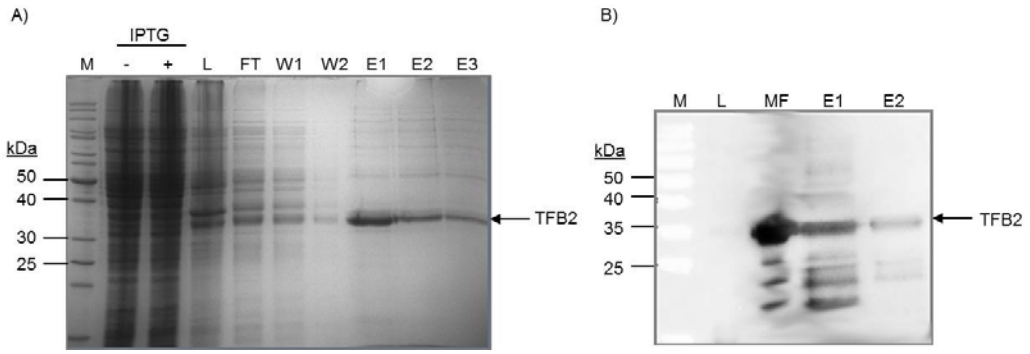


Fig. 4: Purification of recombinant TFB2 from *S. acidocaldarius* using denaturing immobilized metal ion affinity chromatography (IMAC).
Competent *E. coli* BL21 (DE3) cells were transformed with pET15b containing the codon optimized synthesized DNA sequence of *tfb*2 for expression in *E. coli*. A) The recombinant protein was purified in the presence of 6 M urea using Ni-TED columns. The samples before (-) and after (+) induction, cell lysate after urea treatment (L), flow through (FT), wash (W), and elution fractions (E) were separated on a 12.5% SDS-PAA gel alongside PageRuler™ Unstained Protein Ladder (M, Fermentas). Proteins were visualized by Coomassie Brilliant Blue staining. B) Detection of recombinant TFB2 with N-terminal 6x-His-tag (35 kDa) in cell lysate before urea treatment (L), membrane fraction before addition of urea (MF) and elution fractions of Ni-TED column after urea treatment (E1, E2) with alkaline-phosphatase-linked anti-His-tag antibody (1:1,000), M: PageRuler™ Prestained Protein Ladder (Fermentas). The protein concentration was not determined due to the presence of urea.

TFB3 with N-terminal 6x-His tag has a predicted molecular weight of 23 kDa. Besides other contaminating protein bands a band at 23 kDa was visible after induction, in the cell lysate after urea treatment, the flow through, wash fractions and the elution fractions of the Ni-TED chromatography (Fig. 5 A). Immunodetection of the first elution fraction after affinity chromatography with anti-His-tag antibody revealed a band at approximately 23 kDa and two additional bands between 20 kDa and 10 kDa, suggesting protein degradation.

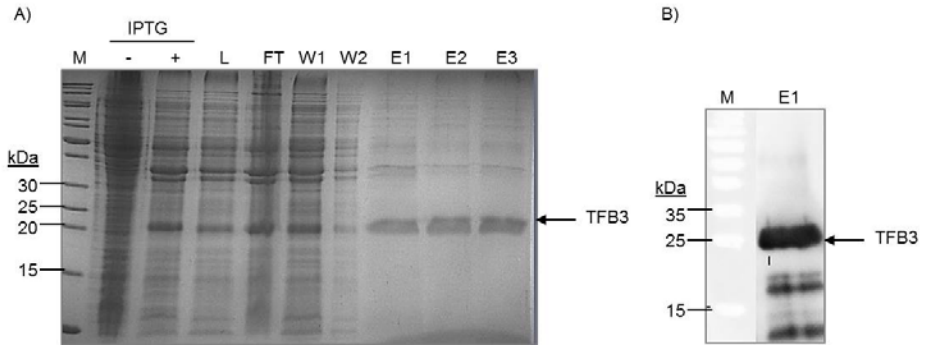


Fig. 5: Purification of recombinant TFB3 from *S. acidocaldarius* under denaturing conditions using IMAC.

Competent *E. coli* BL21 (DE3) cells were transformed with pET15b containing the codon optimized synthesized DNA sequence of *tfb*3 for expression in *E. coli*. A) Recombinant expressed TFB3 was purified in the presence of 6 M urea using Ni-TED columns. The samples before (-) and after (+) induction, cell lysate (L), flow through (FT), wash (W) and elution fractions (E) were separated on a 15% SDS-PAA gel alongside with PageRuler™ Unstained Protein Ladder (M, Fermentas) and visualized with Coomassie Brilliant Blue. B) Immunodetection of recombinant TFB3-N-terminal 6x-His-tag (23 kDa) in elution fraction of Ni-TED column after urea treatment with alkaline-phosphatase-linked anti-His-tag antibody (1:1,000), M: PageRuler™ Prestained Protein Ladder (Fermentas). The protein concentration was not determined due to the presence of urea.

Because the proteins were purified under denaturing conditions they needed to be refolded. First attempts to refold TFB3 failed. A drop by drop dilution with the 30-fold volume of the initial sample volume of TFB3 (3 ml initial sample volume, 1x LEW, 1 mM DTT, 20% (v/v) glycerol) into buffer without urea was tested but the protein precipitated. In addition, it was tried to refold denatured TFB3 on the column by washing it two times with 50 ml LEW buffer without urea. After elution no protein was detectable via SDS-PAGE and corresponding immunoblots in all three elution fractions (data not shown). Even expression of TFB2 and TFB3 in *E. coli* Lemo (DE3) (New England Biolabs) did not increase solubility of these proteins. This expression strain is supposed to accomplish expression of membrane proteins, toxic proteins and proteins prone to insoluble expression by fine tuning of T7 expression. Taken together TFB2 and TFB3 were successfully overexpressed but were insoluble and could only be purified under denaturing conditions. Refolding of TFB2 could not be tested and TFB3 could not be refolded under the tested conditions.

The TBP encoding sequence (Saci_1336) was amplified from genomic DNA, restricted with *Nde*I and *Bam*HI and ligated into pET11c (no tag). The primer sequences used for amplification of TBP are listed in Table 5. The plasmid was sequenced and transformed into the expression strain *E. coli* Rosetta (DE3). Transformed cells were grown in LB-medium at 37°C until reaching OD_{600} 0.5 and induced by addition of 1 mM IPTG. The cultures were incubated for three hours at 37°C and cells were collected by centrifugation (14,800 x g, 4°C, 15 min) (section 2.9.4.). TBP was soluble and 0.35 mg of protein could be purified from 12 g cells (wet weight) by heat-precipitation (80°C, 20 min) followed by ion-exchange chromatography (ResourceQ, GE Healthcare) and gel filtration (HiLoad Superdex 26/60 75 prep grade, GE Healthcare). The purified protein revealed a molecular weight of 19 kDa in a SDS-PAA gel which is similar to the calculated molecular weight of 18.9 kDa (Fig. 6).

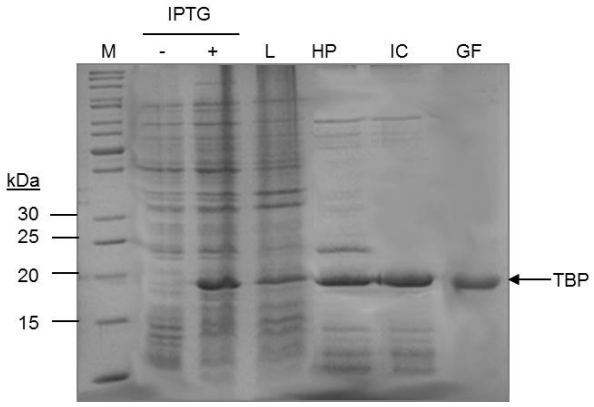


Fig. 6: Purification steps of recombinant TBP from *S. acidocaldarius* analyzed by SDS-PAGE.
E. coli Rosetta (DE3) cells were transformed with pET11c-*tbp* and expressed TBP (19 kDa) was purified from crude extract by heat-precipitation (80°C, 20 min), ion-exchange chromatography and subsequent gel filtration. Samples before (-) and after induction (+), cell lysate (L; 20 µg) and fractions after heat precipitation (HP; 15 µg), ion exchange chromatography (IC; 5 µg) and gel filtration (GF; 2 µg) were separated via a 15% SDS-PAA gel alongside PageRuler™ Unstained Protein Ladder (M; Fermentas). Proteins were stained with Coomassie Brilliant Blue.

To complete the set of transcription factors involved in transcription initiation, the TFE (Saci_0652) encoding sequence was amplified from genomic DNA, restricted with *Nde*I and *Xho*I and ligated into pET30a (C-terminal His-tag). Primer sequences are listed in Table 7. The DNA sequence was confirmed by sequencing and the plasmid was introduced into *E. coli* Rosetta (DE3) for overexpression. As described in section 2.9.5 cultures were incubated at 37°C until reaching an OD_{600} of 0.5. Expression of TFE was induced by addition of 200 µM IPTG and incubation over night at 22°C. Buffers used for disruption of the cells and chromatography were supplemented with 2 mM DTT. The protein was soluble and could be successfully purified via heat-precipitation (75°C, 15 min), IMAC (ResourceQ, GE Healthcare) and gel filtration (HiLoad Superdex 26/60 75 prep grade, GE Healthcare) (Fig. 7). 2.1 mg protein was purified from 18 g cells (wet weight). The fraction after gel filtration contained two adjacent bands with different intensities. The dominant, upper band seemed to be TFE and the faint band directly below might be a degradation product or derives from different post-translational modifications (Fig. 7 A). The calculated size of TFE is 21 kDa and the SDS-PAGE analysis of TFE purification demonstrated in Fig. 7 A showed a slightly smaller size (19 kDa) whereas immunodetection revealed one band at the correct size (21 kDa) (Fig. 7 B).

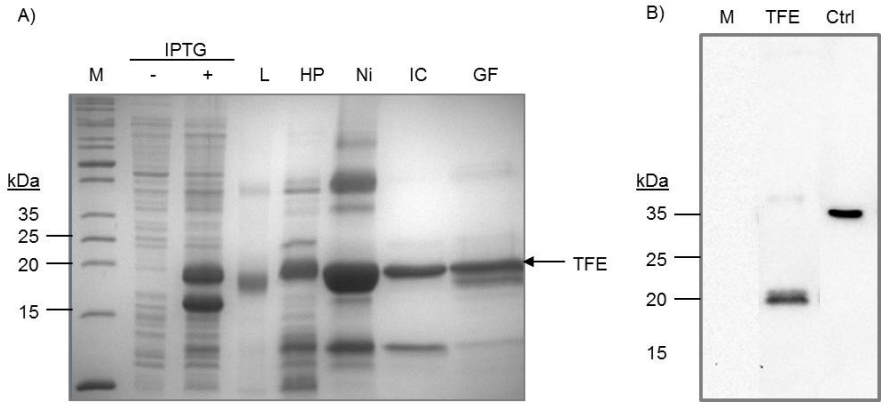


Fig. 7: Purification of recombinant TFE from *S. acidocaldarius* analyzed by SDS-PAGE and immunodetection.
Competent *E. coli* Rosetta (DE3) cells were transformed with pET30a-*tfe*. A) Recombinant expressed TFE (21 kDa) was purified via heat-precipitation, Ni-TED affinity chromatography, ion-exchange chromatography and gel filtration. Samples before (-) and after induction (+), cell lysate (L; 20 µg) and fractions after heat precipitation (HP; 15 µg), Ni-TED purification (Ni; 2 µg), ion-exchange chromatography (IC; 2 µg) and gel filtration (GF; 2 µg); were separated on a 15% SDS-PAA gel alongside with PageRuler™ Unstained Protein Ladder (M; Fermentas). SDS-PAA gels were stained with Coomassie Brilliant Blue. B) Immunodetection of recombinant TFE with C-terminal 6x-His-tag (TFE; 21 kDa) and control multi-tag protein (Ctrl, 32 kDa, Roth) with alkaline-phosphatase-linked anti-His-tag antibody (1:1,000).

For development of antibodies against one of the RNAP subunits the encoding sequence of RpoG (Saci_0661) was cloned into the vector pET30a containing a C-terminal His-tag. The encoding sequence was amplified from genomic DNA using primers listed in Table 7 and restricted with *Nde*I and *Xho*I. After confirmation of the sequence by sequencing the plasmid was transformed into competent *E. coli* Rosetta (DE3) cells. According to section 2.9.6 cultures were grown at 37°C and 180 rpm until reaching an OD_{600} of 0.5. Expression of RpoG was induced by addition of 200 µM IPTG and cultures were transferred to 22°C for over night incubation. The protein was soluble and could be purified to homogeneity by heat precipitation (75 °C, 15 min), Ni-TED purification (Macherey-Nagel) and gel filtration (HiLoad Superdex 26/60 75 prep grade, GE Healthcare) (Fig. 8). From 20 g of cells (wet weight) 1.2 mg protein was purified. Buffers for disruption of the cells and chromatography contained KCl instead of NaCl and 0.5 mM DTT (section 2.9.6). SDS-PAGE analysis and immunodetection with anti-His-tag antibody revealed a pure protein. The molecular weight of RpoG corresponds to the calculated molecular weight of 15.5 kDa (Fig. 8).

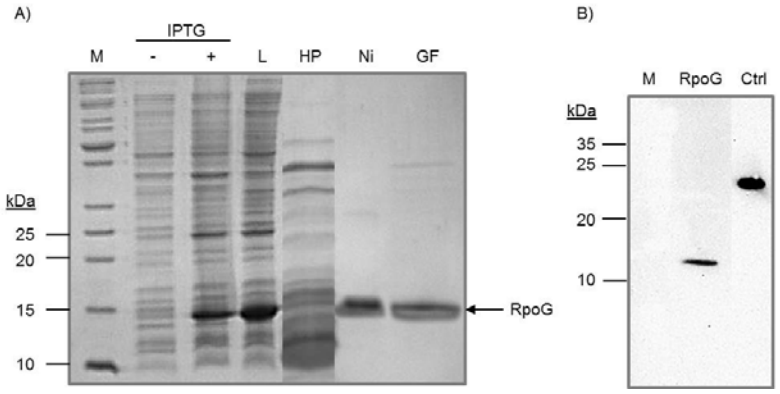


Fig. 8: Purification of recombinant RpoG from *S. acidocaldarius* analyzed by SDS-PAGE and immunodetection.
Competent *E. coli* Rosetta (DE3) cells were transformed with pET30a-*rpoG* (C-terminal 6x-His-tag) and expressed RpoG (15 kDa) was purified by heat precipitation, IMAC and gel filtration. A) Cells before (-) and after induction (+), cell lysate (L; 20 µg), fractions after heat precipitation at 75°C (HP; 15 µg), Ni-TED purification (Ni; 2 µg) and gel filtration (GF; 2 µg) were separated on a 15% SDS-PAA gel alongside PageRuler™ Unstained Protein Ladder (M; Fermentas) and visualized by staining with Coomassie Brilliant Blue. B) Immunodetection of recombinant RpoG with a C-terminal 6x-His-tag (RpoG, 15 kDa) and control multi-tag protein (Ctrl; 32 kDa, Roth) with alkaline-phosphatase-linked anti-His-tag antibody (1:1,000).

Taken together all GTFs from *S. acidocaldarius* namely TFB1, TFB2, TFB3, TBP and TFE as well as RpoG could be recombinant overexpressed in *E. coli* (Fig. 3-8) and all proteins, except TFB2 and TFB3, could be purified to homogeneity (Fig. 3-8). TFB2 and TFB3 were purified under denaturing conditions and therefore need to be refolded. First attempts to refold TFB3 failed because the protein precipitated.

3.1.2. Antibody generation for GTFs TFB1 and TFE

Antibodies were generated against purified recombinant 6x-His-TFB1 and 6x-His-TFE. 1 mg of each protein was shipped to Eurogentec and during 28-days two rabbits were immunized with four boosts of protein to develop polyclonal antibodies. Different dilutions (1:1,000, 1:500, 1:250) of antisera obtained from both animals were tested for application in immunoblots. The optimal antiserum dilution was determined to be 1:500 for both antibodies and they were able to recognize TFB1 and TFE, respectively, as pure recombinant proteins as well as in *S. acidocaldarius* MW001 crude extract (Fig. 9).

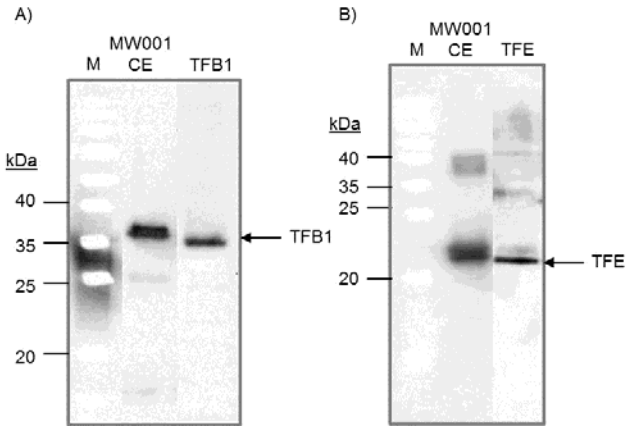


Fig. 9: Immunodetection of *S. acidocaldarius* TFB1 (A) and TFE (B) with generated anti-TFB1 and anti-TFE polyclonal antibodies from rabbit, using MW001 crude extract as well as recombinant proteins.
S. acidocaldarius MW001 crude extract and recombinant TFB1 and TFE were separated via 12.5% SDS-PAA gels and blotted onto PVDF membranes. A) Immunodetection of TFB1 was performed with polyclonal TFB1 antibodies (1:500) generated in this study and anti-rabbit IgG conjugated with alkaline-phosphatase (1:10,000) was used as secondary antibody B) Immunodetection of TFE with generated polyclonal TFE antibodies (1:500) using anti-rabbit IgG conjugated with alkaline-phosphatase (1:10,000) as secondary antibody. M, PageRuler™ Prestained Protein Ladder (Fermentas); MW001 CE, MW001 crude extract; TFB1, recombinant TFB1 (37 kDa); TFE, recombinant TFE (21 kDa).

As shown in Fig. 9 the generated polyclonal antibodies of TFB1 and TFE were successfully applied in immunoblots. For some applications, like ChIP experiments, it is advantageous to work with pure antibodies. For antibody purification from serum an ammonium sulfate precipitation was performed followed by affinity purification using protein-A-agarose beads (Roche) according to the manufacturer's instructions which is described in part 2.12.2. Elution fractions of the antibody were analyzed via SDS-PAGE (Fig. 10). The antibody was free of other protein contaminants and the heavy (approximately 50 kDa) and the light chain (around 18 kDa) expected for rabbit antibody could be detected. The concentration of the elution fraction shown in Fig. 10 was 0.1 mg/ml.

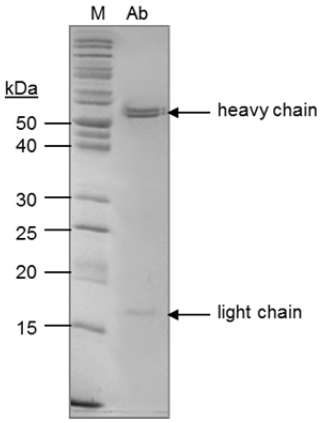


Fig. 10: Purification of *S. acidocaldarius* anti-TFE antibody from polyclonal rabbit antiserum.
Purification was performed via ammonium sulfate precipitation and subsequent purification via protein-A-agarose beads. The elution fraction was separated via a 12.5% SDS-PAA gel and stained with Coomassie Brilliant Blue. M: PageRuler™ Unstained Protein Ladder (Fermentas), Ab: antibody, elution fraction of TFE antibody.

3.1.3. Homologous expression

Heterologous expression of GTFs from *S. acidocaldarius* described above was difficult for TFB1, TFB2 and TFB3. Attempts to overexpress the proteins using different expression plasmids and *E. coli* expression hosts failed. It was only possible to express the proteins in *E. coli* after gene synthesis and codon adaptation. Therefore homologous expression was performed as alternative strategy to purify the GTFs from *S. acidocaldarius*. An advantage of homologous expression compared to heterologous expression is that native protein modifications will be achieved.

Primer sequences and overexpression details can be found in 2.10 in Table 8. The sequences of *tfb*1, *tfb*2, *tfb*3 and *tbp* were amplified from genomic DNA and the PCR products were introduced into the pre-vector pPREX by restriction with *Nco*I and *Bam*HI in order to fuse the proteins to a Strep-His-tag. Afterwards the coding sequences were restricted from the pre-vector (*Nco*I/*Eag*I) and ligated into the overexpression vector pSVA1450. The plasmid pSVA1450 contains a maltose inducible promoter and the downstream positioned LacS encoding sequence of *S. solfataricus*. The *lac*S sequence is replaced by the transcription factors from *S. acidocaldarius* to gain controlled overexpression. Successful cloning into pPREX and pSVA1450 was analyzed by colony PCR and sequencing. The proteins were fused to a Strep-His-tag for a simplified purification of the encoded proteins. In this expression system *S. acidocaldarius* MW001 is used as expression host and induction is achieved by addition of maltose to the growth medium directly after inoculation of the main culture. *S. acidocaldarius* MW001 cells harboring the transcription factors under control of the maltose inducible promoter (pSVA1450) were grown in 8 L overexpression medium that already contained maltose at 78°C. At OD_{600} between 0.6 and 0.8 the cells were collected by centrifugation (14,800 x g, 4°C, 15 min) and disrupted by sonication. Cell-free extracts were applied on Strep-tag columns for purification.

In Fig. 11 and Fig.12 Strep-tag purification steps of homologous expressed TFB2 and TFB3 are depicted and similar results were obtained for TFB1 and TBP (data not shown). The transcription factors were neither detected in the flow through or wash fraction nor in the corresponding immunoblots using anti-Strep-tag antibody (Fig. 11, Fig. 12). In case of Strep-tag purification six elution fractions were collected and the vast majority of proteins eluted between fractions two and five (Fig. 11, Fig. 12). TFB2 and TFB3 have theoretical sizes of 35.6 kDa and 25.6 kDa, respectively. Unspecific protein bands were visible in SDS-PAA gels in all elution fractions of TFB2 and TFB3 at 50 kDa and 15 kDa and the latter band was also detected using Strep-tag antibodies (Fig. 11, Fig. 12). Alternatively, also purification via Ni-TED columns was performed. However, with another matrix the transcription factors could not be purified (data not shown). Immunoblots using antibody against TBP of the Crenarchaeon *Thermoproteus tenax* were performed with whole cells from *S. acidocaldarius* MW001 and cells from the overexpression experiment (MW001-pSVA1450-*tbp*). The anti-*Ttx*-TBP antibody specifically recognized recombinant TBP from *T. tenax* and was also able to recognize pure, recombinant TBP from *S. acidocaldarius* although with less intensity (data not shown). The presence of TBP in *S. acidocaldarius* MW001 cells and in samples of the overexpression experiment could not be verified by immunodetection with anti-*Ttx*-TBP antibody (data not shown). This could be due to the different TBP concentrations in whole

cells compared to purified recombinant protein as well as the lower binding affinity of the antibody to TBP from *S. acidocaldarius*.

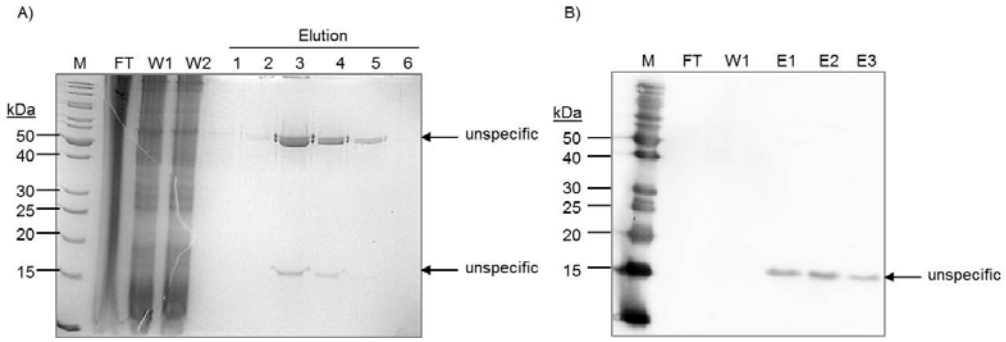


Fig. 11: Homologous expression of TFB2 in *S. acidocaldarius* using Strep-tag columns and immunodetection.
The encoding sequence of *tfb*2 was cloned via the pre-vector pPREX in the overexpression vector pSVA1450. Expression of TFB2 (35.6 kDa) was induced by addition of maltose. A) Flow through (FT), wash (W) and elution fractions (1-6) of Strep-tag purification were separated via 12.5% SDS-PAA gels alongside PageRuler™ Unstained Protein Ladder (M; Fermentas) and were stained with Coomassie Brilliant Blue (B) Immunodetection with anti-Strep-tag antibody (1:4000) linked to alkaline-phosphatase of flow through (FT), wash step 1 (W1) and elution fractions E1 to E3; PageRuler™ Prestained Protein Ladder (M, Fermentas).

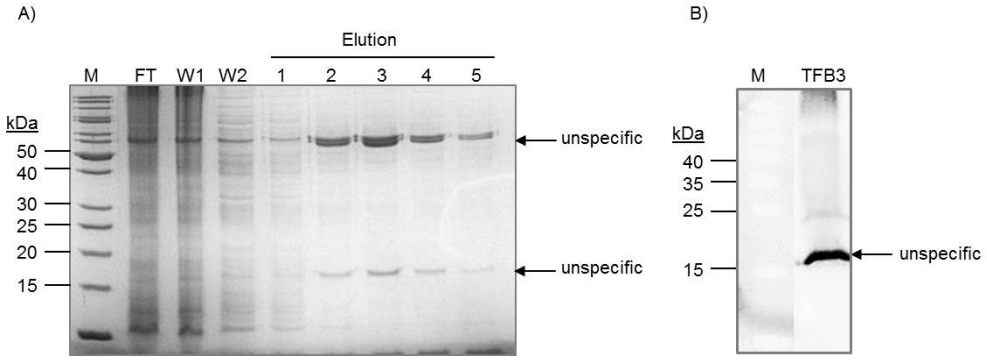


Fig. 12: Homologous expression of TFB3 in *S. acidocaldarius* using Strep-tag columns and immunodetection.
The encoding sequence of *tfb*3 was cloned via the pre-vector pPREX in the overexpression vector pSVA1450. Homologous expression of TFB3 (25.6 kDa) was analyzed by Strep-tag purification and separation of flow through (FT), wash (W) and elution fractions (E) alongside with PageRuler™ Unstained Protein Ladder (M; Fermentas) on a 12.5% SDS-gel which was stained with Coomassie Brilliant Blue B) Immunodetection of homologous expressed TFB3 in elution fraction E3 was performed with anti-Strep-tag antibody (1:4000) linked to alkaline-phosphatase.

Strikingly, homologous overexpression in *S. acidocaldarius* MW001 and purification of all tested GTFs (TFB1, TFB2, TFB3 or TBP) was not possible.

3.1.4. Knock-in for purification of GTFs

Since no homologous overexpression using *S. acidocaldarius* as expression host was possible, a third approach, genomic tagging of endogenous protein encoding sequences was performed. In addition purification of proteins using the homologous host could offer advantages over heterologous expression of recombinant proteins in regard to improved protein folding and

undesirable post-translational modifications. For purification of GTFs from *S. acidocaldarius* MW001 a C- or N-terminal Strep-FLAG tag was introduced into the 5'- or 3'-genomic region of *tfb*1, *tfb*2, *tfb*3 and *tfe* by homologous recombination (sections 2.14.1 and 2.14.2). Briefly, the tag was introduced via PCR mutagenesis and the PCR products were restricted with *Nde*I and *Bam*HI and ligated into pSVA406. Primer sequences are listed in Table 12 and Table 13. The vector was methylated and transformed into competent cells of the uracil auxotroph mutant *S. acidocaldarius* MW001 (lacks *pyrEF* cassette). After first selection, without uracil supplementation, the plasmid integrates into the genome. At this point only cells which produce uracil can grow since containing the marker cassette (*pyrEF*) which encodes enzymes involved in uracil synthesis, i. e. orotate phosphoribosyltransferase and orotidine-5'-monophosphate decarboxylase (Grogan and Gunsalus, 1993). During second selection (+ uracil, + 5-FOA) homologous regions of plasmid and genomic DNA recombine and the plasmid pops-out of the genome. Only non-producing uracil cells survive because the substrate-analogon 5-FOA is converted to the toxic compound 5-fluorouracil, which is an indicator that homologous recombination occurred successfully. DNA of colonies grown on second selection plates was isolated and tested for insertion of the tag by PCR. PCR products of positive clones were sequenced and analyzed to confirm the presence of the tags.

The genetically modified *S. acidocaldarius* MW001 strains were incubated at 78°C until reaching OD_{600} 0.6-0.8 and collected by centrifugation. Cells (8 g wet weight) were disrupted and cell-free extracts were applied on Strep-tag columns (section 2.11). TFB1 and TFE could be purified via Strep-tag columns (Fig. 13). Whereas for TFB1 (37 kDa) significant enrichment was achieved (Fig. 13 A) only a faint band was visible for TFE (22 kDa) after SDS-PAGE analysis (Fig. 13 B).

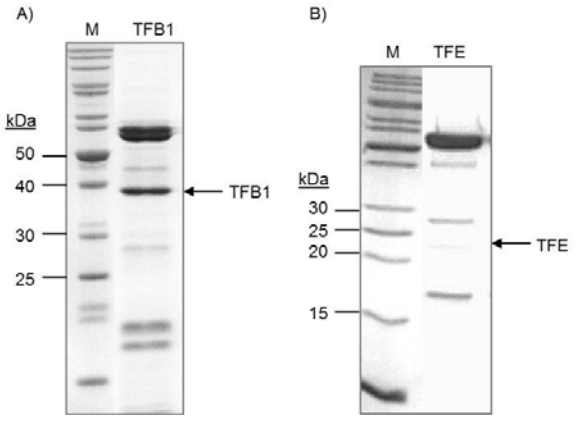


Fig. 13: Purification of TFB1 and TFE with C-terminal Strep-FLAG tags from *S. acidocaldarius* MW001 cells.
Genomically modified *S. acidocaldarius* MW001 cells harboring a C-terminal Strep-FLAG tag at TFB1 or TFE encoding sequences were grown to OD_{600} of 0.8, collected and disrupted by French press. Cell free-extracts were prepared by centrifugation and applied onto Strep-tag columns for purification. Elution fraction 3 of TFB1 (TFB1, 37 kDa, (A)) and TFE (TFE, 22 kDa, (B)) were separated via 12.5% (A) and 15% SDS-PAA gels (B) alongside with PageRuler™ Unstained Protein Ladder (M; Fermentas) and the SDS-PAA gels were stained with Coomassie Brilliant Blue.

For immunodetection of the proteins with anti-Strep-tag or anti-FLAG-tag antibody 15 ml of the culture (OD_{600} 0.6-0.8) were centrifuged and cells were resuspended in 50 µl 4 x SDS-PAGE loading buffer and boiled at 95°C for 10 min and applied on SDS-PAA gels (Fig. 14 B). In case of TFB1 and TFE also purified proteins were used for immunoblot detection. Both proteins could be detected either with anti-FLAG-tag antibody (Fig. 14 A) or Strep-tag antibody (Fig. 14 C) in crude extract of the knock-in strains (data not shown) as well as after the purification steps (Fig. 14 A, C). In case of TFB2 and TFB3 no signal could be detected in immunodetection using anti-Strep-tag antibody (Fig. 14 B). Only an unspecific band around 15 kDa that is always visible in MW001 crude extracts when Strep-tag antibody is used for immunodetection.

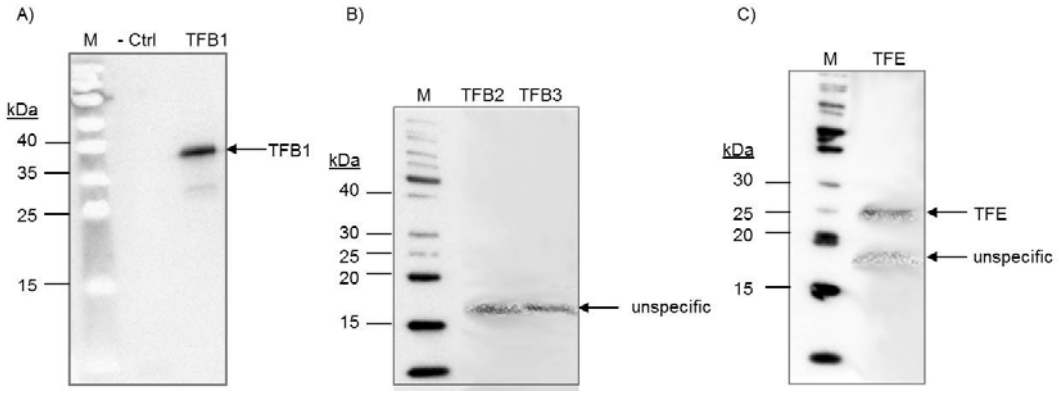


Fig. 14: Immunodetection of genomic Strep-FLAG tagged GTFs.
S. acidocaldarius MW001 cells (control) or genetically modified strains (MW001::*tfb*1C, MW001::*tfb*2N, MW001::*tfb*3N and MW001::*tfeC*) were grown at 78°C until reaching OD_{600} 0.6-0.8 and harvested by centrifugation. Proteins were purified via Strep-tag columns as outlined above (TFB1, TFE) or 15 ml culture was harvested and pellets were resuspended in 50 µl 4x SDS-PAGE loading buffer (TFB2, TFB3). The purified proteins or whole cell extracts were separated via SDS-PAA gels (12.5%) and blotted onto PVDF membranes over night at 4°C. A) Immunodetection with anti-FLAG tag antibody (1:1,000, Sigma-Aldrich) using *S. acidocaldarius* MW001 (-Ctrl) and Strep-tag purified protein from the strain MW001::*tfb*1C (TFB1, 37 kDa); anti-rabbit IgG linked with alkaline-phosphatase (1:10,000) served as secondary antibody. M, PageRuler™ Prestained Protein Ladder. B) and C) Immunodetection using anti-Strep-tag antibody (1:4000) linked to alkaline-phosphatase. Whole cells of the strains MW001**::***tfb*2N (B, TFB2: 35 kDa), MW001::*tfb*3N (B, TFB3, 25 kDa) and purified protein from MW001::*tfeC* (C, TFE, 22 kDa) were analyzed by immunodetection. M, PageRuler™ Unstained Protein Ladder (Fermentas).

3.1.5. Knock-in of a 8x-His-tag for purification of RNAP

For a simplified purification of the RNAP of *S. acidocaldarius* for future applications such as *in vitro* transcription assays an 8x-His-tag was inserted into the 3'-genomic region of the subunit *rpoG* with the same method as described above. 10 L Brock medium supplemented with uracil were inoculated with a pre-culture of *S. acidocaldarius* MW001::*rpoG*-His cells. The culture was grown until OD_{600} of 0.8 and the cells were collected by centrifugation. The pellet was resuspended in 1x LEW buffer containing 0.5 mM DTT, and 300 mM KCl instead of NaCl and protease inhibitor. Cells were disrupted by French press (threefold passage) and

clarified crude extract was obtained by centrifugation. The crude extract was applied to Ni-TED columns and RNAP mainly eluted in fractions 1 and 2 (Fig. 15 A). Imidazole was removed by dialysis against 1x LEW buffer containing 0.5 mM DTT, 300 mM KCl. For storage at -80°C 10% (v/v) glycerol was added (section 2.14.3). RNAP containing fractions were analyzed by SDS-PAGE and immunodetection with anti-His-tag antibody (Fig. 15 B) and anti-TFB1 antibody (Fig. 15 C). The RNAP complex could be purified in one step using Ni-TED columns. All twelve subunits could be detected using SDS-PAGE (Fig. 15 A). From 18 g cells (wet weight) 2.52 mg protein were retrieved. Immunodetection with anti-His-tag antibody confirmed the presence of the RpoG subunit in the purified polymerase complex (Fig. 15 B). Three additional faint bands were visible in Fig. 15 B between 16 kDa and 25 kDa probably due to unspecific binding of anti-His tag antibody. In addition immunodetection with anti-TFB1 antibody which was generated in this study revealed that TFB1 co-purifies with the RNAP. Therefore, for subsequent experiments like electrophoretic mobility shift assay or *in vitro* transcription, further purification steps, e. g. gel filtration, are required to prepare TFB1-free RNAP.

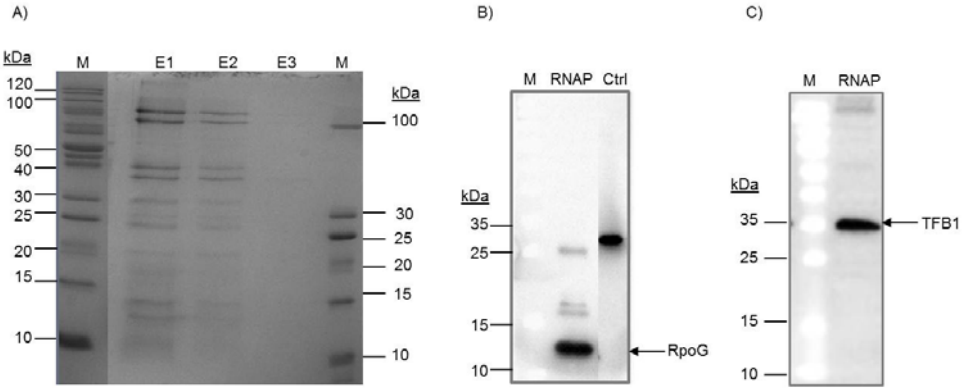


Fig. 15: Purification of genomic His-tagged RNAP from *S. acidocaldarius.*
Samples were analyzed by SDS-PAGE (A) and immunodetection with anti-His-tag (B) and anti-TFB1 antibody (C). The strain MW001::*rpo*G-His was grown in Brock medium containing uracil. Cells were harvested at OD_{600} of 0.8. Crude extracts were prepared by cell disruption with French press and the RNAP was purified via Ni-TED columns.
A) The elution fractions after Ni-TED purification were analyzed via a 12.5% SDS-PAA gel and proteins were stained with Coomassie Brilliant Blue; M, PageRuler™ Unstained Protein Ladder (Fermentas); E (1-3), elution fractions of Ni-TED column; B) M, PageRuler™ Unstained Low Range Protein Ladder (Fermentas); Rpo subunits: A1, 99.8 kDa; A2, 44.4 kDa; B, 122.5 kDa; D, 29.8 kDa; E, 20.5 kDa; F, 13 kDa; G, 14.4 kDa; H, 9.5 kDa; K, 10.2 kDa; L, 10.0 kDa; M, 11.5 kDa; P, 5.6 kDa; B) Immunodetection with alkaline-phosphatase linked anti-His tag antibody (1:1,000) showing the tagged RpoG subunit in elution fraction E1. M, PageRuler™ Prestained Protein Ladder (Fermentas); RNAP, elution fraction E1; Ctrl, multi-tag protein (32 kDa, Roth); C) Immunodetection with anti-TFB1 antibody (1:500, TFB1: 37 kDa), anti-rabbit IgG-antibody linked with alkaline-phosphatase served as secondary antibody (1:10,000). RNAP, elution fraction E1; M, PageRuler™ Prestained Protein Ladder (Fermentas).

In summary the established protocols for purification of genomically tagged transcription factors TFB1 and TFE as well as RNAP will provide the basis for further experiments like electrophoretic mobility shift assays, *in vitro* transcription and immunoblots for comparison of protein expression levels.

3.2. Genome organization of *tfb*2

3.2.1. Genomic context analysis of *tfb*2 (Saci_1341) and sequence analysis

In the previous described experiments TFB2 could neither be overexpressed in *S. acidocaldarius* nor could it be detected after genomic integration of an N-terminal Strep-Flag-tag encoding sequence. Therefore the genomic context and operon structure were analyzed.

The physiological role of TFB2 from *S. acidocaldarius* is still not known and so far only microarray studies suggest a possible function in cell-cycle regulation (Lundgren and Bernander, 2007). Investigation of the genomic context of *tfb*2 might provide some hints for putative functions. In most Crenarchaeotes the genomic context of *tfb*2 is conserved and it seems to be positioned in an operon. To investigate whether other Archaea share a conserved genomic neighborhood of *tfb*2 the genomic context was analyzed by using the IMG cassette find search and SyntTax and results of these analyses are illustrated in Fig. 16. An overview of additional genomic contexts is shown in Fig. 36 in the supplement and it is necessary to mention that the figure only comprises genes that are part of a conserved genomic neighborhood. Signs behind the species name indicate the presence of *gar*1, *tfb*2 or the gene encoding the RPC34 homolog elsewhere in the genome if they do not belong to the conserved genomic context. In case of the order Halobacteriales it was not possible to precisely distinguish which of the TFB proteins corresponds to TFB2 from *S. acidocaldarius*. Halobacteriales belong to the Euryarchaeota and possess more TFB proteins than *S. acidocaldarius* and additional copies of TFBs. Therefore it was problematic to identify *tfb*2 homologs by sequence comparison. Blast searches with the protein sequence of TFB2 from *S. acidocaldarius* as input data resulted in many hits with E-values ranging from 6e-57 and 4e-42 and identities between 32% and 37% making it difficult to determine a precise homolog of TFB2 in Halobacteriales.

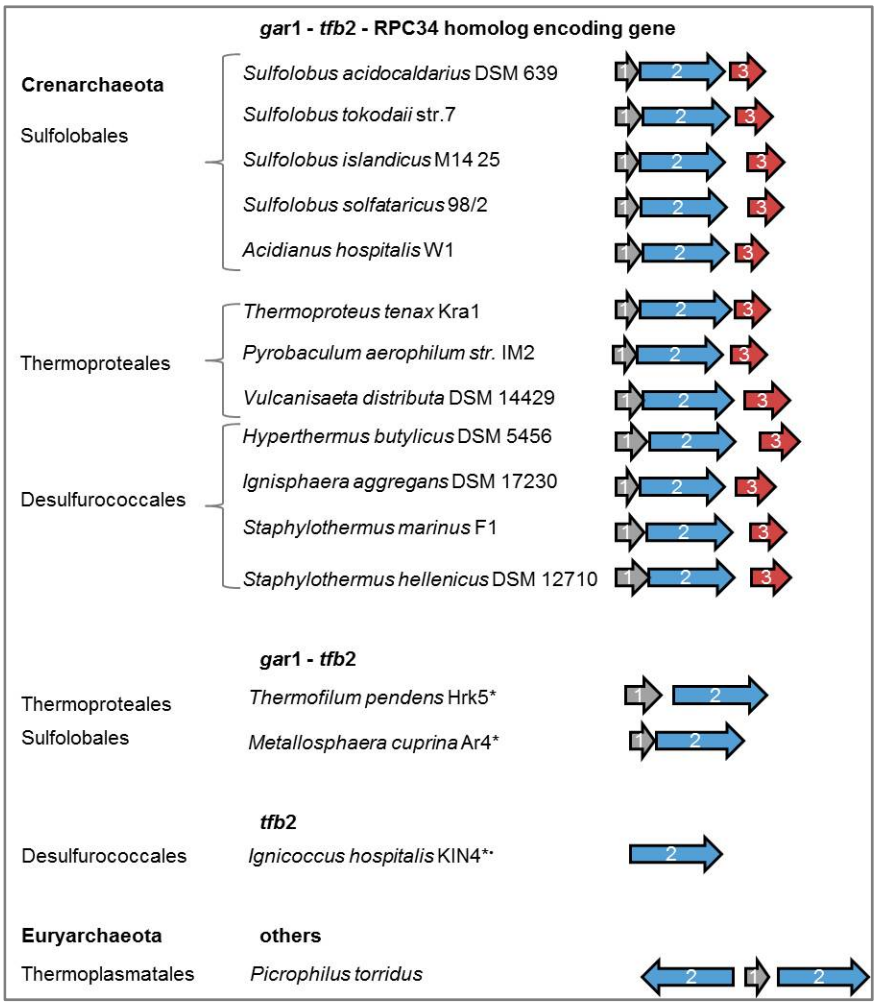


Fig. 16: Conserved genomic neighborhood of *gar*1, *tfb*2 and the RPC34 homolog encoding gene.
The amino acid sequences of Gar1 (Saci_1340), TFB2 (Saci_1341) and the RPC34 homolog (Saci_1342) were analyzed by the program SyntTax (http://archaea.u-psud.fr/synttax/) for analysis of the genomic context. 1, H/ACA ribonucleoprotein *gar*1; 2, transcription factor *tfb*2; 3, RPC34 homolog encoding gene. An asterisk behind the species name indicates the presence of a gene encoding the RPC34 homolog and a point represents a Gar1 encoding gene elsewhere in the genome.

In *S. acidocaldarius* and some other organisms the H/ACA-ribonucleoprotein encoding gene *gar*1 which is involved in RNA modification (Baker et al., 2005, Reichow et al., 2007, Gurha et al., 2007, Hamma and Ferré-D´Amaré, 2010, Kiss et al., 2010, Kamalampeta and Kothe, 2012, Watkins and Bohnsack, 2012), is positioned upstream of the *tfb*2 gene. The gene encoding the homolog of the eukaryotic RPC34 subunit of the RNAPIII, probably *tfe*-β (Blombach et al., 2009), is located downstream of *tfb*2 in Crenarchaeotes. This genomic context is conserved in Crenarchaeota in the orders Sulfolobales, Thermoproteales and Desulfurococcales (Fig. 16). In general RPC34 orthologs were identified by sequence comparisons in all sequenced crenarchaeal and thaumarchaeal as well as in some euryarchaeal genomes (e. g. *Archaeoglobus fulgidus* DSM 4304, *Methanosarcina mazei* Gö1, *Haloferax volcanii* DS2) (Blombach et al., 2009). In addition a large group of organisms found in Euryarchaeota as well as Crenarchaeota possess a genomic conservation of *gar*1 and *tfb*2 (Fig. 16). The euryarchaeal organism *Picrophilus torridus* comprises a unique genomic organization of *gar*1 which is flanked by two *tfb*2 genes that are oriented in divergent

directions. In some euryarchaeal strains the genomic context shows a replacement of the gene encoding the RPC34 homolog in the genomic neighborhood. In case of Methanosarcinales a lysE type translocator encoding gene is positioned downstream of *tfb*2, a metallophosphoesterase encoding gene is present in Thermococcales and some Methanomicrobiales encode tRNA Alanine. The genomic context of Methanococcales like *Methanococcus jannaschii* DSM2661 shows a conservation of *gar*1, *tfb*2 and a type II secretion system protein (Fig. 36 supplement). *Pyrococcus horikoshii* OT3, *Pyrococcus furiosus* DSM 3638 and *Ignicoccus hospitalis* KIN4 harbor a single *tfb*2 gene without any conserved genomic neighborhood. The gene encoding the RPC34 homolog can be positioned downstream of *tfb*2 or separated like in *Aeropyrum pernix* K1, one representative of Desulfurococcales, and *Fervidicoccus fontis* Kam940 (Fig. 36 supplement). In the Thaumarchaeotes *Nitrososphaera gargensis* Ga9 and *Nitrosopumilus maritimus* SCM1 a methyltransferase is found downstream of the gene encoding the RPC34 homolog. In Halobacteriales, like *Haloferax volcanii* DS2, the *gar*1 gene is embedded between a signal recognition particle S19 encoding gene and a hypothetical gene of unknown function.

The start and stop codon of *gar*1 and *tfb*2 overlap by the nucleotides ATGA in most organisms possessing the conserved genomic neighborhood of *gar*1, *tfb*2 and the gene encoding the RPC34 homolog. Whereas in some organisms like *Vulcanisaeta distributa* (ATGGTCACGGCAAATAA) or *S. tokodaii* str. 7 (ATGGAGAAAATAA) the sequences overlap by more nucleotides. In few organisms like *Ignispheara aggregans* both genes, *gar*1 and *tfb*2, are separated by seven nucleotides (TTTTCCT).

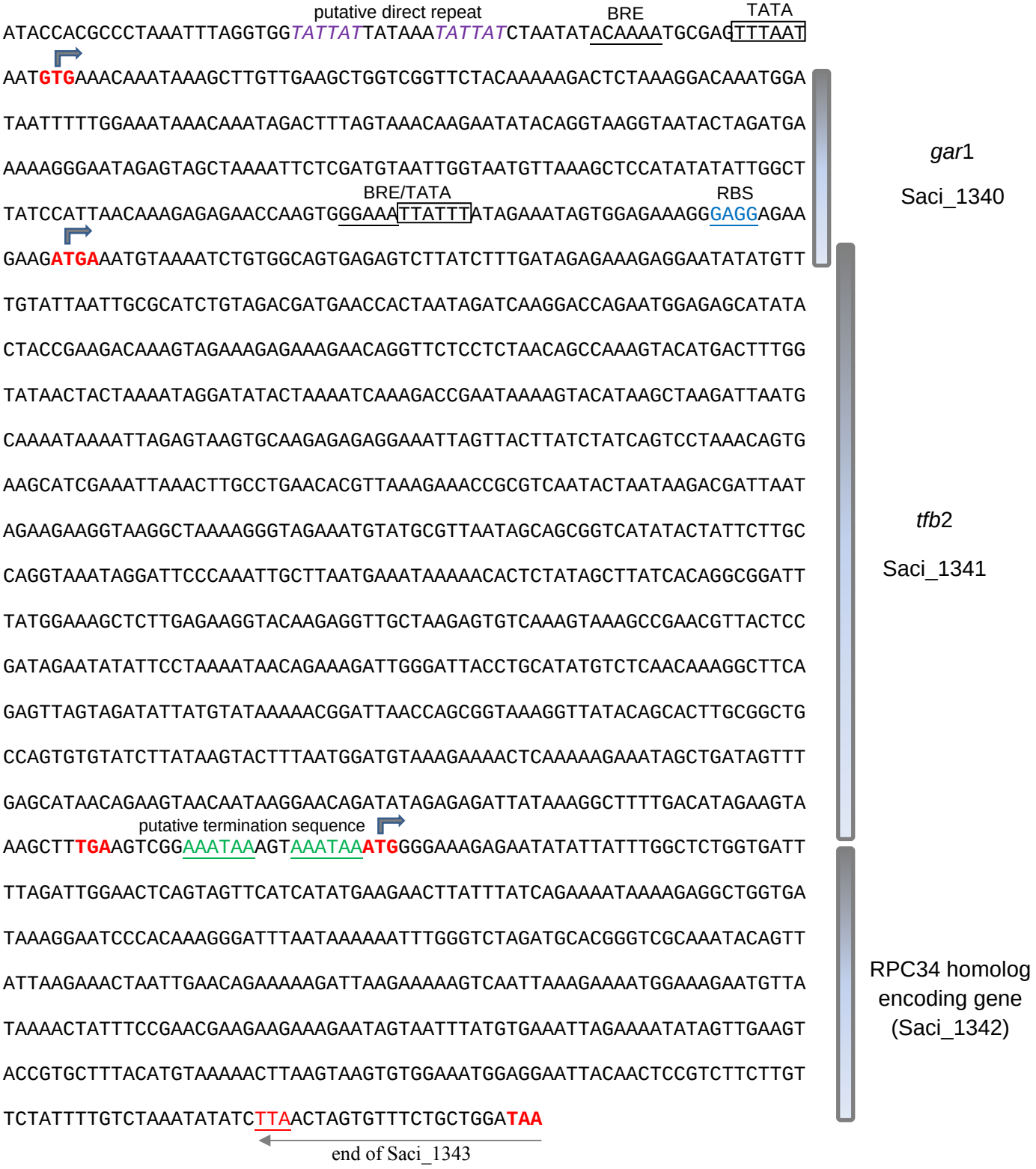


Fig. 17: Analysis of promoter elements, start and stop codons, direct repeats, termination sites and ribosomal binding sites of the *S. acidocaldarius gar*1, *tfb*2 and Saci_1342 gene cluster.
Italic and purple, a direct repeat (TATTAT) probably serving as upstream activation site; underlined, TFB responsive element (BRE); black box, TATA-box; red, start and stop codons; blue arrows, putative transcription start sites; blue and underlined, ribosomal binding site (RBS); green and underlined, putative AT-rich termination signal; red and underlined: inverted sequence of the stop-codon of Saci_1343.

The start codon of *S. acidocaldarius tfb*2 (Saci_1341) overlaps with the stop codon of *gar*1 (Saci_1340) by four nucleotides (Fig. 17). Twenty-one nucleotides are positioned between *tfb*2 and the RPC34 homolog encoding gene (Saci_1342). The upstream sequences of *gar*1 and *tfb*2 contain a BRE-site and a TATA-box (Fig. 17). In case of Saci_1342 no obvious promoter element could be found matching conserved promoter sequences. Interestingly, *tfb*2 contains a ribosomal binding site (RBS) with the consensus 5'-GAGG-3' that is typical for genes that are positioned in an operon (Tolstrup et al., 2000, Slupska et al., 2001, Ma et al., 2002, Toranisson et al., 2005, Chang et al., 2006). A possible AT-rich termination sequence was identified in the region between *tfb*2 and Saci_1342 (AAATAA, Fig.17) that might serve as transcription termination signal. A direct repeat (TATTAT) was detected upstream of the *gar*1 start codon which might serve as regulatory DNA element like upstream activation site or silencing site.

A sequence comparison of the -1 to -80 bp regions of the *gar*1 and *tfb*2 genes from fourteen crenarchaeal organisms (List of organisms with genomic context *gar*1-*tfb*2-Saci_1342 from Fig. 17, Fig. 36) was used to derive a DNA sequence motif using the program WebLogo. The height of the letters depicts their relative frequency in the aligned sequences (Fig. 18). The unit bits reflect the information content. In case of *gar*1 promoters putative BRE and TATA-box consensus sequences as well as two direct repeats present upstream of the core promoter elements could be derived (Fig. 18). For *tfb*2 promoters putative consensus motif for BRE and TATA-box were identified (Fig. 18). In addition a highly conserved clustering of G and A nucleotides directly upstream of the start codon was observed (Fig. 18). This sequence motif seems to represent the ribosomal binding site (RBS).

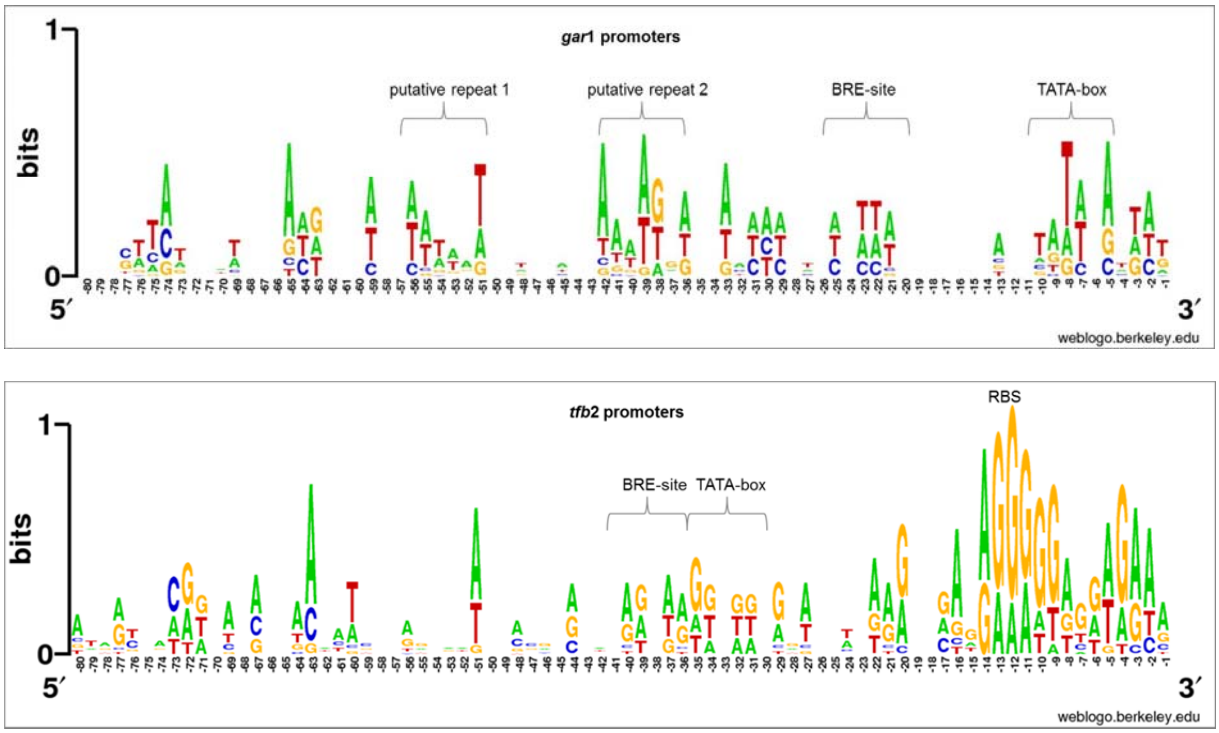


Fig. 18: Sequence logos of the -1 to -80 bp regions of aligned *gar*1 and *tfb*2 sequences from fourteen crenarchaeal organisms.
The sequence logos were generated using the program WebLogo (http://weblogo.berkeley.edu/logo.cgi). -1 to -80 bp upstream sequences of *gar*1 and *tfb*2 from *Acidianus hospitalis* (Ahos_1394, Ahos_1393), *S. acidocaldarius* (Saci_1340, Saci_1341), *S. tokodaii str.* 7 (St1272, St1271), *S. islandicus* M14 (M1425_1280, M1425_1281), *S. solfataricus* 98/2 (SSOl_1922, SSOl_1921), *Thermoproteus tenax* (TTX_2086, TTX_2085), *Thermoproteus neutrophilus* (Tneu_1564, Tneu_1563), *Pyrobvaculum aerophilum* IM2 (PAE3330, PAE3329), *Vulcanisaeta distribute* (Vdis_0754, Vdis_0755), *Vulcanisaeta moutnovskia* (Vmut_1604, Vmut_1605), *Hyperthermus butylicus* (Hbut_0457, Hbut_0458), *Ignisphaera aggregans* (Igag_0016, Igag_0017), *Staphylothermus marinus* (Smar_1249, Smar_1250), *Staphylothermus hellenicus* (Shell_1204,Shell_1203) were used as input data for generation of this logo.

3.2.2. Experimental analysis of genome organization of *tfb*2

S. acidocaldarius DSM 639 wild type and *S. acidocaldarius* MW001 (Δ*pyrEF*) were cultured to early-log, mid-log and stationary phase. RNA was extracted from these samples and transcribed into cDNA using reverse transcriptase. These cDNA samples were used as templates in PCR reactions with different primer combinations. These experiments should elucidate if *gar*1, *tfb*2 and Saci_1342 are co-transcribed and form an operon at the tested growth conditions.

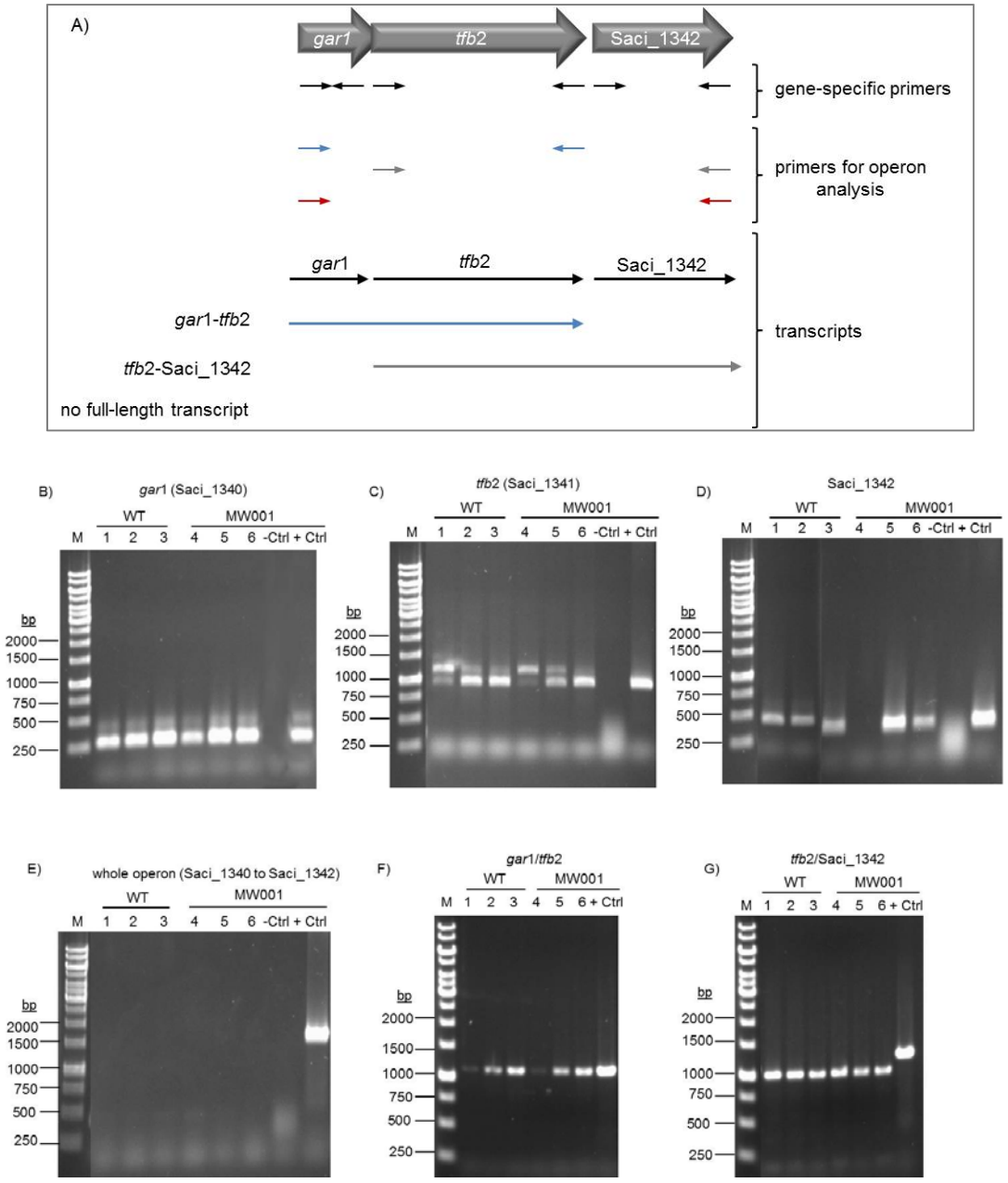


Fig. 19: Gene cluster analysis of the *gar*1 (Saci_1340), *tfb*2 (Saci_1341) and RPC34 homolog encoding gene (Saci_1342) by cDNA amplification at different growth phases.
Total RNA was extracted from *S. acidocaldarius* DSM 639 and *S. acidocaldarius* MW001 cells grown to three different growth phases. Isolated samples were treated with DNaseI and cDNA was generated from RNA using reverse transcriptase. The cDNA was used for PCR amplification with different primer combinations.
A schematic overview of genome organization of *tfb*2 in *S. acidocaldarius*, primers used for PCR and amplified transcripts (A). PCRs using gene-specific primers for *gar*1 (B), *tfb*2 (C), Saci_1342 (D), the whole gene cluster of *gar*1-Saci_1342 (E), *gar*1-*tfb*2 (F) and *tfb*2-Saci_1342 (G) were performed. RNA of *S. acidocaldarius* DSM639 grown to early-log phase (1), log phase (2) and stationary phase (3), RNA of *S. acidocaldarius* MW001 grown to early-log phase (4), log phase (5) and stationary phase (6), -Ctrl: RNA of *S. acidocaldarius* MW001, +Ctrl: DNA of *S. acidocaldarius* MW001. Samples were loaded onto an 1% (w/v) agarose gel for separation and DNA was visualized by ethidiumbromide. M: GeneRuler™ 1 kb DNA Ladder (Fermentas).

All three genes, *gar*1, *tfb*2 and Saci_1342 were transcribed at the tested conditions using primers for single gene amplification (Fig. 19, Table 2). In all PCR reactions with *gar*1

primers a band at 279 bp was visible for all tested growth-phases. The positive control reaction with genomic DNA revealed a PCR product at the same size as the samples. No PCR product was obtained with RNA from *S. acidocaldarius* MW001 as negative control confirming successful DNA digestion. The PCR product of *tfb*2 has a theoretical size of 876 bp and some samples show two adjacent bands probably derived from unspecific primer binding (Fig. 19 C). The amplification product of Saci_1342 gene product had a size of 417 bp and is found in all reactions except in sample 4 and the lane of the negative control (Fig. 19 D). To check whether *gar*1, *tfb*2 and Saci_1342 form an operon, various primer combinations were tested. As shown in Fig. 19 E no-full-length transcript of theoretical size of 1566 bp occurred. Interestingly, a PCR with primer combinations of Saci_1340 and Saci_1341 (1145 bp) as well as Saci_1341 and Saci_1342 (1314 bp) revealed that *gar*1-*tfb*2 and *tfb*2-Saci_1342 were co-transcribed (Fig. 19 F, G). In case of the *gar*1-*tfb*2 transcript an increasing intensity of PCR bands was observed from early-log phase towards exponential growth phase and stationary phase in both strains (Fig. 19 F). The transcript of *tfb*2-Saci_1342 was present at every tested growth condition (Fig. 19 G). These experimental data confirm that the *tfb*2 gene is indeed embedded in two operons and is co-transcribed with *gar*1 or with Saci_1342. Surprisingly, *tfb*2 also possesses typical archaeal promoter elements (BRE/TATA) in addition to a Shine-Dalgarno motif (Fig. 18).

3.3. Reporter gene assays

Transcription factors TFB2 and TFB3 could neither be expressed via homologous expression nor purified from *S. acidocaldarius* after integration of a Strep-Flag tag encoding sequence under standard growth conditions. This experiment will help to identify growth conditions at which the promoters of *tfb*2 and *tfb*3 and other GTFs (*tfb*1, *tbp*, *tfe*) as well as *gar*1 are active and respond to selected stress conditions. The LacS encoding sequence from *S. solfataricus* was cloned under the control of six different promoter sequences of *tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe* and *gar*1 into the plasmid pSVA1450.

3.3.1. Analysis of promoter elements

Upstream regions (-81 to +16 bp) of the transcription factors (*tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe*) and *gar*1 were aligned using the program ClustalW and analyzed for the presence of promoter elements. These results are shown in Fig. 20.

Fig. 20: Comparison of upstream sequences of transcription factors *tfb*1, *tfb*2, *tfb*3. *tbp*, *tfe* and *gar*1. A ClustalW alignment from - 81 to + 16 bp regions is shown.
Underlined, TFB-responsive element (BRE); black box: TATA-box, blue, start codon; asterisks, identical nucleotides; red, direct repeats; green, inverted repetitive sequences.

Upstream sequences of the transcription factors *tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe* and *gar*1 were analyzed for typical promoter elements using the consensus sequences of crenarchaeal BRE-site and TATA-box (RNWAWW and TTTAWA, respectively (R: any, W: A/T)) (Reiter et al., 1990, Hausner et al., 1991, Qureshi and Jackson, 1998, Bell et al., 1999a, Soppa et al., 1999b, Qureshi 2006). No match for the BRE consensus sequences could be identified in case of *tfb*1 and *tbp* and the TATA-element sequence of *tfe* seems to be more diverse. As indicated in Fig. 20 repetitive sequences, direct or inverted repeats, were found upstream of the promoter elements of *tfb*1 (CTTAA, direct repeat), *tfb*2 (AAGA inverted repeat), *tfb*3 (ATCTA, direct repeat), *tbp* (TATA, inverted repeat) and *gar*1 (TATTAT, direct repeat). These sequences might serve as additional regulatory DNA elements like upstream activation sites or inhibitory elements as they were found in other Archaea (Cohen-Kupiec et al., 1997, Gregor and Pfeifer, 2005, Bauer et al., 2008, Peng et al., 2009).

For generation of reporter gene constructs promoter regions of *tfb*1, *tfb*3, *tbp*, *tfe* and *gar*1 (303 bp) and *tfb*2 (269 bp) containing core promoter elements and upstream sequences were amplified from genomic DNA. The PCR products were cloned into pSVA1450. The *mal* promoter of pSVA1450 was thereby replaced by the promoters of the transcription factors. The methylated plasmids were introduced into *S. acidocaldarius* MW001 cells by electroporation and these cells were spread on first selection plates. Colonies were picked and transferred to liquid first selection medium. DNA was extracted from these cultures and the presence of the plasmid was confirmed by PCR using plasmid specific primers.

3.3.2. Determination of β-galactosidase activity

Expression of the reporter gene encoding LacS is controlled by promoters of the GTFs or *gar*1. As initial experiment it was tested if the promoter constructs of *tfb*1, *gar*1 and *tfb*3 influence the expression of the reporter gene and if β-galactosidase activity could be detected. For reporter gene assays cultures were grown to an OD_{600} of about 0.6 and were exposed to cold-shock at 60°C. Ten ml of cell suspension were harvested and the pellet was resuspended in 1x Z buffer (40 mM NaH_2PO_4, 60 mM Na_2HPO_4, 10 mM KCl, $MgSO_4*7H_2O$ 1 mM, 0.5% (v/v) Triton X-100, PMSF and protease inhibitor (Roche), pH 7.0). The samples were

disrupted by sonication and centrifuged to remove cell debris (described in sections 2.15.2-2.15.3). The cell-free extracts were used for enzyme assays and the β-galactosidase activity was determined at 75°C by conversion of ONPG (410 nm). Initial experiments with *tfb*1, *gar*1 and *tfb*3 promoters revealed β-galactosidase activity and confirmed that sample preparation and assay conditions were functional. The colorless substrate ONPG was converted to galactose and the yellow product ortho-nitrophenol (Fig. 21).

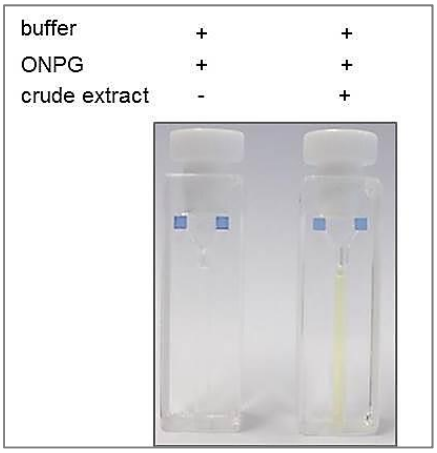


Fig. 21: Example of reporter gene assay before and after addition of crude extract and incubation at 75°C. Crude extracts of cold shocked *S. acidocaldarius* MW001 cells harboring the plasmid pSVA1450-*tfb*1 promoter were prepared by sonication and centrifugation. For determination of β-galactosidase activity, buffer and the substrate ONPG were pre-incubated at 75°C in the photometer. When the photometric assay at 410 nm showed a constant baseline the reaction was started by addition of crude extract. In this figure the reaction mixture containing buffer and substrate is shown before (-) and after addition of crude extract (+).

Preliminary experiments with *tfb*1, *gar*1 and *tfb*3 promoter regions fused to the *lac*S gene revealed conversion of ONPG to the yellow substrate orthonitrophenol in all three samples (Fig. 21, Fig. 22). *S. acidocaldarius* MW001 cells harboring the plasmids pSVA1450 with promoter regions of *tfb*1, *gar*1 and *tfb*3 were grown at standard conditions until reaching an OD_{600} of 0.6. Afterwards cells were collected and crude extracts from these cells were analyzed for β-galactosidase activity. The highest reporter gene expression, i.e. β-galactosidase activity, was obtained with the *tfb*1 promoter (18,205 ± 2370 mod. Miller Units) and decreased for the *gar*1 promoter (9375 ± 1089 mod. Miller units) and the *tfb*3 promoter (4366 ± 185 mod. Miller units, 0 min) (Fig. 22). After cold shock at 60°C the β-galactosidase activity under control of *tfb*1 promoter increased and reached its highest activity after 60 min (26,830 ± 220 mod. Miller units) and afterwards the activity declined again (20,348 ± 215.4 mod. Miller Units, 120 min). In case of the *gar*1 promoter the highest activity was reached after 90 min (16,567 ± 914 mod Miller units) and the lowest activity was determined at 240 min (11,500 ± 0 mod. Miller units) (Fig. 22). The β-galactosidase activity under control of the *tfb*3 promoter seemed not to be affected by cold shock. Taken together the results indicate that the *tfb*1 gene might play a role at standard growth conditions and under cold-shock response.

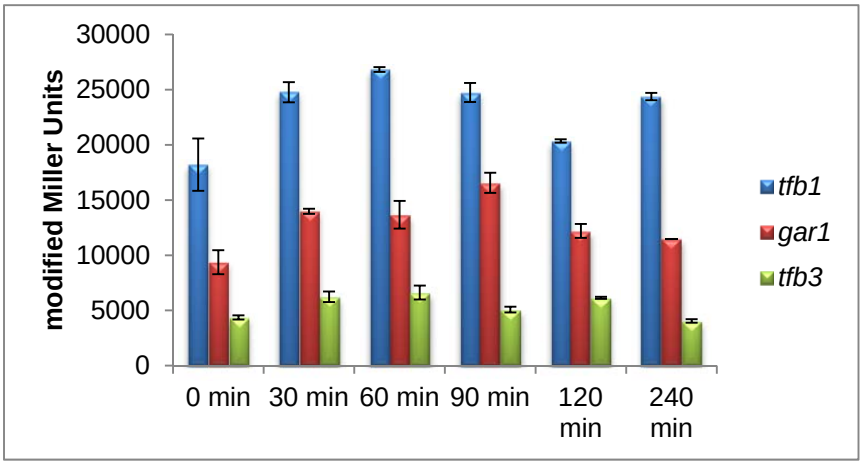


Fig. 22: Effect of cold-shock on *tfb*1, *gar*1 and *tfb*3 promoters analyzed by reporter gene assays.
S. acidocaldarius MW001 strains comprising the *lacS* gene under control of the promoters of *tfb*1, *gar*1 and *tfb*3 were grown in first selection medium to an OD_{600} of 0.6 at 78°C and afterwards the cultures were transferred to 60°C (cold-shock) in an incubator with shaking. Samples of 10 ml were directly removed (0 min) and collected every 30 min until 240 min and samples were stored at -80°C. Cell-free extracts were analyzed for β-galactosidase activity and the conversion of ONPG to orthonitrophenol was monitored at 75°C in a spectrophotometer (410 nm).

In future experiments stress response of promoters of the transcription factors (*tfb*1, *tfb*2, *tfb*3, *tbp*, *tfe*) and *gar*1 of *S. acidocaldarius* will be analyzed by determination of the β-galactosidase activity. Reporter gene assays combined with qRT-PCR in respect to stress response will be a powerful tool for investigation of the role of the transcription factors and especially of *tfb*2.

3.4. Protein-protein interaction studies of RNAP and GTFs from *S. acidocaldarius* assayed in the yeast two-hybrid system

Archaea use one multi-subunit RNAP for transcription which is similar to the eukaryal RNAP II (Zillig et al., 1979, Langer, 1995, Werner, 2007, Werner and Grohmann, 2011). A comparison of RNAPs of Archaea, Bacteria and RNAPII of Eukaryotes can be seen in Fig. 1. All archaeal RNAP subunits except RpoM possess a homologous eukaryotic RNAPII subunit. In case of eukaryotic RPB9 no archaeal or bacterial counterpart is present. Archaeal RpoA, RpoB, RpoD and RpoK are homologous to bacterial β', β, α and ω, respectively. A homolog of RpoM could not be found in any eukaryotic RNAP making this subunit unique in the archaeal lineage and this subunit has been identified among Sulfolobales and Desulfurococcales (Korkhin et al., 2009). By studying the crystal structures of RNAP from *S. shibatae* it can be speculated that Rpo13 could play a role in transcription initiation and elongation by forming the transcription bubble (Korkhin et al., 2009). It has been hypothesized that this subunit was introduced into the RNAP as an evolutionary benefit (Jun et al., 2011).

The close homology of archaeal and eukaryotic RNAPs further supports the hypothesis that both derived from a less complex ancestor before the archaeal and eukaryotic lineages diverged.

The RNAP of *S. acidocaldarius* consists of twelve subunits and during transcription initiation TFB recruits and docks to the RNAP with its N- terminal Zn-ribbon domain. To determine the binding partners, protein-protein interaction studies with all twelve RNAP subunits and transcription factors TFB1, TFB2 and TFB3 from *S. acidocaldarius* were performed. In addition it was tested if all three TFBs bind to TBP because TBP-independent binding of *T. tenax* TFB2 has been observed in EMSA-studies (Marrero et al., 2013 in prep.). Interaction of TBP with RNAP subunits was not tested because the archaeal RNAP lacks the CTD domain which contacts TBP in Eukaryotes. Accordingly, TBP-binding to the RNAP in Archaea should not be possible and previous experiments with *Sso*TBP and RNAP subunits revealed no interaction (Magill et al., 2001).

3.4.1. Interaction studies of RNAP subunits and TFBs from *S. acidocaldarius*

After co-transformation of the prey vector containing the RNAP subunits (pGADT7::*rpo* subunits) and the bait vector encoding the TFBs from *S. acidocaldarius* (pGBKT7::*tfb*1, pGBKT7::*tfb*2, pGBKT7::*tfb*3) into the yeast strain *S. cerevisiae* AH109, cells were plated on SD -Leu/-Trp plates to ensure the success of transformation by first selection. After incubation for three days at 30°C, liquid SD -Leu/-Trp was inoculated with one colony or in case of tiny colonies with two to three colonies and incubated over night at 30°C. After reaching an OD_{600} of 1.5 the cell suspension was diluted to an OD_{600} of 0.1 and 5 µl were spotted onto selective SD -Leu/-Trp/-Ade/-His/+Xα-Gal plates as described by Paytubi and White (2009) (sections 2.16.4, 2.16.5). An interaction of two proteins leads to expression of the reporter genes *ADE*1, *HIS3* and *MEL*1. Yeast clones with interacting partner proteins grow on the selective plates and turn blue because of the α-galactosidase activity encoded by *MEL*1. As shown in Fig. 23 an interaction between TFB2 with RpoE and TFB3 with RpoE and RpoK was visible. RpoE and RpoK are auxiliary proteins in the RNAP and RpoE is part of the protruding stalk that differentiates the archaeal RNAP from the bacterial RNAP. Unexpectedly, TFB1 showed no interaction with any of the RNAP subunits (Fig. 23). Table 16 gives an overview of the yeast two-hybrid results and demonstrates the function of the RNAP subunits.

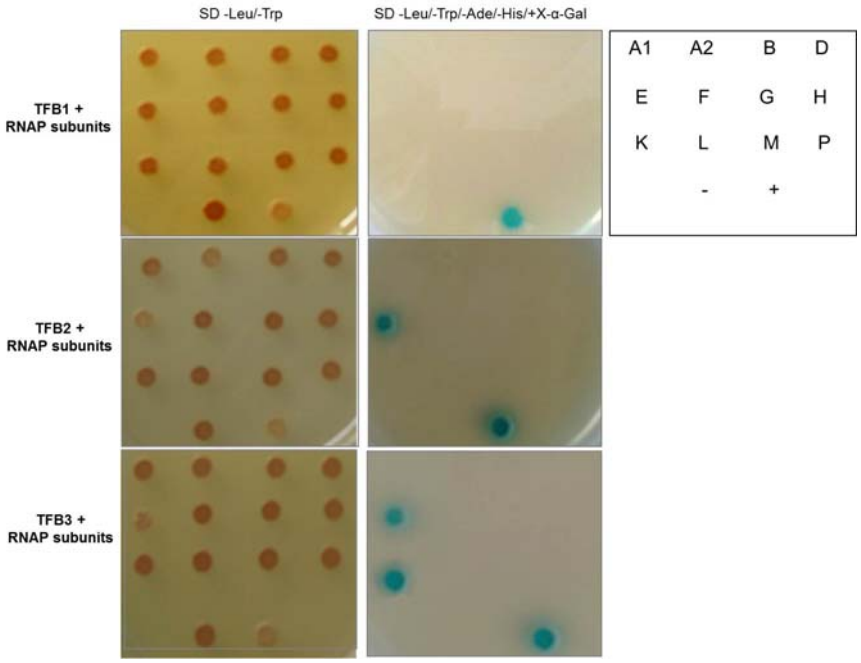


Fig. 23: Yeast two-hybrid interaction studies of *S. acidocaldarius* RNAP subunits with GTFs from *S. acidocaldarius*.
Yeast cells were transformed with plasmids containing the RNAP subunit and transcription factors in different combinations as indicated in the scheme. Samples were incubated on SD -Leu/-Trp medium and interacting medium SD -Leu/-Trp/-Ade/-His/+X-α-Gal for 3 days at 30°C and the results (growth, α-galactosidase activity) were documented. As negative control empty pGADT7 and pGBKT7 (-) were co-transformed into yeast cells. Cells containing pGADT7::Sso*rpo*K and pGBKT7::Sso*tfb*3 served as positive control (+) (Paytubi and White, 2009).

Table 16: Interaction of transcription factors TFB1, TFB2 and TFB3 with RNAP subunits. The function of the different RNAP subunits is indicated.

RNAP subunits (rpo)	Function of RNAP subunit	TFB1	TFB2	TFB3
A1	catalysis	-	-	-
A2	catalysis	-	-	-
B	catalysis	-	-	-
D	assembly	-	-	-
E	auxiliary	-	+	+
F	auxiliary	-	-	-
G	auxiliary	-	-	-
H	auxiliary	-	-	-
K	auxiliary	-	-	+
L	assembly	-	-	-
M	assembly	-	-	-
P	assembly	-	-	-

-: no interaction, +: interaction

3.4.2. Interaction studies of GTFs from *S. acidocaldarius*

The bait vectors containing the TFBs from *S. acidocaldarius* (pGBKT7::*tfb*1, pGBKT7::*tfb*2, pGBKT7::*tfb*3) were co-transformed with the prey vector encoding TBP (pGADT7::*tbp*) into competent *S. cerevisiae* AH109 cells. Cell growth and interaction studies were performed as described before in 3.4.1. The yeast two-hybrid experiment revealed a weak interaction between TFB3 and TBP (Fig. 24). No interaction could be observed for TFB1 and TFB2 with TBP (Fig. 24). TFB1 is the transcription factor that is present under normal growth conditions and is known to interact with TBP (Bell and Jackson, 2000a, Soppa, 2001, Hickey et al., 2002, Geiduscheck and Ouhammouch, 2005). An interaction between TBP and TFB1 might only occur in order to build the ternary complex composed of TBP, DNA and TFB. Regarding TFB2, EMSA experiments revealed that *T. tenax* TFB2 binds to DNA in a TBP-independent manner (Marrero et al., 2013, in prep.) it might be possible that also *S. acidocaldarius* TFB2 binds to DNA without TBP.

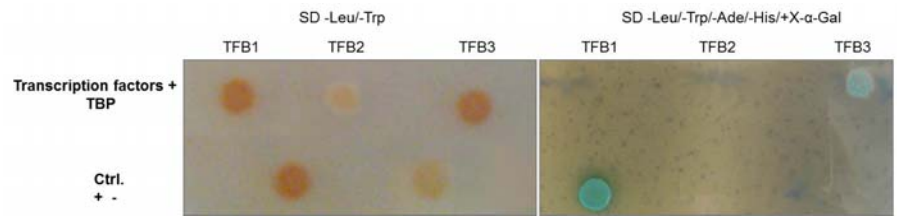


Fig. 24: Interaction studies of transcription factors TFB1, TFB2 and TFB3 with TBP from *S. acidocaldarius.*
Competent *S. cerevisiae* AH109 cells were co-transformed with the prey vector pGADT7::*tbp* and the bait vectors pGBKT7::*tfb*1, pGBKT7::*tfb*2, pGBKT7::*tfb*3, respectively. As positive control pGADT7::Sso*rpoK* and pGBKT7::Sso*tfb*3 were co-transformed into yeast cells and empty vectors pGADT7 and pGBKT7 served as negative control. Samples were selected on SD -Leu/-Trp and SD -Leu/-Trp/-Ade/-His/+X-α-Gal-medium.

To ensure that the proteins fused to the activating domain, encoded in the prey vector pGADT7, do not bind to the DNA without a binding domain fusion partner, empty bait vector pGBKT7 was co-transformed with pGADT7::*rpoE*, pGADT7::*rpoK*, pGADT7::*tbp* respectively. A 1:10 and 1:100 dilution of cells was plated on SD -Leu/-Trp/+Xα-Gal and SD -Leu/-Trp/-Ade/-His/+Xα-Gal agar plates (described in 2.16.6). Only white colonies grew on SD -Leu/-Trp/+Xα-Gal agar demonstrating that there was no expression of reporter genes due to weak unspecific binding (Fig. 25). On the higher selective stringency medium no colonies grew because only strong interactions between proteins lead to expression of all reporter genes (Fig. 25). These results confirm that the tested interactions were true positives and that two interacting partner proteins need to be present for sufficient expression of the reporter genes in the yeast two-hybrid system.

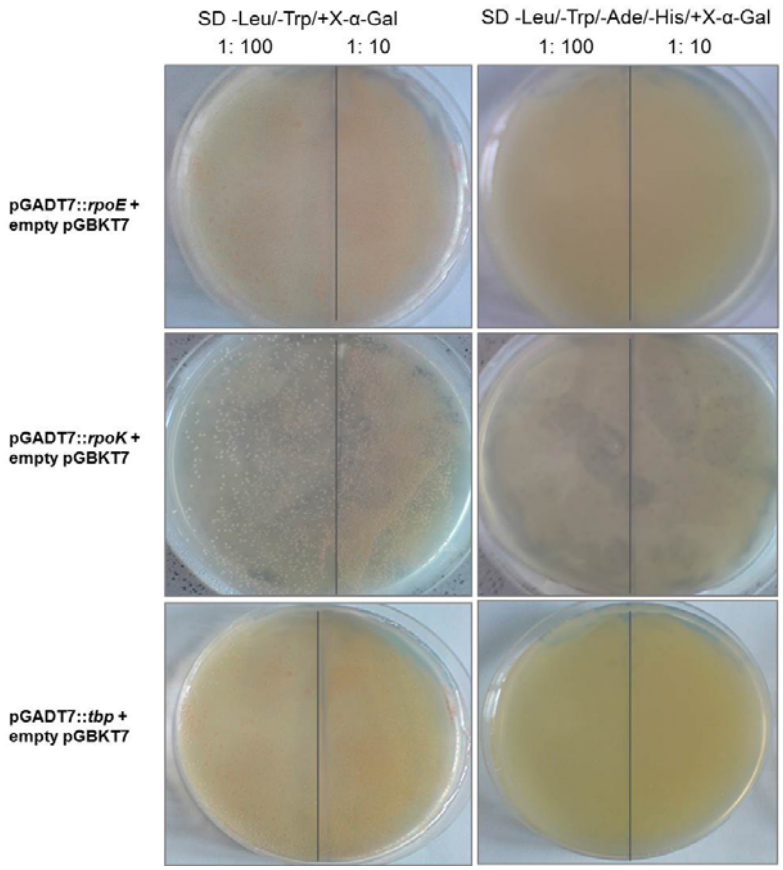


Fig. 25: Analysis for false positive reactions by co-transformation of pGADT7::*rpoE*, pGADT7::*rpoK* and pGADT7::*tbp* with empty pGBKT7.
Competent yeast cells were transformed with empty bait vector pGBKT7 and pGADT7::*rpoE*, pGADT7::*rpoK* or pGADT7::*tbp* and plated on SD -Leu/-Trp medium. A 1:10 and 1:100 dilution from over night cultures were spread onto SD -Leu/-Trp/+X-α-Gal and SD -Leu/-Trp/-Ade/-His/+X-α-Gal plates and incubated for 3 days at 30°C.

In order to proof expression of *S. acidocaldarius* proteins in yeast, immunoblots with anti-c-myc antibody and anti-HA-tag antibody were performed. Various protocols for yeast cell extract preparation were tried like the TCA method, urea/SDS method (described in 2.16.7-2.16.9) or just disruption of the cells by adding SDS page loading buffer and subsequent boiling of the cells. Also different western blot membranes like PVDF and nitrocellulose (0.2 µm or 0.45 µm pore size) and different antibodies like monoclonal anti-c-myc or HA-tag antibody linked with AP (abcam, host: mouse) or a HRP-conjugated secondary anti-mouse antibody (Sigma-Aldrich) were used. In addition several immunoblotting conditions were tested e. g. incubation of the antibodies at RT for one or two hours or incubation over night at 4°C, different amounts of blocking reagent (1% - 3% (w/v) milk powder in antibody solution) or different dilutions of antibodies (1:500, 1:1,000, 1:10,000). Despite intense efforts only RpoA2 and RpoB could be detected (Fig. 26). Even expression of the already published controls *Sso*RpoK and *Sso*TFB3 could not be verified (Magill et al., 2001, Paytubi and White, 2009). With the TCA and urea/SDS method soluble as well as membrane-bound proteins should be detectable. Therefore, even if some Rpo subunits or transcription factors were insoluble after expression in yeast they should be detectable with this method. Taken together

the yeast two-hybrid experiment revealed interactions between RNAP subunits and GTF (Fig. 23, Fig. 24) but expression of these proteins could not be proven. Therefore future studies are required to confirm protein-protein interaction studies.

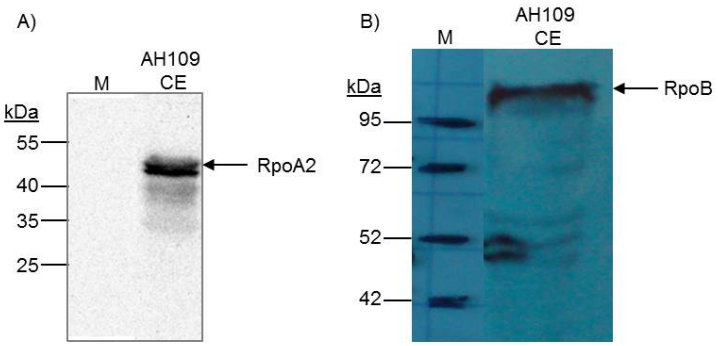


Fig. 26: Analysis of expression of RpoA2 and RpoB from *S. acidocaldarius* in yeast.
Yeast cell extracts of strains transformed with pGADT7::*rpoA*2 and pGADT7::*rpoB* were prepared according to the TCA method described in 2.16.9. 20 µl of the protein extract was separated on a 12.5% SDS-PAA gel (A) and 10% SDS-PAA gel (B) and blotted onto a PVDF membrane with 0.45 µm pore size over night at 4°C and 25 mA. A) The membrane was blocked for 1 hour in TBST with 5% (w/v) milk powder and afterwards incubated over night at 4°C with the anti-HA-tag antibody linked with AP (1:1,000, abcam). Unbound antibody was removed by washing with TBST and TBS. Detection was performed with CDP-Star® (Sigma-Aldrich). M, PageRuler™ Prestained Protein Ladder (Fermentas); RpoA2, yeast cell extract containing RpoA2 (45 kDa); B) The PVDF membrane with a pore size of 0.45 µm was blocked for 1 hour with 5% (w/v) milk powder. The HA-tag antibody was diluted 1:1,000 (Sigma-Aldrich) in TBST with 3% (w/v) milk powder and incubated for 1 hour at RT. The membrane was washed with TBST and incubated with the secondary anti-mouse antibody linked with HRP for 90 min. Before detection by enhanced chemiluminescent (ECL) substrate (Pierce) the membrane was washed with TBST and TBS. M, PageRuler™ Prestained Protein Ladder (Fermentas); RpoB, yeast cell extract containing RpoB (122 kDa).

3.5. Addressing *in vivo* function of GTFs by construction of deletion strains

For further investigation of the cellular function of TFB1, TFB2 and TFB3 the aim was to generate knock-out mutants using two different established methods (Wagner et al., 2009, Sakofsky et al., 2011). On the one hand the pop-in/pop-out strategy was used to delete the encoding genes by homologous recombination via integration of plasmids (Fig. 27) and on the other hand disruption of the genomic region was accomplished by insertion of the *Sso-pyrEF* marker cassette (Fig. 29). For generation of a markerless knock-out mutant the upstream and downstream flanking regions of the gene of interest were amplified from genomic DNA. These regions had compatible overhangs for a following overlap-extension PCR. The resulting PCR product was restricted and ligated into the *E. coli-Sulfolobus* shuttle vector pSVA406. The plasmids were transformed into competent *E. coli* DH5α cells. After successful sequencing the plasmids were transformed into *E. coli* ER1821 cells for methylation. Finally, the plasmids were transformed into *S. acidocaldarius* MW001 cells and plated on first selection plates lacking uracil. Afterwards, colonies grown on plates lacking uracil were transferred into liquid first selection medium and plated on second selection plates (+ 5-FOA, + uracil). Generation of a knock-out was analyzed by PCR and sequencing. The great advantage of a markerless knock-out is the possibility to complement the absent gene by

transformation of a plasmid overexpressing the gene of interest (e. g. pSVA1450 containing a transcription factor).

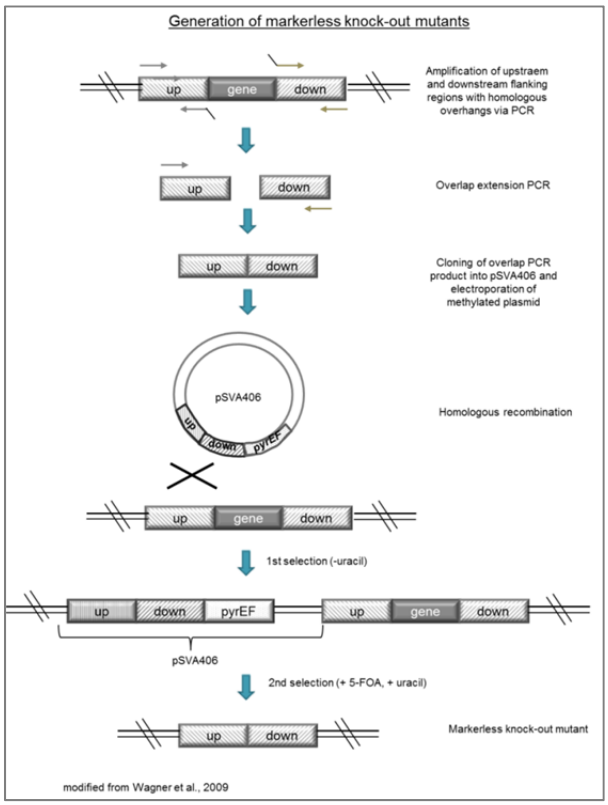


Fig. 27: Generation of markerless knock-out mutants in *S. acidocaldarius* via pop-in/pop-out strategy (mod. from Wagner et al., 2009).
Upstream and downstream flanking regions of the knock-out gene were amplified via PCR using primers to generate compatible overhangs. In an overlap extension PCR the upstream and downstream region **were** combined and digested for cloning into pSVA406. After electroporation of the methylated plasmid into competent *S. acidocaldarius* MW001 cells were spread on first selection plates (- uracil). Colonies grown on these plates were transferred to liquid first selection medium and spread on second selection plates (+5-FOA, + uracil).

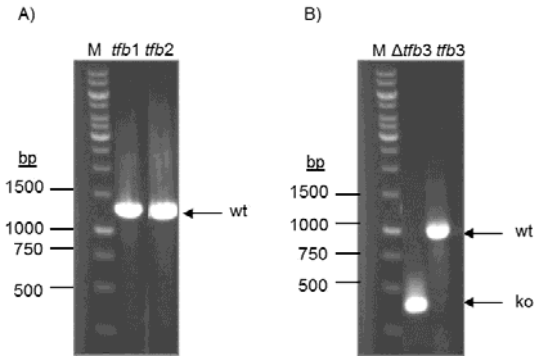


Fig. 28: Analysis of markerless disruption mutants of the *tfb*1, *tfb*2 and *tfb*3 gene by PCR.
Genomic DNA was purified from colonies grown on second selection plates and subsequent PCR was performed with primers annealing up- and downstream of the genes of interest to confirm gene disruption. Primers are listed in Table 9. PCR products were separated on an 1% (w/v) agarose gel and stained with ethidiumbromide. A) PCR products obtained with genomic DNA from the putative *tfb*1 knock-out clone (1338 bp) and putative *tfb*2 knock-out clone (1312 bp) were analyzed alongside with the GeneRuler™ 1 kb DNA Ladder (M, Fermentas). B) PCR product of putative *tfb*3 knock-out clone before (Δ*tfb*3, 330 bp) and after three rounds of selection in second selection medium showing wild-type (wt) pattern (898 bp).

For verification of deletion of the genes of interest by the pop-in/pop-out strategy, PCR was performed with primers annealing to the upstream and downstream flanking regions of the respective genes (Table 9). Fig. 28 illustrates that *tfb*1 and *tfb*2 were still present in the genome. Wild-type PCR products of *tfb*1 and *tfb*2 had sizes of 1338 bp and 1312 bp, respectively. A knock-out of the genes would result in product sizes of 359 bp (*tfb*1) and 439 bp (*tfb*2). A markerless *tfb*3 knock-out mutant (Δ*tfb*3) was generated as indicated by the PCR product of 330 bp obtained from genomic DNA of the deletion strain (Fig. 28). However, after three rounds of selection by growth in second selection medium the wild-type PCR product was present (898 bp) indicating that the knock-out was not stable.

In addition the approach to generate *tfb*1, *tfb*2, *tfb*3, *gar*1, *tfe*, Saci_1342 disruption mutants by insertion of the *pyrEF* cassette was used. In contrast to the markerless knock-out, the selection cassette remains integrated into the genome by the second method illustrated in Fig. 29. For this PCR-based gene disruption method primers contain genomic regions of the gene of interest and parts of the *pyrEF* cassette. The genomic regions and the marker cassette were amplified from pSVA406 DNA. The PCR product is directly introduced into *S. acidocaldarius* MW001 cells by electroporation. To confirm the generation of knock-out mutants colonies grown on first selection plates were analyzed for presence of the *pyrEF* cassette by PCR and sequencing.

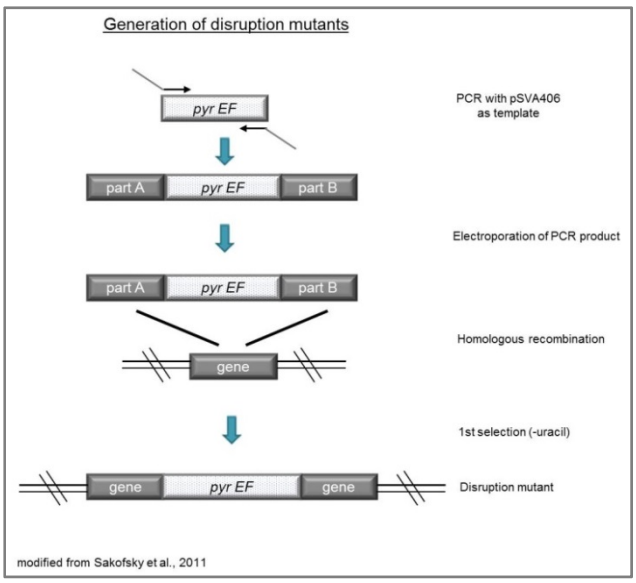


Fig. 29: Generation of disruption mutants in *S. acidocaldarius* by insertion of the *Sso-pyrEF* cassette (mod. from Sakofsky et al., 2011).
The *pyrEF* cassette of *S. solfataricus* was amplified from the plasmid pSVA406 with primers listed in Table 10. The PCR product comprising the *pyrEF* cassette and parts of the gene of interest were introduced into *S. acidocaldarius* MW001 cells via electroporation. The transformed cells were selected by growth on plates lacking uracil to verify the presence of the *pyrEF* cassette which integrates into the genome by homologous recombination.

Several times it was tried to integrate the *pyrEF* cassette into *tfb*1 (Saci_0866), *tfb*2 (Saci_1341), *tfe* (Saci_0652), *gar*1 (Saci_1340) and Saci_1342 genomic sequences by the PCR-based disruption method. Despite enormous efforts by repetition of electroporation and selection knock-out mutants of the aforementioned genes could not be obtained with either of

the methods described before. It is most likely that these proteins are essential for cell viability.

Only for *tfb*3 a successful insertion of the *pyrEF* cassette was achieved (*tfb*3 disruption: 2223 bp) compared to the parent strain *S. acidocaldarius* MW001 (757 bp). Additionally, the PCR product shown in Fig. 30 was sequenced and the insertion of the cassette was confirmed. Following experiments were performed with the disruption mutant (*tfb*3::*pyrEF*). A PCR with genomic DNA from *tfb*3::*pyrEF* strain confirmed that the insertion of the marker cassette was stable (data not shown) after repeated growth (three times) in first selection medium..

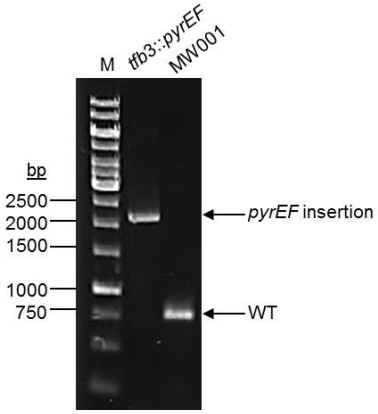


Fig. 30: Analysis of the *S. acidocaldarius* MW001 *tfb*3::*pyrEF* disruption strain.
The *tfb*3::*pyrEF* disruption strain was generated by the PCR-based disruption method and insertion of the *pyrEF* cassette was analyzed by PCR and agarose gel electrophoresis. Genomic DNA was purified from the parent strain MW001 and the *S. acidocaldarius* MW001 *tfb*3::*pyrEF* strain. PCR was performed using primers that annealed upstream and downstream of the insertion position of the selection cassette. The PCR product of the *tfb*3 mutant was excised from the agarose gel and sequenced. PCR products using genomic DNA from *tfb*3::*pyrEF* (2223 bp) and MW001 (757 bp) were separated on a 1% (w/v) agarose gel alongside with GeneRuler™ 1 kb DNA Ladder (M, Fermentas) and DNA was visualized by staining with ethidiumbromide**.**

Because *tfb*3 was shown to be one of the most-upregulated transcripts (7-fold, $\log_2$ expression ratio of 2.82) after UV-exposure in microarray studies (Fröls et al., 2007, Götz et al., 2007), growth of the disruption mutant was compared to the parent strain MW001 after irradiation with different UV-doses (50, 75, 200 J/m^2). Cells of the parent strain MW001 and the *tfb*3::*pyrEF* mutant strain were grown to an OD_{600} of 0.3 at 78°C with shaking. Then the 50 ml cultures were split into half and poured into pre-warmed petri-dishes. The dishes were placed into a Stratalinker® (Stratagene) and irradiated with three defined UV-doses (50, 75, 200 J/m^2). Corresponding cultures were combined again in an Erlenmeyer flask and incubated at room temperature for 15 min in the dark to prevent light-induced repair by photolyase. Afterwards the flasks were incubated in the dark at 78°C with shaking. This procedure has been described by Fröls and colleagues and an overview is given in Fig. 31 (Fröls et al., 2007, Götz et al., 2007). Growth behavior after UV-irradiation was monitored by measuring the optical density at 600 nm. As control both strains were incubated without UV-treatment (Ctrl). As additional control (Ctrl 200) cultures were treated the same way as the UV-irradiated samples but without application of a UV-dose and the incubation time equaled the duration to apply the highest dose of 200 J/m^2. All samples were grown in triplicates and the

experiments were repeated. The arithmetic mean of both experiments was calculated and is shown in Fig. 32.

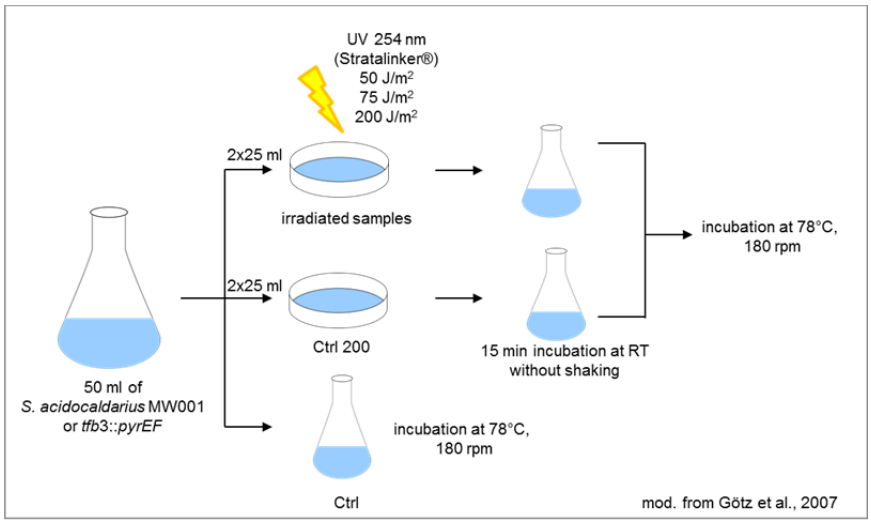


Fig. 31: Description of the UV-irradiation experiment for growth analysis.
Parent and mutant strain cells were grown at 78°C until reaching OD_{600} 0.3. For irradiation the cultures were split into half and each half was poured into pre-warmed petri dishes. Doses of 50, 75 and 200 J/m^2 were applied and the cultures were incubated at RT in the dark for 15 min before they were replaced to 78°C. Ctrl 200 samples were handled according to the irradiated cells but without application of a dose. These cultures were incubated the same time as needed to apply a dose of 200 J/m^2. As additional control (Ctrl) cultures were grown at 78°C for the whole duration of the experiment.

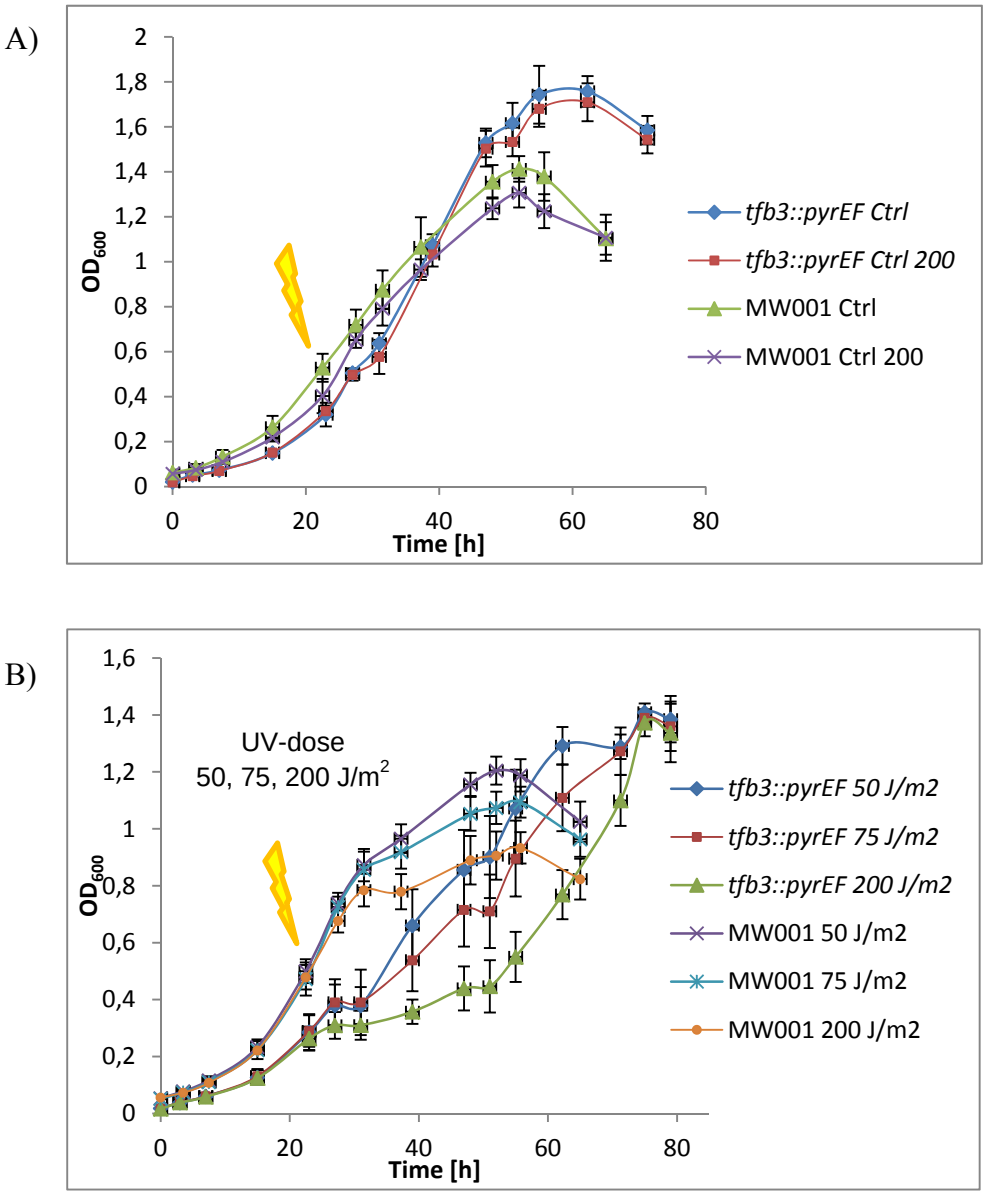


Fig. 32: Effect of *tfb*3 disruption on UV-response of *S. acidocaldarius*.
Cultures of the parent strain *S. acidocaldarius* MW001 and the *tfb*3::*pyrEF* disruption strain were treated as described in Fig. 31. The optical density of control samples (Ctrl, Ctrl 200, A) and irradiated samples (50, 75 and 200 J/m^2, B) was determined at 600 nm. Ctrl: parent strain or *tfb*3::*pyrEF* grown in shaker at 78°C without UV-shock. Ctrl 200 and cultures of the parent strain or *tfb*3::*pyrEF* were treated the same way as samples without application of a dose and incubation time in the Stratalinker® corresponded to the time needed to apply 200 J/m^2.

Both control reactions (MW001 Ctrl, MW001 Ctrl 200) of the parent strain MW001 showed similar growth behavior and reflect that the sample preparation and temperature drop did not interfere with the results obtained after UV-exposure (Fig. 32 A). The control samples of the disruption mutant *tfb*3::*pyrEF* (*tfb*3::*pyrEF* Ctrl, *tfb*3::*pyrEF* Ctrl 200) grew without any obvious difference indicating no influence on the growth behavior by sample preparation or temperature drop (Fig. 32 A). In this experiment the control samples of the *tfb*3::*pyrEF* (Ctrl,

Ctrl 200) disruption mutant reached a higher maximal optical density than all cultures of the parent strain MW001.

Directly after UV-shock, growth of irradiated MW001 cells was not affected by any applied dose (Fig. 32 B). However, first growth defects occurred 15 hours after treatment of the cultures reflected by a slightly reduced growth rate. All irradiated samples of MW001 were affected by the tested UV-doses (50, 75, 200 J/m^2) and a higher dose led to a more severe growth deficiency (Fig. 32 B). All cultures of the parent strain MW001 reached stationary phase 29.5 hours after UV-treatment. Growth defects of the irradiated samples of *tfb*3::*pyrEF* (50, 75, 200 J/m^2) occurred directly after UV-shock with the tendency that a higher dose led to more severe growth defects. The disruption mutant *tfb*3::*pyrEF* showed an adaptation period of at least 16 hours after UV-irradiation at every dose (Fig. 32 B). But the mutant strain could restore its ability to grow 16 hours after UV-shock. A plateau, indicating stationary phase, was reached faster by the non-irradiated control samples of *tfb*3::*pyrEF* (32.5 hours after UV-shock) rather than by the irradiated cultures (52.5 hours after treatment).

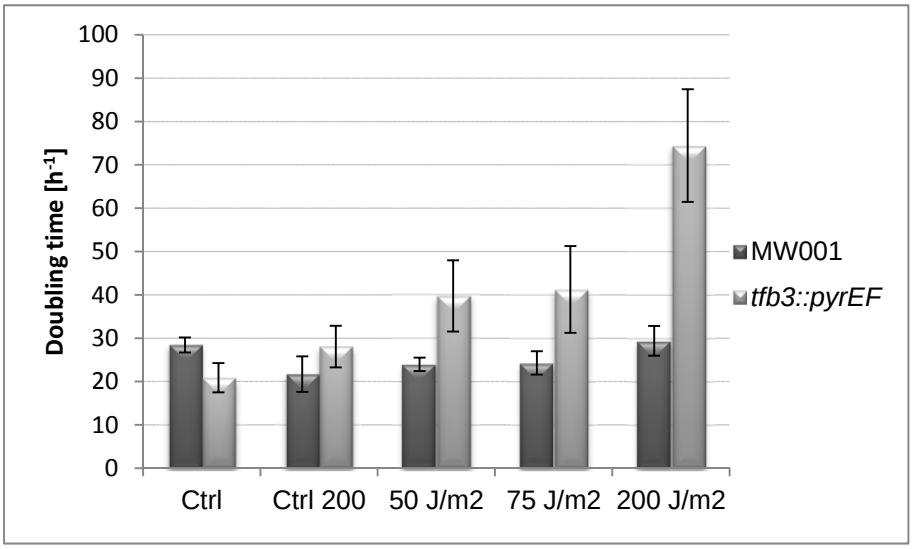


Fig. 33: Effect of UV-irradiation on the doubling time of *S. acidocaldarius* MW001 and the *tfb*3::*pyrEF* disruption mutant.
Doubling times of non-irradiated (Ctrl, Ctrl 200) and irradiated (50, 75, 200 J/m^2) *S. acidocaldarius* MW001 and the *tfb*3::*pyrEF* cells were calculated for the time within eight hours after UV-irradiation.

Growth rates and resulting doubling times between the time point of irradiation and the first eight hours after irradiation were determined. Calculations of the doubling time of the parent strain MW001 and the *tfb*3::*pyrEF* disruption mutant were performed as described in section 2.13.1. and the arithmetic mean was used as basis for Fig. 33. No difference was observed for the controls of the parent strain MW001 (Ctrl, Ctrl 200) and the UV-exposed cultures of MW001 (dark grey bars Fig. 33). The controls of the *tfb*3::*pyrEF* disruption mutant (Ctrl, Ctrl 200) revealed similar doubling times (21 h^{-1} ± 3.4 h^{-1} and 28 h^{-1} ± 4.8 h^{-1}), whereas the UV-irradiated cells (50 J/m^2, 75 J/m^2, 200 J/m^2) showed dramatically higher doubling times and thus reduced growth. No significant difference was observed between the doubling times of *tfb*3::*pyrEF* cells irradiated with 50 J/m^2 and 75 J/m^2 doses. Compared to the control samples the doubling time of the mutant strain increased almost two-fold at 50 J/m^2 (40 h^{-1} ± 8 h^{-1}) and

75 J/m^2 (41 h^{-1} ± 10 h^{-1}) and four-fold at a dose of 200 J/m^2 (75 h^{-1} ± 13 h^{-1}). The results indicate that the *tfb*3::*pyrEF* disruption mutant requires a longer adaptation phase than the parent strain after UV-exposure. The phenotype of the disruption mutant suggests a role of TFB3 in early UV-response and supports previous microarray data on UV-response (75 J/m^2, 200 J/m^2) (Fröls et al., 2007, Götz et al., 2007).

4. Discussion

4.1. Investigation of GTFs from *S. acidocaldarius* by expression and purification

The archaeal transcription apparatus comprises a minimal set of transcription factors compared to Eukaryotes. For transcription initiation only TFB, TBP and RNAP are required (Bell and Jackson, 1998a, b, Bell et al., 1999a, Bell and Jackson, 2000a). TBP recognizes the TATA-box and binds to this promoter element. Upon binding of TBP to the DNA TFB is recruited which contacts the BRE-promoter element. As last step during the initiation process the RNAP is bound by the N-terminus of TFB (Thomm et al., 1996, Bell and Jackson, 1998a, b, Bell et al., 1999a, Bell and Jackson, 2001, Soppa, 2001, Geiduschek and Ouhammouch, 2005, Werner and Grohmann, 2011). Although various transcriptional regulators, mostly of the bacterial-type, are present in Archaea it seems that the multiplicity of TFB and TBP proteins additionally enables the organisms to adapt to quickly changing environments. It has been previously proposed that multiple TFB and TBP proteins have a comparable role as bacterial sigma factors (Baliga, 2000, Micorescu et al., 2008, Lu et al., 2008, Peng et al., 2009, Paytubi and White, 2009, Turkarslan et al., 2011). This work focused on the elucidation of the role of transcription factors TFB1, TFB2 and TFB3 from the thermoacidophilic Crenarchaeon *S. acidocaldarius*.

4.1.1. Recombinant expression of GTFs from *S. acidocaldarius*

To investigate the function of the GTFs from *S. acidocaldarius,* recombinant proteins were generated. The recombinant transcription factor TFB1 was solely expressed after codon-usage adaptation as well as expression with less IPTG and lower temperature. The protein was purified to homogeneity as confirmed via SDS-PAGE and immunodetection (Fig. 3). In Fig. 3 A a double band was visible after affinity chromatography and gel filtration. This additional band was not observed in freshly prepared samples (data not shown) and therefore it seems to be a degradation product. In immunoblots with anti-His-tag antibody only one band at the correct size of 37 kDa was visible (Fig. 3 B). The protein samples for the purification overview were stored at -80°C for over two months. It seems that TFB1 was not stable at - 80°C although glycerol was added for stabilization of the protein. Therefore it has to be noted that TFB1 should be purified shortly before performance of EMSA studies or *in vitro* transcription experiments.

The encoding sequence of TFB2 was optimized by gene synthesis. TFB2 was expressed as insoluble protein although the expression strain was incubated at lower temperature with less IPTG, which are growth conditions that are supposed to support production of soluble protein. Even expression in another overexpression strain (*E. coli* Lemo21 (DE3)) did not increase solubility. This was rather surprising because transcription factors are cytoplasmic proteins and normally not membrane bound. Nevertheless, the protein could be purified under denaturing conditions and was detected in SDS-PAA gels and immunoblot with anti-His-tag antibody (Fig. 4). A prominent band corresponding to TFB2 was visible after SDS-PAGE (Fig. 4 A). Since affinity chromatography was the only purification step, additional proteins were purified which attached to the Ni-column (Fig. 4). Therefore TFB2 purification requires additional optimization.

Also transcription factor TFB3 was expressed as insoluble protein after codon-usage adaptation and expression over night at 22°C. Like for TFB2, solubility could not be increased by cold-expression or using another expression strain (data not shown). TFB3 was purified under denaturing conditions from the membrane fraction using IMAC and after SDS-PAGE analysis a dominant band corresponding to the calculated size of TFB3 was visible (Fig. 5). In the SDS-PAA gel shown in Fig. 5 A additional faint protein bands were visible above TFB3. Since impurities were still present in the elution fractions after Ni-TED purification, further purification steps are required. In immunoblots two strong signals were detected below TFB3, which might correspond to degradation products of TFB3 although protease inhibitor for disruption of cells and 20% (v/v) glycerol were added for purification (Fig. 5 B). Since the protein was purified under denaturing conditions it needs to be refolded. First attempts to refold the protein by drop by drop dialysis or refolding of the protein on the column failed. Hence further purification steps like gel filtration are required after improvement of refolding of the protein.

Recombinant TBP was expressed at standard conditions as soluble protein and was purified to homogeneity. The protein separated on SDS-PAA gels corresponded to the same size as the calculated molecular weight of 18.9 kDa (Fig. 6).

The transcription factor TFE was expressed as soluble protein and was purified. A double band was visible in Fig. 7 A which probably corresponds to a degradation product or derives from different post-translational modifications. TFE revealed a protein size of 19 kDa after SDS-PAGE compared to its expected size at 21 kDa. The lower molecular weight might result from post-translational modification and/or the amount of hydrophobic amino acids in the protein (35%), because proteins composed of many hydrophobic residues bind more detergent (SDS) and show a lower molecular weight in SDS-PAA gels (Hames, 1998).

The RNAP subunit RpoG, which will serve for antibody generation, was soluble expressed and purified to homogeneity. SDS-PAGE analysis revealed that the purified protein corresponded to its theoretical size (Fig. 8 A) and it was visible in immunodetection using anti-His-tag antibody (Fig. 8 B).

In this study polyclonal antibodies were raised against TFB1 and TFE providing a tool for protein detection in *S. acidocaldarius* (Fig. 9). The antibodies specifically recognized purified, recombinant proteins as well as the respective endogenous proteins in MW001 crude extracts (Fig. 9). For some applications like ChIP-Seq experiments it is advantageous to use pure antibody and not antiserum. Therefore antibodies were purified from antiserum by ammonium sulfate precipitation followed by affinity purification using protein A-agarose. The antibodies were successfully purified and only the heavy and the light chain of the antibody were visible in SDS-PAA gels (Fig. 10).

In other studies TFB1 from *S. acidocaldarius* has been cloned into pET30a by using *Nde*/*Xho*I restriction sites and was expressed as soluble protein using BLR (DE3) pLysS cells for overexpression (Bell and Jackson, 2000). Recombinant TFB1 from *S. solfataricus* was reported to be expressed as soluble protein after cloning the encoding sequence into the plasmid pDEST and overexpression Rosetta (DE3) cells (Götz et al., 2007, Abella et al.,

2007, Paytubi and White, 2009). Qureshi and colleagues cloned the *tfb*1 encoding sequence from *S. shibatae* into the plasmid pGEX-2TKP and it was successfully overexpressed as soluble protein and purified via glutathione-agarose beads (Qureshi et al., 1995b). TFB1 from *T. tenax* was reported to be expressed (pET11c, BL21 (DE3) pRIL) and needed to be purified from inclusion bodies under denaturing conditions and was successfully renatured by repeated dialysis in buffer without urea (Marrero et al., 2013, in prep.).

The transcription factor TFB2 has not been overexpressed and purified in Crenarchaeotes so far, except from *T. tenax* (Marrero et al., 2013, in prep.). TFB2 from *T. tenax* was cloned into pET302 with an N-terminal His-tag, expressed as soluble protein and purified to homogeneity (Marrero et al., 2013 in prep.).

The *tfb*3 gene from *S. solfataricus* has been cloned into the vector pDEST and expression of TFB3 was induced by addition of 200 µM IPTG and over night expression. The protein was successfully recombinant expressed and soluble (Götz et al., 2007).

The encoding sequence of *tbp* was amplified from genomic DNA from *S. acidocaldarius* and introduced into the plasmid pET30a via its *Nde*I/*Xho*I restriction sites. The recombinant protein was soluble expressed using BLR (DE3) pLysS cells and purified via a C-terminal His-tag (Bell and Jackson, 2000a). TBP from *S. solfataricus* and *S. shibatae* were also expressed as soluble proteins (Abella et al., 2007, Iqbal and Qureshi, 2010, Qureshi et al., 1995a). TBP from the Crenarchaeote *T. tenax* was also expressed as soluble protein using pET11c as overexpression vector and Rosetta (DE3) cells for overexpression (Marrero et al., 2013 in prep.).

As previously described by Bell and co-workers TFE from *S. solfataricus* was amplified from genomic DNA and ligated into pET30a. Overexpression was performed in BLR (DE3) cells harboring the pRIL plasmid and the protein was soluble expressed (Bell et al., 2001). Ten years later Iqbal and Qureshi reported expression of TFE from *S. solfataricus* using the plasmid pQE30Xa (Iqbal and Qureshi, 2010).

The RNAP subunit RpoG is conserved among hyperthermophilic Crenarchaeota and one Korarchaeote (*Korarchaeum* sp.) (Koonin et al., 2007). Overexpression of the RpoG subunit from the RNAP of *S. acidocaldarius* or other Crenarchaeota has not been reported in the literature so far. Therefore this is the first report of expression and purification of RpoG in Crenarchaeota.

The genes *tfb*1, *tfb*2 and *tfb*3 from *S. acidocaldarius* were cloned into pET30a in this study and were transformed into the overexpression strain BLR (DE3) pLysS according to the method described by Bell and Jackson, 2000. Surprisingly, TFB1, TFB2 and TFB3 expression using BLR (DE3) pLysS cells could not be confirmed with the standard overexpression procedure (LB-medium, 1 mM IPTG, 3 h induction, 37°C). Mutations in the cloned PCR products of *tfb*1, *tfb*2 and *tfb*3 were excluded by sequencing. Probably different growth media compositions and/or overexpression parameters could have led to this discrepancy because the authors refer to the Novagen manual for expression conditions and do not mention overexpression details (Bell and Jackson, 2000).

Transcription factors from *S. acidocaldarius* and other Crenarchaeota have been reported to be recombinant expressed in different *E. coli* strains except RpoG and in case of TFB2 expression has only been observed for *T. tenax* TFB2. These GTFs were mostly expressed as soluble proteins without codon-adapted sequences for expression in *E. coli.*

In summary, the transcription factors TFB1, TBP, TFE and the RNAP subunit RpoG from *S. acidocaldarius* were successfully purified to homogeneity. TFB2 and TFB3 were expressed as recombinant protein, subsequently purified under denaturing conditions and refolding conditions need to be optimized for these proteins. Specific anti-TFB1 and anti-TFE antibodies were generated in this study and are now available for determination of endogenous protein level of these transcription factors in response to different stress parameters.

4.1.2. Homologous expression of GTFs in *S. acidocaldarius*

Homologous expression of TFB1, TFB2, TFB3 and TBP of *S. acidocaldarius* was used as second approach to purify these proteins. *S. acidocaldarius* was used as expression host but expression was not possible although homologous expression by pSVA1450 in *S. acidocaldarius* has been reported for Agl3, a UDP sulfoquinovose synthase (Meyer et al., 2011), MalA from the maltose/maltodextrin operon (Choi et al., 2013), Lrs14 a transcriptional regulator (Orell et al., 2013) and proteins involved in flagella assembly (Lassak et al., 2012) as well as heterologous expression of proteins from *S. solfataricus* (ABCE1) and *T. tenax* (phosphofructokinase) (Berkner et al., 2010, Wagner et al., 2013 submitted). The close relative *S. solfataricus* has been used as expression host for various heterologous (e. g. sulfur oxygenase/reductase from *Acidianus ambivalens,* alcohol dehydrogenase from *Bacillus stearothermophilus*) and homologous proteins (e.g. ABCE1, β-galactosidase, ATPase) (Contursi et al., 2003, Jonuscheit et al., 2003, Angelov and Liebl, 2010). In Halobacteria genetic manipulation was established much earlier and homologous (halorhodopsin, a light-driven anion pump from *Halobacterium halobium*) as well as heterologous expression of bacterial (*Zymomonas mobilis* pyruvate decarboxylase) and human proteins (β_2-adrenoceptors) has been demonstrated (Heymann et al., 1993, Kaczowka et al., 2005, Söhlemann et al., 1997). In the Euryarchaeote *P. furiosus* a subcomplex of hydrogenase has been successfully homologous expressed (Hopkins et al., 2011). Nevertheless, homologous overexpression of transcription factors (TFB1, TFB2, TFB3, TBP) from *S. acidocaldarius* or other Archaea has not been reported previously.

Expression of TFB1, TFB2, TFB3 or TBP was not detectable in the flow-through and wash or elution fractions of the Strep-tag purification shown in Fig. 11 and Fig. 12 or of Ni-TED columns (data not shown). It is unlikely that none of the proteins was able to bind to the columns used for purification. Sequencing of the expression plasmids confirmed that the sequences were successfully cloned and were free of mutations. Two proteins at around 50 kDa and 15 kDa were present in SDS-PAA gels in the elution fractions of TFB2 (Fig. 11 A) and TFB3 (Fig 12 A). In corresponding immunoblots only the 15 kDa band was detected by anti-Strep-tag antibody (Fig. 11 B, Fig. 12 B). Since both proteins were present after Strep-tag purification of other samples in SDS-PAA gels and the 15 kDa protein was additionally identified by immunodetection it is assumed that both proteins are unspecific proteins rather

than degradation products. The presence of the 50 kDa band has been reported for protein purification from *S. solfataricus* and *S. acidocaldarius* cell extracts via Strep-tag columns previously (Albers et al., 2006, Wagner et al., 2012). Both unspecific proteins were not detected using His-tag antibody (data not shown). None of the TFB proteins or TBP could be purified via His-tag or Strep-tag columns or was identified by immunodetection (Fig. 11, Fig. 12). Even by immunodetection with anti-*Ttx*-TBP antibodies it was not possible to detect the corresponding *Sulfolobus* protein although the antibody was able to recognize heterologous expressed protein. The reason for this is unclear but it might be possible that overexpression of an endogenous protein interferes with internal regulatory processes. Maybe the overexpressed transcription factors were quickly degraded after synthesis although it is not clear why only plasmid-encoded tagged proteins would be degraded. In addition, a higher culture volume than 8 L resulting in more cell mass might be required for protein purification compared to *E. coli*.

4.1.3. Knock-in of Strep-FLAG tags for purification of GTFs from *S. acidocaldarius*

Purification of proteins from the host organism might be advantageous compared to heterologous expression in respect to protein folding or modification. Therefore, integration of genomic tags was used as an alternative strategy to purify proteins from *Sulfolobus* cells. TFB1 and TFE were successfully purified from *S. acidocaldarius* cells by integration of a C-terminal Strep-FLAG tag (Fig. 13, Fig. 14). After Strep-tag purification a significant enrichment was observed in the elution fractions of TFB1 and TFE, however other proteins especially a thick band at 55 kDa were still visible (Fig. 13). This band was always present when Strep-tag material has been used for protein purification from *S. acidocaldarius* cells. Nevertheless, after immunodetection with anti-FLAG-tag and anti-Strep-tag antibody only one band corresponding to TFB1 and TFE was detectable (Fig. 14 A, C). The protein levels of TFB1 and TFE were high enough for purification using standard growth conditions confirming previous reports that TFB1 is the typical TFIIB homolog present at normal growth conditions (Bell et al., 1998a, b, Bell and Jackson, 2000a) and TFE is not required for transcription initiation but it facilitates transcription by supporting TBP binding in cases of imperfect TATA-boxes (Bell et al., 2001, Qureshi et al., 1997). TFB2 and TFB3 could not be purified via an N-terminal Strep-FLAG tag. Only an unspecific band at around 15 kDa was visible in immunodetection due to unspecific binding of the anti-Strep-tag antibody (Fig. 14 B). Therefore TFB2 and TFB3 are either not present at standard growth conditions or the amount is too low and the culture volume needs to be scaled up for purification.

In other studies two proteins from *S. acidocaldarius*, namely AglB (Saci_1274), a membrane protein and the hypothetical cytosolic protein Saci_1210 were genomically tagged with a His-tag or a tandem Strep-His-tag (Wagner et al., 2012). However, genomic integration of a tag for purification or detection of transcription factors has not been reported for *S. acidocaldarius* or Crenarchaeota. In the Euryarchaeote *H. salinarum* NRC-1 a HA-tag has been genomically integrated at the *tfb*D gene for determination of genome-wide binding sites by CHIP-seq (Wilbanks et al., 2012).

Genomic integration of a FLAG-tag as described in this study has not been published so far and therefore demonstrates that FLAG-tagged proteins can be successfully detected and

purified even at elevated temperatures of 78°C. The use of genomic tags not only enables protein purification. In addition immunoprecipitation experiments for finding unknown binding partners or the *in vivo* protein level can be determined by immunodetection.

For future experiments like *in vitro* transcription it was necessary to purify the RNAP from *S. acidocaldarius*. The RNAP holoenzyme was successfully purified from *S. acidocaldarius* cells by Ni-affinity chromatography after integration of a C-terminal His-tag at the RpoG subunit (Fig. 15 A). The RpoG subunit was detected in the RNAP holoenzyme in immunoblots using anti-His-tag antibody (Fig. 15 B). In addition immunodetection with the in this study generated anti-TFB1 antibody revealed a co-purification of TFB1 (Fig. 15 C). Co-purification of *S. solfataricus* TFB1 and TFB3 with RNAP has been reported previously (Götz et al., 2007, Paytubi and White, 2009). Immunoprecipitation was performed with RNAP and recombinant transcription factors (TFB1, TFB3) from *S. solfataricus* using RpoB antibody. Afterwards co-immunoprecipitation could be detected in SDS-gels and by immunodetection (Götz et al., 2007). Endogenous RNAP was purified from *S. solfataricus* cells after two-step purification using heparin and gel filtration. Samples obtained by gel filtration were analyzed for the presence of TFB1, TFB3 and TBP with protein-specific antibodies by immunodetection and revealed a co-purification of TFB1 and TFB3 (Paytubi and White, 2009).

4.2. Genome context and promoter analyses and evidence for co-transcription

Genomic context analysis can be used for functional prediction of encoded proteins. A genomic conservation of *gar*1, *tfb*2 and the gene encoding the RPC34 homolog in the division of Crenarchaeota (Sulfolobales, Thermoproteales and Desulfurococcales) was reported previously (Marrero, 2009, Fig. 16). In addition other conserved genomic contexts were found in some Crenarchaeota and mainly Euryarchaeota (see Fig. 36 supplement). Gar1 is a part of the H/ACA ribonucleoprotein complex found in Archaea and Eukaryotes. Small nucleolar ribonucleoprotein (snoRNP, snRNP in Archaea) complexes are important for rRNA processing and H/ACA snRNPs mediate pseudouridylation of tRNA for stabilization (Baker et al., 2005, Reichow et al., 2007, Hamma and Ferré-D´Amaré, 2010, Watkins and Bohnsack, 2012, Kiss et al., 2010, Kamalampeta and Kothe, 2012). Recent sequence comparisons found orthologs of the RPC34 subunit of the RNAP among Thaumarchaeota, Crenarchaeota and only few Euryarchaeota (Blombach et al., 2009). Blombach and colleagues analyzed the genomic context of the RPC34 orthologs and found the same genomic organization as was shown in this report (Fig. 16, Fig. 36). The genes encoding *gar*1 and *tfb*2 are co-localized with the gene encoding the RPC34 ortholog among *Sulfolobus* species, *S. maritimus* and *H. butylicus* (Blombach et al., 2009). The authors suggest a role of the RPC34 ortholog in transcription of rRNA and tRNA genes because of the genomic neighborhood described before (Blombach et al., 2009). In Eukaryotes the third nuclear RNAPIII transcribes tRNA genes, 5S rRNA and some snRNA genes. The RPC34 subunit interacts with TFIIIB70, the TFIIB homolog, (Brun et al., 1997, Werner et al., 1993, Wu et al., 2011) and forms a subcomplex with RPC31-RPC34-RPC82 that facilitates transcription initiation (Werner et al., 1993, Brun et al., 1997, Proshkina et al., 2006, Wu et al., 2011).

In some archaeal genomes the sequences of *gar*1 and *tfb*2 partially overlap. In *S. acidocaldarius* the start and stop codon of *gar*1 (Saci_1341) and *tfb*2 (Saci_1341) overlap (ATGA). Alignment of *gar*1 and *tfb*2 upstream sequences (-1 to -80 bp) from fourteen crenarchaeal organisms was the basis for generation of a sequence logo (Fig. 18). Typical promoter elements, BRE-site and TATA-element, were detected in close proximity to the predicted transcription start site of *gar*1. In addition, putative regulatory AT-rich repeats were found upstream of the core promoter elements (Fig. 17, Fig. 18). The *tfb*2 gene possesses two inverted repeats upstream of the putative TATA-box and BRE-element that could serve as regulator binding sites. In addition to core promoter elements *tfb*2 harbors a RBS sequence whereas *gar*1 does not have an anchoring sequence for ribosomes (Fig. 18). In Archaea many genes have leaderless transcripts lacking any ribosomal binding site for translation initiation. According to this the first gene of an operon misses such a ribosomal binding sequence whereas genes within an operon have a Shine-Dalgarno sequence upstream of the start-codon (Tolstrup et al., 2000, Slupska et al., 2001, Ma et al., 2002, Toranisson et al., 2005, Chang et al., 2006, Quax et al., 2013). Therefore two different translation initiation mechanisms seem to occur (Quax et al., 2013).

To analyze whether *gar*1, *tfb*2 and Saci_1342 are transcribed as one unit, RNA was extracted and cDNA was synthesized by reverse transcriptase. This cDNA served as template for PCR with various primer combinations (Fig. 19 A). As depicted in Fig. 19 B, C and D all three genes were transcribed as single genes when gene-specific primers were used. In general the amplification products had the expected size (Fig. 19 B, C, D, E, F, G), however in case of *tfb*2 in some lanes double bands occurred in close proximity which might be due to unspecific primer annealing. In Fig. 19 D in lane 4 no amplification product for Saci_1342 was visible possibly due to low cDNA content.

The experiments in this study demonstrated that *tfb*2 is co-transcribed with *gar*1 in a growth-phase-dependent manner or co-transcribed with Saci_1342, but no full-length transcript occurred (Fig. 19 E, F, G). The band intensity of the *gar*1-*tfb*2 PCR product increased from early-log to exponential growth phase towards stationary phase (Fig. 19 F). The *tfb*2-Saci_1342 transcript was present at every growth-phase without any obvious preference (Fig. 19 G). A growth-phase dependent transcription of *gar*1-*tfb*2 prompts to the question of regulatory effects that influence the presence of this transcript.

4.3. Reporter gene assays

To investigate at which growth and stress conditions GTFs from *S. acidocaldarius* are present, promoter sequences of the transcription factors (*tfb*1, *tfb*2, *tfb*3, *tfe*, *tbp*) and the H/ACA ribonucleoprotein *gar*1, controlling the expression of the *lacS* gene, were cloned. Promoter regions of these genes were analyzed for typical promoter elements. Not all promoters contain conserved BRE-sites or TATA-box sequences such as *tfb*1, *tbp* and the *tfe* promoter (Fig. 20). In addition, all promoter sequences, except *tfe*, exhibited putative regulatory sequences (Fig. 20). First experiments revealed that the reporter assay system is functional and preliminary experiments with the promoter regions of *tfb*1, *gar*1 and *tfb*3 were analyzed for reporter gene expression changes in response to cold shock (60°C) (Fig. 21,

Fig. 22). It seemed that the *gar1* (+ 4296 mod. Miller units after 60 min) and *tfb*3 (+ 2259 mod. Miller units after 60 min) promoter regions did not significantly alter the expression level of the reporter gene whereas LacS activity increased when the *tfb*1 promoter region (+ 8625 mod. Miller units after 60 min) was tested (Fig. 22). At the time point before cold-shock (0 min) it was shown that *tfb*1 promoter driven expression is higher than that of *gar*1 and *tfb*3 (Fig. 22). This indicates that TFB1 is the major transcription factor present at normal growth conditions and an upregulation (1.47-fold) was also observed in response to cold-shock.

In *S. acidocaldarius*, *S. solfataricus* and *S. islandicus* promoters of the maltose-binding protein (*mal*) and the chaperone TF55 (*tf*55α) as well as the arabinose binding protein (*ara*S) have been investigated (Jonuscheit et al., 2003, Berkner et al., 2010, Peng et al., 2012, Wagner et al., 2013 submitted). In Euryarchaeota promoter-driven expression of a reporter gene has been investigated for a transcriptional regulator (*glp*R) and other promoters but not for transcription factors (Rawls et al., 2010, Holmes and Dyall-Smith, 2000, Sniezko et al., 1998, Thompson and Daniels, 1998, Kuo et al., 1997, Cohen-Kupiec et al., 1997, Palmer and Daniels, 1994).

In the future the study of expression of all promoter constructs under different stress conditions will provide great advance for understanding the role of transcription factors in *S. acidocaldarius* and probably Crenarchaeota.

4.4. Interaction studies between RNAP and GTFs from *S. acidocaldarius*

The archaeal RNAP resembles the eukaryotic RNAPII in structure and function (Zillig et al., 1979, Best and Olsen, 2001, Werner, 2007). Several hypotheses address the question of RNAP evolution and it was discussed that the primordial RNAP was a ribozyme or that transcription regulators were incorporated into the RNAP (Iyer et al., 2003, Werner, 2008, Grohmann and Werner, 2010, Werner and Grohmann, 2011). Eukaryotes comprise several RNAPs for transcription whereas Bacteria and Archaea use only one enzyme. The structures of RNAPs in all three domains have been solved. Yeast RNAPII consists of twelve subunits with high similarity to the archaeal subunits (Fig. 1). All RNAPs in the three domains of life have a crab-claw-like structure (Hirata et al., 2008b, Cramer et al., 2001, Kim et al., 1995, Zhang et al., 1999, Tsuji et al., 1981, Vassylyev et al., 2002, Murakami, 2013). The first crystal structure has been solved by Cramer and colleagues in 2000 (Cramer et al., 2000) and structures including TFIIB or showing DNA-RNAP contacts were analyzed afterwards (Bushnell et al., 2004, Chen et al., 2004, Chen and Hampsey, 2004, for review see Hahn, 2004). The structures of bacterial RNAP from *Thermus aquaticus* (Kim et al., 1995, Zhang et al., 1999), *Thermus thermophilus* (Tsuji et al., 1981, Vassylyev et al., 2002) and *E. coli* (Murakami, 2013) gained further insights into the transcription apparatus of Bacteria. In the archaeal kingdom the structures of *S. shibatae*, *S. solfataricus* and *P. furiosus* RNAP have been solved (Korkhin et al., 2009, Wojtas et al., 2012, Hirata et al., 2008b, Hirata and Murakami, 2009, Kusser et al., 2008).

RNAP is one of the core aspects of transcription. Several structural analyses have shed light on subunit composition and architecture. Studies of RNAP subunit interaction (Dieci et al.,

1995, Zaychikov et al., 1995, Werner and Weinzierl, 2002, Goede et al., 2006, Best and Olsen, 2001) and interaction with transcription factors augment understanding of transcription (Magill et al., 2001, Goede et al., 2006, Paytubi and White 2009). Yeast two-hybrid experiments were performed to investigate which RNAP subunits bind to the GTFs from *S. acidocaldarius*. The present study revealed an interaction of *Sa*TFB2 and *Sa*TFB3 with RpoE and for *Sa*TFB3 an additional interaction with RpoK (Fig. 23, Table 16). RpoE and RpoK have an auxiliary function and RpoE builds a heterodimeric complex with RpoF similar as eukaryotic Rbp4 and Rbp7 (Werner and Weinzierl, 2005, Werner, 2007, Grohmann and Werner, 2011b). The RpoE/RpoF subcomplex is involved in all three steps of transcription. It facilitates open-complex formation during initiation and is involved in elongation and termination of transcription (Werner and Weinzierl, 2005, Werner, 2007, Grohmann and Werner, 2011b). The interactions illustrated in Fig. 25 seem to be true positives as they do not drive transcription of reporter genes in the yeast two-hybrid system without a binding partner. This is the first report of protein interaction studies with TFB2 from Crenarchaeota. An interaction of TFB1 with RNAP subunits was not observed. Possibly TFB1 requires specific DNA contacts for sufficient binding to the RNAP like for building the ternary complex. However, TFB1 should be able to bind to the RNAP because it co-purified with the RNAP as shown in this and previous work (Fig. 15 C, Paytubi and White, 2009).

In addition to TFB-RNAP interaction, binding of TFB proteins from *S. acidocaldarius* to TBP has been investigated in this study. Only TFB3 showed a weak interaction with TBP (Fig. 24). Also here no interaction of TFB1 and TBP could be detected. Intriguingly, also for TFB2 no binding to TBP was observed (Fig. 24).

Additional experiments to verify protein-protein interaction like immunoprecipitation with recombinant overexpressed proteins or pull-down assays as well as photometric determination of α-galactosidase activity would further support the yeast two-hybrid experiment.

To verify expression of the proteins from the yeast two-hybrid experiments, immunoblots with anti-c-myc-tag and anti-HA-tag antibody were performed. However, only the expression of two RNAP subunits (RpoA2 and RpoB) could be confirmed despite testing various protein extraction and immunoblotting methods (Fig. 26). Immunodetection conditions need to be optimized for confirmation of the results obtained from the yeast two-hybrid experiment.

Interaction studies between all twelve RNAP subunits from *S. solfataricus* and *Sso*TFB1 as well as *Sso*TFB3 revealed an interaction of both transcription factors with RpoE and RpoK in yeast two-hybrid assays (Paytubi and White, 2009) and the authors suggest that TFB1 and TFB3 share common RNAP binding sites. In that study endogenous RNAP from *S. solfataricus* was purified and immunodetection with TFB1 and TFB3 specific antibodies revealed that TFB1 and TFB3 co-purified with RNAP (Paytubi and White, 2009). IN another study EMSA experiments with *Ttx*TFB2 and *Ttx*TBP demonstrated a TBP-independent binding of TFB2 to different tested promoters (Marrero et al., 2013, in prep.). The binding sites of TFB2 were determined by ExoIII-footprinting and TFB2 bound upstream of the core promoter to an AT-rich region and it was speculated that TFB2 might act as an activator or a repressor of transcription initiation (Marrero et al., 2013, in prep.). It has been reported previously that archaeal TFB and TBP from *P. woesei*, *S. shibatae* and human interact by the

C-terminal core domain of TFB containing two cyclin repeats in the presence of DNA (Kosa et al., 1997, Qureshi and Jackson, 1998, Littlefield et al., 1999, Nikolov et al., 1995). Crystal structures of a complex of *P. woesei* TBP, the C-terminal part of TFB and promoter DNA showed an interaction (Littlefield et al., 1999, Kosa et al., 1997). In *H. salinarum* NRC-1 TBPe co-imunoprecipitated with TFBb, TFBc, TFBd, TFBe and TFBg and for TFBc an interaction with TBPd and TFBa with TBPc was demonstrated (Facciotti et al., 2007).

It was previously reported that the N-terminal domain of *Sso*TFB1 (1-109) showed a stronger interaction to RpoK than the full-length protein (Magill et al., 2001, Paytubi and White, 2009). Mutagenesis of residues in the N-terminal Zn-ribbon of *Sso*TFB1 (Cys31 and Leu38) negatively affected binding of *Sso*TFB1 to the RNAP (Magill et al., 2001). In addition an impaired interaction between an *Sso*TFB3 double mutant (C22S, C25S) and RpoK has also been observed (Paytubi and White, 2009). These data suggest that *Sso*TFB1 and *Sso*TFB3 interact with RNAP from *S. solfataricus* via their N-terminal Zn-ribbons (Magill et al., 2001, Paytubi and White, 2009).

It is tempting to speculate that interaction studies between the N-terminal part of *Sa*TFB1 and the RNAP from *S. acidocaldarius* might show stronger binding of TFB1 compared to the full-length protein. Archaeal and eukaryotic TFB proteins comprise a Zn-ribbon motif (CHCC or CCCC) at their N-termini. A corresponding metal-binding sequence is found in *P. furiosus* TFB and TFIIB from yeast to humans (Pinto et al., 1992, Zhu et al., 1996, Chen et al., 2000, Hahn and Roberts, 2000). In *S. acidocaldarius* and *S. solfataricus* two cysteine residues are replaced by serine and threonine (Qureshi et al., 1995b, Micorescu et al., 2008). TFB1 and TFB2 of *S. acidocaldarius* are composed of an N-terminal Zn-ribbon domain, a B-finger and a C-terminal DNA binding domain separated by a flexible linker region (Fig. 34) (Bell and Jackson, 2000a, Bartlett et al., 2005). The N-terminal part of TFB1 is necessary for binding to the RNAP (Bell and Jackson, 2000a, Magill et al., 2001). Several studies demonstrated that truncation of the TFB N-terminus of *S. acidocaldarius*, *S. solfataricus* and human results in inability of RNAP recruitment (Buratowski and Zhou, 1993, Barberis et al., 1993, Bell and Jackson, 2000a, Magill et al., 2001). In Eukaryotes it was previously shown that the N-terminal domain of TFIIB is crucial for RNAP binding (Ha et al., 1993, Chen and Hampsey, 2004, Pardee et al., 1998). The B-finger domain of archaeal transcription factors can be mutated without alteration of start site selection (Bell and Jackson, 2000a, Magill et al., 2001). In contrast, mutations in the yeast B-finger region led to use of different start sites (Pinto et al., 1994) and mutations in the B-finger of TFIIB from human did not affect transcription activity but shifted start site selection at some tested promoters (Hawkes and Roberts, 1999, Fairley et al., 2002, Thompson et al., 2009) whereas the deletion of the whole B-finger domain resulted in inability of transcription (Tran and Gralla, 2008). The B-finger domain is necessary for stabilization of the pre-initiation complex and functions as post-recruitment factor in Archaea (Werner and Weinzierl, 2005). Deletion of the linker domain of TFB (MJ0782) from *M. janaschii* negatively affected the RNAP activity (Weinzierl and Wiesler, 2011). TFB3 proteins from organisms belonging to Sulfolobales and Thermoprotelaes, including *S. acidocaldarius*, *S. solfataricus* and *T. tenax,* possess a Zn-ribbon motif but lack the B-finger and the first cyclin-repeat of the DNA-binding domain (Fig. 34) (Bartlett et al., 2005, Marrero et al., 2013 in prep.).

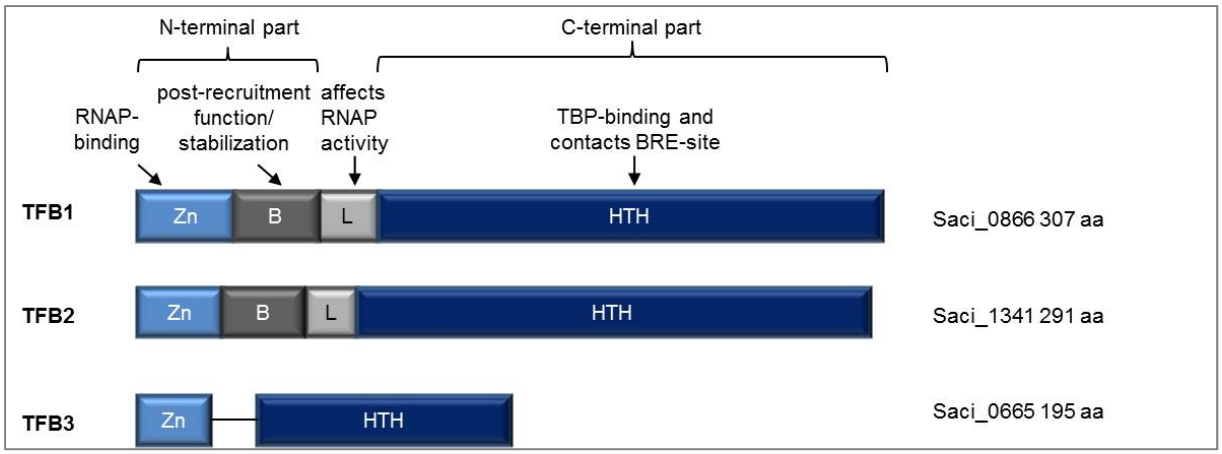


Fig. 34: Domain structure of TFB proteins from *S. acidocaldarius.*
TFB1 and TFB2 are full-length proteins comprising an N-terminal and C-terminal part connected by a flexible linker region (L). The Zn-ribbon (Zn) is the RNAP-docking domain and the linker region stimulates RNAP activity (Weinzierl and Wiesler, 2011). TFB3 is shorter and lacks the B-finger (B) which stabilizes the RNAP-DNA complex (Werner and Weinzierl, 2005) and the first cyclin-fold of the helix-turn-helix domain (HTH).

In conclusion interactions between TFB2 and RpoE as well as TFB3 with RpoE and RpoK, two subunits with an auxiliary function, were detected in yeast two-hybrid experiments as well as an interaction between TFB3 and TBP. TFB1 seems to require specific DNA contacts for binding to the RNAP and binding studies only with the N-terminal part of TFB1 comprising the Zn-ribbon would probably exhibit an interaction. It might be possible that TFB2 from *S. acidocaldarius* binds to DNA in a TBP-independent manner according to the reaction reported for TFB2 from *T. tenax*. Protein expression needs to be verified except for RpoA2 and RpoB by immunodetection.

4.5. Generation of transcription factor deletion strains in *S. acidocaldarius*

To investigate the *in vivo* role of transcription factors, the aim of this study was to generate deletion strains of the transcription factors *tfb*1, *tfb*2, *tfb*3 and *tfe* as well as *gar*1 and the gene encoding the RPC34 homolog (Saci_1342) in *S. acidocaldarius*. First attempts to generate disruption mutants of the genes *gar*1, Saci_1342 and *tfe* failed. Gar1 and the RPC34 homolog have been shown to be essential in *S. cerevisiae* (Girard et al., 1992, Stettler et al., 1992) and might be essential in *S. acidocaldarius* as well. These data suggest that *tfe* is required for cell survival although it might be dispensable because it is not necessarily required for transcription initiation *in vitro* (Bell and Jackson, 2000a, Bell et al., 2001, Hanzelka et al., 2001, Hirata et al., 2008a). Until now no knock-out mutants of *tfb*1, *tfb*2 or *tfb*3 have been reported in Crenarchaeota. Deletion of *tfb*1 or *tfb*2 genes was not possible, neither with the in-frame markerless knock-out mutant method (Wagner et al., 2009, Fig. 27, Fig. 28 A) nor with the PCR-based disruption method (Sakofsky et al., 2011, Fig. 29). It is most likely that TFB1 and TFB2 are essential for the cell. As mentioned before, TFB1 was reported to be the transcription factor at standard growth conditions possibly acting for transcription of housekeeping genes like σ^{70} (Bell and Jackson, 2000a, Paytubi and White, 2009). In case of TFB2 only microarray-studies exist assuming a role in cell-cycle regulation (Lundgren and Bernander, 2007). These microarray data and qPCR experiments with synchronized

S. acidocaldarius cells showed that *tfb*2 transcripts are induced at the transition from G_1 to S-phase suggesting a cell-cycle dependent role (Lundgren and Bernander, 2007). A markerless knock-out of *tfb*3 could be generated but seemed not to be stable after repeated growth in second selection medium (Fig. 28 B). The reason for this instability of the deletion is not clear. However, a disruption mutant could be successfully generated by integration of the *pyrEF* cassette into the genomic region of the *tfb*3 gene (*tfb*3::*pyrEF*) (Fig. 30).

Investigation of *in vivo* function of transcription factors by deletion of the gene of interest has been used mainly in the euryarchaeal branch. Several knock-out mutants were reported in *H. salinarum* NRC-1 (Δ*tfb*A, *tbp*D, Δ*tfb*C) (Kaur et al., 2010, Coker and DasSarma, 2007), *T. kodakaraensis* (Δ*tfb*1, Δ*tfb*2 both mutants are viable) (Santangelo et al., 2007) and *M. mazei* (Δ*tbp1* essential, *Δtbp*2 and *Δtbp*3 dispensable) (Reichlen et al., 2010). Whereas deletion of transcription factors has been reported in some Euryarchaeota, knock-out of transcription factors from organisms belonging to the phylum Crenarchaeota has only been documented for *S. islandicus* (Zhang et al., 2010). Deletion of *S. islandicus tfb*2 and *tfb*3 was tested via the pop-in/pop-out technique but no deletion mutants were obtained.

A *tfb*3 deletion mutant was not obtained with the pop-in/pop-out technique by Zhang and co-workers in *S. islandicus* or in *S. acidocaldarius* as in the present study. The *tfb*3 deletion was not stable and it reverted to the wild-type situation (Fig. 28 B). The revertant might be explained by homologous recombination events but in that case a genomically unmodified DNA template from the parent strain MW001 would have to be present. This is unlikely because the culture was not contaminated with foreign DNA as has been proven by one single band after PCR analysis (data not shown). Therefore it is not clear why the knock-out was not stable. Nevertheless, after integration of the *Sso-pyrEF* marker cassette and resulting gene disruption a stable mutant was generated and tested for its response to UV-irradiation (*tfb*3::*pyrEF*). Therefore this is the first report of a viable *tfb*3 mutant in *S. acidocaldarius* and in the phylum Crenarchaeota.

4.6. TFB2: a cell-cycle regulator?

Microarray studies showed a cyclic-induction of *gar*1 and *tfb*2 at the transition from G_1 to the replicative S-phase proposing a co-transcription of these genes (Lundgren and Bernander, 2007). The assumed co-transcription of *gar*1-*tfb*2 has been confirmed in the present study (Fig. 19 F). The stability of mRNAs in *S. solfataricus* has been investigated by two independent research groups (Andersson, 2006, Bini et al., 2002). They showed that *tfb*1 mRNA (half-life: 2 h in Bini et al., 2002, 9.12 min in Andersson et al., 2006) is more stable compared to *tfb*2 (half-life: 13.5 min in Bini et al., 2002, 3.32 min in Andersson et al., 2006) or *tfb*3 mRNA (half-life of 7.76 min in Andersson et al., 2006) indicating a regulatory function of the alternative transcription factors TFB2 and TFB3. EMSA binding studies revealed a TBP-independent binding of TFB2 to AT-rich regions upstream of the core promoter in *T. tenax* (Marrero et al., 2013, in prep.) pointing to an activating or inhibitory mode of action. Yeast two-hybrid experiments performed in the present study revealed an interaction between TFB2 and RNAP subunit RpoE but not with TBP (Fig. 24). In the Euryarchaeote *P. furiosus* a role of TFB2 in heat-shock response has been implicated.

Interestingly, TFB2 from this organism lacks the B-finger domain which has a stabilizing function during transcription (Shockley et al., 2003, Micorescu et al., 2008, Werner and Weinzierl, 2005). However, a mutational approach (Δ*tfb*1 and Δ*tfb*2) revealed no obvious functional difference of TFB1 and TFB2 in *T. kodakaraensis,* a close relative of *P. furiosus,* under standard conditions (Santangelo et al., 2007). *Tk*TFB1 (Tko_1280) and *Tk*TFB2 (Tko_2287) share almost 60% sequence identity and identical residues are present in the C-terminal DNA-binding domain which might explain that *T. kodakaraensis* deletion mutants show no obvious phenotype with only one TFB protein present in the knock-out strains. The TFBs from *T. kodakaraensis* seem to bind to the same promoters and the authors suggest that more divergent GTFs might drive transcription from specific promoters (Santangelo et al., 2007). In *S. acidocaldarius* no knock-out mutant of *tfb*2 could be generated indicating an essential role for cell viability (Fig. 28 A). However, the role of TFB2 in *S. acidocaldarius* still remains to be elucidated. Low mRNA half-lives of *tfb*2 in *S. solfataricus* and *S. acidocaldarius* cells, the genomic neighborhood and the essentiality of the *tfb*2 gene point to a regulatory role of TFB2. TFB2 is able to contact RNAP (RpoE) and may not require TBP to bind to DNA. It is possible that TFB2 binds TBP-independent to upstream regions of promoters as has been demonstrated for *T. tenax* (Marrero et al., 2013, in prep.) and might serve as activator or inhibitor of transcription of specific genes during certain cell-cycle phases and/or it could alter rRNA modification or protein biosynthesis (genome context) in response to different growth-phases. DNA-protein binding studies and *in vitro* transcription with the purified RNAP could provide further insight into the function of TFB2 as well as the generated promoter constructs.

4.7. TFB3 and its implication in UV-response

Microarray studies with *S. solfataricus* and *S. acidocaldarius* revealed that the *tfb*3 transcript was highly upregulated after UV-exposure. In these studies *S. solfataricus* and *S. acidocaldarius* were exposed to doses of 200 J/m^2 and 75 J/m^2, respectively (Götz et al., 2007, Fröls et al., 2007). To examine if the *tfb*3 disruption mutant (*tfb*3::*pyrEF*) created in this study is affected by different UV-doses, growth experiments were performed. In general, UV-irradiation had an influence on growth behavior of both strains, the parent strain MW001 and the mutant strain *tfb*3::*pyrEF*. The effect was dose dependent and a higher dose led to decreased growth of both strains (Fig. 32 B). Directly after UV-shock the mutant strain showed a significantly reduced growth compared to the parent strain (Fig. 32 B, Fig. 33). The doubling time decreased approximately two-fold at doses of 50 J/m^2 and 75 J/m^2 and four-fold for 200 J/m^2 compared to the control samples of the disruption mutant (Fig. 33). After 16 hours upon UV-shock *tfb*3::*pyrEF* cells restored their ability to grow, but reached stationary phase 20 hours later than the control reactions without irradiation (Fig. 32 B). The parent strain MW001 showed first growth defects 15 hours after UV-treatment (Fig. 32 B). As shown in Fig. 32 the mutant strain reached a higher optical density in stationary phase than the parent strain MW001. The growth medium for MW001 was supplemented with uracil because this strain is an uracil auxotroph mutant. The medium for the disruption strain was not supplemented with uracil, because the *pyr*EF cassette was inserted into the genome to disrupt the gene of interest. Therefore the *tfb*3::*pyrEF* mutant strain produces uracil and is not

dependent on uracil addition to the medium. Furthermore addition of uracil could be disadvantageous for the *tfb*3::*pyrEF* disruption strain because the *pyrEF* marker cassette might be lost due to less selective pressure. Therefore growth of the MW001 strain might be limited by uracil depletion, resulting in lower maximal optical densities compared to the disruption mutant. Taken together these data confirm a possible role of *tfb*3 in early-UV response.

The microarray experiment from Götz et al., 2007 was performed with a UV-dose of 200 J/m^2 and samples for RNA extraction were taken immediately after UV-treatment and every 30 min afterwards until two hours after UV-treatment. *Tfb*3 transcripts were upregulated (7-fold for *S. acidocaldarius*, $\log_2$ expression ratio of 2.82) shortly after UV-irradiation and also within these two hours a dramatically reduced growth of the *tfb*3::*pyrEF* strain was observed in the present study (Götz et al., 2007). *In vitro* transcription assays with strong promoters (T6, *ssb*) and promoters that were known to be activated (*dps*, *thsB*) or repressed after UV-irradiation (*sta1*) were performed (Paytubi and White, 2009). *In vitro* transcription was possible with all tested promoters and TFB3 binding to promoters was only identified in the presence of TFB1 and in addition a competition for binding of TFB1 and TFB3 to the RNAP has been observed (Götz et al., 2007, Paytubi and White, 2009). Immunoprecipitation and *in vitro* transcription studies revealed that binding of TFB3 to the TBP-DNA complex necessitates the presence of TFB1 for transcription (Paytubi and White, 2009). A concentration dependent stimulatory effect on transcription by TFB3 has been demonstrated and the authors suggest a role of TFB3 as transcription activator. No interaction between TFB1 and TFB3 was detectable by immunoprecipitation; however, *in vivo* interactions when TFB1 is already bound to DNA cannot be excluded (Paytubi and White, 2009). The *tfb*3::*pyrEF* mutant generated in this study was not lethal but showed a decreased growth rate upon UV-exposure (Fig. 32, Fig. 33). TFB1 is still present in the *tfb*3::*pyrEF* mutant strain and supports basal transcription rates. With regard to the results obtained in this study and from Paytubi and White (2009) TFB3 indeed might act as an activator for transcription in the presence of TFB1.

Especially organisms exposed to extreme environments, like many Archaea, suffer from various conditions inducing DNA damage (e.g. heat, oxygen, UV). Interestingly, the spontaneous mutation rate of *S. solfataricus* is comparable to that of Bacteria (Martusewitsch et al., 2000, Grogan et al., 2001, Grogan 2004). Several studies were performed on UV-damage and DNA damage-response in the thermoacidophiles *S. acidocaldarius* and *S. solfataricus* and the formation of photoproducts (cyclobutane pyrimidine dimer (CPD), 6,4 pyrimidine-pyrimidones (6-4 PP)) was observed as well as double strand breaks (Fröls et al., 2007). Homologous proteins involved in nucleotide excision repair, base excision repair and homologous recombination repair were found in Archaea, whereas mismatch repair has not been reported yet (Grogan, 2004). Archaea have a photolyase for blue-light dependent photoreactivation (Kiener et al., 1989, Fujihashi et al., 2006). For bacterial-type NER only few species encoding UvrABC homologs are found in the archaeal domain (e. g. *Methanobacterium thermoautotrophicum*, *Thermoplasma acidophilum*, *Methanosarcina mazei*) (Grogan, 2004). Mainly homologs of the eukaryal type XPB, XPD, XPF and XPG were identified. But the recognition proteins, XPA and XPC, are absent in Archaea,

suggesting another damage recognition mechanism. Another possibility of UV-induced DNA-damage repair is the homologous recombination repair in Archaea mainly mediated by Mre11 and Rad50 (Baliga et al., 2004, McCready et al., 2005, Delmas et al. 2009, Rolfsmeier et al., 2010).

Two independent microarray studies of *S. solfataricus* exposed to two different doses, 75 J/m^2 and 200 J/m^2, respectively, showed 90% overlap of upregulated genes. A UV-dose dependent difference in the microarrays from Götz et al. (2007) and Fröls et al. (2007) was observed (Fröls et al., 2009). In both studies no significant effect on genes encoding proteins, possibly acting in DNA damage response, was detected. For example neither *dpo*IV, the photolyase gene (*phr*B) nor NER genes were affected by UV-exposure. Instead genes that were up-regulated could be grouped to cell-cycle (*cdc*6-2, only upregulated in *S. solfataricus*), replication (*dpo*II), transcription (*tfb*3), pili-formation (*ups*-operon) and recombination (*mre*11, *rad*50, *rec*B-like nuclease) (Götz et al., 2007, Fröls et al., 2007, Fröls et al., 2009). The photolyase gene *phr*B was not affected by UV-irradiation because the cultures were incubated in the dark after UV-treatment (Fröls et al., 2007, Fröls et al., 2009). Transcript levels of genes implicated in NER were not changed upon UV-irradiation. Taken together it seems most likely that in the presence of light photoreactivation by photolyase (light-repair) is the favored repair pathway in *S. acidocaldarius* and *S. solfataricus*. In the absence of light DNA-damage seems to be rather curated by homologous recombination repair than by NER.

In addition to changing transcript levels upon UV-treatment it was observed that UV-irradiation of *S. solfataricus* cells lead to aggregation and pili-formation (Fröls et al., 2007, Fröls et al., 2009). This mechanism could facilitate DNA transfer and thus provide a mechanism for genome propagation. Fröls and co-workers investigated a dose-dependent cell aggregation following UV-exposure with seven different doses ranging from 5 J/m^2 to 1000 J/m^2. The lowest dose of 5 J/m^2 induced cellular aggregation and the highest aggregation state could be detected with 50 J/m^2 and 75 J/m^2, whereas doses of 200 J/m^2 and 1000 J/m^2 yielded in very low or even no aggregation (Fröls et al., 2007, Fröls et al., 2008). In *H. volcanii* and *S. solfataricus* cellular aggregation and subsequent DNA transfer has been reported (Rosenshine et al., 1989, Schleper et al., 1995, Fröls et al., 2008). Pili-formation and cellular aggregation were dependent actions. The *ups*-operon (UV-inducible type IV pili operon of *Sulfolobus*) encodes proteins for pili-assembly. This operon consists of a hypothetical protein UpsX, a secretion-ATPase UpsE, a transmembrane protein UpsF and two pilins UpsAB (Fröls et al., 2008, Vassart et al., 2013). In *S. acidocaldarius* it is assumed that four genes downstream of this operon might be involved in DNA repair following UV-irradiation. These genes encode an endonuclease (Saci_1497), a ParB-like protein (Saci_1498), a glycosyltransferase (Saci_1499) and a helicase (Saci_1500) (Vassart et al., 2013). Indeed endonucleases, glycosyltransferases and helicases are key players in DNA repair. Deletion of *ups*EF in *S. acidocaldarius* resulted in a decreased survival after UV-irradiation (Ajon et al., 2011). Cells not only aggregate, they also dissociate eight to ten hours post UV-treatment (Ajon et al., 2011). Chromosomal marker exchange has been reported in *S. acidocaldarius* and a conjugation-like system which requires cell-cell contacts was proposed (Grogan, 1996b, Ghané and Grogan, 1998, Grogan, 2009). Later on, a UV-inducible marker exchange has been observed in this organism (Wood et al., 1997, Schmidt et al.,

1999). The *ups* operon was shown to be a binding target of the transcriptional regulator Lrp from *S. acidocaldarius* (Saci_1588, *Sa*Lrp) (Vassart et al., 2013). A band shift was observed in EMSA studies using *Sa*Lrp and control regions of the genes *ups*X (Saci_1492), *ups*E (Saci_1493), *ups*A (Saci_1495), Saci_1497, Saci_1498 and Saci_1500. A *Sa*Lrp deletion mutant exhibited a reduced number of cells per aggregate and it can be speculated that DNA might be transferred via this pili and homologous recombination repair could maintain cell survival (Vassart et al., 2013).

It is known that the *tfb*3 transcript is up-regulated after UV-irradiation. UV-light induces a cascade of cellular processes to overcome DNA damage in *S. acidocaldarius*. TFB3 could probably bind to TBP, RNAP and maybe TFB1 upon UV-shock thereby activating transcription of genes involved in pili formation, homologous recombination repair and replication. A global model summarizing the cellular response to UV-damage in *S. acidocaldarius* is illustrated in Fig. 35.

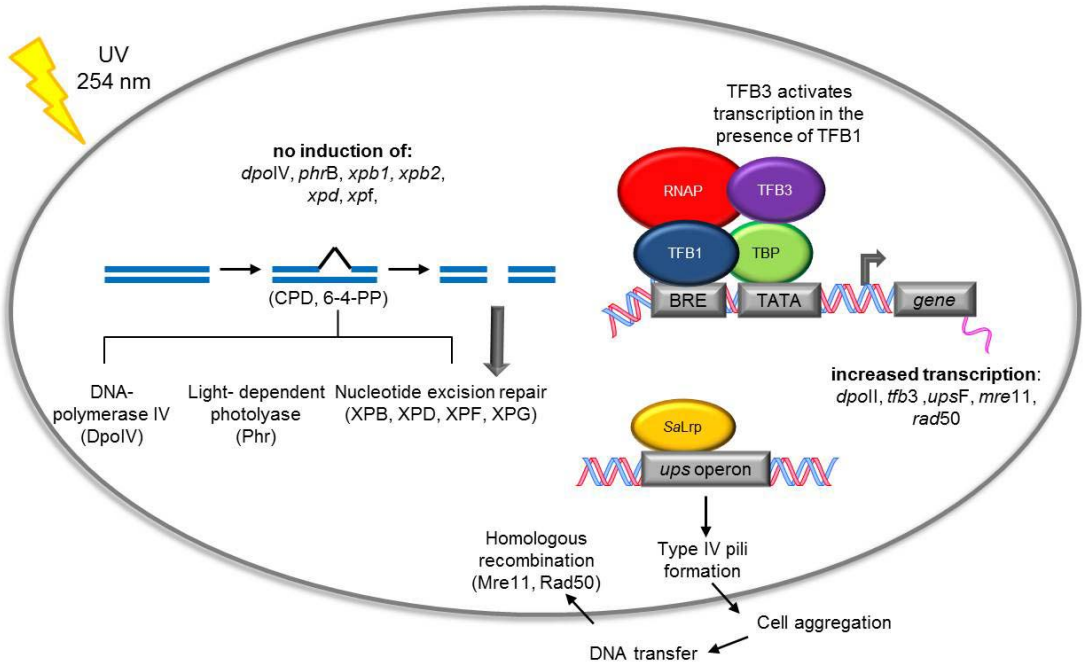


Fig. 35: Model of general UV-response in *S. acidocaldarius*.
UV-light induces formation of cyclobutane pyrimidine dimer (CPD) and 6,4 pyrimidine-pyrimidones (6-4-PP) photoproducts. If the photoproducts are not repaired before one round of replication double strand breaks occur. In general DNA-polymerase IV helps to bypass lesions. The photolyase solves dimer formation by using light-energy. Nucleotide excision repair proteins facilitate repair of CPDs and 6-4 photoproducts as well as double-strand breaks. In *S. acidocaldarius* pili-formation is induced by UV-light and the transcription regulator *Sa*Lrp promotes expression of the *ups*-operon (Saci_1492-1496) and adjacent genes downstream of the operon (Saci_1497-Saci_1500) by binding to an AT-rich sequence upstream of its promoter sequences. Pili formation results in cell aggregation and DNA can be transferred by a conjugation-like mechanism. This transferred DNA is then used in homologous recombination events to survive DNA damage. TFB3 probably binds to TBP and contacts the RNAP at the same docking sites as TFB1 and activates transcription in the presence of TFB1. It might also be possible that TFB3 contacts TFB1.

4.8. Multiplicity of general transcription factors in Archaea: A similar function as bacterial σ-factors?

Archaea possess homologs of the eukaryotic transcription factors TFIIB, TBP and TFIIE (Marsh et al., 1994, Langer et al., 1995, Kyrpides and Ouzounis 1999, Bell and Jackson, 2001, Hickey et al., 2002). Multiple homologs of TFB and TBP are found and depending on the organism and function of these transcription factors differs. Table 17 summarizes the number of TFB and TBP proteins in Archaea.

Table 17: Number of TFB and TBP paralogs of some representatives of all archaeal phyla

Phylum	Organism	TFB paralogs	TBP paralogs	Literature
Crenarchaeota	*Ignicoccus hospitalis*	2	1	Podar et al., 2008
	Sulfolobus acidocaldarius DSM 639	3	1	Chen et al., 2005
	Sulfolobus solfataricus P2	3	1	She et al., 2001
	Sulfolobus tokodaii str. 7	3	1	Kawarabayasa et al., 2001
	Metallosphaera sedula	2	1	Auernik KS, et al. 2008
	Pyrobaculum aerophilum	3	1	Fitz-Gibbon, et al. 2002
	Thermoproteus tenax Kra1	4	1	Siebers et al., 2011
Euryarchaeota	*Halobacterium salinarum* NRC-1	7	6	Ng et al., 2000
	Methanosarcina mazei Gö1	1	3	Deppenmeier et al., 2002
	Methanosarcina acetivorans	1	3	Galagan et al., 2002
	Pyrococcus abysii	2	1	Cohen et al., 2003
	Pyrococcus furiosus	2	1	Robb et al., 2001
	Thermococcus kodakarensis KOD1	2	1	Fukui et al., 2005
	Methanocaldococcus janaschii	1	1	Bult et al., 1996
	Haloferax volcanii DS2	8	4	Hartman et al., 2010
	Natrialba magadii ATCC 43099	7	1	Siddaramappa et al., 2012
Korarchaeota	*Korarchaeum cryptophilum*	2	1	Elkins et al., 2008
Nanoarchaeota	*Nanoarchaeumn equitans*	1	1	Waters et al., 2003
Thaumarchaeota	*Nitrosopumilus maritimus*	8	2	Walker et al., 2010
	Nitrososphaera gargensis	11	1	Spang et al., 2012
	Cenarchaeum symbiosum	5	3	Hallam et al., 2006

In the bacterial lineage adaptation to fast changing environmental conditions is mediated by transcription regulators and σ-factors. The bacterial core-RNAP consists of four subunits. Different σ-factors facilitate recognition and specific positioning of the RNAP to certain promoters. Sigma70 is the σ-factor found in all bacterial genomes and is sufficient for transcription of housekeeping genes (Paget and Helmann 2003, Gourse et al., 2006, Sachdeva et al., 2010). *E. coli* possesses up to six alternative σ-factors for adaptation for example to heat shock conditions (σ^{32}), stationary phase (σ^{S}) or extracytoplasmic stresses (σ^{E}) (Gourse et al., 2006). In *Streptococcus pyogenes* two additional σ-factors exist (Sachdeva et al., 2010). *Corynebacterium glutamicum* encodes seven σ-factors dedicated to different functions. But it has been observed that two or even more promoters are controlled by the same σ-factor in a number of genes (Patek and Nésvera, 2011). *Mycobacterium tuberculosis* has thirteen σ-factors and *Streptococcus coelicolor* up to sixty. A correlation between the number of σ-factors and the genome size was identified (Sachdeva et al., 2010). Transcription of *E. coli* genes is controlled by seven σ-factors and 300 transcriptional regulators (Ishihama, 2012). A

comparison of eukaryal and bacterial transcriptional response to stress demonstrates that the eukaryal transcription apparatus is much more complex than the bacterial transcription machinery (Sadhale et al., 2007). The author suggests that sigma factor-like regulatory role is distributed to many more proteins involved in transcription in Eukaryotes.

The phenomenon of multiple *tfb* and *tbp* paralogs is not only restricted to the domain of Archaea. Even in Eukaryotes several TFIIB-like and TBP-related factors have been investigated. Eukaryotes exhibit an immense number of transcription factors for sufficient transcription using all RNAPs. And some of these factors have specialized functions as can be seen for archaeal transcription factors and bacterial σ-factors. In addition to TFIIB, TFIIB-related factors (BRF) or plant-specific TFIIB-related proteins (BRP1, BRP2) or pollen expressed transcription factor (PTF2) have been investigated (Knutson, 2013, Cavel et al., 2011). BRF splicing variants play a role in RNAPIII transcription in yeast by targeting different promoters (McCulloch et al., 2000). *Arabidopsis thaliana* possesses up to fourteen different TFIIB-like factors (Knutson, 2013). BRP1 is a transcription factor probably involved in organelle-specific tasks and interacts with RNAPI (Lagrange et al., 2003, Imamura et al., 2008, Cavel et al., 2011). The transcription factor BRP2 may control endosperm growth because its expression is only detectable in reproductive organs (Cavel et al., 2011). *A. thaliana* PTF2 was found to bind to TBP2, RNAP and dsDNA and thereby seems to regulate pollen germination and embryogenesis (Niu et al., 2013). Early studies with *A. thaliana* and *Zea mays* revealed that two *tbp* genes are present. These two genes differed in their tissue accumulation pattern which indicated a tissue-specific expression (Vogel et al., 1993). Following experiments with *Arabidopsis* shoot tissues confirmed a tissue-specific expression of TBP2 suggesting a role in shoot production (Li et al., 2001). In *Drosophila melanogaster* the transcription factor TBP-related factor-1 (TRF1) has been shown to promote transcription of the *tudor* gene (Holmes and Tjian, 2000). The authors propose that these additional transcription factors have evolved for a tissue-specific regulation to control complex gene expression. Later on it was shown that TRF-1 is only present in *Drosophila* neuronal and germ cells (Persengiev et al., 2003). TRF-1 is able to bind TFIIA and TFIIB as well as recognizes and contacts the TATA-box (Crowley et al., 1993, Hochheimer and Tjian, 2003). Another transcription factor, TRF-2, has been identified in Metazoa (Rabenstein et al., 1999). It binds to TFIIA and TFIIB but fails to bind to the TATA-box (Rabenstein et al., 1999). The authors suggest that the evolution of multiple TBP proteins was necessary to orchestrate the amount of genes and complex expression patterns. An interesting hypothesis has been stated by Hochheimer and Tjian by suggesting that the TBP-related factors might act according to bacterial σ-factors (Hochheimer and Tjian, 2003). TRF-3 (TBP2) is a vertebrate-specific TBP-related protein which is able to interact with TFIIA, TFIIB and the TATA-box, like TBP and TRF-1. ChIP-experiments with mouse embryonic stem cells revealed a role in early mouse development of TBP2 (Yang et al., 2006). Studies with *Xenopus* and mouse showed higher expression levels of TBP2 in the oocytes compared to TBP (Persengiev et al., 2003, Gazdag et al., 2007, Akhtar and Veenstra, 2009, Müller and Tora, 2009).

But what is the role of a set multiple transcription factors in Archaea and do they act in stress response? The Euryarchaeon *H. salinarum* NRC-1 possesses seven TFBs and six TBPs that are encoded on the chromosome (*tfb*A, B, D, F, G; *tbp*E, F) or on the two minichromosomes

pNRC-100 (*tbp*A, B, C, D) and pNRC-200 (*tfb*C, E) (Baliga et al., 2000, Teufel et al., 2008). Some transcription factors of *H. salinarum* NRC-1 have multiple gene copies like *tbp*A (2), *tbp*B (4) and *tbp*C (2) on both minichromosomes and they differ in sequence from the transcription factors encoded on the chromosome. These alterations are mainly due to insertion-elements (Teufel et al., 2008). In theory these transcription factors could interact in fourty-two different combinations (Baliga et al., 2000). Knock-out strains of *tfb*A and *tbp*D resulted in reduced growth rate of the mutants at elevated temperatures. Microarray studies with these mutants showed a downregulation of genes involved in heat-shock response (Coker and DasSarma, 2007). Facciotti and colleagues mapped binding sites of *H. salinarum* NRC-1 TFB and TBP by ChIP-Chip (Facciotti et al., 2007). They demonstrated that more than half of the target genes were associated with more than one TFB and TBP. Nevertheless, the authors observed a preferential binding of TFBf to promoters of genes involved in ribosome biogenesis and TFBg to genes implied in phosphate transport and phototrophy related proteins. They hypothesize that the transcription factors compete for DNA-binding and thereby cause a negative regulatory effect. Deletion of the transcription factor *tfb*C from *H. salinarum* NRC-1 implicated a role in oxidative stress response (Kaur et al., 2010). For TFBe a function in adaptation to low temperatures was suggested (Turkarslan et al., 2011). Transcription of the *tfb*D, *tfb*A and *tfb*E genes was repressed by high oxygen concentration in *H. salinarum* NRC-1 (Schmid et al., 2007). Transcriptional analysis using qRT-PCR and growth experiments with three different *H. salinarum* strains (NRC-1, PHH1, PHH4) gained further insight into the variety of halobacterial transcription factors. The three different strains differ in the amount of *tfb* and *tbp* genes that are encoded on the chromosome and plasmids. Cold-shock reduced the amounts of all *tfb* genes except *tfb*F transcripts. Heat-shock upregulated the transcription of *tfb*A and *tfb*B in PHH1 and PHH4 dramatically but not in the strain NRC-1. *H. salinarum* NRC-1 showed a down regulation of these genes but an increase in *tfb*G transcript levels has been observed (Bleiholder et al., 2012). The strains *H. salinarum* NRC-1, PHH1 and PHH4 were tested for different mRNA levels of *tbp* genes during growth at aerobic and anaerobic conditions. In both strains *tbp*E was the dominant transcription factor at logarithmic growth phase (Teufel et al., 2008). PHH4 only possesses *tbp*E and showed a decreased growth rate during anaerobic growth conditions in stationary phase (Teufel et al., 2008). Another halophile, *Haloferax volcanii,* harbors six *tfb* genes and four *tbp* genes. Transcript levels of TFB2 increased upon heat shock at 60°C. Moreover, immunoblots with anti-TFB1 and anti-TFB2 antibody revealed a slight increase at protein level as well (Thompson et al., 1999).

In the methanogenic strain *M. mazei* it was shown that TBP bound stronger to heat-shock promoters during heat-shock conditions. The authors assumed that heat stress changes the TBP-binding activity (De Biase et al., 2002). The methanogen *M. acetivorans* harbors one TFB and three TBP proteins and qRT-PCR with cells grown on methanol, acetate or trimethylamine showed highest transcript levels for *tbp*1, followed by *tbp*2 and *tbp*3 (Reichlen et al., 2010). Transcript abundance of *tfb*1 was 20-fold greater compared to that of *tfb*2 for all three tested substrates and 100-fold to 200-fold higher than the transcript level of *tfb*3 grown on methanol or trimethylamine and acetate, respectively (Reichlen et al., 2010). TBP1 seemed to be essential but the *tbp*2 and *tbp*3 gene could be deleted. Growth experiments revealed a longer lag-phase of the mutants compared to the wild-type strain. Microarray experiments

with these mutant strains showed an altered regulation of 92 and 77 genes possibly acting in adaptation to environmental conditions (Reichlen et al., 2010). Transcription abundance of 28 of these genes was altered in both mutant strains. The rest of the genes could be grouped into many functional categories like amino acid transport and metabolism or encode hypothetical proteins. One example of the first category is the gene MA3046 which encodes an ABC-binding protein, a part of an ABC-transporter with a fold-change of 4.97. In the Δ*tbp*3 deletion mutant mostly genes encoding hypothetical proteins are upregulated (e. g. MA0546 11.36-fold change). The Euryarchaeon *P. furiosus* encodes two *tfb* genes and one *tbp* gene. Northern blot analysis revealed that the transcript levels of *tfb*2 (PF0687) arose following heat shock whereas the levels of TFB1 remained unaffected (Shockley et al., 2003). TFB2 functions poorly in transcription activation from strong promoters (Micorescu et al., 2008). Transcription studies with promoters known to be involved in heat-shock response demonstrated that TFB2 is not selective for heat-shock conditions (Micorescu et al., 2008). *T. kodakaraensis* harbors two TFB proteins and one TBP. Both *tfb* genes could be deleted implicating that these genes are not essential. These observations demonstrate that the TFB proteins from *T. kodakaraensis* are not restricted to several stress conditions and that they compensate their function (Santangelo et al., 2007).

Whereas a plethora of TFB and TBP homologs has been investigated in the euryarchaeal branch the function of transcription factors from Crenarchaeota remains poorly understood. Microarray studies with synchronized *S. acidocaldarius* cells showed a cyclic-induction of *tfb*2 at G_1/S-phase transition pointing to regulation of the cell-cycle (Lundgren and Bernander, 2007). The stability of mRNAs of *S. solfataricus tfb* genes differs greatly. The mRNA of *tfb*1 is more stable than that of *tfb*2 and *tfb*3 at standard conditions (Bini et al., 2002, Andersson et al., 2006). These data point to a specialized role of TFB2 and TFB3 in *S. solfataricus* but it cannot be excluded that these transcription factors might have overlapping functions. In *S. acidocaldarius* and *S. solfataricus* TFB1 seems to have a similar function like σ^{70} in Bacteria for regulation of housekeeping genes and TFB2 and TFB3 might have specialized functions like alternative transcription factors. As discussed above microarray studies with *S. acidocaldarius* and *S. solfataricus* revealed that transcript levels of TFB3 increased upon UV-irradiation (Götz et al., 2007, Fröls et al., 2007) pointing to a possible role of TFB3 from *Sulfolobales* in UV-response. Further on Paytubi and White (2009) demonstrated that TFB3 acts as an activator of transcription in the presence of TFB1. In this thesis the first *tfb*3 insertion mutant in Crenarchaeota has been generated. The mutant revealed a dose-dependent growth defect as early UV-response supporting the idea that TFB3 is a transcriptional activator.

Fifty-six out of over eighty-two sequenced archaeal genomes possess two or more TFB or TBP homologs (Turkarslan et al., 2011). A lineage specific expansion of general transcription factors as adaptation and specialization to the environment can be observed in some organisms. But the data also demonstrate that roles of transcription factors partly overlap and it is suggested that those that are not essential for cell viability were acquired by gene duplication (Turkarslan et al., 2011). To modulate transcription not only *tfb* and *tbp* proteins are required as well as a multitude of transcription regulators. The multiplicity of transcription factors probably arose mainly by gene duplication and horizontal gene transfer to survive in

challenging environments. Comparison of GTFs in all three domains revealed that the amount of transcription factors correlates with genome size (Pèrez-Rueda et al., 2004, Coulson et al., 2007, Charoensawan et al., 2010, Martínez Núñez et al., 2013). Horizontal gene transfer of housekeeping genes seems to occur more frequently than of genes involved in transcription and translation (Jain et al., 1999). A recent hypothesis assumes that lack of some transcription factors can be bypassed by the presence of histones and histone-like proteins (Pèrez-Rueda et al., 2004). Also transcriptional regulators are of importance by altering the response according to environmental conditions to facilitate survival of the cell. These regulators block access of the RNAP to the promoter or binding of transcription factors or promote transcription by recruiting transcription factors (Ouhammouch, 2004, Bell, 2005, Geiduschek and Ouhammouch, 2005). Even Eukaryotes possess additional multiple GTFs and transcriptional regulators to facilitate tissue-specific gene expression. Just in regard of some specialized roles it is tempting to speculate that multiple GTF homologs act according to specialized sigma factors for an adaptation to survival in different environments. This hypothesis will be investigated by further functional analysis of transcription factors especially in the crenarchaeal branch.

Fine tuning of transcription is not only facilitated by a multiplicity of GTFs and the presence of various transcriptional regulators it is also interplay between mRNA stability, RNA processing and post-translational modification.

Within the last years more studies focused on detection of non-coding RNAs (nc-RNAs). Nc-RNAs are not translated into a protein and lack an open reading frame. Their size ranges from twenty to five-hundred nucleotides. They are involved in gene silencing, RNA processing and chromosomal architecture (Eddy, 2001, Tang et al., 2005, Mattick and Makunin, 2006). Three studies report the presence of non-coding-RNAs (ncRNA) in *S. solfataricus* (Tang et al., 2005, Zago et al., 2005, Wurtzel et al., 2010). RNA-sequencing experiments with four *Pyrobaculum* species (*P. arsenaticum, aerophilum, calidifontis, islandicum*) exhibited three novel cis-antisense non-coding RNAs and one of this new RNAs was encoded opposite to the *tfb*1 gene (Bernick et al., 2012). The role of ncRNAs in transcription regulation is still not fully understood but it might play an important role in fine tuning of transcript levels of GTFs.

Gene regulation of cellular processes can occur on transcript level by mRNA stability and also post-translational at protein level. Some common post-translational modifications are acetylation, glycosylation, methylation, phosphorylation and ubiquitination (Mann and Jensen, 2003). Archaea comprise eukaryotic-like Ser/Thr kinases and phosphatases. Two component systems phosphorylating histidine residues are limited to Euryarchaeota (Smith et al., 1997, Kim and Forst, 2001, Kenelly, 2003, Esser et al., 2012). Studies of the phosphoproteome of *S. solfataricus* grown on glucose and tryptone revealed a phosphorylation of TFB1, TFB3 and TBP when cells were grown with glucose. TFB1 (SSO0446) and TFB3 (SSO0280) harbored two phosphorylation sites at serine and tyrosine within the sequences 1-ML<u>Y</u>L<u>S</u>EENK-9 and 114-<u>S</u>KLGLLI<u>Y</u>ER-123, respectively. TBP was phosphorylated at tyrosine in the peptide 75-K<u>Y</u>GIKIVGKPK-85. Phosphorylation of TFB2 or TFE was not detected in this study (Esser et al., 2012). It is tempting to speculate that the phosphorylation state of TBP regulates transcription by changing the protein conformation

and maybe binding behavior. In *S. acidocaldarius* MW001 two protein phosphatases were deleted and the phosphorylation patterns of all strains were compared (Reimann et al., 2013, submitted). TFB3 was phosphorylated at serine and tyrosine residues (187-VPSQTLKTM-195, 10-EIVWDYK-16) in the *ptp* deletion strain but not in the reference strain MW001 (Reimann et al., 2013, submitted) indicating that TFB3 is post-translationally modified. TFB3 from *S. acidocaldarius* is phosphorylated at other serine or tyrosine residues as *Sso*TFB3. Therefore post-translational modification might play an additional role in regulation of transcription.

In this study GTFs from *S. acidocaldarius* were successfully cloned for heterologous expression in *E. coli* and protocols for overexpression and purification were established. These proteins are the basis for further experiments and can be applied in EMSA experiments, *in vitro* transcription assays or immunoprecipitation. Furthermore in this study generated specific polyclonal anti-TFB1 and anti-TFE antibodies can be applied to determine the endogenous protein level in response to environmental stresses in the future. The RNAP was successfully purified after genomic integration of a His-tag at the RpoG subunit and is a pre-requisite for *in vitro* transcription assays. Protein interactions between TFBs, RNAP and TBP were investigated by yeast-two hybrid analysis. The present study gave putative implications for TFB2 in cell-cycle and rRNA modification as well as protein biosynthesis in response to different growth phases. This idea was derived by genome context analysis and operon determination. Additionally, this study supports the idea that TFB3 from *S. acidocaldarius* is a transcriptional activator which has been shown by *in vivo* functional analysis of the growth behavior of a *tfb*3::*pyrEF* disruption mutant in response to three defined UV-doses. In general the study provides a fundamental basis (e.g. recombinant proteins, His-tagged RNAP, reporter gene constructs) to elucidate the function of TFB homologs in *S. acidocaldarius* in the future by various methods.

5. Summary

The archaeal transcription machinery encompasses one multi-subunit RNA polymerase (RNAP), resembling the RNAPII of Eukaryotes, homologs of the TATA-binding protein (TBP), transcription factor TFIIB (TFB) and TFIIE-α (TFE) (Langer and Zillig, 1993, Marsh et al., 1994, Langer et al., 1995, Kyrpides and Ouzounis 1999, Bell and Jackson 1998a, b, Bell and Jackson, 2001, Hickey et al., 2002, Geiduscheck and Ouhammouch, 2005, Werner and Grohmann, 2011). The current mechanistic understanding of transcription initiation is that TBP binds to the TATA-Box (~25 bp upstream of the transcription start site) whereupon TFB binds to the TBP-DNA complex forming sequence specific contacts with a purine-rich TFB-responsive element (BRE). Subsequently, the N-terminus of TFB recruits the RNAP to build the ternary pre-initiation complex. RNAP, TBP and TFB are solely sufficient for transcription of archaeal promoters *in vitro* (Bell and Jackson 1998a, b, Bell and Jackson, 2001, Hickey et al., 2002, Geiduscheck and Ouhammouch, 2005). Therefore archaeal transcription is generally regarded as a simpler model of the more complex eukaryal processes (Baliga et al. 2000).

Intriguingly, most Archaea possess multiple homologs of general transcription factors (GTFs) however, their number is species-dependent. Whereas the function of multiple GTFs has been addressed in different Euryarchaeota (e.g. *P. furiosus, H. salinarum* NRC-1) (Thomsen et al., 2001, De Biase et al., 2002, Facciotti et al., 2007, Coker and DasSarma, 2007, Santangelo et al., 2007, Micorescu et al., 2008, Turkarslan et al., 2011) the role in Crenarchaeota is still unclear.

The current study focusses on the thermoacidophile Crenarchaeote *Sulfolobus acidocaldarius* which grows optimally at 80°C and a pH of 2-3 (Brock et al., 1972). This organism possesses three TFBs and one TBP (Chen et al., 2005). The aim of this project was to provide all proteins required for EMSAs and *in vitro* transcription system and finally elucidate the function of multiple transcription factors, by the creation of reporter gene constructs and knock-out mutants. In order to address the question if TFB1, TFB2 and TFB3 might be involved in stress response in *S. acidocaldarius* as has been previously proposed for bacterial σ-factors (Baliga, 2000, Micorescu et al., 2008, Lu et al., 2008, Peng et al., 2009, Paytubi and White, 2009, Turkarslan et al., 2011).

The transcription factors (TFB1 (codon adaptation), TBP, TFE) and RNAP subunit RpoG could be successfully overexpressed in *E. coli* and purified to homogeneity. TFB2 and TFB3 (both codon-adapted) were expressed as insoluble proteins in *E. coli* and were purified under denaturing conditions; suitable conditions for refolding still need to be established. Polyclonal antibodies (rabbit) were generated against recombinant TFB1 and TFE and are now available for determination of endogenous protein levels and ChIP-Seq experiments. Unexpectedly, homologous expression of TFB1, TFB2, TFB3 and TBP in *S. acidocaldarius* using the *S. acidocaldarius* expression system (pSVA1450, maltose-inducible promoter) was not detected, which might be due to instable mRNA (transcript level) or low cellular protein level. As alternative approach a Strep-FLAG tag was integrated into the 5'- or 3'-genomic region of TFB1, TFB2, TFB3 and TFE for purification of the transcription factors from *S. acidocaldarius*. Endogenous TFB1 and TFE were successfully purified via the integrated

genomic tags and these mutant strains are now available for studies of protein levels under different stress conditions *in vivo*. Interestingly, TFB2 and TFB3 could not be detected after integration of a genomic tag suggesting low expression rates under the tested growth conditions.

A pre-requisite for future experiments like *in vitro* transcription is the availability of RNAP from *S. acidocaldarius*. The RNAP was successfully purified from *Sulfolobus* cells after integration of a His-tag on the *rpoG* subunit and a co-purification of TFB1 with RNAP was observed using in this study generated anti-TFB1 antibody. Therefore all components of the transcription apparatus, the GTFs and the RNAP of *S. acidocaldarius*, are now available for future experiments.

In order to investigate a possible function of TFB2 the genomic context was analyzed. A conserved gene context of *tfb*2 was observed in many Archaea and for *S. acidocaldarius* the formation of an operon with *gar*1 (Saci_1340), *tfb*2 (Saci_1341) and Saci_1342 was predicted. Comparison of upstream sequences of *gar*1 and *tfb*2 from fourteen sequenced crenarchaeal organisms revealed the presence of typical promoter elements (BRE/TATA) upstream of *gar*1 and *tfb*2 and an additional conserved ribosomal binding site (GAGG) for *tfb*2. RNA was isolated from two *S. acidocaldarius* strains (DSM 639, MW001) and cDNA was generated using reverse transcriptase. PCR with this cDNA revealed no full-length transcript of the whole operon (*gar*1-*tfb*2-RPC34 homolog encoding gene). However, co-transcription with its upstream gene *gar*1 (*gar*1-*tfb*2) in a growth-phase dependent manner and alternatively with Saci_1342 (*tfb*2-Saci_1342) was observed. This suggests that *tfb*2 possesses an additional promoter although it is part of an operon. The genomic context and co-transcription imply a possible role in transcription as well as rRNA modification.

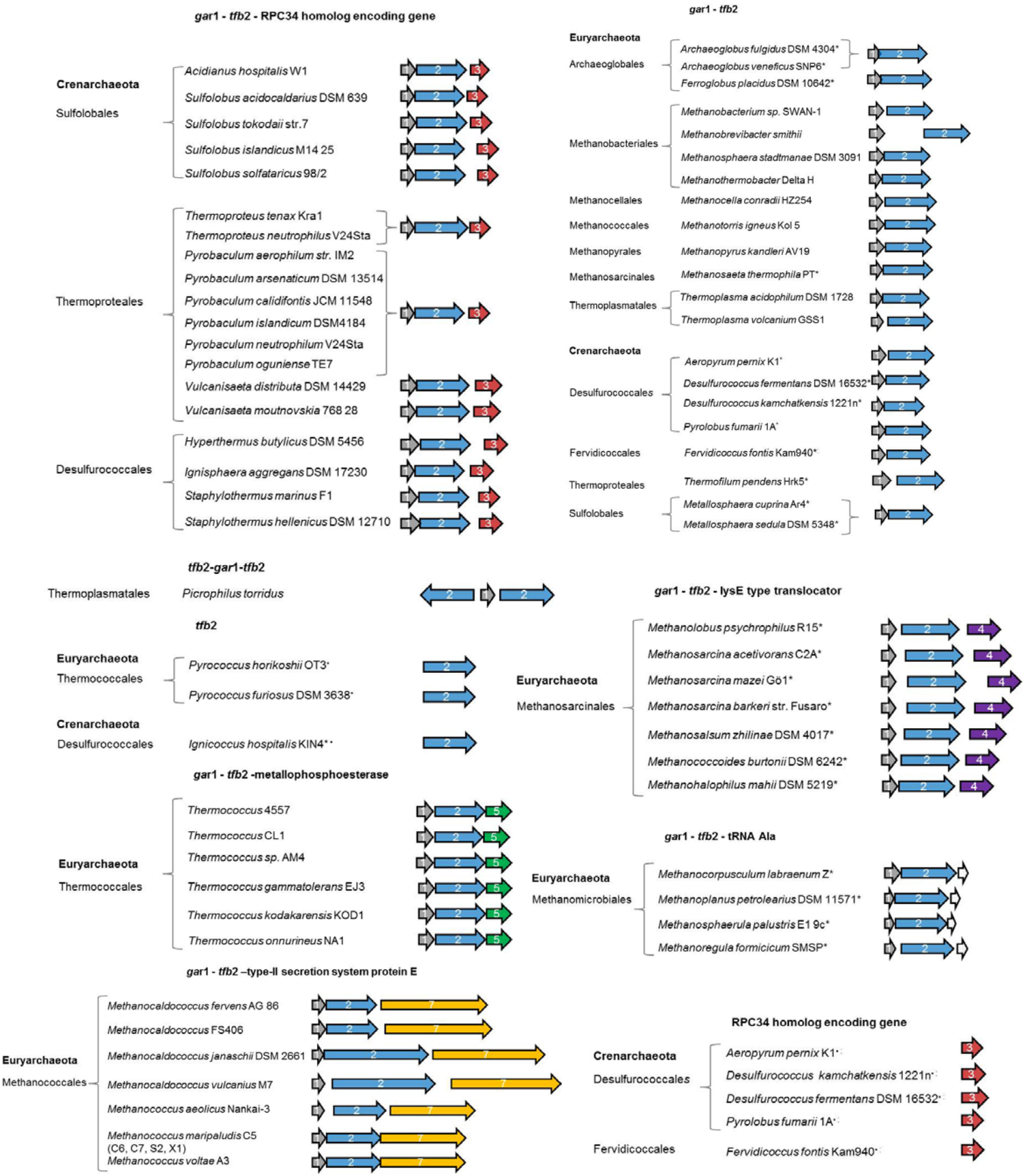
Because endogenous TFB2 and TFB3 could not be purified via homologous expression or by in-genome tagging it seemed most likely that these proteins are not present at standard growth conditions. To elucidate growth conditions where TFB2 and TFB3 are present, reporter gene constructs of promoter regions of the GTFs *tfb*1, *tfb*2, *tfb*3, *tbp*, and *tfe* from *S. acidocaldarius* and *gar*1 (-1 to -303 bp, for *tfb*2 -1 to -269 bp), controlling β-galactosidase activity, were generated. First reporter gene assays with the *tfb*1, *gar*1 and *tfb*3 promoter region in response to cold shock revealed that only expression of the *tfb*1 promoter was changed and lead to higher β-galactosidase activity (+ 8625 mod. Miller units after 60 min). However, also expression of the *gar*1 (+ 4296 mod. Miller units after 60 min) and *tfb*3 promoter (+ 2259 mod. Miller units after 60 min), although reduced compared to *tfb*1, was detected. Future reporter gene assays with promoter constructs generated in this study will expand the knowledge of function of multiple transcription factors in *S. acidocaldarius*.

To study protein-protein interactions between GTFs with RNAP subunits and GTFs with TBP of *S. acidocaldarius* yeast two-hybrid experiments were performed. Protein-protein interactions between TFB2 and RpoE and TFB3 with RpoE, RpoK and TBP were observed. However, protein expression could only be verified for RpoA2 and RpoB by immunodetection with anti-HA-tag antibody.

Finally, to unravel the *in vivo* function of the general transcription factors the aim was to generate deletion mutants of *tfb*1, *tfb*2 and *tfb*3 by two different strategies (pop-in/pop-out (Wagner et al., 2009, Wagner et al., 2012), PCR-based disruption (Sakofsky et al., 2011)). Despite all efforts no *tfb*1 and *tfb*2 deletion strains could be generated. It is most likely that these genes play an essential role in *S. acidocaldarius* and deletion is lethal. For *tfb*3 a knock-out by the pop-in/pop-out strategy was not stable, but a *tfb*3 disruption mutant (*tfb*3::*pyrEF*) was generated. Growth experiments with the parent strain MW001 and the mutant strain *tfb*3::*pyr*EF after irradiation with defined UV-doses (50, 75, 200 J/m^2) showed effects in early UV-response but growth was restored resembling the parent strain after 16 hours upon UV-treatment. These data, combined with results obtained from previous studies in *S. acidocaldarius* and *S. solfataricus* (Fröls et al., 2007, Götz et al., 2007, Paytubi and White, 2009), point to a role of TFB3 as a transcriptional activator by enhancing transcription in the presence of TFB1 by binding to TBP and RNAP.

Taken together this study provides an important basis (e.g. recombinant proteins, His-tagged RNAP, reporter gene constructs) for further experiments to unravel the function of GTFs in *S. acidocaldarius* and in Crenarchaeota.

6. Supplement



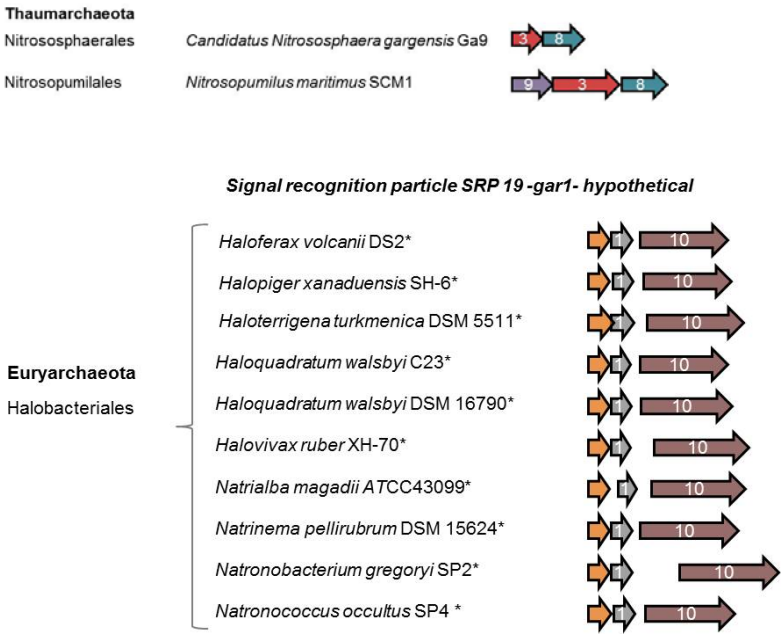


Fig. 36: Conserved genomic neighborhood of *gar*1, *tfb*2 and RPC34 homolog encoding gene (complete version).
The amino acid sequences of Gar1 (Saci_1340), TFB2 (Saci_1341) and the RPC34 homolog (Saci_1342) were analyzed by the program SyntTax (http://archaea.u-psud.fr/synttax/) for analysis of the genomic context. 1, H/ACA ribonucleoprotein *gar*1; 2, transcription factor *tfb*2; 3, RPC34 homolog encoding gene; 4, lysE type translocator; 5, metallophosphoesterase; white arrows, tRNA Alanine; 7, type-II secretion system protein; 8, methyltransferase; 9, hypothetical protein; 10, hypothetical protein; orange arrow, signal recognition particle SRP19. An asterisk behind the species name indicates the presence of a gene encoding the RPC34 homolog and a point represents a Gar1 encoding gene elsewhere in the genome. TFB2 homologs are represented by a colon. It was not possible to distinguish which halobacterial TFB is homologous to TFB2 from *S. acidocaldarius* by sequence comparison.

Codon optimized sequence of *tfb*1 for expression in *E. coli*, 933 bp, in pUC57, underlined: *Nde*I, *Bam*HI

CATATGGTGGAACAGTCTAAAGTTCCGAGCTCTAGTCTGTGCCCGCCGGATAAAATTATC
TTTGATGAAGAACGTGGCGAATATATCTGTACCGAAACGGGTGAAGTGATTGAAGAACGT
ATTATCGATCAGGGTCCGGAATGGCGTGCATTCACCCCGGAAGAAAAAGAAAAACGTAGC
CGTGTGGGCGGTCCGCTGAACCAGACCATTCATGATATGGGCATTAGCACGGTTATCGAT
TGGAAAGATAAAGATGCCATGGGTCGCTCTCTGGACCCGAAACGTCGCCTGGAAGTGCTG
CGTTGGCGCAAATGGCAGATTCGTACCCGCATCCAGAGCTCTATTGATCGTAACCTGGCA
CAGGCGATGAATGAACTGGAACGCATCGGTAACCTGCTGAATCTGCCGAAAGCCGTTAAA
GATGAAGCGGCCCTGATTTATCGTAAAGCAGTGGAAAAAGGCCTGGTTCGTGGTCGCTCT
ATCGAAAGTGTGGTTGCAGCGAGCATTTACGCCGCATGCCGTCGCATGAAAATGGCGCGC
ACCCTGGATGAAATCGCGCAGTTTACGAAAGCCAACCGTAAAGAAGTGGCCCGTTGTTAT
CGCCTGATCCTGCGCGAACTGGATATTGAAGTGCCTGTTAGTGATCCGAAAGATTACGTT
ACCCGTATCGGCACGCTGCTGAGTCTGAGCGGTATTACCATGAAACACGCGGCCGAAATT
ATCGAAAAAGCGAAAAATAGCGGCCTGACGGCCGGTAAAGATCCGGCCGGTCTGGCAGCC
GCAGCAATTTATATCGCGGCCCTGCTGAACGATGAACGTCGCACCCAGAAAGAAATTGCC
CAGGTGGCAGGTGTTACCGAAGTGACGGTTCGTAATCGCTACAAAGAACTGACGCAGGAA
CTGAAAATTCAGATCCCGAGCCAGTAAGGATCC

Alignment of codon optimized sequence of *tfb*1 (Sbjct) and the original genomic (Query) sequence

```
 Score =  480 bits (532),  Expect = 1e-139
 Identities = 661/924 (72%), Gaps = 0/924 (0%)
 Strand=Plus/Plus

Query  1    ATGGTGGAACAGTCTAAAGTTCCGAGCTCTAGTCTGTGCCCGCCGGATAAAATTATCTTT  60
            ||||| |||||   ||| |||||    ||    || || || || || || || || |||
Sbjct  1    ATGGTAGAACAAAGTAAGGTTCCTTCTTCCTCCCTATGTCCACCTGACAAGATAATATTT  60

Query  61   GATGAAGAACGTGGCGAATATATCTGTACCGAAACGGGTGAAGTGATTGAAGAACGTATT  120
            ||||||||||||||| |||||||| ||||| || || |||||||| || |||||| | ||
Sbjct  61   GATGAAGAACGTGGAGAATATATTTGTACTGAGACAGGTGAAGTAATAGAAGAAAGAATA  120

Query  121  ATCGATCAGGGTCCGGAATGGCGTGCATTCACCCCGGAAGAAAAAGAAAAACGTAGCCGT  180
            || ||||| ||||| || ||| | ||||| || || ||||| ||||| ||  | ||  |
Sbjct  121  ATTGATCAAGGTCCTGAGTGGAGGGCATTTACTCCAGAAGAGAAAAGAGAAGAGGAGTAGA  180

Query  181  GTGGGCGGTCCGCTGAACCAGACCATTCATGATATGGGCATTAGCACGGTTATCGATTGG  240
            || || || ||  | || || || || |||||||||||| ||    || ||||| ||||||
Sbjct  181  GTAGGTGGACCTTTAAATCAAACAATCCATGATATGGGAATATCTACTGTTATAGATTGG  240

Query  241  AAAGATAAAGATGCCATGGGTCGCTCTCTGGACCCGAAACGTCGCCTGGAAGTGCTGCGT  300
            ||||| |||||||| |||||  | ||| | ||||| ||  |  |  | ||||| ||  |
Sbjct  241  AAAGACAAAGATGCTATGGGAAGGTCTTTAGACCCTAAGAGGAGATTAGAAGTACTAAGA  300

Query  301  TGGCGCAAATGGCAGATTCGTACCCGCATCCAGAGCTCTATTGATCGTAACCTGGCACAG  360
            ||| | ||||||||||  | ||  | || ||   ||| || ||  | ||||| |||||
Sbjct  301  TGGAGAAAATGGCAGATAAGGACAAGGATTCAATCCTCAATAGACAGAAACCTAGCACAA  360

Query  361  GCGATGAATGAACTGGAACGCATCGGTAACCTGCTGAATCTGCCGAAAGCCGTTAAAGAT  420
            || ||||| || || ||  | || |||||  | || ||| |||| ||||| ||||| ||
Sbjct  361  GCTATGAACGAGCTTGAGAGAATTGGTAATTTACTCAATTTGCCTAAAGCAGTTAAGGAC  420

Query  421  GAAGCGGCCCTGATTTATCGTAAAGCAGTGGAAAAAGGCCTGGTTCGTGGTCGCTCTATC  480
            ||||| || || |||||  | ||||| || || ||||| || ||  | ||  |   |||
Sbjct  421  GAAGCTGCTCTAATTTACAGAAAAGCCGTAGAGAAAGGACTAGTCAGAGGAAGAAGTATT  480

Query  481  GAAAGTGTGGTTGCAGCGAGCATTTACGCCGCATGCCGTCGCATGAAAATGGCGCGCACC  540
            || ||||| ||||||||    ||||| || |||||  |  | ||||||||||  | ||
Sbjct  481  GAGAGTGTCGTTGCAGCTTCAATTTATGCAGCATGTAGAAGAATGAAAATGGCTAGAACA  540

Query  541  CTGGATGAAATCGCGCAGTTTACGAAAGCCAACCGTAAAGAAGTGGCCCGTTGTTATCGC  600
            || || ||||| || || || || || || ||  | |||||||| ||| | |||||| |
Sbjct  541  CTAGACGAAATAGCCCAATTCACTAAGGCTAATAGAAAAGAAGTTGCCAGATGTTATAGG  600

Query  601  CTGATCCTGCGCGAACTGGATATTGAAGTGCCTGTTAGTGATCCGAAAGATTACGTTACC  660
            || || ||| | ||| | || || || || || ||||||||||| || ||||||||||
Sbjct  601  CTAATACTGAGAGAATTAGACATAGAGGTACCAGTTAGTGATCCTAAGGATTACGTTACT  660

Query  661  CGTATCGGCACGCTGCTGAGTCTGAGCGGTATTACCATGAAACACGCGGCCGAAATTATC  720
             | || || || ||  | ||| | || ||||| || ||||| || || || || || ||
Sbjct  661  AGAATAGGTACTCTATTAAGTTTAAGTGGTATAACAATGAAGCATGCCGCAGAGATAATA  720

Query  721  GAAAAAGCGAAAAATAGCGGCCTGACGGCCGGTAAAGATCCGGCCGGTCTGGCAGCCGCA  780
            ||||| || || |||   ||  | || || |||||||||| || ||||| || || ||
Sbjct  721  GAAAAGGCTAAGAATTCAGGATTAACTGCTGGTAAAGATCCCGCGGGTCTAGCGGCTGCT  780

Query  781  GCAATTTATATCGCGGCCCTGCTGAACGATGAACGTCGCACCCAGAAAGAAATTGCCCAG  840
            |||||||||| || ||  |  |||| |||||| |  | || ||||| |||||||| |||
Sbjct  781  GCAATTTATATAGCAGCATTATTGAATGATGAAAGAAGAACACAGAAGGAAATTGCACAG  840

Query  841  GTGGCAGGTGTTACCGAAGTGACGGTTCGTAATCGCTACAAAGAACTGACGCAGGAACTG  900
            || || || || || ||||| |||||| | ||| | || |||||  | || || ||| |
Sbjct  841  GTAGCTGGCGTAACGGAAGTTACGGTTAGGAATAGATATAAAGAGTTAACTCAAGAATTA  900

Query  901  AAAATTCAGATCCCGAGCCAGTAA  924
            || || ||||| || || || |||
Sbjct  901  AAGATACAGATTCCAAGTCAATAA  924
```

Codon optimized sequence of *tfb*2 for expression in *E. coli*, 885 bp, in pUC57, underlined: *Nde*I, *Bam*HI

CATATGAAATGCAAAATCTGTGGCAGCGAATCTCTGATTTTTGATCGTGAACGCGGTATC
TACGTGTGCATTAACTGTGCCTCTGTTGATGATGAACCGCTGATTGATCAGGGCCCGGAA
TGGCGTGCATATACCACGGAAGATAAAGTGGAACGTGAACGCACCGGTAGCCCGCTGACG
GCAAAAGTTCATGATTTCGGCATCACCACGAAAATTGGTTACACCAAAATCAAAGATCGC
ATTAAAGTGCACAAACTGCGTCTGATGCAGAATAAAATCCGCGTTAGCGCGCGTGAACGC
AAACTGGTGACCTATCTGTCTGTTCTGAACAGCGAAGCATCTAAACTGAATCTGCCGGAA
CATGTGAAAGAAACGGCGAGTATTCTGATCCGTCGCCTGATTGAAGAAGGCAAAGCGAAA
CGTGTTGAAATGTACGCCCTGATCGCGGCCGTGATTTATTACAGCTGCCAGGTTAACCGC
ATCCCGAAACTGCTGAATGAAATTAAAACCCTGTATAGTCTGAGCCAGGCAGATCTGTGG
AAAGCGCTGGAAAAAGTGCAGGAAGTTGCCAAAAGCGTGAAAGTTAAACCGAACGTGACC
CCGATTGAATATATCCCGAAAATTACCGAACGTCTGGGTCTGCCGGCCTACGTGAGCACC
AAAGCAAGCGAACTGGTTGATATCATGTATAAAAATGGTCTGACCAGCGGCAAAGGTTAT
ACGGCGCTGGCAGCGGCCTCTGTGTACCTGATCAGTACCCTGATGGATGTGAAGAAAACC
CAGAAAGAAATTGCGGATTCTCTGAGTATCACCGAAGTGACGATTCGTAACCGCTACCGT
GAAATCATCAAAGCCTTCGATATCGAAGTTAAACTGTAAGGATCC

Alignment of codon optimized sequence of *tfb*2 (Sbjct) and the original genomic (Query) sequence

```
 Score =  486 bits (538),  Expect = 3e-141
 Identities = 633/873 (73%), Gaps = 2/873 (0%)
 Strand=Plus/Plus

Query  1    ATGAAATGCAAAATCTGTGGCAGCGAATCTCTGATTTTTGATCGTGAACGCGGTATCTAC  60
            ||||||||| ||||||||||||||| ||   ||| || |||||| | ||| | || || ||
Sbjct  1    ATGAAATGTAAAATCTGTGGCAGTGAGAGTCTTATCTTTGATAGAGAAAGAGGAATATAT  60

Query  61   GTGTGCATTAACTGTGCCTCTGTTGATGATGAACCGCTGATTGATCAGGGCCCGGAATGG  120
            || || ||||| || || ||||| || |||||||| || || ||||| || || ||||||
Sbjct  61   GTTTGTATTAATTGCGCATCTGTAGACGATGAACCACTAATAGATCAAGGACCAGAATGG  120

Query  121  CGTGCATATACCACGGAAGATAAAGTGGAACGTGAACGCACCGGTAGCCCGCTGACGGCA  180
             | |||||||| || ||||| ||||| ||| | ||| | || |||   || || || ||
Sbjct  121  AGAGCATATACTACCGAAGACAAAGTAGAAAGAGAAAGAACAGGTTCTCCTCTAACAGCC  180

Query  181  AAAGTTCATGATTTCGGCATCACCACGAAAATTGGTTACACCAAAATCAAAGATCGCATT  240
            ||||| ||||| || || || || || ||||| || || || ||||||||||| || ||
Sbjct  181  AAAGTACATGACTTTGGTATAACTACTAAAATAGGATATACTAAAATCAAAGACCGAATA  240

Query  241  AAAGTGCACAAACTGCGTCTGATGCAGAATAAAATCCGCGTTAGCGCGCGTGAACGCAAA  300
            ||||| || || ||   | ||||| ||||||||  | || || ||  | ||  | |||
Sbjct  241  AAAGTACATAAGCTAAGATTAATGCAAAATAAAATTAGAGTAAGTGCAAGAGAGAGGAAA  300

Query  301  CTGGTGACCTATCTGTCTGTTCTGAACAGCGAAGCATCTAAACTGAATCTGCCGGAACAT  360
             | || || ||||| || || || ||||| |||||||| ||| | ||  |||| |||||
Sbjct  301  TTAGTTACTTATCTATCAGTCCTAAACAGTGAAGCATCGAAATTAAACTTGCCTGAACAC  360

Query  361  GTGAAAGAAACGGCGAGTATTCTGATCCGTCGCCTGATTGAAGAAGGCAAAGCGAAACGT  420
            || |||||||| |||   || || ||  | ||  | || |||||||| || || ||| |
Sbjct  361  GTTAAAGAAACCGCGTCAATACTAATAAGACGATTAATAGAAGAAGGTAAGGCTAAAAGG  420

Query  421  GTTGAAATGTACGCCCTGATCGCGGCCGTGATTTATTACAGCTGCCAGGTTAACCGCATC  480
            || |||||||| ||  | || || || || || || ||    |||||||| ||  | ||
Sbjct  421  GTAGAAATGTATGCGTTAATAGCAGCGGTCATATACTATTCTTGCCAGGTAAATAGGATT  480

Query  481  CCGAAACTGCTGAATGAAATTAAAACCCTGTATAGTCTGAGC-CAGGCAGATCTGTGGAA  539
            || ||| |||| |||||||| ||||| || ||||| || | | ||||| ||| | |||||
Sbjct  481  CCCAAATTGCTTAATGAAATAAAAACACTCTATAG-CTTATCACAGGCGGATTTATGGAA  539

Query  540  AGCGCTGGAAAAAGTGCAGGAAGTTGCCAAAAGCGTGAAAGTTAAACCGAACGTGACCCC  599
            ||| || || || || || || ||||| || || || ||||| || |||||||| || ||
Sbjct  540  AGCTCTTGAGAAGGTACAAGAGGTTGCTAAGAGTGTCAAAGTAAAGCCGAACGTTACTCC  599

Query  600  GATTGAATATATCCCGAAAATTACCGAACGTCTGGGTCTGCCGGCCTACGTGAGCACCAA  659
            ||| |||||||| || ||||| || ||| |  ||||  | || || || ||    || ||
Sbjct  600  GATAGAATATATTCCTAAAATAACAGAAAGATTGGGATTACCTGCATATGTCTCAACAAA  659

Query  660  AGCAAGCGAACTGGTTGATATCATGTATAAAAATGGTCTGACCAGCGGCAAAGGTTATAC  719
             ||    ||  | || ||||| |||||||||| ||  | |||||||| ||||||||||
Sbjct  660  GGCTTCAGAGTTAGTAGATATTATGTATAAAAACGGATTAACCAGCGGTAAAGGTTATAC  719

Query  720  GGCGCTGGCAGCGGCCTCTGTGTACCTGATCAGTACCCTGATGGATGTGAAGAAAACCCA  779
             || || || || |||  |||||| || || |||||  | |||||||| ||||||||  ||
Sbjct  720  AGCACTTGCGGCTGCCAGTGTGTATCTTATAAGTACTTTAATGGATGTAAAGAAAACTCA  779

Query  780  GAAAGAAATTGCGGATTCTCTGAGTATCACCGAAGTGACGATTCGTAACCGCTACCGTGA  839
             |||||||| || |||  | |||| || || ||||| || ||  | ||| | ||  | ||
Sbjct  780  AAAAGAAATAGCTGATAGTTTGAGCATAACAGAAGTAACAATAAGGAACAGATATAGAGA  839

Query  840  AATCATCAAAGCCTTCGATATCGAAGTTAAACT  872
             || || || || || || || ||||| || ||
Sbjct  840  GATTATAAAGGCTTTTGACATAGAAGTAAAGCT  872
```

Codon optimized sequence of *tfb*3 for expression in *E. coli*, 597 bp, in pUC57, underlined: *Nde*I, *Bam*HI

CATATGGAATGCCCGGAATGTAAATCTCGTGAAATTGTGTGGGATTATAAATGCGGCAAC
CTGGTGTGCAGTAATTGTGGTCTGGTTCTGGATAAAATCTACGATGATCATAACTACATC
GATAACGAAATCATGATCAAAATCCAGAGCACCTTCACGAACGTTACCATTCTGACGTAC
AAAGATAAAATCGAAAAAATCGATAAAATTATCAAATTCAACAGCAAACTGTCTAAAAAA
CAGCTGCCGAAATCTCGTAAACTGCTGAACTACAACAACGCAATCATCCGCAGCTCTAGT
GCGACCATCATCAAATACCTGGATTTCAACGAAAAACTGCTGCTGGTGTACGATATTATC
GATACGATCCCGATCCTGAACAACATCAGCATCAAATACAAAGTTGCGCTGGCCATGTAC
TTCTACGATAAGAAAACCTTCAACAAAATCATGAACAATCTGGAAATCAGTAACAAATAC
TTCAACAAAATCCTGCGTCGCCTGAACAGCAAAGAAAAAATGATCATCATGGAAAAAGTG
ATCAATCTGCTGGAACAGCGCGTTCCGAGCCAGACCCTGAAAACGATGTAAGGATCC

Alignment of codon optimized sequence of *tfb*3 (Sbjct) and the original genomic (Query) sequence

```
 Score =  425 bits (470),  Expect = 4e-123
 Identities = 447/588 (77%), Gaps = 0/588 (0%)
 Strand=Plus/Plus

Query  1    ATGGAATGCCCGGAATGTAAATCTCGTGAAATTGTGTGGGATTATAAATGCGGCAACCTG  60
            ||||||||  ||  ||  ||||||   | |  |||||  ||  ||||||||  ||  |||||  ||   |
Sbjct  1    ATGGAATGTCCCGAGTGTAAAAGTAGAGAAATAGTATGGGATTACAAGTGCGGAAATTTA  60

Query  61   GTGTGCAGTAATTGTGGTCTGGTTCTGGATAAAATCTACGATGATCATAACTACATCGAT  120
            ||  ||  ||||||||  ||   |  |||  |  |||||  ||||||||||||||||||||  ||||||
Sbjct  61   GTTTGTAGTAATTGCGGATTAGTTTTAGATAAGATCTACGATGATCATAACTATATCGAT  120

Query  121  AACGAAATCATGATCAAAATCCAGAGCACCTTCACGAACGTTACCATTCTGACGTAcaaa  180
            ||  |||||  |||||  |||||  ||    |||  ||  ||  |||||  ||  |||  |  ||  ||  |||
Sbjct  121  AATGAAATTATGATAAAAATACAATCCACTTTTACTAACGTCACAATTTTAACTTATAAA  180

Query  181  gataaaatcgaaaaaatcgataaaattatcaaattcaacagcaaactgtctaaaaaaCAG  240
            ||||||||  ||  |||||  ||||||||  ||  |||||  ||  ||  |||||  ||  ||||||||
Sbjct  181  GATAAAATTGAGAAAATTGATAAAATAATAAAATTTAATAGTAAACTATCAAAAAAACAA  240

Query  241  CTGCCGAAATCTCGTAAACTGCTGAACTACAACAACGCAATCATCCGCAGCTCTAGTGCG  300
            ||  ||  |||   |  |  |||  |   |  ||  ||  ||  ||||||||  ||   |    |||||||||
Sbjct  241  CTTCCTAAAAGTAGAAAATTATTAAATTATAATAACGCAATAATTAGATCCTCTAGTGCT  300

Query  301  ACCATCATCAAATACCTGGATTTCAACGAAAAACTGCTGCTGGTGTACGATATTATCGAT  360
            ||  ||  ||  ||||||  |||||||  ||  |||||  ||   |  ||  ||  ||  ||  |||||  |||
Sbjct  301  ACAATTATAAAATACTTGGATTTTAATGAAAAGCTATTACTTGTATATGACATTATTGAT  360

Query  361  ACGATCCCGATCCTGAACAACATCAGCATCAAATACAAAGTTGCGCTGGCCATGTACTTC  420
            ||  ||  ||  ||   ||||  ||  ||     ||  ||||||||||||||   |  ||||||||  |||
Sbjct  361  ACTATTCCAATTTTGAATAATATTTCAATTAAATACAAAGTTGCATTAGCCATGTATTTC  420

Query  421  TACGATAAGAAAACCTTCAACAAAATCATGAACAATCTGGAAATCAGTAACAAATACTTC  480
            ||||||||||||||  ||  ||||||||  |||||  |||  |  ||  ||  ||  |||||  ||  ||
Sbjct  421  TACGATAAGAAAACTTTTAACAAAATTATGAATAATTTAGAGATAAGCAACAAGTATTTT  480

Query  481  AACAAAATCCTGCGTCGCCTGAACAGCAAAGAAAAAATGATCATCATGGAAAAAGTGATC  540
            |||||  ||  ||   |   |  ||  |||||  ||  |||||  |||||  ||  ||||||||||||  ||
Sbjct  481  AACAAGATACTAAGAAGACTTAACAGTAAGGAAAAGATGATAATAATGGAAAAAGTCATT  540

Query  541  AATCTGCTGGAACAGCGCGTTCCGAGCCAGACCCTGAAAACGATGTAA  588
            |||||  ||  |||||   |  ||  ||  |||||  ||   |  ||||||||||||
Sbjct  541  AATCTACTAGAACAAAGAGTACCTAGCCAAACGTTAAAAACGATGTAA  588
```

Fig. 37: Codon optimized sequences of *tfb*1, *tfb*2 and *tfb*3 from *S. acidocaldarius* for expression in *E. coli* and sequence alignments of optimized (Sbjct) and original gene (Query) sequence.

7. Literature

ABELLA, M., RODRÍGUEZ, S., PAYTUBI, S., CAMPOY, S., WHITE, M.F., BARBÉ, J. (2007). The *Sulfolobus solfataricus* radA paralogue SSO0777 is DNA damage inducible and positively regulated by the Sta1 protein. Nucleic Acids Res. 35(20):6788-97.

ADACHI, **N.**, NATSUME, R., SENDA, M., MUTO, S., SENDA, T., HORIKOSHI, M. (2004). Purification, crystallization and preliminary X-ray analysis of *Methanococcus jannaschii* TATA box-binding protein (TBP). Acta Crystallogr D Biol Crystallogr. 60(Pt 12 Pt 2):2328-31.

ADACHI, N., SENDA, M., NATSUME, R., SENDA, T., HORIKOSHI, M. (2008). Crystal structure of *Methanococcus jannaschii* TATA box-binding protein. Genes Cells. 13(11):1127-40. doi: 10.1111/j.1365-2443.2008.01233.x.

AHMED, H., ETTEMA, T.J., TJADEN, B., GEERLING, A.C., VAN DER OOST, J., SIEBERS, B. (2005). The semi-phosphorylative Entner-Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. Biochem J 390(Pt 2): 529-540.

AJON, M., FRÖLS, S., VAN WOLFEREN, M., STOECKER, K., TEICHMANN, D., DRIESSEN, A.J., GROGAN, D.W., ALBERS, S.-V., SCHLEPER, C. (2011). UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. Mol Microbiol 82(4): 807-817.

AKCA, E., CLAUS, H., SCHULTZ, N., KARBACH, G., SCHLOTT, B., DEBAERDEMAEKER, T., DECLERCQ, J.P., KÖNIG, H. (2002). Genes and derived amino acid sequences of S-layer proteins from mesophilic, thermophilic, and extremely thermophilic methanococci. Extremophiles. 6(5):351-8.

AKHTAR, W., VEENSTRA, G.J.. (2009). TBP2 is a substitute for TBP in *Xenopus* oocyte transcription. BMC Biol. 7:45. doi: 10.1186/1741-7007-7-45.

ALBERS, S.-V., JONUSCHEIT, M., DINKELAKER, S., URICH, T., KLETZIN, A., TAMPÉ, R., DRIESSEN, A.J., SCHLEPER, C. (2006). Production of recombinant and tagged proteins in the hyperthermophilic archaeon *Sulfolobus solfataricus.* Appl Environ Microbiol. 72(1):102-11.

ALTSCHUL, S.F., GISH. W,, MILLER, W., MYERS, E.W., LIPMAN, D.J. (1990). Basic local alignment search tool. J Mol Biol 215:403–10.

AMMAR, R., TORTI, D., TSUI, K., GEBBIA, M., DURBIC, T,. BADER, G.D., GIAEVER, G., NISLOW, C. (2012). Chromatin is an ancient innovation conserved between Archaea and Eukarya. eLife. 2012; 1: e00078. doi: 10.7554/eLife.00078.

ANDERSSON, A.F., LUNDGREN, M., ERIKSSON, S., ROSENLUND, M., BERNANDER, R., NILSSON, P.. (2006). Global analysis of mRNA stability in the archaeon *Sulfolobus*. Genome Biol. 7(10): R99.

ANGELOV, A., LIEBL, W. (2010). Heterologous gene expression in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Methods Mol Biol. 668:109-16. doi: 10.1007/978-1-60761-823-2_7. Retraction in: Methods Mol Biol. 2010;668:E1.

AO, X., LI, Y., WANG, F., FENG, M., LIN, Y., ZHAO, S., LIANG, Y., PENG, N. (2013). *Sulfolobus* Initiator element: An important contributor to promoter strength. J Bacteriol. [Epub ahead of print]

ARAVIND, L., KOONIN, E.V. (1999). DNA-binding proteins and evolution of transcription regulation in the archaea. Nucleic Acids Res. 27(23) 4658–4670.

AUERNIK, K.S., MAEZATO, Y., BLUM, P.H., KELLY, R.M. (2008). The genome sequence of the metal-mobilizing, extremely thermoacidophilic archaeon *Metallosphaera sedula* provides insights into bioleaching-associated metabolism. Appl Environ Microbiol. 74(3):682-92.

BAKER, D.L., YOUSSEF, O.A., CHASTKOFSKY, M.I., DY, D.A., TERNS, R.M., TERNS, M.P. (2005). RNA-guided RNA modification: functional organization of the archaeal H/ACA RNP. Genes Dev. 19(10): 1238-1248.

BALIGA, N.S., GOO, Y.A., NG, W.V., HOOD, L., DANIELS, C.J., DASSARMA, S. (2000). Is gene expression in *Halobacterium* NRC-1 regulated by multiple TBP and TFB transcription factors? Mol Microbiol. 36(5):1184-5

BALIGA, N,S,, BJORK, S.J., BONNEAU, R., PAN, M., ILOANUSI, C., KOTTEMANN, M.C., HOOD, L., DIRUGGIERO, J. (2004). Systems level insights into the stress response to UV radiation in the halophilic archaeon *Halobacterium* NRC-1. Genome Res. 14(6):1025-35.

BANERJEE, S., CHALISSERY, J., BANDEY, I., SEN, R. (2006). Rho-dependent transcription termination: more questions than answers. J Microbiol. 44(1): 11–22.

BARBERIS, A., MÜLLER, C.W., HARRISON, S.C., PTASHNE, M. (1993). Delineation of two functional regions of transcription factor TFIIB. Proc Natl Acad Sci U S A. 90(12):5628-32.

BARTLETT, M.S. (2005). Determinants of transcription initiation by archaeal RNA polymerase. Curr Opin Microbiol. 8(6):677-84.

BAUER, M., MARSCHAUS, L., REUFF, M., BESCHE, V., SARTORIUS-NEEF S., PFEIFER, F. (2008). Overlapping activator sequences determined for two oppositely oriented promoters in halophilic Archaea Nucl. Acids Res. 36 (2): 598-606. doi:10.1093/nar/gkm1077.

BELL, S.D., JACKSON, S.P. (1998a). Transcription in Archaea. Cold Spring Harb Symp Quant Biol. 63:41-51.

BELL, S.D., JACKSON, S.P. (1998b). Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol. 6(6):222-8.

BELL, S.D., KOSA, P.L., SIGLER, P.B., JACKSON, S.P. (1999a). Orientation of the transcription preinitiation complex in archaea. Proc Natl Acad Sci U S A. 96(24):13662-7.

BELL, S.D., CAIRNS, S.S., ROBSON, R.L., JACKSON, S.P. (1999b). Transcriptional regulation of an archaeal operon *in vivo* and *in vitro*. Mol Cell. 4(6):971-82

BELL, S.D., JACKSON, S.P. (2000a). The role of transcription factor B in transcription initiation and promoter clearance in the archaeon *Sulfolobus acidocaldarius*. J Biol Chem. 275(17): 12934-12940.

BELL, S.D., JACKSON, S.P. (2000b). Mechanism of autoregulation by an archaeal transcriptional repressor. J Biol Chem. 275(41):31624-9.

BELL, S.D., BRINKMAN, A.B., VAN DER OOST, J., JACKSON, S.P. (2001). The archaeal TFIIEalpha homologue facilitates transcription initiation by enhancing TATA-box recognition. EMBO Rep. 2(2):133-8.

BELL, S.D,, JACKSON, S.P. (2001). Mechanism and regulation of transcription in archaea. Curr Opin Microbiol. 4(2):208-13.

BELL, S.D. (2005). Archaeal transcriptional regulation-variation on a bacterial theme? Trends Microbiol. 13(6): 262-265.

BERKNER, S., WLODKOWSKI, A., ALBERS, S.-V, LIPPS, G. (2010). Inducible and constitutive promoters for genetic systems in *Sulfolobus acidocaldarius*. Extremophiles 14(3): 249-259.

BERNICK, D.L., DENNIS, P.P., LUI, L.M., LOWE, T.M. (2012). Diversity of antisense and other non-coding RNAs in Archaea revealed by comparative small RNA sequencing in four *Pyrobaculum* species. Front Microbiol. 3:231. doi: 10.3389/fmicb.2012.00231.

BEST, A.A., OLSEN, G.J. (2001). Similar subunit architecture of archaeal and eukaryal RNA polymerases. FEMS Microbiol Lett. 195(1):85-90.

BINI, E., DIKSHIT, V., DIRKSEN, K., DROZDA, M., BLUM, P. (2002). Stability of mRNA in the hyperthermophilic archaeon *Sulfolobus solfataricus*. RNA. 8(9):1129-36.

BIENIOSSEK, C., PAPAI, G., SCHAFFITZEL, C., GARZONI, F., CHAILLET, M., SCHEER, E., PAPADOPOULOS, P., TORA, L., SCHULTZ, P., BERGER, I. (2013). The architecture of human general transcription factor TFIID core complex. Nature. 493(7434): 699-702.

BLEIHOLDER, A., FROMMHERZ, R., TEUFEL, K., PFEIFER, F. (2012). Expression of multiple *tfb* genes in different *Halobacterium salinarum* strains and interaction of TFB with transcriptional activator GvpE. Arch Microbiol. 194(4): 269-279.

BLOMBACH, F., MAKAROVA, K.S., MARRERO, J., SIEBERS, B., KOONIN, E.V., VAN DER OOST, J. (2009). Identification of an ortholog of the eukaryotic RNA polymerase III subunit RPC34 in Crenarchaeota and Thaumarchaeota suggests specialization of RNA polymerases for coding and non-coding RNAs in Archaea. Biol Direct. 4: 39.

BRADFORD, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 72. 248-254.

BRINKMAN, A.B., BELL, S.D., LEBBINK, R.J., DE VOS, W.M., VAN DER OOST, J. (2002). The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. J Biol Chem. *2*77, 29537-29549.

BRINKMAN, A.B., ETTEMA, T.J.G, DE VOS, W.M., VAN DER OOST, J. (2003). The Lrp family of transcriptional regulators. Mol Microbiol. 48(2), 287–294.

BROCK, T.D., BROCK, K.M., BELLY, R.T., WEISS, R.L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Arch Mikrobiol. 84(1):54-68.

BROWN, R.E., JARVIS, K.L., HYLAND, K.J. (1989). Protein measurement using bicinchoninic acid: elimination of interfering substances. Anal Biochem. 180(1).136–139

BULT, C.J., WHITE, O., OLSEN, G.J., ZHOU, L., FLEISCHMANN, R.D., SUTTON, G.G., BLAKE, J.A., FITZGERALD, L.M., CLAYTON, R.A., GOCAYNE, J.D., KERLAVAGE, A.R., DOUGHERTY, B.A., TOMB, J.F., ADAMS, M.D., REICH, C.I., OVERBEEK, R., KIRKNESS, E.F., WEINSTOCK, K.G., MERRICK, J.M., GLODEK, A., SCOTT, J.L., GEOGHAGEN, N.S., VENTER, J.C. (1996). Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. Science. 273(5278):1058-73

BURATOWSKI, S., ZHOU, H. (1993). Functional domains of transcription factor TFIIB. Proc Natl Acad Sci U S A. 90(12):5633-7.

BURLEY, S.K., ROEDER, R.G. (1996). Biochemistry and structural biology of transcription factor IID (TFIID). Annu Rev Biochem. 65:769-99.

BURNS, C.M., RICHARDSON, L.V., RICHARDSON, J.P. (1998). Combinatorial effects of

NusA and NusG on transcription elongation and Rho-dependent termination in *Escherichia coli*. J Mol Biol. 278(2):307-16.

BUSHNELL, D.A., WESTOVER, K.D., DAVIS, R.E., KORNBERG, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. Science. 303(5660): 983-988.

CARROLL, K.L., PRADHAN, D.A., GRANEK, J.A., CLARKE, N.D., CORDEN, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol. 24(14):6241-52.

CAVEL, E., PILLOT, M., PONTIER, D., LAHMY, S., BIES-ETHEVE, N., VEGA, D., GRIMANELLI, D., AGRANGE, T. (2011). A plant-specific transcription factor IIB-related protein, pBRP2, is involved in endosperm growth control. PLoS One. 6(2):e17216. doi: 10.1371/journal.pone.0017216.

CHANG, B., HALGAMUGE, S., TANG, S. (2006). Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. Gene. 373:90–99.

CHAROENSAWAN, V., WILSON, D., TEICHMANN, S.A. (2010). Genomic repertoires of DNA-binding transcription factors across the tree of life. Nucleic Acids Res. 38(21): 7364-7377.

CHASMAN, D.I., FLAHERTY, K.M., SHARP, P.A., KORNBERG, R.D. (1993). Crystal structure of yeast TATA-binding protein and model for interaction with DNA. Proc Natl Acad Sci U S A. 90(17):8174-8.

CHEN, H.T., LEGAULT, P., GLUSHKA, J., OMINCHINSKI, J.G., SCOTT, R.A. (2000). Structure of a (Cys_3His) zinc ribbon, a ubiquitous motif in archaeal and eukaryal transcription. Protein Science. 9:1743-1752.

CHEN, B.S., MANDAL, S.S., HAMPSEY, M. (2004). High-resolution protein-DNA contacts for the yeast RNA polymerase II general transcription machinery. Biochemistry. 43(40):12741-9.

CHEN, B.S., HAMPSEY, M. (2004). Functional interaction between TFIIB and the Rpb2 subunit of RNA polymerase II: implications for the mechanism of transcription initiation. Mol Cell Biol. 24(9):3983-91.

CHEN, L., BRÜGGER, K., SKOVGAARD, M., REDDER, P., SHE, Q., TORARINSSON, E., GREVE, B., AWAYEZ, M., ZIBAT, A., KLENK, H.P., GARRETT, R.A. (2005). The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. J Bacteriol. 187(14):4992-9.

CHOI KH, HWANG S, CHA J. (2013). Identification and characterization of MalA in the maltose/maltodextrin operon of *Sulfolobus acidocaldarius* DSM639. J Bacteriol. 195(8):1789-99. doi: 10.1128/JB.01713-12.

CIAMPI, M,S. (2006). Rho-dependent terminators and transcription termination. Microbiol. 152(Pt 9):2515-28.

CLAUS, H., AKÇA, E., DEBAERDEMAEKER, T., EVRARD, C., DECLERCQ, J.P., KÖNIG, H. (2002). Primary structure of selected archaeal mesophilic and extremely thermophilic outer surface layer proteins. Syst Appl Microbiol. 25(1):3-12.

COHEN, G.N., BARBE, V., FLAMENT, D., GALPERIN, M., HEILIG, R., LECOMPTE, O., POCH, O., PRIEUR, D., QUÉRELLOU, J., RIPP, R., THIERRY, J.C., VAN DER OOST, J., WEISSENBACH, J., ZIVANOVIC, Y., FORTERRE, P. (2003). An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. Mol Microbiol. 47(6):1495-512.

COHEN-KUPIEC, R., BLANK, C., LEIGH, J.A. (1997). Transcriptional regulation in

Archaea: *in vivo* demonstration of a repressor binding site in a methanogen. Proc Natl Acad Sci U S A. 94(4):1316-20.

COKER, J.A., DASSARMA, S. (2007). Genetic and transcriptomic analysis of transcription factor genes in the model halophilic Archaeon: coordinate action of TbpD and TfbA. BMC Genet. 8: 61.

CONAWAY, R.C., CONAWAY, J.W. (1993). General initiation factors for RNA polymerase II. Annu Rev Biochem. 62:161-90.

CONAWAY, R.C., CONAWAY, J.W. (2013). The Mediator complex and transcription elongation. Biochim Biophys Acta 1829(1): 69-75.

CONTURSI, P., CANNIO, R., PRATO, S., FIORENTINO, G., ROSSI, M., BARTOLUCCI, S. (2003). Development of a genetic system for hyperthermophilic Archaea: expression of a moderate thermophilic bacterial alcohol dehydrogenase gene in *Sulfolobus solfataricus*. FEMS Microbiol Lett. 218(1):115-20.

COULSON, R.M., TOUBOUL, N., OUZOUNIS, C.A. (2007). Lineage-specific partitions in archaeal transcription. Archaea. 2:117-125.

CRAMER, P., BUSHNELL, D.A., FU, J., GNATT, A.L., MAIER-DAVIS, B., THOMPSON, N.E., BURGESS, R.R., EDWARDS, A.M., DAVID, P.R., KORNBERG, R.D. (2000). Architecture of RNA polymerase II and implications for the transcription mechanism. Science. 288(5466):640-9.

CRAMER, P., BUSHNELL, D.A., KORNBERG, R.D. (2001). Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. Science. 292(5523): 1863-1876.

CRICK, F.H.C. (1958). The biological replication of macromolecules. XII 188. In Symp. Sec. Exp. Biol.

CRICK, F.H.C. (1970). Central dogma of molecular biology**.** Nature. 227(5258):561-3.

CROWLEY, T.E., HOEY, T., LIU, J.K., JAN, Y.N., JAN, L.Y., TJIAN, R. (1993). A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila***.** Nature. 361(6412):557-61.

ČUBOŇOVÁ, L., SANDMAN, K., HALLAM, S.J., DELONG, E.F., REEVE, J.N. (2005). Histones in crenarchaea. J Bacteriol. 2005 Aug;187(15):5482-5.

DAHLKE, I., THOMM, M. (2002). A *Pyrococcus* homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. Nucleic Acids Res. 30(3): 701–710.

DARST, S.A. (2001). Bacterial RNA polymerase. Curr Opin Struct Biol. 11(2):155-62.

DE BIASE, A., MACARIO, A.J., CONWAY DE MACARIO, E. (2002)**.** Effect of heat stress on promoter binding by transcription factors in the cytosol of the archaeon *Methanosarcina mazeii***.** Gene. 282(1-2):189-97.

DEDECKER, B.S., O'BRIEN, R., FLEMING, P.J., GEIGER, J.H., JACKSON, S.P., SIGLER, P.B. (1996). The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. J Mol Biol. 264(5):1072-84

DELMAS, S., SHUNBURNE, L., NGO, H.P., ALLERS, T. (2009). Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon *Haloferax volcanii* by restraining homologous recombination. PLoS Genet. 5(7):e1000552. doi: 10.1371/journal.pgen.1000552.

DENG, W., ROBERTS, S.G. (2007). TFIIB and the regulation of transcription by RNA polymerase II. Chromosoma. 116(5):417-29.

DEPPENMEIER, U., JOHANN, A., HARTSCH, T., MERKL, R., SCHMITZ, R.A., MARTINEZ ARIAS, R., HENNE, A., WIEZER, A., BÄUMER S., JACOBI, C., BRÜGGEMANN, H., LIENARD, T., CHRISTMANN, A., BÖMEKE, M,, STECKEL, S,, BHATTACHARYYA, A., LYKIDIS, A., OVERBEEK, R., KLENK, H.P., GUNSALUS, R.P., FRITZ, H.J., GOTTSCHALK, G. (2002). The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol. 4(4):453-61.

DIECI, G., HERMANN,-L.E. DENMAT, S., LUKHTANOV, E., THURIAUX, P., WERNER, M., SENTENAC, A. (1995). A universally conserved region of the largest subunit participates in the active site of RNA polymerase III. EMBO J. 14(15):3766-76.

DIXIT, V., BINI, E., DROZDA, M., BLUM, P. (2004). Mercury inactivates transcription and the generalized transcription factor TFB in the archaeon *Sulfolobus solfataricus*. Antimicrob Agents Chemother 48(6): 1993-1999.

DORAZI, R., GÖTZ, D., MUNRO, S., BERNANDER, R., WHITE, M.F. (2007). Equal rates of repair of DNA photoproducts in transcribed and non-transcribed strands in *Sulfolobus solfataricus*. Mol Microbiol. 63(2): 521-529.

EDDY, S.R. (2001). Non-coding RNA genes and the modern RNA world. Nat Rev Genet. 2001 Dec;2(12):919-29.

EGLOFF, S., MURPHY, S. (2008). Cracking the RNA polymerase II CTD code. Trends Genet. 24(6):280-8. doi: 10.1016/j.tig.2008.03.008.

EICHLER, J. (2003). Facing extremes: archaeal surface-layer (glyco)proteins. Microbiol. 149(12): 3347-3351.

ESSER, D., PHAM, T.K., REIMANN, J., ALBERS, S.-V., SIEBERS, B., WRIGHT, P.C. (2012). Change of carbon source causes dramatic effects in the phospho-proteome of the archaeon *Sulfolobus solfataricus*. J Proteome Res. 11(10):4823-33. doi: 10.1021/pr300190k.

FACCIOTTI, M.T., REISS, D.J., PAN, M., KAUR, A., VUTHOORI, M., BONNEAU, R., SHANNON, P., SRIVASTAVA, A., DONOHOE, S.M., HOOD, L.E., BALIGA, N.S. (2007). General transcription factor specified global gene regulation in archaea. Proc Natl Acad Sci U S A. 104(11):4630-5.

ELKINS, J.G., PODAR, M., GRAHAM, D.E., MAKAROVA, K.S., WOLF, Y., RANDAU, L., HEDLUND, B.P., BROCHIER-ARMANET, C., KUNIN, V., ANDERSON, I., LAPIDUS, A., GOLTSMAN, E., BARRY, K., KOONIN, E.V., HUGENHOLTZ, P., KYRPIDES, N., WANNER, G., RICHARDSON, P., KELLER, M., STETTER, K.O. (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. Proc Natl Acad Sci U S A. 105(23):8102-7. doi: 10.1073/pnas.0801980105.

FAIRLEY, J.A., EVANS, R., HAWKES, N.A., ROBERTS, S.G. (2002). Core Promoter-Dependent TFIIB Conformation and a Role for TFIIB Conformation in Transcription Start Site Selection. Mol Cell Biol. 22(19): 6697-6705.

FIKES, J.D., BECKER, D.M., WINSTON, F., GUARENTE, L. (1990). Striking conservation of TFIID in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae.* Nature. 346(6281):291-4.

FIORENTINO, G., CANNIO, R., ROSSI, M., BARTOLUCCI, S. (2003). Transcriptional regulation of the gene encoding an alcohol dehydrogenase in the archaeon *Sulfolobus solfataricus* involves multiple factors and control elements. J Bacteriol. 185(13):3926-34.

FIORENTINO, G., RONCA, R., CANNIO, R., ROSSI, M., BARTOLUCCI, S. (2007).

MarR-like transcriptional regulator involved in detoxification of aromatic compounds in *Sulfolobus solfataricus*. J Bacteriol. 189(20):7351-60.

FIORENTINO, G., DEL GIUDICE, I., BARTOLUCCI, S., DURANTE, L., MARTINO, L., DEL VECCHIO, P. (2011). Identification and physicochemical characterization of BldR2 from *Sulfolobus solfataricus*, a novel archaeal member of the MarR transcription factor family. Biochem. 50(31): 6607-6621.

FITZ-GIBBON ST, LADNER H, KIM UJ, STETTER KO, SIMON MI, MILLER JH. (2002). Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*. Proc Natl Acad Sci U S A. 99(2):984-9.

FRÖLS, S., GORDON, P.M., PANLILIO, M.A., DUGGIN, I.G., BELL, S.D., SENSEN, C.W., SCHLEPER, C. (2007). Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. J Bacteriol. 189(23): 8708-8718.

FRÖLS, S., AJON, M., WAGNER, M., TEICHMANN, D., ZOLGHADR, B., FOLEA, M., BOEKEMA, E.J., DRIESSEN, A.J., SCHLEPER, C., ALBERS, S.-V. (2008). UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. Mol Microbiol. 70(4): 938-952.

FRÖLS, S., WHITE, M.F., SCHLEPER, C. (2009). Reactions to UV damage in the model archaeon *Sulfolobus solfataricus*. Biochem Soc Trans. 37(Pt 1): 36-41.

FUJIHASHI, M., NUMOTO, N., KOBAYASHI, Y., MIZUSHIMA, A., TSUJIMURA, M., NAKAMURA, A., KAWARABAYASI, Y., MIKI, K. (2006). Crystal structure of archaeal photolyase from *Sulfolobus tokodaii* with two FAD molecules: implication of a novel light-harvesting cofactor. J Mol Biol. 365(4):903-10.

FUKUI, T., ATOMI, H., KANAI, T., MATSUMI, R., FUJIWARA, S., IMANAKA, T. (2005). Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. Genome Res. 15(3):352-63.

GALAGAN, J.E., NUSBAUM, C., ROY, A., ENDRIZZI, M.G., MACDONALD, P., FITZHUGH, W., CALVO, S., ENGELS, R., SMIRNOV, S., ATNOOR, D., BROWN, A., ALLEN, N., NAYLOR, J., STANGE-THOMANN, N., DEARELLANO, K., JOHNSON, R., LINTON, L., MCEWAN, P., MCKERNAN, K., TALAMAS, J., TIRRELL, A., YE, W., ZIMMER, A., BARBER, R.D., CANN, I., GRAHAM, D.E., GRAHAME, D.A., GUSS, A.M., HEDDERICH, R., INGRAM-SMITH, C., KUETTNER, H.C., KRZYCKI, J.A., LEIGH, J.A., LI, W., LIU, J., MUKHOPADHYAY, B., REEVE, J.N., SMITH, K., SPRINGER, T.A., UMAYAM, L.A., WHITE, O., WHITE, R.H., CONWAY DE MACARIO, E., FERRY, J.G., JARRELL, K.F., JING, H., MACARIO, A.J., PAULSEN, I., PRITCHETT, M., SOWERS, K.R., SWANSON, R.V., ZINDER, S.H., LANDER, E., METCALF, W.W., BIRREN, B. (2002). The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Res. 12(4):532-42.

GAZDAG, E., RAJKOVIC, A., TORRES-PADILLA, M.E., TORA, L. (2007). Analysis of TATA-binding protein 2 (TBP2) and TBP expression suggests different roles for the two proteins in regulation of gene expression during oogenesis and early mouse development. Reproduction. 134(1): 51-62.

GEIDUSCHEK, E.P., OUHAMMOUCH, M. (2005). Archaeal transcription and its regulators. Mol Microbiol. 56(6):1397-1407.

GHANÉ, F., GROGAN, D.W. (1998). Chromosomal marker exchange in the thermophilic 4 archaeon *Sulfolobus acidocaldarius*: physiological and cellular aspects. Microbiol. 144(6) 1649-1657.

GIRARD, J.P., LEHTONEN, H., CAIZERGUES-FERRER, M., AMALRIC, F., TOLLERVEY, D., LAPEYRE, B.
(1992). GAR1 is an essential small nucleolar RNP protein required for pre-rRNA processing in yeast. EMBO J. 11(2):673-82.

GOEDE, B., NAJI, S., VON KAMPEN, O., ILG, K., THOMM, M. (2006). Protein-protein interactions in the archaeal
transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors. J Biol Chem. 281(41): 30581-30592.

GÖTZ, D., PAYTUBI, S., MUNRO, S., LUNDGREN, M., BERNANDER, R., WHITE, M.F. (2007). Responses of hyperthermophilic crenarchaea to UV irradiation. Genome Biol. 8(10):R220.

GOURSE, R.L., ROSS, W., RUTHERFORD, S.T. (2006). General pathway for turning on promoters transcribed by RNA polymerases containing alternative sigma factors. J Bacteriol. 188(13): 4589-4591.

GREGOR, D., PFEIFER, F. (2005). *In vivo* analyses of constitutive and regulated promoters in halophilic archaea. Microbiol. 151, 25–33.

GROGAN, D.W. (1989). Phenotypic characterization of the archaebacterial genus *Sulfolobus*: comparison of five wild-type strains. J Bacteriol. 171(12). 6710-6719

GROGAN, D.W., GUNSALUS, R.P. (1993). *Sulfolobus acidocaldarius* synthesizes UMP via a standard *de novo* pathway: results of biochemical-genetic study. J Bacteriol. 175(5):1500-7.

GROGAN, D.W. (1996a). Organization and interactions of cell envelope proteins of the extreme thermoacidophile *Sulfolobus acidocaldarius.* Can J Microbiol. 42(11), 1163-1171.

GROGAN, D.W. (1996b). Exchange of genetic markers at extremely high temperatures in the archaeon *Sulfolobus acidocaldarius*. J Bacteriol. 178(11):3207-11.

GROGAN, D.W, CARVER, G.T., DRAKE, J.W. (2001). Genetic fidelity under harsh conditions: analysis of spontaneous mutation in the thermoacidophilic archaeon *Sulfolobus acidocaldarius*. Proc Natl Acad Sci USA 98(14): 7928-7933.

GROGAN, D.W. (2004). Stability and repair of DNA in hyperthermophilic Archaea. Curr Issues Mol Biol. 6(2):137-44.

GROGAN, D.W. (2009). Homologous recombination in *Sulfolobus acidocaldarius*: genetic assays and functional properties. Biochem Soc Trans. 37(Pt 1): 88-91.

GROHMANN, D., NAGY, J., CHAKRABORTY, A., KLOSE, D., FIELDEN, D., EBRIGHT, R.H., MICHAELIS, J., WERNER, F. (2011). The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. Mol Cell. 43(2):263-74. doi: 10.1016/j.molcel.2011.05.030.

GROHMANN, D., WERNER, F. (2010). Hold on!: RNA polymerase interactions with the nascent RNA modulate transcription elongation and termination. RNA Biol. 7(3):310-5.

GROHMANN, D., WERNER, F. (2011a). Recent advances in the understanding of archaeal transcription. Curr Opin Microbiol. 14(3): 328-334.

GROHMANN, D., WERNER, F. (2011b). Cycling through transcription with the RNA polymerase F/E (RPB4/7) complex: structure, function and evolution of archaeal RNA polymerase. Res Microbiol. 162(1):10-8.

GROSSMAN, A.D., STRAUS, D.B., WALTER, W.A., GROSS, C.A. (1987). Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. Genes & Dev. 1: 179-184.

GRÜNBERG, S., BARTLETT, M.S., NAJI, S., THOMM, M. (2007). Transcription factor E

is a part of transcription elongation complexes. J Biol Chem. 282(49):35482-90. Epub 2007 Oct 5

GUILLIÈRE, F., DANIOUX, C., JAUBERT, C., DESNOUES, N., DELEPIERRE, M., PRANGISHVILI, D., SEZONOV, G., GUIJARRO, J.I. (2013). Solution structure of an archaeal DNA binding protein with an eukaryotic zinc finger fold. PLoS One 8(1): e52908.

GURHA, P., JOARDAR, A., CHAURASIA, P., GUPTA, R. (2007). Differential roles of archaeal box H/ACA proteins in guide RNA-dependent and independent p pseudouridine formation. RNA Biology. 4:2, 101-109.

HA, I., ROBERTS, S., MALDONADO, E., SUN, X., KIM, L.U., GREEN, M., REINBERG, D. (1993). Multiple functional domains of human transcription factor IIB: distinct interactions with two general transcription factors and RNA polymerase II. Genes Dev. **7**(6): 1021-1032.

HAAG, J.R., PIKAARD, C.S. (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol. 12(8):483- 92. doi: 10.1038/nrm3152.

HAHN, S,, ROBERTS, S. (2000). The zinc ribbon domains of the general transcription factors TFIIB and Brf: conserved functional surfaces but different roles in transcription initiation. Genes Dev. 14(6): 719–730.

HAHN, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nat Struct Mol Biol. 11(5): 394-403.

HALLAM, S,J,, KONSTANTINIDIS, K.T., PUTNAM, N., SCHLEPER, C., WATANABE, Y., SUGAHARA, J., PRESTON, C., DE LA TORRE, J., RICHARDSON, P.M., DELONG, E.F. (2006). Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. Proc Natl Acad Sci U S A. 103(48):18296-301.

HAMES, B.D. (1998). A practical approach. 3. Oxford University; Gel electrophoresis of proteins.

HAMMA, T., FERRÉ-D'AMARÉ, A.R. (2010). The box H/ACA ribonucleoprotein complex: interplay of RNA and protein structures in post-transcriptional RNA modification. J Biol Chem. 285(2): 805-809.

HANAHAN, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 166(4):557-80.

HANZELKA, B.L., DARCY, T.J., REEVE, J.N. (2001). TFE, an archaeal transcription factor in *Methanobacterium thermoautotrophicum* related to eucaryal transcription factor TFIIEalpha. J Bacteriol. 183(5): 1813-1818.

HARLEY, C.B., REYNOLDS, R.P. (1987). Analysis of *E. coli* promoter sequences. Nucleic Acids Res. 15(5):2343-61.

HARTMAN, A.L., NORAIS, C., BADGER, J.H., DELMAS, S., HALDENBY, S., MADUPU, R., ROBINSON, J., KHOURI, H., REN, Q., LOWE, T.M., MAUPIN-FURLOW, J., POHLSCHRODER, M., DANIELS, C., PFEIFFER, F., ALLERS, T., EISEN, J.A. (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. PLoS One. 5(3):e9605. doi: 10.1371/journal.pone.0009605.

HAUSNER, W., FREY, G., THOMM, M. (1991). Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA(Val) gene of *Methanococcus vannielii*. J Mol Biol. 222(3):495-508

HAUSNER, W., LANGE, U., MUSFELDT, M. (2000), Transcription factor S, a cleavage induction factor of the archaeal RNA polymerase. M. J Biol Chem. 275(17):12393-9.

HAWKES, N.A., ROBERTS, S.G. (1999). The role of human TFIIB in transcription start site selection *in vitro* and *in vivo*. J Biol Chem. 274(20):14337-43.

HENKIN, T.M. (2000). Transcription termination control in bacteria. Curr Opin Microbiol. 3:149-153.

HEYMANN, J.A., HAVELKA, W.A., OESTERHELT, D. (1993). Homologous overexpression of a light-driven anion pump in an archaebacterium. Mol Microbiol. 7(4):623-30.

HICKEY, A.J., CONWA DE MACARIO, E., MACARIO, A.J. (2002). Transcription in the archaea: basal factors, regulation, and stress-gene expression. Crit Rev Biochem Mol Biol. 37(6):537-99.

HIRATA, A., KANAI, T., SANTANGELO, T.J., TAJIRI, M., MANABE, K., REEVE, J.N., IMANAKA, T., MURAKAMI, K.S. (2008a). Archaeal RNA polymerase subunits E and F are not required for transcription in vitro, but a *Thermococcus kodakarensis* mutant lacking subunit F is temperature-sensitive. Mol Microbiol. 70(3):623-33. doi: 10.1111/j.1365-2958.2008.06430.x.

HIRATA, A., KLEIN, B.J., MURAKAMI, K.S. (2008b). The X-ray crystal structure of RNA polymerase from Archaea. Nature. 451(7180):851-4. doi: 10.1038/nature06530. Erratum in: Nature. 452(7184):248.

HIRATA, A., MURAKAMI, K.S. (2009). Archaeal RNA polymerase. Curr Opin Struct Biol. 19(6):724-31. doi: 10.1016/j.sbi.2009.10.006.

HIROSE, Y., OHKUMA, Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the integrated events of eucaryotic gene expression. J Biochem. 141(5):601-8.

HIRTREITER, A., DAMSMA, G.E., CHEUNG, A.C., KLOSE, D., GROHMANN, D., VOJNIC, E., MARTIN, A.C., CRAMER, P., WERNER, F. (2010). Spt4/5 stimulates transcription elongation through the RNA polymerase clamp coiled-coil motif. Nucleic Acids Res. 38(12):4040-51. doi: 10.1093/nar/gkq135.

HOCHHEIMER, A., Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev. 17(11): 1309-1320.

HOFACKER, A., SCHMITZ, K.M., CICHONCZYK, A., SARTORIUS-NEEF S., PFEIFER, F. (2004). GvpE- and GvpD-mediated transcription regulation of the p-gvp genes encoding gas vesicles in *Halobacterium salinarum*. Microbiol. 150(Pt 6):1829-38.

HOLMES, M.C., TJIAN, R. (2000). Promoter-selective properties of the TBP-related factor TRF1. Science. 288(5467):867-70.

HOLMES, M.L., DYALL-SMITH, M.L. (2000). Sequence and expression of a halobacterial beta-galactosidase gene. Mol Microbiol. 36(1):114-22.

HOPKINS, R.C., SUN, J., JENNEY FE J.R., CHANDRAYAN, S.K., MCTERNAN, P.M., ADAMS, M.W. (2011). Homologous expression of a subcomplex of *Pyrococcus furiosus* hydrogenase that interacts with pyruvate ferredoxin oxidoreductase. PLoS One. 6(10):e26569. doi: 10.1371/journal.pone.0026569.

HSIN,, J.P., MANLEY, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. 26(19): 2119-2137.

HUGHES, K.T., MATHEE, K. (1998). The anti-sigma factors. Annu Rev Microbiol. 52:231-86.

IMAMURA S, HANAOKA M, TANAKA K. (2008). The plant-specific TFIIB-related protein, pBrp, is a general transcription factor for RNA polymerase I. EMBO J. 27(17):2317-27. doi: 10.1038/emboj.2008.151.

IMAMURA, S., HANAOKA, M., TANAKA, K. (2008). The plant-specific TFIIB-related protein, pBrp, is a general transcription factor for RNA polymerase I. EMBO J. 27(17):2317-27. doi: 10.1038/emboj.2008.151

INOUE, H., NOJIMA, H., OKAYAMA, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. Gene. 96(1):23-8.

IQBAL, J., QURESHI, S.A. (2010). Selective depletion of *Sulfolobus solfataricus* transcription factor E under heat shock conditions. J Bacteriol. 192(11): 2887-2891.

ISHIHAMA, A. (2012). Prokaryotic genome regulation: A revolutionary paradigm. Proc Jpn Acad Ser B Phys Biol Sci, Series B 88(9): 485-508.

IYER, L.M., KOONIN, E.V., ARAVIND, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. BMC Struct Biol. 3:1.

JAIN, R., RIVERA, M.C., LAKE, J.A. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. Proc Natl Acad Sci U S A. 96(7):3801-6.

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):1241-52.

JUN, S.H., REICHLEN, M.J., TAJIRI, M., MURAKAMI, K.S. (2011). Archaeal RNA polymerase and transcription regulation. Crit Rev Biochem Mol Biol. 46(1): 27-40.

KACZOWKA, S.J., REUTER, C.J., TALARICO, L.A., MAUPIN-FURLOW, J.A. (2005). Recombinant production of *Zymomonas mobilis* pyruvate decarboxylase in the haloarchaeon *Haloferax volcanii*. Archaea. 1(5):327-34.

KAMALAMPETA, R., KOTHE, U. (2012). Archaeal proteins Nop10 and Gar1 increase the catalytic activity of Cbf5 in pseudouridylating tRNA. Sci Rep 2: 663.

KAUR, A., VAN, P.T., BUSCH, C.R., ROBINSON, C.K., PAN, M., PANG, W.L., REISS, D.J., DIRUGGIERO, J., BALIGA, N.S. (2010). Coordination of frontline defense mechanisms under severe oxidative stress. Mol Syst Biol. 6:393. doi: 10.1038/msb.2010.50.

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., KIKUCHI, H. (2001). Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, *Sulfolobus tokodaii* strain7. DNA Res. 8(4):123-40.

KENNELLY, P.J. (2003). Archaeal protein kinases and protein phosphatases: insights from genomics and biochemistry. Biochem J. 370(Pt 2):373-89.

KIENER, A., HUSAIN, I., SANCAR, A., WALSH, C. (1989). Purification and properties of *Methanobacterium thermoautotrophicum* DNA photolyase. J Biol Chem. 264(23):13880-7.

KIM, D., FORST, S. (2001). Genomic analysis of the histidine kinase family in bacteria and archaea. Microbiol. 147(Pt 5):1197-212.

KIM, T.K., ROEDER, R.G. (1994). Involvement of the basic repeat domain of TATA-binding protein (TBP) in transcription by RNA polymerases I, II, and III. J Biol Chem. 269(7):4891-4.

KIM, Y., EOM, S.H., WANG, J., LEE, D.S., SUH, S.W., STEITZ, T.A. (1995). Crystal structure of *Thermus aquaticus* DNA polymerase. Nature. 376(6541):612-6.

KISS, T., FAYET-LEBARON, E., JÁDY, B.E. (2010). Box H/ACA small ribonucleoproteins. Mol Cell. 37(5): 597-606.

KNUTSON, B.A. (2013). Emergence and expansion of TFIIB-like factors in the plant kingdom.Gene. 526(1):30-8. doi: 10.1016/j.gene.2013.04.022.

KOONIN, E.V., MAKAROVA, K.S., ELKINS, J.G. (2007). Orthologs of the small RPB8 subunit of the eukaryotic RNA polymerases are conserved in hyperthermophilic Crenarchaeota and "Korarchaeota". Biol Direct. 14;2:38.

KORKHIN, Y., UNLIGIL, U.M., LITTLEFIELD, O., NELSON, P.J., STUART, D.I.,

SIGLER, P.B., BELL, S.D., ABRESCIA, N.G. (2009). Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. PLoS Biol. 7(5):e1000102. doi: 10.1371/journal.pbio.1000102.

KOSA, P.F., GHOSH, G., DEDECKER, B.S., SIGLER, P.B. (1997). The 2.1-A crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box. Proc Natl Acad Sci U S A. 94(12):6042-7.

KOSTREWA, D., ZELLER, M.E., ARMACHE, K.J., SEIZL, M., LEIKE, K., THOMM, M., CRAMER, P. (2009). RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature. 462(7271): 323-330.

KRÜGER, K., HERMANN, T., ARMBRUSTER, V., PFEIFER, F. (1998). The transcriptional activator GvpE for the halobacterial gas vesicle genes resembles a basic region leucine-zipper regulatory protein. J Mol Biol. 279(4):761-71.

KUO, Y.P., THOMPSON, D.K., ST JEAN, A., CHARLEBOIS, R.L., DANIELS, C.J. (1997). Characterization of two heat shock genes from *Haloferax volcanii*: a model system for transcription regulation in the Archaea. J Bacteriol. 179(20):6318-24.

KUSSER, A.G., BERTERO, M.G., NAJI, S., BECKER, T., THOMM, M., BECKMANN, R., CRAMER, P. (2008). Structure of an archaeal RNA polymerase. J Mol Biol. 376(2):303-7. doi: 10.1016/j.jmb.2007.08.066.

KYRPIDES, N.C., OUZOUNIS, C.A. (1999). Transcription in archaea. Proc Natl Acad Sci U S A. 96(15):8545-50.

LAEMMLI, U.K. (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. 227 , 680-685.

LAGRANGE, T., HAKIMI, M.A., PONTIER, D., COURTOIS, F., ALCARAZ, J.P., GRUNWALD, D., LAM, E., LERBS-MACHE, S. (2003). Transcription Factor IIB (TFIIB)-Related Protein (pBrp), a Plant-Specific Member of the TFIIB-Related Protein Family. Mol Cell Biol. 23(9): 3274-3286.

LANGER, D., ZILLIG, W. (1993). Putative *tfIIs* gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase. Nucleic Acids Res. 21(9):2251.

LANGER, D., HAIN, J., THURIAUX, P., ZILLIG, W. (1995). Transcription in archaea: similarity to that in eucarya. Proc Natl Acad Sci U S A. 92(13):5768-72.

LANZENDOERFER, M., LANGER, D., HAIN, J., KLENK, H.P., HOLZ, I., ARNOLD-AMMER, I., ZILLIG, W: Structure and function of the DNA-dependent RNA polymerase of *Sulfolobus*. Syst Appl Microbiol. 16:656-664.

LASSAK, K., GHOSH, A., ALBERS, S.-V. (2012). Diversity, assembly and regulation of archaeal type IV pili-like and non-type-IV pili-like surface structures. Res Microbiol. 163(9-10):630-44. doi: 10.1016/j.resmic.2012.10.024.

LECHNER, J., SUMPER, M. (1987). The primary structure of a procaryotic glycoprotein. Cloning and sequencing of the cell surface glycoprotein gene of halobacteria J Biol Chem. 262(20):9724-9.

LEONARD, P.M., SMITS, S.H.J., SEDELNIKOVA, S.E., BRINKMAN, A.B., DE VOS, W.M., VAN DER OOST, J., RICE, D.W., RAFFERTY, J.B. (2001). Crystal structure of the Lrp-like transcriptional regulator from the archaeon *Pyrococcus furiosus*. EMBO J. 20(5): 990–997.

LI, Y.F., DUBOIS, F., ZHOU, D.X. (2001). Ectopic expression of TATA box-binding protein induces shoot proliferation in *Arabidopsis*. FEBS Lett. 489(2-3):187-91.

LIE, T.J., WOOD, G.E., LEIGH, J.A. (2005). Regulation of *nif* expression in *Methanococcus maripaludis*: roles of the euryarchaeal repressor NrpR, 2-oxoglutarate, and two operators. J Biol Chem. 280(7):5236-41.

LITTLEFIELD, O., KORKHIN, Y., SIGLER, P.B. (1999). The structural basis for the oriented assembly of a TBP/TFB/promoter complex. Proc Natl Acad Sci U S A. 96(24):13668-73.

LOWERY-GOLDHAMMER, C. , RICHARDSON, J,P. (1974). An RNA-dependent nucleoside triphosphate phosphohydrolase (ATPase) associated with rho termination factor. Proc Natl Acad Sci U S A. 71(5):2003-7.

LU, Q., HAN, J., ZHOU, L., COKER, J.A., DASSARMA, P., DASSARMA, S., XIANG, H. (2008). Dissection of the regulatory mechanism of a heat-shock responsive promoter in Haloarchaea: a new paradigm for general transcription factor directed archaeal gene regulation. Nucleic Acids Res. 36(9): 3031-3042.

LUNDGREN, M., BERNANDER, R. (2007). Genome-wide transcription map of an archaeal cell cycle. Proc Natl Acad Sci U S A. 104(8): 2939-2944.

LUO, X., SCHWARZ-LINEK, U., BOTTING, C.H., HENSEL, R., SIEBERS, B., WHITE, M.F. (2007). CC1, a novel crenarchaeal DNA binding protein. J Bacteriol. 189(2):403-9.

LUSE, D.S. (2013). Promoter clearance by RNA polymerase II. Biochim Biophys Acta 1829(1): 63-68.

MA, J., CAMPBELL, A., KARLIN, S. (2002). Correlations between Shine-Dalgarno Sequences and gene features such as predicted expression levels and operon structures. J Bacteriol. 184(20): 5733-5745.

MAGILL, C.P., JACKSON, S.P., BELL, S.D. (2001). Identification of a conserved archaeal RNA polymerase subunit contacted by the basal transcription factor TFB. J Biol Chem. 276(50): 46693-46696.

MANN, M., JENSEN, O.N. (2003). Proteomic analysis of post-translational modifications. Nat Biotechnol. 21(3):255-61.

MARKOWITZ, V.M., CHEN, I.M.A., PALANIAPPAN, K., CHU, K., SZETO, E., GRECHKIN, Y., RATNER, A., JACOB, B., HUANG, J., WILLIAMS, P., HUNTEMANN, M., ANDERSON, I., MAVROMATIS, K., IVANOVA, N.N., KYRPIDES, N.C. (2012). IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res. 40(D1): D115–D122.

MARMUR, J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J Mol Biol. 3(2), 208–218, IN1.

MARTINEZ-RUCOBO, F.W., SAINSBURY, S., CHEUNG, A.C., CRAMER, P. (2011). Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. EMBO J. 30(7):1302-10. doi: 10.1038/emboj.2011.64.

MARTÍNEZ-NÚÑEZ, M.A., POOT-HERNANDEZ, A.C., RODRÍGUEZ-VÁZQUEZ, K., PEREZ-RUEDA, E. (2013). Increments and duplication events of enzymes and transcription factors influence metabolic and regulatory diversity in prokaryotes. PLoS One. 8(7):e69707. doi: 10.1371/journal.pone.0069707.

MARTUSEWITSCH, E., SENSEN, C.W., SCHLEPER, C. (2000). High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements. J Bacteriol. 182(9):2574-81.

MARRERO COTO, J. (2009). Transcription regulation in the hyperthermophilic crenarchaeon *Thermoproteus tenax* Kra1. PhD thesis.

MARRERO, J., BLOMBACH, F., RAUCH, B., BRINKMANN, H., SIEBERS, B. (2013, in prep.). Phylogeny and genome context analysis of Transcription Factor B in Archaea: evidence of a core of TFB orthologs of Euryarchaea and Crenarchaea sharing a common history.

MARSH, T.L., REICH, C.I., WHITELOCK, R.B., OLSEN, G.J. (1994). Transcription factor IID in the Archaea: sequences in the *Thermococcus celer* genome would encode a product closely related to the TATA-binding protein of eukaryotes. Proc Natl Acad Sci U S A. 91(10):4180-4.

MATSUDA, T., MORIKAWA, M., HARUKI, M., HIGASHIBATA, H., IMANAKA, T., KANAYA, S. (1999). Isolation of TBP-interacting protein (TIP) from a hyperthermophilic archaeon that inhibits the binding of TBP to TATA-DNA. FEBS Lett. 457(1):38-42.

MATSUDA, T., FUJIKAWA, M., HARUKI, M., TANG, X.F., EZAKI, S., IMANAKA, T., MORIKAWA, M., KANAYA, S. (2001). Interaction of TIP26 from a hyperthermophilic archaeon with TFB/TBP/DNA ternary complex. Extremophiles. 5(3):177-82.

MATTICK, J.S., MAKUNIN, I.V. (2006). Non-coding RNA. Hum Mol Genet. 15 Spec No 1: R17-29.

MCCREADY, S., MÜLLER, J.A., BOUBRIAK, I., BERQUIST, B.R., NG, W.L., DASSARMA, S. (2005). UV irradiation induces homologous recombination genes in the model archaeon *Halobacterium* sp. NRC-1. Saline Systems 1: 3.

MCCULLOCH, V., HARDIN, P., PENG, W., RUPPERT, J.M., LOBO-RUPPERT, S.M. (2000). Alternatively spliced hBRF variants function at different RNA polymerase III promoters. EMBO J. 19(15):4134-43.

MEINHART, A., BLOBEL, J., CRAMER, P. (2003). An extended winged helix domain in general transcription factor E/IIE alpha. J Biol Chem. 278(48):48267-74.

MEYER, B.H., ZOLGHADR, B., PEYFOON, E., PABST, M., PANICO, M., MORRIS, H.R., HASLAM, S.M., MESSNER, P., SCHÄFFER, C., DELL, A., ALBERS, S.-V. (2011). Sulfoquinovose synthase - an important enzyme in the N-glycosylation pathway of *Sulfolobus acidocaldarius*. Mol Microbiol. 82(5): 1150-1163.

MICORESCU, M., GRÜNBERG, S., FRANKE, A., CRAMER, P., THOMM, M., BARTLETT, M. (2008). Archaeal transcription: function of an alternative transcription factor B from *Pyrococcus furiosus*. J Bacteriol. 190(1): 157-167.

MILLER, J.H. (1972). Assay of β-galactosidase. Experiments in molecular genetics. p. 352-355. CSH Laboratory press, Cold Spring Harbor N.Y.

MINAKHIN, L., BHAGAT, S., BRUNNING, A., CAMPBELL, E.A., DARST, S.A., EBRIGHT, R,H,, SEVERINOV, K. (2001). Bacterial RNA polymerase subunit omega and eukaryotic RNA polymerase subunit RPB6 are sequence, structural, and functional homologs and promote RNA polymerase assembly. Proc Natl Acad Sci U S A. 98(3):892-7

MISCHO, H.E., PROUDFOOT, N.J. (2013). Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. Biochim Biophys Acta. 1829(1): 174-185.

MÜLLER, F., TORA, L. (2009). TBP2 is a general transcription factor specialized for female germ cells. J Biol. 2009;8(11):97. doi: 10.1186/jbiol196.

MURAKAMI, K.S. (2013). X-ray crystal structure of *Escherichia coli* RNA polymerase σ70 holoenzyme. J Biol Chem. 288(13):9126-34. doi: 10.1074/jbc.M112.430900.

NAJI, S., GRÜNBERG, S., THOMM, M. (2007). The RPB7 orthologue E' is required for transcriptional activity of a reconstituted archaeal core enzyme at low temperatures and stimulates open complex formation. J Biol Chem. 282(15):11047-57.

NG, W.V., KENNEDY, S.P., MAHAIRAS, G.G., BERQUIST, B., PAN, M., SHUKLA, H.D., LASKY, S.R., BALIGA, N.S., THORSSON, V., SBROGNA, J., SWARTZELL, S., WEIR, D., HALL, J., DAHL, T.A., WELTI, R., GOO, Y.A., LEITHAUSER, B., KELLER, K., CRUZ, R., DANSON, M.J., HOUGH, D.W., MADDOCKS, D.G., JABLONSKI, P.E., KREBS, M.P., ANGEVINE, C.M., DALE, H., ISENBARGER, T.A., PECK, R.F., POHLSCHRODER, M., SPUDICH, J.L., JUNG, K.W., ALAM, M., FREITAS, T., HOU, S., DANIELS, C.J., DENNIS, P.P., OMER, A.D., EBHARDT, H., LOWE, T.M., LIANG, P., RILEY, M., HOOD, L., DASSARMA, S. (2000). Genome sequence of *Halobacterium* species NRC-1. Proc Natl Acad Sci U S A. 97(22):12176-81.

NAPOLI, A., VAN DER OOST, J., SENSEN, C.W., CHARLEBOIS, R.L., ROSSI, M., CIARAMELLA, M. (1999). An Lrp-like protein of the hyperthermophilic archaeon *Sulfolobus solfataricus* which binds to its own promoter. J Bacteriol. 181(5):1474-80.

NIKOLOV, D.B., HU, S.H., LIN, J., GASCH, A., HOFFMANN, A., HORIKOSHI, M., CHUA, N.H., ROEDER, R.G., BURLEY, S.K. (1992). Crystal structure of TFIID TATA-box binding protein. Nature. 360(6399):40-6.

NIKOLOV, D.B., CHEN, H., HALAY, E.D., USHEVA, A.A., HISATAKE, K., LEE, D.K., ROEDER, R.G., BURLEY, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. Nature. 377(6545):119-28.

NIKOLOV, D.B., CHEN, H., HALAY, E.D., HOFFMAN, A., ROEDER, R.G., BURLEY, S.K. (1996). Crystal structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad Sci U S A. 93(10):4862-7.

NIU, Q.K., LIANG, Y., ZHOU, J.J., DOU, X.Y., GAO, S.C., CHEN, L.Q., ZHANG, X.Q., YE, D. (2013). Pollen-expressed transcription factor 2 encodes a novel plant-specific TFIIB-related protein that is required for pollen germination and embryogenesis in *Arabidopsis*. Mol Plant. 6(4):1091-108. doi: 10.1093/mp/sst083.

OCHS, S.M., THUMANN, S., RICHAU, R., WEIRAUCH, M.T., LOWE, T.M., THOMM, M., HAUSNER, W. (2012). Activation of archaeal transcription mediated by recruitment of transcription factor B. J Biol Chem. 287(22): 18863-18871.

OHKUMA, Y., SUMIMOTO, H., HORIKOSHI, M., ROEDER, R.G. (1990). Factors involved in specific transcription by mammalian RNA polymerase II: purification and characterization of general transcription factor TFIIE. Proc Natl Acad Sci U S A. 87(23):9163-7.

OHKUMA, J. (1997). Multiple Functions of General Transcription Factors TFIIE and TFIIH in Transcription: Possible Points of Regulation by Trans-Acting Factors. J Biochem. 122 (3): 481-489.

ORELL, A., PEETERS, E., VASSEN, V., JACHLEWSKI, S., SCHALLES, S., SIEBERS, B., ALBERS, S.-V. (2013). Lrs14 transcriptional regulators influence biofilm formation and cell motility of Crenarchaea. ISME J. doi: 10.1038/ismej.2013.68.

ORPHANIDES, G., LAGRANGE, T., REINBERG, D. (1996). The general transcription factors of RNA polymerase II. Genes Dev. 10(21): 2657-2683.

OUHAMMOUCH, M., GEIDUSCHEK, E.P. (2001). A thermostable platform for transcriptional regulation: the DNA-binding properties of two Lrp homologs from the hyperthermophilic archaeon *Methanococcus jannaschii*. EMBO J. 20(1-2):146-56.

OUHAMMOUCH, M., DEWHURST, R.E., HAUSNER, W., THOMM, M., GEIDUSCHEK,

E.P. (2003). Activation of archaeal transcription by recruitment of the TATA-binding protein. Proc Natl Acad Sci U S A. 100(9):5097-102.

OUHAMMOUCH, M., WERNER, F., WEINZIERL, R.O., GEIDUSCHEK, E.P. (2004). A fully recombinant system for activator-dependent archaeal transcription. J Biol Chem. 279(50):51719-21.

OUHAMMOUCH, M. (2004). Transcriptional regulation in Archaea. Curr Opin Genet Dev. 14(2):133-8.

OUHAMMOUCH, M., LANGHAM, G.E., HAUSNER, W., SIMPSON, A.J., EL-SAYED, N.M., GEIDUSCHEK, E.P. (2005). Promoter architecture and response to a positive regulator of archaeal transcription. Mol Microbiol. 56(3):625-37.

OUHAMMOUCH, M., HAUSNER, W., GEIDUSCHEK, E.P. (2009). TBP domain symmetry in basal and activated archaeal transcription. Mol Microbiol. 71(1):123-31. doi: 10.1111/j.1365-2958.2008.06512.x.

PAGET, M.S., HELMANN, J.D. (2003). The sigma70 family of sigma factors. Genome Biol. 4(1):203.

PALMER, J.R., DANIELS, C.J. (1994). A transcriptional reporter for *in vivo* promoter analysis in the archaeon *Haloferax volcanii*. Appl Environ Microbiol. 60(10):3867-9.

PARDEE, T.S., BANGUR, C.S., PONTICELLI, A.S. (1998). The N-terminal region of yeast TFIIB contains two adjacent functional domains involved in stable RNA polymerase II binding and transcription start site selection. J Biol Chem. 273(28): 17859-17864.

PÁTEK, M., NEŠVERA, J. (2011). Sigma factors and promoters in *Corynebacterium glutamicum*. J Biotechnol. 154(2-3):101-13. doi: 10.1016/j.jbiotec.2011.01.017.

PAYTUBI, S., WHITE, M.F. (2009). The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. Mol Microbiol. 72(6):1487-99. doi: 10.1111/j.1365-2958.2009.06737.x.

PENG, N., DENG, L., MEI, Y., JIANG, D., HU, Y., AWAYEZ, M., LIANG, Y., SHE, Q. (2012). A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon *Sulfolobus islandicus*. Appl Environ Microbiol. 78(16):5630-7. doi: 10.1128/AEM.00855-12.

PÉREZ-RUEDA. E., COLLADO-VIDES, J., SEGOVIA, L. (2004). Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. Comput Biol Chem. 28(5-6):341-50.

PERSENGIEV, S.P., ZHU, X., DIXIT, B.L., MASTON, G.A., KITTLER, E.L., GREEN, M.R. (2003). TRF3, a TATA-box-binding protein-related factor, is vertebrate-specific and widely expressed. Proc Natl Acad Sci U S A. 100(25): 14887-14891.

PESTER, M., SCHLEPER, C., WAGNER, M. (2011). The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. Curr Opin Microbiol. 14(3):300-6. doi: 10.1016/j.mib.2011.04.007.

PEETERS, E., THIA-TOONG, T.L., GIGOT, D., MAES, D., CHARLIER, D. (2004). Ss-LrpB, a novel Lrp-like regulator of *Sulfolobus solfataricus* P2, binds cooperatively to three conserved targets in its own control region. Mol Microbiol. 54(2), 321-336.

PEETERS, E., CHARLIER, D. (2010). The Lrp family of transcription regulators in Archaea. Archaea, vol. 2010, Article ID 750457, 10 pages, 2010. doi:10.1155/2010/750457

PENG, N., XIA, Q., CHEN, Z., LIANG, Y.X., SHE, Q. (2009). An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Mol Microbiol. 74(4) 928–939.

PETERS, J.M., VANGELOFF, A.D., LANDICK, R. (2011). Bacterial transcription

terminators: the RNA 3'-end chronicles. J Mol Biol. 412(5):793-813. doi: 10.1016/j.jmb.2011.03.036.

PFEIFER, F., ZOTZEL, J., KURENBACH, B., RÖDER, R., ZIMMERMANN, P. (2001). A p-loop motif and two basic regions in the regulatory protein GvpD are important for the repression of gas vesicle formation in the archaeon *Haloferax mediterranei*. Microbiol. 147(Pt 1):63-73.

PINTO, I., WARE, D.E., HAMPSEY, M. (1992). The yeast *sua7* gene encodes a homolog of human transcription factor TFIIB and is required for normal start site selection *in vivo*. Cell. 68(5):977-88.

PINTO, I., WU, W.H., NA, J.G., HAMPSEY, M. (1994). Characterization of *sua*7 mutations defines a domain of TFIIB involved in transcription start site selection in yeast. J Biol Chem. 269(48):30569-73.

PLÖßER, P., PFEIFER, F. (2002). A bZIP protein from halophilic archaea: structural features and dimer formation of cGvpE from *Halobacterium salinarum.* Mol Microbiol 45(2), 511–520.

PODAR, M., ANDERSON, I., MAKAROVA, K.S., ELKINS, J.G., IVANOVA, N., WALL, M.A., LYKIDIS, A., MAVROMATIS, K., SUN, H., HUDSON, M.E., CHEN, W., DECIU, C., HUTCHISON, D., EADS, J.R., ANDERSON, A., FERNANDES, F., SZETO, E., LAPIDUS, A., KYRPIDES, N.C., SAIER, M.H. JR, RICHARDSON, P.M., RACHEL, R., HUBER, H., EISEN, J.A., KOONIN, E.V., KELLER, M., STETTER, K.O. (2008). A genomic analysis of the archaeal system *Ignicoccus hospitalis-Nanoarchaeum equitans*. Genome Biol. 2008;9(11):R158. doi: 10.1186/gb-2008-9-11-r158.

PONTIER, D., YAHUBYAN, G., VEGA, D., BULSKI, A., SAEZ-VASQUEZ, J., HAKIMI, M.A., LERBS-MACHE, S., COLOT, V., LAGRANGE, T. (2005). Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis.* Genes Dev. 19(17):2030-40.

PRIBNOW, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. Proc Natl Acad Sci U S A. 72(3):784-8.

PROSHKINA, G.M., SHEMATOROVA, E.K., PROSHKIN, S.A., ZAROS, C., THURIAUX, P., SHPAKOVSKI, G.V. (2006). Ancient origin, functional conservation and fast evolution of DNA-dependent RNA polymerase III. Nucleic Acids Res. 34(13): 3615-3624.

QUAX, T.E., WOLF, Y.I., KOEHORST, J.J., WURTZEL, O., VAN DER OOST, R., RAN, W., BLOMBACH, F., MAKAROVA, K.S., BROUNS, S.J., FORSTER, A.C., WAGNER, E.G., SOREK, R., KOONIN, E.V., VAN DER OOST, J. (2013). Differential translation tunes uneven production of operon-encoded proteins. Cell Rep. 4(5):938-44. doi: 10.1016/j.celrep.2013.07.049.

QURESHI, S.A., BAUMANN, P., ROWLANDS, T., KHOO, B., JACKSON, S.P. (1995a). Cloning and functional analysis of the TATA binding protein from *Sulfolobus shibatae*. Nucleic Acids Res. 23(10):1775-81.

QURESHI, S.A., KHOO, B., BAUMANN, P., JACKSON, S.P. (1995b). Molecular cloning of the transcription factor TFIIB homolog from *Sulfolobus shibatae*. Proc Natl Acad Sci U S A. 92(13):6077-81.

QURESHI, S.A., BELL, S.D., JACKSON, S.P. (1997). Factor requirements for transcription in the Archaeon *Sulfolobus shibatae*. EMBO J. 16(10):2927-36.

QURESHI, S.A., JACKSON, S.P. (1998). Sequence-specific DNA binding by the *S. shibatae* TFIIB homolog, TFB, and its effect on promoter strength. Mol Cell. 1(3):389-400.

RABENSTEIN, M.D., ZHOU, S., LIS, J.T., TJIAN, R. (1999). TATA box-binding protein (TBP)-related factor 2 (TRF2), a third member of the TBP family. Proc Natl Acad Sci U S A. 96(9):4791-6.

RAINA, S., MISSIAKAS, D., GEORGOPOULOS, C. (1995). The *rpo*E gene encoding the σ^E (σ^{24}) heat shock sigma factor of *Escherichia coli.* EMBO J. 14(5) 1043-1055.

RAWLS, K.S., YACOVONE, S.K., MAUPIN-FURLOW, J.A. (2010). GlpR represses fructose and glucose metabolic enzymes at the level of transcription in the haloarchaeon *Haloferax volcanii*. J Bacteriol. 192(23):6251-60. doi: 10.1128/JB.00827-10.

REAM, T.S., HAAG, J.R., WIERZBICKI, A.T., NICORA, C.D., NORBECK, A.D., ZHU, J.K., HAGEN, G., GUILFOYLE, T.J., PASA-TOLIĆ, L., PIKAARD, C.S. (2009). Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol Cell. 33(2): 192-203.

REEVE, J.N. (2003). Archaeal chromatin and transcription. Mol Microbiol. 48(3):587-98.

REICHLEN, M.J., MURAKAMI, K.S., FERRY, J.G. (2010). Functional analysis of the three TATA binding protein homologs in *Methanosarcina acetivorans*. J Bacteriol. 192(6): 1511-1517.

REICHOW, S.L., TOMOKO, .HAMMA, T., FERRE´-D'AMARE, A.R., VARANI, G. (2007). The structure and function of small nucleolar Ribonucleoproteins. Nucleic Acids Res, 35(5): 1452–1464. doi:10.1093/nar/gkl1172.

REIMANN, J., ESSER, D., ORELL, A., AMMAN, F., PHAM, T.K., NOIREL, J., LINDÅS, A.-C., BERNANDER, R., WRIGHT, P.C., SIEBERS, B., ALBERS, S.-V. (2013). Archaeal Signal Transduction: Impact of Protein Phosphatase Deletions on Cell Size, Motility and Energy Metabolism in *Sulfolobus acidocaldarius.* Submitted.

REINES, D., CHAMBERLIN, M.J., KANE, C.M. (1989). Transcription elongation factor SII (TFIIS) enables RNA polymerase II to elongate through a block to transcription in a human gene *in vitro*. J Biol Chem. 264(18):10799-809.

REINES, D., CONAWAY, R.C., CONAWAY, J.W. (1999). Mechanism and regulation of transcriptional elongation by RNA polymerase II. Curr Opin Cell Biol. 11(3):342-6.

REITER, W.-D., HÜDEPOHL, U., ZILLIG, W. (1990). Mutational analysis of an archaebacterial promoter: Essential role of a TATA box for transcription efficiency and start site selection *in vitro.* Proc Natl Acad Sci USA. 87: 9509-9513.

RICHARDSON, J.P. (2002). Rho-dependent termination and ATPases in transcript termination. Biochim Biophys Acta. 1577 (2002) 251–260.

ROBB, F.T., MAEDER, D.L., BROWN, J.R., DIRUGGIERO, J., STUMP, M.D., YEH, R.K., WEISS, R.B., DUNN, D.M. (2001). Genomic sequence of hyperthermophile, *Pyrococcus furiosus*: implications for physiology and enzymology. Methods Enzymol. 330:134-57.

ROBERTS, J.W., SHANKAR, S., FILTER, J.J. (2008). RNA polymerase elongation factors. Annu Rev Microbiol. 62: 211-233.

ROEDER, R.G. (1996). The role of general initiation factors in transcription by RNA polymerase II. Trends Biochem Sci. 21(9):327-35.

ROLFSMEIER, M.L., LAUGHERY, M.F., HASELTINE, C.A. (2010). Repair of DNA double-strand breaks following UV damage in three *Sulfolobus solfataricus* strains. J Bacteriol. 192(19):4954-62. doi: 10.1128/JB.00667-10.

ROSENSHINE, I., TCHELET, R., MEVARECH, M.. (1989**)**. The mechanism of DNA transfer in the mating system of an archaebacterium. Science. 245(4924):1387-9.

SACHDEVA, P., MISRA, R., TYAGI, A.K., SINGH, Y. (2010). The sigma factors of *Mycobacterium tuberculosis*: regulation of the regulators. FEBS J. 277(3): 605-626.

SADHALE, P., VERMA, J., NAOREM, A. (2007). Basal transcription machinery: role in r egulation of stress response in eukaryotes. J Biosci. 32(3):569-78.

SAINSBURY, S., NIESSER, J., CRAMER, P. (2013). Structure and function of the initially transcribing RNA polymerase II-TFIIB complex. Nature. 493(7432):437-40. doi: 10.1038/nature11715.

SAKOFSKY, C.J., RUNCK, L.A., GROGAN, D.W. (2011). *Sulfolobus* mutants, generated via PCR products, which lack putative enzymes of UV photoproduct repair. Archaea. 2011: 864015.

SAMBROOK, J., FRITSCH, E.F., MANIATIS, T. (1989), Molecular cloning: a laboratory manual. 2. ed., 3. vol., Cold Spring Harbor Laboratory Press, New York, 253p.

SANDMAN, K., REEVE, J.N. (2005). Archaeal chromatin proteins: different structures but common function? Curr Opin Microbiol. 8(6):656-61.

SANGER, F., NICKLEN, S., COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc.Natl. Acad. Sci: USA. 74(12). 5463-5467.

SANTANGELO, T.J., REEVE, J.N. (2006). Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. J Mol Biol. 355(2):196-210.

SANTANGELO, T.J., CUBONOVÁ, L., JAMES, C.L., REEVE, J.N. (2007)**.** TFB1 or TFB2 is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription *in vitro*. J Mol Biol. 367(2):344-57.

SANTANGELO, T.J., CUBONOVÁ, L., SKINNER, K.M., REEVE, J.N. (2009). Archaeal intrinsic transcription termination *in vivo*. J Bacteriol. 191(22): 7102-7108.

SATO, T., ATOMI, H. (2011). Novel metabolic pathways in Archaea. Curr Opin Microbiol. 14(3): 307-314.

SCHALLER, H., GRAY, C., HERRMANN, K. (1975) Nucleotide sequence of an RNA polymerase binding site from the DNA of bacteriophage fd. Proc Natl Acad Sci U S A. 72(2):737-41.

SCHEUCH, S., MARSCHAUS, L., SARTORIUS-NEEF, S., PFEIFER, F. (2008). Regulation of *gvp* genes encoding gas vesicle proteins in halophilic Archaea. Arch Microbiol. 190(3), 333-339.

SCHLEPER, C., HOLZ, I., JANEKOVIC, D., MURPHY, J., ZILLIG, W. (1995). A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. J Bacteriol. 177(15):4417-26.

SCHMID, A.K., REISS, D.J., KAUR, A., PAN, M., KING, N., VAN, P.T., HOHMANN, L., MARTIN, D.B., BALIGA, N.S. (2007). The anatomy of microbial cell state transitions in response to oxygen. Genome Res. 17(10):1399-413.

SCHMIDT, K.J., BECK, K.E., 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(4):1407-15.

SEDELNIKOVA, S.E., SMITS, S.H., LEONARD, P.M., BRINKMAN, A.B., VAN DER OOST, J., RAFFERTY, J.B., RICE, D.W. (2001). Crystallization and quaternary structure analysis of an Lrp-like regulatory protein from the hyperthermophile *Pyrococcus furiosus*. Acta Crystallogr D Biol Crystallogr. 57(Pt 6):886-8.

SHE, Q., SINGH, R.K., CONFALONIERI, F., ZIVANOVI,C 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., VAN DER OOST, J. (2001), The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. Proc Natl Acad Sci U S A. 98(14):7835-40.

SHIBATA, R., BESSHO, Y., SHINKAI, A., NISHIMOTO, M., FUSATOMI, E., TERADA, T., SHIROUZU, M., YOKOYAMA, S. (2007). Crystal structure and RNA-binding analysis of the archaeal transcription factor NusA. Biochem Biophys Res Commun. 355(1):122-8.

SHOCKLEY, K.R., WARD, D.E., CHHABRA, S.R., CONNERS, S.B., MONTERO, C.I., KELLY, R.M. (2003). Heat shock response by the hyperthermophilic archaeon *Pyrococcus furiosus*. Appl Environ Microbiol. 69(4): 2365-2371.

SIDDARAMAPPA, S., CHALLACOMBE, J.F., DECASTRO, R.E., PFEIFFER, F., SASTRE, D.E., GIMÉNEZ, M.I., PAGGI, R.A., DETTER, J.C., DAVENPORT, K.W., GOODWIN, L.A., KYRPIDES, N., TAPIA, R., PITLUCK, S., LUCAS, S., WOYKE, T., MAUPIN-FURLOW, J.A. (2012). A comparative genomics perspective on the genetic content of the alkaliphilic haloarchaeon *Natrialba magadii* ATCC 43099T. BMC Genomics. 13:165. doi: 10.1186/1471-2164-13-165.

SIEBERS, B., SCHÖNHEIT, P. (2005). Unusual pathways and enzymes of central carbohydrate metabolism in Archaea. Curr Opin Microbiol. 8(6):695-705.

SIEBERS, B., ZAPARTY, M., RADDATZ, G., TJADEN, B., ALBERS, S.-V, BELL, S.D., BLOMBACH, F., KLETZIN, A., KYRPIDES, N., LANZ, C., PLAGENS, A., RAMPP, M., ROSINUS, A., VON JAN, M., MAKAROVA, K.S., KLENK, H.P., SCHUSTER, S.C., HENSEL, R. (2011). The complete genome sequence of *Thermoproteus tenax*: a physiologically versatile member of the Crenarchaeota. PLoS One. 2011;6(10):e24222. doi: 10.1371/journal.pone.0024222.

SMALE, S.T., KADONAGA, J.T. (2003). The RNA polymerase II core promoter. Annu Rev Biochem. 72:449-79.

SMITH, S.C., KENNELLY, P.J., POTTS, M. (1997) Protein-tyrosine phosphorylation in the Archaea. J. Bacteriol. 179*:*2418–2420.

SLUPSKA, M.M., KING, A.G., FITZ-GIBBON, S., BESEMER, J., BORODOVSKY, M., MILLER, J.H. (2001). Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum.* J Mol Biol. 309(2):347-60.

SNIEZKO, I., DOBSON-STONE, C., KLEIN, A. (1998). The *tre*A gene of *Bacillus subtilis* is a suitable reporter gene for the archaeon *Methanococcus voltae*. FEMS Microbiol Lett. 164(2):237-42.

SÖHLEMANN, P., SOPPA, J., OESTERHELT, D., LOHSE, M.J. (1997). Expression of beta 2-adrenoceptors in halobacteria. Naunyn Schmiedebergs Arch Pharmacol. 355(2):150-60.

SOPPA, J. (1999a). Transcription initiation in Archaea: facts, factors and future aspects. Mol Microbiol. 31(5):1295-305.

SOPPA, J. (1999b). Normalized nucleotide frequencies allow the definition of archaeal promoter elements for different archaeal groups and reveal base-specific TFB contacts upstream of the TATA box. Mol Microbiol. 31(5):1589-92.

SOPPA, J. (2001). Basal and regulated transcription in archaea. Adv Appl Microbiol. 50:171-217.

SPAIN, M.M., GOVIND, C.K. (2011). A role for phosphorylated Pol II CTD in modulating

transcription coupled histone dynamics. Transcription. 2(2):78-81.

SPANG, A., POEHLEIN, A., OFFRE, P., ZUMBRÄGEL, S., HAIDER, S., RYCHLIK, N., NOWKA, B., SCHMEISSER, C., LEBEDEVA, E.V., RATTEI, T., BÖHM, C., SCHMID, M., GALUSHKO, A., HATZENPICHLER, R., WEINMAIER, T., DANIEL, R., SCHLEPER, C., SPIECK, E., STREIT, W., WAGNER, M. (2012). The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. Environ Microbiol. 14(12):3122-45. doi: 10.1111/j.1462-2920.2012.02893.x.

STETTLER, S., MARIOTTE, S., RIVA, M., SENTENAC, A., THURIAUX, P. (1992). An essential and specific subunit of RNA polymerase III (C) is encoded by gene RPC34 in *Saccharomyces cerevisiae*. J Biol Chem. 267(30):21390-5.

STUDIER, F.W., MOFFATT, B.A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 189(1):113-30.

TANG, T.H., POLACEK, N., ZYWICKI, M., HUBER, H., BRUGGER, K., GARRETT, R., BACHELLERIE, J.P., HÜTTENHOFER, A. (2005). Identification of novel non-coding RNAs as potential antisense regulators in the archaeon *Sulfolobus solfataricus.* Mol Microbiol. 55(2):469-81.

TEUFEL, K., BLEIHOLDER, A., GRIESBACH, T., PFEIFER, F. (2008). Variations in the multiple *tbp* genes in different *Halobacterium salinarum* strains and their expression during growth. Arch Microbiol. 190(3):309-18. doi: 10.1007/s00203-008-0383-5.

THOMM, M. (1996). Archaeal transcription factors and their role in transcription initiation.FEMS Microbiol Rev. 18(2-3):159-71.

THOMSEN, J., DE BIASE, A., KACZANOWSKI, S., MACARIO, A.J., THOMM, M., ZIELENKIEWICZ, P., MACCOLL, R., CONWAY DE MACARIO, E. (2001). The basal transcription factors TBP and TFB from the mesophilic archaeon *Methanosarcina mazeii*: structure and conformational changes upon interaction with stress-gene promoters. J Mol Biol. 309(3):589-603.

THOMPSON, D.K., DANIELS, C.J. (1998). Heat shock inducibility of an archaeal TATA-like promoter is controlled by adjacent sequence elements. Mol Microbiol. 27(3):541-51.

THOMPSON, D.K., PALMER, J.R., DANIELS, C.J. (1999). Expression and heat-responsive regulation of a TFIIB homologue from the archaeon *Haloferax volcanii*. Mol.Microbiol. 33: 1081-1092.

THOMPSON, N.E., GLASER, B.T., FOLEY, K.M., BURTON, Z.F., BURGESS, R.R. (2009). Minimal promoter systems reveal the importance of conserved residues in the B-finger of human transcription factor IIB. J Biol Chem. 284(37): 24754-24766.

THOMPSON, W. (2003). Gibbs Recursive Sampler: finding transcription factor binding sites. Nucleic Acids Res. 31(13): 3580-3585.

TOLSTRUP, N., SENSEN, C.W., GARRETT, R.A., CLAUSEN, I.G. (2000). Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. Extremophiles. 4(3):175-9.

TORARINSSON, E., KLENK, H.P., GARRETT, R.A. (2005). Divergent transcriptional and translational signals in Archaea. Environ Microbiol. 7(1):47-54.

TRAN, K, GRALLA, J.D. (2008). Control of the timing of promoter escape and RNA catalysis by the transcription factor IIb fingertip. J Biol Chem. 283(23): 15665-15671.

TSAI, F.T., SIGLER, P.B. (2000). Structural basis of preinitiation complex assembly on human pol II promoters. EMBO J. 19(1):25-36.

TSUJI, S., IMAHORI, K., NONOMURA, Y. (1981). The quaternary structure of DNA-

dependent RNA polymerase. J Biochem. 89(6):1903-12.

TURKARSLAN, S., REISS, D.J., GIBBIN, S.G., SU, W.L., PAN, M., BARE, J.C., PLAISIER, .C.L., BALIGA, N.S. (2011). Niche adaptation by expansion and reprogramming of general transcription factors. Mol Syst Biol. 7:554. doi: 10.1038/msb.2011.87.

VAN DE VOSSENBERG, J.L., DRIESSEN, A.J., KONINGS, W.N. (1998). The essence of being extremophilic: the role of the unique archaeal membrane lipids. Extremophiles. 2(3):163-70.

VAN DE WERKEN, H.J., VERHEES, C.H., AKERBOOM, J., DE VOS, W.M., VAN DER OOST, J. (2006). Identification of a glycolytic regulon in the archaea *Pyrococcus* and *Thermococcus*. FEMS Microbiol Lett. 260(1):69-76.

VAN WOLFEREN, M., AJON, M., DRIESSEN, A.J., ALBERS, S.-V. (2013). How hyperthermophiles adapt to change their lives: DNA exchange in extreme conditions. Extremophiles. 17(4): 545-563.

VASSART, A., VAN WOLFEREN, M., ORELL, A., HONG, Y., PEETERS, E., ALBERS, S.-V., CHARLIER, D. (2013). Sa-Lrp from *Sulfolobus acidocaldarius* is a versatile, glutamine-responsive, and architectural transcriptional regulator. Microbiologyopen 2(1): 75-93.

VASSYLYEV, D.G., SEKINE, S., LAPTENKO, O., LEE, J., VASSYLYEVA, M.N., BORUKHOV, S., YOKOYAMA, S. (2002). Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. Nature. 417(6890):712-9.

VILLAFANE, A., VOSKOBOYNIK, Y., RUHL, I., SANNINO, D., MAEZATO, Y., BLUM, P., BINI, E. (2011). CopR of *Sulfolobus solfataricus* represents a novel class of archaeal-specific copper-responsive activators of transcription. Microbiol. 157(Pt 10): 2808-2817.

VOGEL, J.M., ROTH, B., CIGAN, M., FREELING, M. (1993). Expression of the two maize TATA binding protein genes and function of the encoded TBP proteins by complementation in yeast. Plant Cell. 5(11):1627-38.

WAGNER, M., BERKNER, S., AJON, M., DRIESSEN, A.J., LIPPS, G., ALBERS, S.-V. (2009). Expanding and understanding the genetic toolbox of the hyperthermophilic genus *Sulfolobus*. Biochem Soc Trans. 37(Pt 1):97-101. doi: 10.1042/BST0370097.

WAGNER, M., VAN WOLFEREN, M., WAGNER, A., LASSAK, K., MEYER, B.H., REIMANN, J., ALBERS, S.-V. (2012). Versatile Genetic Tool Box for the Crenarchaeote *Sulfolobus acidocaldarius*. Front Microbiol. 3:214. doi: 10.3389/fmicb.2012.00214.

WAGNER, M., WAGNER, A., MA, X., KORT, J., GHOSH, A., RAUCH, B., SIEBERS, B. ALBERS, S.-V. (2013). MalR, the positive regulator of the *Sulfolobus acidocaldarius mal* operon. submitted.

WALKER, C.B., DE LA TORRE, J.R., KLOTZ, M.G., URAKAWA, H., PINEL, N., ARP, D.J., BROCHIER-ARMANET, C., CHAIN, P.S., CHAN, P.P., GOLLABGIR, A., HEMP, J., HÜGLER, M., KARR, E.A., KÖNNEKE, M., SHIN, M., LAWTON, T.J., LOWE, T., MARTENS-HABBENA, W., SAYAVEDRA-SOTO, L.A., LANG, D., SIEVERT, S.M., ROSENZWEIG, A.C., MANNING, G., STAHL, D.A. (2010). *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. Proc Natl Acad Sci U S A. 107(19):8818-23. doi: 10.1073/pnas.0913533107.

WANG, Y., FAIRLEY, J.A., ROBERTS, S.G. (2010). Phosphorylation of TFIIB links transcription initiation and termination. Curr Biol. 20(6): 548-553.

WANG, Y., ROBERTS, S.G. (2010). New insights into the role of TFIIB in transcription initiation. Transcription. 1(3):126-129.

WATERS, E., HOHN, M.J., AHEL, I., GRAHAM, D.E., ADAMS, M.D., BARNSTEAD, M., BEESON, K.Y., BIBBS, L., BOLANOS, R., KELLER, M., KRETZ, K., LIN, X., MATHUR, E., NI, J., PODAR, M., RICHARDSON, T., SUTTON, G.G., SIMON, M., SOLL, D., STETTER, K.O., SHORT, J.M., NOORDEWIER, M. (2003). The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. Proc Natl Acad Sci U S A. 100(22):12984-8.

WATKINS, N.J., BOHNSACK, M.T. (2012). The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. Wiley Interdiscip Rev RNA. 3(3): 397-414.

WEINZIERL, R,O. (2010). Nanomechanical constraints acting on the catalytic site of cellular RNA polymerases. Biochem Soc Trans. 38(2):428-32. doi: 10.1042/BST0380428

WEINZIERL, R.O., WIESLER, S.C. (2011). Revealing the functions of TFIIB. Transcription. 2(6): 254-257.

WERNER, M., CHAUSSIVERT, N., WILLIS, I.M., SENTENAC, A. (1993) Interaction between a complex of RNA polymerase III subunits and the 70-kDa component of transcription factor IIIB. J Biol Chem. 268:20721–20724.

WERNER, F. (2007). Structure and function of archaeal RNA polymerases. Mol Microbiol. 65(6):1395-404.

WERNER, F. (2008). Structural evolution of multisubunit RNA polymerases. Trends Microbiol. 16(6):247-50. doi: 10.1016/j.tim.2008.03.008.

WERNER, F. (2013). Molecular mechanism of transcription elongation in Archaea. Chem. Rev. Article ASAP

WERNER, F., WEINZIERL, R.O. (2002). A recombinant RNA polymerase II-like enzyme 4 capable of promoter-specific transcription. Mol Cell. 10(3):635-46.

WERNER, F., WEINZIERL, R.O. (2005). Direct modulation of RNA polymerase core functions by basal transcription factors. Mol Cell Biol. 25(18): 8344-8355.

WERNER, F., GROHMANN, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. Nat Rev Microbiol. 9(2):85-98. doi: 10.1038/nrmicro2507.

WIESLER, S.C., WEINZIERL, R.O. (2011). "The linker domain of basal transcription factor TFIIB controls distinct recruitment and transcription stimulation functions." Nucleic Acids Res. 39(2): 464-474.

WILBANKS, E.G., LARSEN, D.J., NECHES, R.Y., YAO, A.I., WU, C.Y., KJOLBY, R.A., FACCIOTTI, M.T. (2012). A workflow for genome-wide mapping of archaeal transcription factors with ChIP-seq. Nucleic Acids Res. 40(10):e74. doi: 10.1093/nar/gks063.

WOJTAS, M., PERALTA, B., ONDIVIELA, M., MOGNI, M., BELL, S.D., ABRESCIA, N.G. (2011). Archaeal RNA polymerase: the influence of the protruding stalk in crystal packing and preliminary biophysical analysis of the Rpo13 subunit. Biochem Soc Trans. 39(1):25-30. doi: 10.1042/BST0390025.

WOJTAS, M.N., MOGNI, M., MILLET, O., BELL, S.D., ABRESCIA, N.G. (2012). Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. Nucleic Acids Res.40(19): 9941-9952.

WOOD, E.R., GHANE, F., GROGAN, D.W. (1997) Genetic responses of the thermophilic

archaeon *Sulfolobus acidocaldarius* to short-wavelength UV light. J Bacteriol. 179:5693–5698.

WU, C.C., LIN, Y.C., CHEN, H.T. (2011). The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIIC, Bdp1, and the RNA polymerase III active center. Mol Cell Biol. 31(13): 2715-2728.

WURTZEL, O., SAPRA, R., CHEN, F., ZHU, Y., SIMMONS, B.A., SOREK, R. (2010). A single-base resolution map of an archaeal transcriptome. Genome Res. 20(1): 133-141.

YAMAGUCHI, Y., SHIBATA, H., HANDA, H. (2013). Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. Biochim Biophys Acta. 1829(1): 98-104.

YAMAMOTO, T., MATSUDA, T., INOUE, T., MATSUMURA, H., MORIKAWA, M., KANAYA, S., KAI, Y. (2006). Crystal structure of TBP-interacting protein (Tk-TIP26) and implications for its inhibition mechanism of the interaction between TBP and TATA-DNA. Protein Sci. 15(1): 152-161.

YANG, Y., CAO, J., HUANG, L., FANG, H.Y., SHENG, H.Z. (2006). Regulated expression of TATA-binding protein-related factor 3 (TRF3) during early embryogenesis. Cell Res. 16(7): 610-621.

YOKOYAMA, K., ISHIJIMA, S.A., KOIKE, H., KURIHARA, C., SHIMOWASA, A., KABASAWA, M., KAWASHIMA, T., SUZUKI, M. (2007). Feast/Famine regulation by transcription factor FL11 for the survival of the hyperthermophilic archaeon *Pyrococcus* OT3. Structure. 15(12),1542-1554.

ZAGO, M.A., DENNIS, P.P., OMER, A.D. (2005). The expanding world of small RNAs in the hyperthermophilic archaeon *Sulfolobus solfataricus*. Mol Microbiol. 55(6):1812-28.

ZAYCHIKOV, E., DENISSOVA, L., HEUMANN, H. (1995). Translocation of the *Escherichia coli* transcription complex observed in the registers 11 to 20: "jumping" of RNA polymerase and asymmetric expansion and contraction of the "transcription bubble". Proc Natl Acad Sci U S A. 1995 Feb 28;92(5):1739-43.

ZHANG, G., CAMPBELL, E.A., MINAKHIN, L., RICHTER, C., SEVERINOV, K., DARST, S.A. (1999). Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 A resolution. Cell. 98(6):811-24.

ZHANG, C., GUO, L., DENG, L., WU, Y., LIANG, Y., HUANG, L., SHE, Q. (2010). Revealing the essentiality of multiple archaeal pcna genes using a mutant propagation assay based on an improved knockout method. Microbiol. 156(Pt 11):3386-97. doi: 10.1099/mic.0.042523-0.

ZHANG, Z., GUO, L., HUANG, L. (2012). Archaeal chromatin proteins. Sci China Life Sci 55(5): 377-385.

ZHU, W., ZENG, Q., COLANGELO, C.M., LEWIS, M., SUMMERS, M.F., SCOTT, R.A. (1996). The N-terminal domain of TFIIB from *Pyrococcus furiosus* forms a zinc ribbon. Nat Struct Biol. 3(2):122-4.

ZILLIG, W., STETTER, K.O., JANEKOVIĆ, D**.** (1979). DNA-dependent RNA polymerase from the archaebacterium *Sulfolobus acidocaldarius*. Eur J Biochem. 96(3):597-604.

8. Appendix

8.1. List of abbreviations

µ	mikro
5-FOA	5-Fluoroorotic acid
Ade	Adenine
bp	base pairs
BLAST	Basic local alignment search tool
BRE	TFB-responsive element
bZIP	Basic leucine zipper
cDNA	copy DNA
CTD	Carboxy terminal domain
DNA	Desoxy ribonucleic acid
DPE	Downstream promoter element
ECL	Enhanced chemiluminescent substrate
e.g.	for example
EDTA	Ethylenediaminetetraacetic acid
et al.	*et alii*: and others
g	Gravitational acceleration
GTF	General transcription factor
h	hour
HABA	2-[4'-hydroxy-benzeneazo]benzoic acid
His	Histidine
HRP	Horseradish peroxidase
HTH	Helix-turn-helix
i.e.	*id est*: that is
Inr	Initiator element
IPTG	Isopropyl-β-D-thiogalactopyranosid
kDa	kilo Dalton
L	liter
LB	Lysogeny broth (also known as Luria-Bertani or Lennox broth)
Leu	Leucine
LRP	Leucine responsive regulatory protein
M	molar (mol/l)
mA	milli Ampere
MD	maltodextrin
min	minute
mod.	modified
mRNA	messenger RNA
MW	Molecular weight
n	nano
ncRNA	non-coding RNA
OD	Optical density
ONPG	Ortho nitrophenyl β galactosid

PAA	Polyacrylamide
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
pH	negative decadic logarithm of the hydrogen ion concentration [H^+]
PIC	Pre-initiation complex
PMSF	Phenyl methane sulfonyl fluoride
RBS	Ribosomal-binding site
RNA	Ribonucleic acid
RNAP	RNA polymerase
RPM	Rounds per minute
RT	Room temperature
SDS	Sodium dodecyl sulfate
siRNA	silencing RNA
snoRNA	small nucleolar RNA
TBP	TATA-binding protein
TEMED	NN,N´,N´-Tetramethylethylenediamine
TFB	Transcription factor B, homolog of eukaryotic TFIIB in Archaea
TFE	Transcription factor E, homolog of eukaryotic TFIIE-α in Archaea
Tris	Tris-(hydroxymethyl)aminomethane
tRNA	transfer RNA
Trp	Tryptophane
U	Unit of enzyme activity
UAS	Upstream activating site
UV	Ultra violet
V	Volt
X-α-Gal	5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside
Zn	Zinc

8.2. Publikationsliste

in preparation — Marrero, J., Blombach, F., Rauch, B., Brinkmann, H., Siebers, B.
Phylogeny and genome context analysis of Transcription Factor B in Archaea: evidence of a core of TFB orthologs of Euryarchaea and Crenarchaea sharing a common history

submitted — Bräsen, C., Esser, D., Rauch, B., Siebers, B. (2013) (MMBR review)
Unique features of archaeal carbohydrate metabolism

submitted — Wagner, M., Wagner, A., Ma, X., Kort, J., Ghosh, A., Rauch, B., Siebers, B., Albers, S.-V. (2013) (ASM)
MalR, the positive regulator of the *Sulfolobus acidocaldarius* mal operon

10/2010 Rauch, B., Pahlke, J., Schweiger, P., Deppenmeier, U. **Characterization of enzymes involved in the central metabolism of *Gluconobacter oxydans*,** Appl Microbiol Biotechnol. 2010 Oct;88(3):711-8.

8.2.1. Poster und Vorträge

10.-12.04.2013 **Functional analysis of multiple transcription factors in *S. acidocaldarius*,** Annual GRK Retreat, Ostbevern (Vortrag)

03.-04.05.2011 **Hot transcription: Functional analysis of multiple transcription factors in *S. acidocaldarius***, Annual GRK Retreat, Dormagen-Zons (Vortrag)

07.-10.07.2010 **Transcription in Archaea: A hot mosaic**, Chromatin and Epigenetics, Essen (Vortrag)

16.-17.04.2010 **Transcription: The hot way**, Annual GRK Retreat, Düsseldorf (Vortrag)

10.-13.09.2012 **Functional analysis of multiple transcription factors in *S. acidocaldarius*,** Extremophiles, Sevilla (Spanien), FEMS Grant (Poster)

05.-07.09.2012 **Functional analysis of multiple transcription factors in *S. acidocaldarius*,** Chromatin and Epigenetics, Essen (Poster)

02.-04.07.2012 **Functional analysis of multiple transcription factors in *S. acidocaldarius*,** Molecular Biology of Archaea3, Marburg (Poster)

03.-06.04.2011 **Hot transcription: Functional analysis of multiple general transcription factors in Crenarchaeota**, VAAM Jahrestagung, Karlsruhe (Poster)

06.-07.01.2011 **Transcription in Achaea: A hot mosaic**, Genetics, Biochemistry and Molecular Biology of Archaea, St. Andrews (Schottland), Poster prize (Poster)

8.3. Lebenslauf

Der Lebenslauf ist in der Online-Version aus Datenschutzgründen nicht enthalten.

8.4. Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel

„Functional analysis of multiple general transcription factors in *Sulfolobus acidocaldarius*"

selbst verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe, und dass die Arbeit in dieser oder ähnlicher Form noch bei keiner anderen Universität eingereicht wurde.

Essen, September 2013

Unterschrift

8.5. Acknowledgements

I want to thank my supervisor Prof. Dr. Bettina Siebers for giving me the opportunity to work in her lab, providing me with helpful advice and critical reading of this thesis.

I`m deeply grateful to Prof. Dr. Perihan Nalbant for being one of the referees of this thesis.

My special thanks are dedicated to Jeannette Marrero Coto who supported the present work with ideas, experimental advice and for critical reading of this thesis. Many thanks for training me in the methods of yeast two-hybrid, EMSAs and *in vitro* transcription.

I would like to express my gratitude to Dr. Sonja-Verena Albers from the Max-Planck Institute for Terrestrial Microbiology in Marburg and her lab members, especially Michaela, Alexander and Anna-Lena, for providing the plasmids for overexpression/reporter gene assays and genetic manipulation of *Sulfolobus* cells. Thank you very much for supplying the present work with the fundamental genetic tools, supportive advice and instructions. My thanks also go to Prof. Malcolm White from the University of St. Andrews for sending the yeast two-hybrid vectors and the positive controls from the publication Paytubi and White, 2009.

During my time as a PhD student I had the chance to work in an inspiring environment. I thank my colleagues from the MEB for their timely support, friendly discussions, delicious dinners and for having a nice time in the lab. Thanks: Tina, Jeannette, Mella, Christopher, Theresa, Patrick, Anna, Silke, Jens, Sabine, Julia, Kohei, Dominik, Verena, Thomas, Britta and Frank. In addition I want to express my gratitude to my student assistants Marius and Max for supporting my PhD-project with their help.

This project was funded by the DFG and I thank them very much for the financial support. A lot of thanks are dedicated to the members of the graduate college GRK 1431, the PIs as well as the students, for supporting my present work with material, helpful advice and fruitful discussions. In general I want to thank you for having a lot of good times together as GRK member. .

I thank my friends for their encouragement and support during my time as PhD and writing of the thesis. Especially I want to thank Sarah and Steffi for proofreading of this thesis, helpful advice and for having a good time with you.

I express my deep gratitude to my beloved family for their valuable support, understanding and helpful encouragement.