The Transcriptional Regulator TFB-RF1 Activates Transcription of a Putative ABC Transporter in Pyrococcus furiosus

Robert Reichelt, Katharina M. A. Ruperti, Martina Kreuzer, Stefan Dexl, Michael Thomm and Winfried Hausner*

Institute of Microbiology and Archaea Center, University of Regensburg, Regensburg, Germany

Transcription factor B recruiting factor 1 (TFB-RF1; PF1088) is a transcription regulator which activates transcription on archaeal promoters containing weak TFB recognition elements (BRE) by recruiting TFB to the promoter. The mechanism of activation is described in detail, but nothing is known about the biological function of this protein in *Pyrococcus furiosus*. The protein is located in an operon structure together with the hypothetical gene *pf1089* and western blot as well as end-point RT-PCR experiments revealed an extremely low expression rate of both proteins. Furthermore, conditions to induce the expression of the operon are not known. By introducing an additional copy of *tfb-RF1* using a *Pyrococcus* shuttle vector we could circumvent the lacking expression of both proteins under standard growth conditions as indicated by western blot as well as end-point RT-PCR experiments. A ChIP-seq experiment revealed an additional binding site of TFB-RF1 in the upstream region of the *pf1011/1012* operon, beside the expected target of the *pf1089/tfb-RF1* region. This operon codes for a putative ABC transporter which is most-related to a multidrug export system and *in vitro* analysis using gel shift assays, DNase I footprinting and *in vitro* transcription confirmed the activator function of TFB-RF1 on the corresponding promoter. These findings are also in agreement with *in vivo* data, as RT-qPCR experiments also indicate transcriptional activation of both operons. Taken together, the overexpression strategy of *tfb-RF1* enabled the identification of an additional operon of the TFB-RF1 regulon which indicates a transport-related function and provides a promising starting position to decipher the physiological function of the TFB-RF1 gene regulatory network in *P. furiosus*.

Keywords: archaea, *Pyrococcus*, transcriptional activation, TFB-RF1, ABC transporter, ChIP-seq

INTRODUCTION

The core transcription machinery of Archaea resembles a simplified version of the eukaryotic RNA polymerase (RNAP) II system, which consists of only one multi-subunit RNA polymerase (RNAP). For initiation of transcription three general transcription factors (TFs), the TATA-binding protein (TBP), TFB, and TFE are necessary for recognition of the two core promoter elements, the TATA box and the TFB recognition element (BRE; Soppa, 2001; Thomm and Hausner, 2007; Grohmann and Werner, 2011; Weinzierl, 2013; Gehring et al., 2016). The first step of transcriptional initiation
is the interaction of TBP with the TATA box. In the next step TFB stabilizes this complex and also interacts with the BRE upstream and with the transcription start site downstream of the TATA box (Renfrow et al., 2004). Recruitment of the RNAP completes the pre-initiation complex. The third TF, TFE, stimulates open complex formation and stabilizes early elongation complexes (Grüenberg et al., 2007; Naji et al., 2007; Blombach et al., 2015). During elongation, the elongation factor Spt4/5 can displace TFE from the RNAP and stimulate processivity (Grohmann et al., 2011). Recent data indicate that Spt4/5 is recruited proximal to the transcription start site on the majority of transcription units, whereas on a subset of genes Spt4/5 is recruited later during elongation (Smollett et al., 2017).

It is interesting to note that the regulation of this eukaryotic-like transcription machinery is controlled by a system very similar to that found in Bacteria (Aravind and Koonin, 1999; Soppa, 2001; Bell, 2005; Geiduschek and Ouhammouch, 2005; Jun et al., 2011; Peeters et al., 2013; Karr, 2014). Corresponding genes are also frequently organized in operon structures and controlled by single promoters. The transcription rate of such promoters is influenced by transcription regulators, which in many cases consist of a helix-turn-helix DNA-binding motif. One general physiological function of these TFs is to sense signals from outside (like, e.g., availability of substrates or presence of toxic compounds) and translate them into a transcription dependent expression outcome. For this, these regulators often interact as dimers or oligomers with semipalindromic DNA sequences located either upstream, within or downstream of the corresponding promoters. Bound regulators within the promoter interfere with TFB or TBP binding by steric hindrance and if bound further downstream inhibit RNAP recruitment. In both cases the regulators work as repressors and transcription is abolished on the level of initiation (Peeters et al., 2013).

In contrast, activating TFs usually bind upstream of BRE and TATA box and a detailed in vitro analysis of archaeal activation mechanisms revealed so far two basic principles: the activator promotes either binding of TBP or TFB (Ouhammouch et al., 2005; Ochs et al., 2012). In the first case the TATA box has strong deviations from the consensus sequence and in the second case there is an imperfect BRE. A detailed analysis of the gas vesicle regulator GvpE of Halobacterium salinarum revealed that this activator can even interact with both, different TBPs and TFB (Teufel and Pfeifer, 2010; Bleiholder et al., 2012).

In Pyrococcus furiosus, there are several TFs described in more detail: the leucine-responsive regulatory protein, Lrp, which inhibits transcription of the own gene by abrogating RNAP recruitment (Dahle and Thomm, 2002). Phr, the regulator of the archaeal heat shock genes, has a similar mechanism (Vierke et al., 2003). In both cases, the TBP/TFB complex and the corresponding regulators can independently bind to the DNA. SurR can act as transcriptional activator and repressor controlling hydrogen and elemental sulfur metabolism (Lipscomb et al., 2009). TrmB and TrmBL1 are two sugar sensing regulators of ABC transporter genes, which are controlled by different inducers (Lee et al., 2005, 2007; Gindner et al., 2014). A detailed analysis of the Pyrococcus genome revealed more than 80 putative TFs, which represent about 4% of all open reading frames (Pérez-Rueda and Janga, 2010). The physiological function of almost all of these TFs and the corresponding gene regulatory networks are unknown.

One of these candidates with unknown biological function is TFB-RF1, which activates transcription by recruiting TFB to the promoter (Ochs et al., 2012). To shed light on the function of TFB-RF1, we studied the genome-wide binding sites of the regulator. We expressed the protein using a shuttle vector-based genetic system for Pyrococcus (Waage et al., 2010). Applying chromatin immunoprecipitation in combination with high-throughput sequencing (ChiP-seq) we identified a putative ABC transporter as an additional member of the TFB-RF1 regulatory network.

**MATERIALS AND METHODS**

**Strains and Growth Conditions**

*Pyrococcus furiosus* was cultivated under anaerobic conditions at 85°C in 1/2 SME medium as described previously (Waage et al., 2010). Gelrite (1%) was added for solidification of the medium. For the comparison of the two strains, *Pfu* pYS3 containing the empty shuttle vector and *Pfu* pYS5 with *tfb-rf1* under the control of a gluconeogenic promoter, a slightly modified medium without peptonite and with 0.025% yeast extract was used. Growing on starch was performed in the presence of 0.1% starch and for growing on pyruvate, the starch was replaced with 40 mM Na-pyruvate. For the selection of the shuttle vector the antibiotic simvastatin was used at a concentration of 10 μM. Each strain was cultivated with starch or with pyruvate in three independent experiments until a cell density of approximately 1 × 10^8 cells per ml. Cell numbers were analyzed with a Thoma counting chamber (0.02 mm depth; Marienfeld, Lauda-Königshofen, Germany) using phase-contrast microscopy.

**Strains, Plasmids, and Primers**

All used strains, plasmids, and primers are shown in the supplement (Supplementary Table 1).

**Construction of the Shuttle Vector pYS5 and Transformation in *P. furiosus***

Shuttle vector pYS5 was obtained according to the construction of pYS4 (Waage et al., 2010). The promoter sequence of the fructose-1,6 bisphosphatase was amplified from genomic DNA using the primers EcoRV-PP0613Pr-F and PF1088/PF0613Prom-R. For the amplification of *tfb-rf1* the primers PF1088-F and PF1088-His-R were used. The terminating region of the histone A1 gene was obtained using the primer pair His-PF1831Term-F and PF1831T-EcoRV-R. Overlapping parts of fragments were fused by single overlap extension PCR. Finally, the construct was integrated into the pYS3 vector next to the *hmgCoA* reductase gene cassette using the flanking EcoRV sites. The construction of the plasmid was verified by DNA sequencing. Transformation into *P. furiosus* cells was done according to Waage et al. (2010).
ChIP experiments were carried out as described previously (Reichelt et al., 2016). Antibody containing fractions were pooled and dialyzed overnight against phosphate-buffered saline (PBS). For the preparation of cell extracts from the Pyrococcus strains, cell cultures with a volume of 20 ml and a cell density of approximately 1 × 10^8 cells per ml were harvested, washed with 1 ml PBS and resuspended in 300 μl PBS including protease inhibitor mix (cOmplete Ultra Tablets, Roche Applied Science, Mannheim, Germany). Cells lysis was induced by sonication (Sonopuls HD2070, Bandelin Electronics, Berlin, Germany), cell debris was removed by centrifugation and the protein concentrations of the supernatants were determined by Bradford assay. The signals were visualized using a fluorescence image analyzer (FLA-5000, Fuji, Japan). Each experiment was repeated three times and representative data are shown in the manuscript.

**Immunoprecipitation**

ChIP experiments were carried out as described previously (Reichelt et al., 2016). *P. furiosus* cells harboring the pYS5 plasmid were cultured in a fermenter in the presence of 40 mM sodium pyruvate for formaldehyde-crosslinking. For immunoprecipitation 10 μg of purified polyclonal antibodies raised against TFB-RF1 was coupled to 50 μl Dynabeads Protein G (Thermo Fisher Scientific, Waltham, MA, United States) according to the manufacturer's instructions. After elution from the beads DNA was purified and subsequently used for library preparation. Input sample was prepared as described previously (Reichelt et al., 2016) and three replicates of immunoprecipitation (IP1, IP2, and IP3) were used for further analysis.

**Library Preparation and Sequencing**

Library preparations were done according to the NEBNext<sup>®</sup>-ChIP-Seq library prep reagent set for Illumina protocol (New England Biolabs, Ipswich, MA, United States). For multiplex sample preparation the NEBNext<sup>®</sup> Multiplex oligos (primer sets 1 and 2; New England Biolabs, Ipswich, MA, United States) were used and libraries were PCR amplified by the NEBNext<sup>®</sup> High Fidelity Master Mix (New England Biolabs, Ipswich, MA, United States). Libraries were pooled in equimolar ratios and sequenced using the Illumina HiSeq 2000 platform (read length = 50 b) (Illumina, San Diego, CA, United States). For further analysis only the demultiplexed and quality filtered reads were used.

**Data Processing and Peak Calling**

Reads were mapped to the *P. furiosus* DSM3638 genome using Bowtie2 with default settings (Langmead et al., 2009). Aligned and unaligned reads were written to different files. The Integrative Genome Viewer (IGV) was used for visualization of short-reads genome occupancy (Robinson et al., 2011). The Bam file of aligned reads was converted to the Sam format (Li et al., 2009) and from each sample 10 million reads were randomly selected (Blankenberg et al., 2010). After converting back to the Bam format samples were analyzed with the Pique software package (Wilbanks et al., 2012) using the following settings: 'a = 50; l = 300; ‘analysis_region’: 0 1908256 (whole genome); ‘norm_region’: 1660000 1661000 (pfk promoter region); and ‘mask’ 1613140 1629427 (PF1937-PF1952). Called peaks were verified by repeating the analysis using additional genomic regions or no specified genomic region as ‘norm_region.’ Downstream analyses were performed as recommended for the software. Statistical analysis was carried out using SPSS Statistics 21 (IBM). In addition, only binding sites were considered for further analysis, which were present in the three replicates of IP, which do not correspond to the ribosomal gene locus in the *P. furiosus* genome and which show higher enrichment scores than the regions corresponding to genes and promoter or terminator regions present as an additional copy on the plasmid pYS5 (Supplementary Table 2). ChIP-seq raw data are available in the ArrayExpress database<sup>1</sup> under accession number E-MTAB-6618.

**RNA Isolation**

*Pyrococcus furiosus* total RNA was purified using the peqGOLD TriFast<sup>™</sup> reagent (Peqlab, Erlangen, Germany). 10 to 20 ml cell culture was pelleted and cells were lysed by addition of 1 ml TriFast followed by rigorous shaking for 15 min. After adding 0.2 ml 2 M sodium acetate pH 4.0 RNA was isolated according to the manufacturer's instructions. Contaminating DNA was removed via the TURBO DNA-free<sup>™</sup> Kit (Thermo Fisher Scientific, Waltham, MA, United States). The integrity of the total RNA was assessed by formaldehyde agarose gel electrophoresis and purified RNA was stored at −80. Three independent RNA samples were prepared from both culture conditions and strains.

**Reverse Transcriptase PCR (RT-PCR) and Quantitative Real-Time PCR (RT-qPCR)**

Reverse transcription was done using a QuantiTect Reverse Transcription Kit (Qiagen, Hilden, Germany) in combination with a Random Primer Mix (Promega, Madison, WI, United States). For each assay 415 ng total RNA was used. RT-PCR and RT-qPCR primer pairs were designed using the Primer3 software package (Supplementary Table 1; Untergasser et al., 2012). To confirm the absence of genomic DNA, negative control reactions containing RNA and all reagents except for the reverse transcriptase were performed (−RT cDNA).

---

<sup>1</sup>www.usegalaxy.org

<sup>2</sup>www.ebi.ac.uk/arrayexpress
RT-qPCR reactions were assembled as triplicates in a total volume of 10 µl using the SensiMix® SYBR® No-ROX Kit (Bioline, Luckenwalde, Germany). Primers were added to a final concentration of 0.3 µM and the total volume of the DNA samples in each reaction was 4 µl. The efficiency of the primer pairs, determined by serial dilutions, was in the range between 0.9 and 0.95 (Bustin et al., 2009). RT-qPCR reactions were run on a Rotorgene 6000 platform using a three step protocol: 95°C – 10’ for one cycle; 95°C – 30’, 58°C – 30’, 72°C – 30’ for 40 cycles. No template controls were used as negative controls for each primer pair and the specificity of the PCR products was verified by melt curve analysis. Data evaluation was done using the corresponding Rotorgene software package (Qiagen, Hilden, Germany). Relative expression levels are shown as log2 ratios as calibrator. Pf0256 encodes the constitutively expressed RNA polymerase subunit E’. The applicability of pf0256 as calibrator was evaluated using various genes and *P. furiosus* wild type cells grown on pyruvate and starch (Supplementary Figure 1).

## Proteins and DNA Templates

The purification of RNAP and the TFs TBP, TFB, and TFB-RF1 was performed as described previously (Ochs et al., 2012). DNA regions of interest were obtained from genomic DNA by PCR amplification with the Phusion DNA polymerase (New England Biolabs, Ipswich, MA, United States) using corresponding primers (Supplementary Table 1). For electrophoretic mobility shift assay (EMSA) one of the primers was labeled with Cyanine-5 (Cy5) and for DNaseI footprinting one of the primers was labeled with 6-carboxyfluorescein (FAM). For in vitro transcription assays DNA templates were PCR amplified using unmodified primers.

## Electrophoretic Mobility Shift Assay (EMSA)

Cy5-labeled DNA and various amounts of proteins (TBP, TFB, RNAP, and TFB-RF1) were assembled in a 15 µl reaction volume according to Ochs et al. (2012). After incubation at 70°C for 15 min, protein-DNA complexes were analyzed using a non-denaturating 8% polyacrylamide gel. DNA fragments were visualized using a fluor imager (FLA-5000, Fuji, Japan). EMSAs were repeated at least two times and one representative gel is presented.

## DNase I Footprint

FAM-labeled template DNA and TFB-RF1 were incubated under the conditions used for the gel shift assay. DNase I was added for 1 to 5 min at 37°C and the reaction was stopped by the addition of 95% formamide. DNA was ethanol precipitated and resuspended in 3 µl of formamide buffer. A DNA sequencing ladder using the same primer was generated as a molecular mass standard. Samples were loaded onto a 4.5% denaturing polyacrylamide gel and analyzed using an ABI 377 DNA sequencer. DNase I footprints were repeated two times and one representative experiment is shown.

## In Vitro Transcription Assay

Transcription reactions were assembled in a volume of 25 µl in the presence of 10 nM template DNA, 10 nM RNAP, 238 nM TFB, 135 nM TFB, 0.25 mg/ml BSA, 440 µM ATP, 440 µM GTP, 440 µM CTP, 2.7 µM UTP, and 0.049 MBq [α-32P]-UTP (111 TBq/mmol) as described previously (Ochs et al., 2012). After incubation at 70°C for 10 min the RNA transcripts were purified by phenol/chloroform/isoamylalcohol extraction and separated on a denaturing 8% polyacrylamide gel. Radioactive labeled transcription products were visualized using a Phosphorimag (FLA-5000, Fuji, Japan). For quantification of the transcripts we used the AIDA Image Analyzer software (Raytest, Straubenhardt, Germany).

## Bioinformatical Analyses

The genomic sequences of 35 *Thermococcales* species were obtained from NCBI Taxonomy Browser. Homologous protein searches were done by BLAST. TF binding site discovery was performed using the MEME Suite (Bailey et al., 2009). Predictions of transmembrane helices in proteins were done using the TMHMM Server v. 2.0. Sequence logos were created by WebLogo 3 (Crooks et al., 2004).

## RESULTS

The transcriptional regulator TFB-RF1 (PF1088) activates transcription on archael promoters containing weak BRE elements by recruiting TFB to the promoter (Ochs et al., 2012). We have described the mechanism of activation in detail, but so far we have no indication for the biological role of this protein in *P. furiosus*. The corresponding gene is located downstream in an operon structure together with the hypothetical gene pf1089 (Figure 1A) and TFB-RF1 activates transcription of this operon (Ochs et al., 2012). To get an idea about the physiological function of TFB-RF1 we checked the expression of the regulator under glycolytic (starch) as well as gluconeogenic (pyruvate) conditions by western blot analysis. The TFB-RF1-specific polyclonal antibody recognized the recombinant protein (Figure 1B, lanes 1 and 2), but no signals were detected in crude extracts from *P. furiosus* cells grown under both conditions (Figure 1B, lanes 3 to 6, upper part), except control signals with an antibody raised against subunit F of the RNAP (lower part).

---

1. www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/
2. blast.ncbi.nlm.nih.gov/Blast.cgi
3. meme-suite.org
4. www.cbs.dtu.dk/services/TMHMM/
5. weblogo.threeplusone.com/
To check the expression of the pf1089/tfb-rf1 operon on the RNA level, end-point RT-PCR was performed. Using a PCR amplicon spanning both genes (indicated by arrows in Figure 1A), only a weak signal could be detected in cells grown on starch or pyruvate (Figure 1C, lanes 3 and 4) in comparison to the housekeeping gene pf0256 (lanes 5 and 6). The molecular weight marker is shown on the left (M). Primary polyclonal antibodies raised against TFB-RF1 (upper part) and the RNA polymerase subunit E (RpoF; lower part) were used in a 1:1000 dilution. RT-PCR was performed in the presence (+RT) or absence (−RT) of reverse transcriptase. The specificity of the PCR reaction was checked via a no template control (NTC). P0256 encodes the RNAS polymerase subunit E’ and serves as housekeeping gene. The lanes with molecular weight markers are depicted with M.

### Figure 1

**Expression of TFB-RF1 in Pyrococcus furiosus.**

**A** Schematic drawing of the pf1089/tfb-rf1 operon. The direction of transcription and the primers used for RT-PCR are indicated by arrows. The corresponding amplicon is 129 bp and encompasses both genes.

**B** Western blot analysis using recombinantly expressed TFB-RF1 (lanes 1 and 2) and P. furiosus (Pfu) crude extracts from cells grown on starch (lanes 3 and 4) or pyruvate (lanes 5 and 6). The molecular weight marker is shown on the left (M). Primary polyclonal antibodies raised against TFB-RF1 (upper part) and the RNA polymerase subunit E (RpoF; lower part) were used in a 1:1000 dilution.

**C** End-point RT-PCR of the pf1089/tfb-rf1 operon and pf0256 with total RNA obtained from P. furiosus cells grown on starch (lanes 1 to 6) or pyruvate (lanes 7 to 12). RT-PCR was performed in the presence (+RT) or absence (−RT) of reverse transcriptase. The specificity of the PCR reaction was checked via a no template control (NTC). P0256 encodes the RNA polymerase subunit E’ and serves as housekeeping gene. The lanes with molecular weight markers are depicted with M.

**FIGURE 1 |** Expression of TFB-RF1 in Pyrococcus furiosus. (A) Schematic drawing of the pf1089/tfb-rf1 operon. The direction of transcription and the primers used for RT-PCR are indicated by arrows. The corresponding amplicon is 129 bp and encompasses both genes. (B) Western blot analysis using recombinantly expressed TFB-RF1 (lanes 1 and 2) and P. furiosus (Pfu) crude extracts from cells grown on starch (lanes 3 and 4) or pyruvate (lanes 5 and 6). The molecular weight marker is shown on the left (M). Primary polyclonal antibodies raised against TFB-RF1 (upper part) and the RNA polymerase subunit F (RpoF; lower part) were used in a 1:1000 dilution. (C) End-point RT-PCR of the pf1089/tfb-rf1 operon and pf0256 with total RNA obtained from P. furiosus cells grown on starch (lanes 1 to 6) or pyruvate (lanes 7 to 12). RT-PCR was performed in the presence (+RT) or absence (−RT) of reverse transcriptase. The specificity of the PCR reaction was checked via a no template control (NTC). P0256 encodes the RNA polymerase subunit E’ and serves as housekeeping gene. The lanes with molecular weight markers are depicted with M.

### Overexpression of TFB-RF1 Using the Shuttle Vector pYS5

To circumvent the low expression rate of the pf1089/tfb-rf1 operon we used a Pyrococcus/Esherichia coli shuttle vector for overexpression of TFB-RF1 (Waage et al., 2010). The constructed plasmid pYS5 contains an additional copy of tfb-rf1 under the control of a gluconeogenic promoter and was successfully transformed into Pyrococcus (Figure 2A). A growth behavior analysis of Pyrococcus pYS5 under glycolytic as well as gluconeogenic conditions showed only minor differences compared to cells harboring the empty shuttle vector pYS3 as control (Supplementary Figure 2). To test for overexpression of TFB-RF1 we performed western blot experiments with both strains either grown on starch or pyruvate as described before (Figure 2B). Using crude extracts from Pyrococcus pYS3, TFB-RF1 is also not detectable similar to the wild type (Figure 2B, lanes 1 to 4). In contrast, using Pyrococcus pYS5 a weak signal could be detected in crude extract from cells grown with starch and two increased signals in crude extract from pyruvate cells (Figure 2B, lanes 5 to 8). The two signals correspond to the monomer and the dimer of TFB-RF1 and all attempts to achieve complete denaturation of the protein with increased SDS concentrations were not successful (data not shown). It is interesting to note that the ratio between the monomer and the dimer is the other way round in comparison to the recombinant protein produced in E. coli (compare Figure 1B, lane 2 and Figure 2B, lane 8).

To check if the plasmid-expressed TFB-RF1 is able to act in trans on the activation of the genome-located pf1089/tfb-rf1 operon we performed RT-PCR reactions as described before. Using RNA from Pyrococcus pYS5 grown on pyruvate, an enhanced pf1089/tfb-rf1 signal could be detected even stronger than the signal of the house-keeping pf0256 (Figure 2C, lanes 19 and 22). As the promoter is only active under gluconeogenic conditions we obtained wild type results under glycolytic conditions (Figure 2C, lanes 13 and 16). This is also the case for RNA from the control strain Pyrococcus pYS3 (Figure 2C, lanes 1, 4, 7, and 10). Taken together, these results demonstrate that the strategy to bypass the missing natural expression of TFB-RF1 by introducing an additional plasmid-encoded copy of tfb-rf1 is working and TFB-RF1 is able to stimulate transcription on corresponding promoters located in the genome.

### Genome-Wide Identification of TFB-RF1 Bindings Sites in Vivo by ChIP-seq

To investigate if TFB-RF1 recognizes additional binding sites within the Pyrococcus genome we performed ChIP-seq experiments as described previously (Reichelt et al., 2016). We used formaldehyde-fixed Pyrococcus pYS5 cells grown on pyruvate for immunoprecipitation with the TFB-RF1-specific antibody in three replicates. The software package PIQUE was used for peak calling, which is adapted to the requirements of the strategy to bypass the missing natural expression of TFB-RF1 by introducing an additional plasmid-encoded copy of tbf-rf1 is working and TFB-RF1 is able to stimulate transcription on corresponding promoters located in the genome.
pf1089 and proofs that transcriptional activation of the operon (Figure 2C) is correlated with in vivo binding of TFB-RF1 upstream of pf1089 (Figure 3B). The additional signal for tfb-rf1 in both samples, IP and input DNA, resulted from overexpression on the plasmid. The second peak covers the promoter region upstream of the operon pf1011/pf1012, which encode most likely an ABC transporter system (Figure 3C). The ChIP-seq approach indicates that the regulatory function of TFB-RF1 is not only restricted to auto-activation of its own operon, but also controls in addition operon pf1011/ pf1012.

Activation of the pf1011/pf1012 Operon by TFB-RF1 in Vitro and in Vivo

To verify the specific interaction of TFB-RF1 with the promoter region of the newly identified operon pf1011/ pf1012 a gel shift assay was performed (Figure 4A). Using the upstream region of pf1011 as template with increasing concentrations of TFB-RF1 a specific DNA protein complex was formed (Figure 4A, lanes 2 to 5), but there is no interaction with the strong gdh promoter (lanes 7 to 10). To determine the binding site of TFB-RF1 in more detail, we also applied DNase I footprinting experiments (Figure 4B). TFB-RF1 protected a region on the non-template strand from −47 to −70 upstream of the translation start site (Figure 4B, lanes 5 to 8 and Figure 4C). This region shows striking sequence similarities to the one found in the pf1089 promoter region in our previous study (Figure 4C, lower part; Ochs et al., 2012). Both DNA sequences contain the minimal palindromic consensus sequence 5′-TCTG-N6-CAGA-3′ within the footprint, which most likely serves as DNA recognition element for TFB-RF1 in both promoter regions (Figure 4C, blue letters). Additionally, the binding sites predicted with peak calling using PIQUE are located within both footprint regions (Figure 4C, black asterisks and Supplementary Table 2) which shows the high spatial resolution of the technique. The TFB-RF1 binding site in the pf1011 promoter is located just upstream of a presumable BRE and TATA box in accordance to the pf1089 promoter (Figure 4C, red letters). A comparison of the BRE of pf1011 with the Pyrococcus consensus sequence 5′-VRAAA-3′ (van de Werken et al., 2006) also confirms a weak BRE, although with a reduced number of deviations compared to the consensus sequence. Therefore, we assume that pf1011 also requires TFB-RF1 for efficient TFB recruitment and for transcriptional activation.

To test this hypothesis we applied a gel shift assay with the upstream region of pf1011 as DNA template. Using sufficient TBP together with a TFB concentration of 24 nM, no TBP/TFB/DNA shift was formed and even higher amounts (96 nM) did not lead to a strong shift (Figure 5A, lanes 2 and 3). In contrast, in the presence of TFB-RF1 a TBP/TFB/TFB-RF1/DNA complex was efficiently build even under TFB-limiting conditions (Figure 5A, lanes 4 and 5). Furthermore, TFB-RF1 stimulates transcription from the pf1011 promoter in a cell-free transcription assay.
FIGURE 3 | ChIP-seq analysis with Pfu pYS5 grown on pyruvate. Mapping of the TFB-RF1 immunoprecipitation (IP) and input sample reads was visualized using the IGV genome browser (Robinson et al., 2011) for the whole genome (A), the pf1089/tfb-rf1 (B), and pf1011/1012 (C) promoter regions. Prominent peaks are labeled.

(Figure 5B, lanes 1 to 6), whereas transcription from a control promoter is not affected (Figure 5B, lanes 7 to 12). For the pf1011 promoter we observed sixfold activation in comparison to sevenfold activation on the pf1089 promoter (Ochs et al., 2012). We assume that the weak BREs shift TFB recruitment as the rate limiting step on both promoters and the increased activation in the case of the pf1089 promoter is caused by the additional deviations of the BRE in comparison to the consensus sequence (see also Figure 4C). Taken together, the in vitro experiments clearly demonstrate that the newly identified binding site upstream of pf1011 belongs to the TFB-RF1 regulon and also requires this protein for transcriptional activation.

To correlate our in vitro data with the in vivo situation, we isolated RNA from Pyrococcus pYS5 and pYS3 cells grown on starch or pyruvate and analyzed RNA levels by RT-qPCR. Both transcripts, pf1089/tfb-rf1 as well as pf1011/1012, are highly upregulated in the pYS5 strain after growth on pyruvate (Figure 5C, black bars), whereas in the pYS3 strain these transcripts are almost not affected by switching from starch to pyruvate (Figure 5C, white bars). A direct comparison of the RNA levels between pYS5 and pYS3...
strains revealed a 54-fold activation for the pf1089/tfb-rf1 amplicon and 32-fold for the pf1011/1012 amplicon under pyruvate conditions. In the presence of starch the activation is reduced to 4.5-fold for pf1089/tfb-rf1 and 2.6-fold for pf1011/1012.

DISCUSSION

Research on regulatory networks of putative TFs could be very difficult, especially if nothing is known about the physiological conditions under which the regulator is active. One of such candidates is TFB-RF1 of *P. furiosus*. Although the activation mechanism is studied in detail (Ochs et al., 2012), nothing is known about the physiological role of the protein. To bypass the missing knowledge how to induce expression of the protein, we used a *Pyrococcus/E. coli* shuttle vector for overexpression of TFB-RF1 under the control of a gluconeogenic promoter (Waage et al., 2010). In combination with ChiP-seq we were able to identify the pf1011/pf1012 operon as an additional target of the TFB-RF1 regulon. *In vitro* and *in vivo* data clearly demonstrate that the activation mechanism of TFB-RF1 on the pf1011/pf1012 operon is also based on a weak BRE similar to the pf1089/tfb-rf1 operon. A more detailed comparison of the BRE of pf1011 with the *Pyrococcus* consensus sequence 5′-VRAAA-3′ (van de Werken et al., 2006) revealed two deviations for the pf1011 promoter instead of three for the pf1089 promoter. In this context it is interesting to note that the extent of TFB-RF1 induced activation correlates with the strength of the BRE. In both cases, *in vitro* and *in vivo*, the activation-fold is reduced for the pf1011 promoter. This finding is in line with an *in vitro* mutational analysis of the BRE of the pf1089 promoter, which indicated that a consensus BRE even abolishes TFB-RF1 induced activation (Ochs et al., 2012). Furthermore, a spontaneous mutant in the BRE of the fla operon in *Methanococcus maripaludis* also bypassed the need for the transcriptional activator EarA (Ding et al., 2017). Analysis of the DNA sequence from the fla promoter of this spontaneous mutant revealed a deletion of three adenines which converts the weak BRE in a strong one. Both organisms belong to the Euryarchaeota, but the activation of the arabinose S gene of *Sulfolobus solfataricus* (Peng et al., 2009) demonstrates that this type of activation is also present in Crenarchaeota. Overall, it seems that transcriptional activation by a TFB recruitment mechanism evolved in different archaea as well as for different genes.
To analyze in more detail, if this activation mechanism is conserved for the regulation of these two operons, we searched 35 published genome sequences of the Thermococcales for homologous genes.

We identified the pf1089/tfb-rf1 operon in 28 organisms, whereas in 14 organisms both operons were present (Figure 6A and Supplementary Table 3). An alignment of the upstream sequences of the corresponding genes revealed three conserved elements: a bona-fide TATA box, a rather weak BRE compared to the Thermococcus consensus sequence (Jäger et al., 2014) and further upstream in a conserved distance the 13 nt palindromic consensus sequence $5'$-TCTGAANTTCAGA-3'$ as TFB-RF1 Binding Motif (TRBM; Figure 6B). These data indicate a strong phylogenetic conservation of this gene regulatory network and for all of these promoters TFB-RF1 most likely acts as activator of transcription by supporting the recruitment of TFB to an imperfect BRE.

The identification of the pf1011/pf1012 operon as an additional target of the TFB-RF1 regulon may give new implications for unraveling the physiological function. Sequence analysis revealed that the corresponding protein PF1011 contains two Pfam domains, an ATP-binding domain of ABC transporters (ABC_tran/PF00005) in the N-terminal region and an additional domain of unknown function (DUF4162/PF13732) in the C-terminal region (Finn et al., 2016). The ATPase in the N-terminal region is composed of a typical ABC transporter domain, including the Walker A and B motif, the ABC signature motif and the Q-, D-, and H-loop (Rees et al., 2009). The domain of unknown function can be mainly found at the C-terminus of bacterial or archaeal ABC transporter proteins. PF1012 is the transmembrane protein of the transporter with six predicted transmembrane segments (TMS) and belongs to the ABC2 family (ABC2_membrane_3/PF12698). This is a separate group of ABC transporters where the six-TMS topology resulted from intragenic duplication of a primordial three-TMS-encoding genetic element (Wang et al., 2009).

A more detailed sequence similarity search against the Transporter Classification Database (TCDB)\(^8\) using BLAST revealed that the homologs for both proteins with the lowest E-values belong to the drug exporter-1 family (3.A.1.105; Saier et al., 2016). This family includes eukaryotic as well as prokaryotic drug transporters. The latter are mainly found in gram-positive bacteria with DrrAB and LmrA as one of the first discovered members. DrrAB is from the daunorubicin-producing strain of Streptomyces peucetius and LmrA is a multidrug transporter from Lactococcus lactis (van Veen et al., 1996; Kaur, 1997). The eukaryotic family includes the multidrug resistance P-glycoprotein, which confers resistance to anti-cancer drugs in humans (Gottesman et al., 1996), but this pump can also transport 100s of structurally unrelated hydrophobic compounds.

\(^8\)www.tcdb.org/
amphipathic compounds, including therapeutic drugs, peptides and lipid-like compounds (Sharom, 2011). Due to this broad range of possible substrates, the function of the transporter in \textit{P. furiosus} remains speculative. It is interesting to note that a computational analysis of PF1089 also revealed four TMS in its sequence, which denotes it as integral component of the membrane with unknown function. In summary, both so far identified TFB-RF1-regulated operons point to a transport-related function and provide a promising starting position to decipher the physiological function of the TFB-RF1 gene regulatory network.

**AUTHOR CONTRIBUTIONS**

MK constructed the \textit{Pyrococcus} strains. KR did the RNA work and the western blots and SD the \textit{in vitro} transcription experiments. RR performed the ChIP-seq experiments and the characterization of the DNA protein interactions. WH and RR wrote the manuscript. WH, RR, and MT coordinated and supervised the work. All authors agreed to the final version of the manuscript.

**REFERENCES**

Aravind, L., and Koonin, E. V. (1999). DNA-binding proteins and evolution of transcription regulation in the archaea. \textit{Nucleic Acids Res.} 27, 4658–4670. doi: 10.1093/nar/27.23.4658

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. \textit{Nucleic Acids Res.} 37, W202–W208. doi: 10.1093/nar/gkp335

Bell, S. D. (2005). Archaeal transcriptional regulation–variation on a bacterial theme? \textit{Trends Microbiol.} 13, 262–265. doi: 10.1016/j.tim.2005.03.015

Blankenberg, D., Kuster, G., von Coraor, N., Ananda, G., Lazarus, R., and Mangan, M. (2010). Galaxy: a web-based genome analysis tool for experimentalists. \textit{Curr. Protoc. Mol. Biol.} 19, 19–21. doi: 10.1002/0471142727.mb1910s89

Bleholder, A., Frommherz, R., Teufel, K., and Pfeifer, F. (2012). Expression of multiple \textit{fbp} genes in different \textit{Halobacterium salinarum} strains and interaction of TFB with transcriptional activator GvpE. \textit{Arch. Microbiol.} 194, 269–279. doi: 10.1007/s00203-011-0756-z

Blombach, F., Salvadori, E., Fouqueau, T., Yan, J., Reimann, J., Sheppard, C., et al. (2015). Archaeal TF\(\beta\)/\(\beta\) is a hybrid of TFIIE and the RNA polymerase III subcomplex KRPC6/39. \textit{eLife} 4:e08378. doi: 10.7554/eLife.08378

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2005). TrmB, a sugar sensing regulator of ABC transporter genes in \textit{Pyrococcus kodakarense}. \textit{BMC Genomics} 15:684. doi: 10.1186/1471-2164-15-684

Jager, D., Förstner, K. U., Sharma, C. M., Santangelo, T. J., and Reeve, J. N. (2014). Primary transcriptome map of the hyperthermophilic archaeon \textit{Thermococcus kodakarense}. \textit{Mol. Cell} 43, 263–274. doi: 10.1016/j.molcel.2011.05.030

Dahlke, I., and Thomm, M. (2002). A \textit{Pyrococcus} homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment. \textit{Nucleic Acids Res.} 30, 701–710. doi: 10.1093/nar/30.3.701

Geidusche, E. P., and Ohuammouch, M. (2005). Archaeal transcription and its regulators. \textit{Mol. Microbiol.} 56, 1397–1407. doi: 10.1111/j.1365-2958.2005.04627.x

Gindner, A., Hausner, W., and Thomm, M. (2014). The TrmB family: a versatile group of transcriptional regulators in Archaea. \textit{Extremophiles} 18, 925–936. doi: 10.1007/s00792-014-0677-2

Grohmann, D., Nagy, J., Chakraborty, A., Klose, D., Fielden, D., Ebright, R. H., et al. (2011). The initiation factor TFE and the elongation factor Sptri/5 compete for the RNAP clamp during transcription initiation and elongation. \textit{Mol. Cell} 43, 263–274. doi: 10.1016/j.molcel.2011.05.030

Funding

This work was supported by the Institute of Microbiology and Archaea Center of the University of Regensburg and by the German Research Foundation (DFG) with the funding program Open Access Publishing.

Acknowledgments

The authors would like to thank Renate Richau for technical assistance and Thomas Hader and Konrad Eichinger for helping with the fermentation. They also thank Mario Rost for the initial work on expression of TFB-RF1 and the KFB Regensburg for performing Illumina sequencing and technical support.

Supplementary Material

The Supplementary Material for this article can be found online at: \url{https://www.frontiersin.org/articles/10.3389/fmicb.2018.00838/full#supplementary-material}
global transcriptional regulator of the hyperthermophilic archaeon Pyrococcus furiosus. Mol. Microbiol. 65, 305–318. doi: 10.1111/j.1365-2958.2007.05780.x
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
Lipscomb, G. L., Keese, A. M., Cowart, D. M., Schut, G. J., Thomm, M., Adams, M. W. W., et al. (2009). SurR: a transcriptional activator and repressor controlling hydrogen and elemental sulphur metabolism in Pyrococcus furiosus. Mol. Microbiol. 71, 332–349. doi: 10.1111/j.1365-2958.2008.06525.x
Naji, S., Grünberg, S., and 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, 11047–11057. doi: 10.1074/jbc.M611674200
Ochs, S. M., Thumann, S., Richau, R., Weirauch, M. T., Lowe, T. M., Thomm, M., et al. (2012). Activation of archaeal transcription mediated by recruitment of transcription factor B. J. Biol. Chem. 287, 18863–18871. doi: 10.1074/jbc.M112.365742
Ouhoumouch, M., Langham, G. E., Hauser, W., Simpson, A. J., El-Sayed, N. M. A., and Geiduschek, E. P. (2005). Promoter architecture and response to a positive regulator of archaeal transcription. Mol. Microbiol. 56, 625–637. doi: 10.1111/j.1365-2958.2005.04563.x
Peeters, E., Peixeiro, N., and Sezonov, G. (2013). Cis-regulatory logic in archaeal transcription. Biochem. Soc. Trans. 41, 326–331. doi: 10.1042/BST20120312
Peng, N., Xia, Q., Chen, Z., Liang, Y. X., and She, Q. (2009). An upstream activation element exerting differential transcriptional activation on an archaeal promoter. Mol. Microbiol. 74, 928–939. doi: 10.1111/j.1365-2958.2009.06908.x
Pérez-Rueda, E., and Janga, S. C. (2010). Identification and genomic analysis of transcription factors in archaeal genomes exemplifies their functional architecture and evolutionary origin. Mol. Biol. Evol. 27, 1449–1459. doi: 10.1093/molbev/msq033
Rees, D. C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to move. Nat. Rev. Mol. Cell Biol. 10, 218–227. doi: 10.1038/nrm2646
Reichelt, R., Gindner, A., Thomm, M., and Hausner, W. (2016). Genome-wide transcription analysis reveals two modes of Spt4/5 recruitment to the archaeal transcription initiation complex. J. Biol. Chem. 291, 161–178. doi: 10.1074/jbc.M115002900
Scott, R. A. (2004). Transcription factor B contacts promoter DNA near the TATA-box-binding proteins of Halobacterium salinarum. Arch. Microbiol. 180, 134–149. doi: 10.1007/s00203-009-0537-0
Thomm, M., and Hausern, W. (2007). “Transcriptional mechanisms,” in Archaea: Evolution, Physiology and Molecular Biology, eds A. Roger Garrett and K. Hans-Peter (Hoboken, NJ: Blackwell Publishing Ltd).
Vierke, G., Engelmann, A., Hebbeln, C., and Thomm, M. (2003). A novel archaeal transcriptional regulator of heat shock response. J. Biol. Chem. 278, 18–26. doi: 10.1074/jbc.M209250200
Waegeman, I., Schmid, G., Thumann, S., Thomm, M., and Hausern, W. (2010). Shuttle vector-based transformation system for Pyrococcus furiosus. Appl. Environ. Microbiol. 76, 3308–3313. doi: 10.1128/AEM.01951-09
Wang, B., Bukerevech, M., Sun, E. I., Yen, M. R., and Saier, M. H. J. R. (2009). Membrane porters of ATP-binding cassette transport systems are polyphyletic. J. Membr. Biol. 231, 1–10. doi: 10.1007/s00232-009-9200-6
Weinzierl, R. O. J. (2013). The RNA polymerase factory and archaeal transcription. Chem. Rev. 113, 8350–8376. doi: 10.1021/cr400148k
Wilbanks, E. G., Larsen, D. J., Neches, R. Y., Yao, A. I., Wu, C.-Y., Kjolby, R. A. S., et al. (2012). A workflow for genome-wide mapping of archaeal transcription factors with ChIP-seq. Nucleic Acids Res. 40:e74. doi: 10.1093/nar/gks063

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Reichelt, Raperti, Kreuzer, Desl, Thomm and Hausner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.