

A transcriptional factor B paralog functions as an activator to DNA damage-responsive expression in archaea

Xu Feng^{1,2}, Mengmeng Sun², Wenyuan Han², Yun Xiang Liang¹ and Qunxin She^{1,2,*}

¹State Key Laboratory of Agricultural Microbiology and College of Life Science and Technology, Huazhong Agricultural University, 430070 Wuhan, China and ²Archaea Centre, Department of Biology, University of Copenhagen, Ole Maaløes Vej 5, DK-2200 Copenhagen N, Denmark

Received February 27, 2018; Revised March 17, 2018; Editorial Decision March 19, 2018; Accepted March 20, 2018

ABSTRACT

Previously it was shown that UV irradiation induces a strong upregulation of *tfb3* coding for a paralog of the archaeal transcriptional factor B (TFB) in *Sulfolobus solfataricus*, a crenarchaea. To investigate the function of this gene in DNA damage response (DDR), *tfb3* was inactivated by gene deletion in *Sulfolobus islandicus* and the resulting Δ *tfb3* was more sensitive to DNA damage agents than the original strain. Transcriptome analysis revealed that a large set of genes show TFB3-dependent activation, including genes of the *ups* operon and *ced* system. Furthermore, the TFB3 protein was found to be associated with DDR gene promoters and functional dissection of TFB3 showed that the conserved Zn-ribbon and coiled-coil motif are essential for the activation. Together, the results indicated that TFB3 activates the expression of DDR genes by interaction with other transcriptional factors at the promoter regions of DDR genes to facilitate the formation of transcription initiation complex. Strikingly, TFB3 and *Ced* systems are present in a wide range of crenarchaea, suggesting that the *Ced* system function as a primary DNA damage repair mechanism in Crenarchaeota. Our findings further suggest that TFB3 and the concurrent TFB1 form a TFB3-dependent DNA damage-responsive circuit with their target genes, which is evolutionarily conserved in the major lineage of Archaea.

INTRODUCTION

RNA transcription is the first step of decoding genetic information from DNA, and RNA polymerase (RNAP), a multi-protein complex that is responsible for the process, is evolutionarily conserved in all three domains of life. RNAPs in archaea and eukaryotes have 12 or more subunits, which

are larger than the bacterial counterparts consisting of 5 subunits (1). Studies on transcription initiation show that the two types of RNAP use different mechanisms to initiate gene transcription. In bacteria, the bacterial RNAP and sigma factor form the holoenzyme, in which the sigma subunit recognizes promoters and binds there to form a pre-initiation complex (PIC) (2). RNA transcription in archaea is more related to the process by RNAP II, the enzyme that is responsible for synthesis of mRNA in eukaryotes. The archaeal transcription starts with recognition of a promoter by the TATA-binding protein and transcriptional factor B (TFB), and then, RNAP is recruited to the promoter along with the general transcriptional factor E to form the PIC (3,4).

Current investigation on archaeal transcriptional regulation has revealed that the eukaryotic-like transcriptional machinery is primarily controlled by the promoter-centered mode of regulation; transcription factors specifically bind to DNA motifs present on gene promoter regions and regulate the gene transcription by affecting the PIC formation on the promoters (5,6). Transcriptional factors of both bacterial and eukaryotic types have been identified in archaea and function in gene-specific regulation (7–9). In addition, many archaeal genomes encode multiple TBP and/or TFB (10), and for this reason, archaea have the potential to explore these basal transcriptional factors to exert global regulation in analogy to sigma factors in bacterial transcription.

Indeed, early works on two TFB paralogs (TFB1 and TFB2) of *Thermococcus kodakarensis* reveals that each of them can support transcription *in vitro* without any apparent selectivity on promoter, and neither of them is essential for cell growth (11). Nevertheless, *tfb1* is expressed to a higher level in *T. kodakarensis* and supports better cell growth, relative to *tfb2* (12). Furthermore, characterization of *Pyrococcus furiosus* TFB1 and TFB2 shows that the two factors have different capability to further RNA transcription *in vitro* (13). *Sulfolobus solfataricus* and *Sulfolobus acidocaldarius* encode three paralogs of TFB pro-

*To whom correspondence should be addressed. Tel: +45 532 2013; Fax: +45 3532 2128; Email: qunxin@bio.ku.dk

tein (TFB1, -2 and -3) among which TFB3 is a truncated version of archaeal TFBs (14,15). Investigation of their expression in these archaea shows that only *tfb3* is strongly up-regulated by UV irradiation (16,17). *In vitro* experiments indicate that the *S. solfataricus* TFB3, TBP and TFB1 form a complex in the presence of a promoter DNA fragment, and it was thus suggested that TFB3 can function as a general activator to gene transcription in this archaeon (18). Nevertheless, it remains elusive what genes are to be regulated by TFB3 and whether the truncated version of archaeal TFB could regulate UV-responsive expression in these archaea.

UV lights and other DNA damage agents have been used in investigation of DNA damage responsive regulation of genome expression in Bacteria and Eukarya. These studies have revealed a series of coordinated cellular and molecular in these organisms to prevent accumulation of DNA lesions and to facilitate maintenance of genome integrity, and the revealed network of these cellular events are collectively called the DNA damage response (DDR) (19–22). Here *Sulfolobus islandicus*, a genetic model for which a complete genetic toolbox has been developed (23), was used for investigation of TFB3 function in the archaeal DDR. We found that TFB3 is associated with promoters of DDR genes and it plays an essential role in regulating a number of genes involved in cell aggregation or/and intercellular DNA exchange in this archaeon.

MATERIALS AND METHODS

Cell growth and DNA damage treatment

The *S. islandicus* strains, i.e. E233S1, the wild-type (WT) strain and the $\Delta tfb3$ mutant as well as their plasmid-carrying derivatives (Supplementary Table S1) were grown at 78°C in SCV media (Basal media supplemented with 0.2% sucrose, 0.2% Casamino acids and 1% vitamin solution), and uracil was added to 20 µg/ml if required (24).

For DNA damage treatment with chemicals, exponentially growing cultures of *S. islandicus* strains ($OD_{600} = 0.2$) were supplemented with one of the following drugs, including 4-nitroquinoline-1-oxide (NQO), methyl methanesulfonate (MMS), cisplatin and hydroxyurea (HU) at the concentrations indicated in each experiment. UV irradiation was conducted by placing 30 ml of culture in a petri dish of 9 cm in diameter and irradiated with a setting of energy level at 200 J/m² using the UV Stratalinker 1800 (Stratagene, USA). The treated cultures were incubated for the time periods indicated in each experiment, during which cell samples were taken for OD_{600} determination, cell aggregation analysis, extraction of total RNAs and preparation of cell extracts as described in individual experiments.

Construction of a *tfb3* in-frame deletion mutant of *S. islandicus*

The CRISPR-based genetic manipulation recently developed in our laboratory (25) was employed to construct a $\Delta tfb3$ mutant using *S. islandicus* REY15A (26). The genome-editing plasmid carried the mini-CRISPR array containing a spacer derived from the target site in the *tfb3* gene and the donor DNA of the $\Delta tfb3$ gene allele (Supplementary Figure S1). The target site started with the proto-

spacer adjacent motif (CCN) positioned 138 bp after the start codon of the *tfb3* gene, and the immediately adjacent 40-nt sequence was used as the spacer. Oligos designed for construction of the spacer of an artificial mini-CRISPR plasmid (pAC) and the donor DNA of the genome editing plasmid (pGE) were listed in Supplementary Table S2. The pAC-*tfb3* plasmid was constructed as described (27). Then, the donor DNA was generated by splicing and overlapping extension (SOE)-PCR (28) and inserted into pAC-*tfb3*, yielding pGE-*tfb3* plasmid. The genome-editing plasmid was then introduced into the competent cell of E233S1 by electroporation, giving transformants on selective plates, which should carry the designed $\Delta tfb3$ allele. The identity of deletion mutants was revealed by polymerase chain reaction (PCR) amplification of the $\Delta tfb3$ allele and verified by DNA sequencing. Plasmids were cured from deletion mutants by *pyrEF* counter selection with uracil and 5-FOA, yielding $\Delta tfb3$ for further experiment.

Construction of plasmids for expression of mutated *tfb3* genes

The WT *tfb3* gene and the mutant derivative coding for the N-terminal 49 amino acids (Zn ribbon) were directly amplified from the genomic DNA of *S. islandicus* REY15A (26) by PCR using primer set of *tfb3*pro-f/*tfb3*WT-r and *tfb3*pro-f/*tfb3*N-Zn-r, respectively. The resulting DNA fragments contained the native promoter region and the specific coding sequence of WT TFB3 and TFB3N-Zn. The remaining three mutants carried point mutations in *tfb3*, which were obtained by SOE-PCR, following the reported procedure (28). Specifically, *tfb3M1* carrying substitutions in the Coil 1 motif (R145A, K146A) was generated with the primer set of *tfb3*WT-f/*tfb3*CoilM1-SOE-r and *tfb3*CoilM1-SOE-f/*tfb3*WT-r whereas *tfb3M2* carrying the CoilM2 mutation (L148A, K149A, L151A) was amplified with *tfb3*WT-f/*tfb3*CoilM2-SOE-r and *tfb3*CoilM2-SOE-f/*tfb3*WT-r. The *tfb3C3C25* mutant harboring substitution for C3S C25T was generated with the primer set of C3SC25T-f/C3SC25T-SOE-r and C3SC25T-SOE-f/*tfb3*WT-r. All primers for SOE-PCR are listed in Supplementary Table S2. The resulting DNA fragments were cleaved with NdeI and SalI, together with the native *tfb3* promoter DNA amplified with *tfb3*pro-f/*tfb3*pro-r (SphI/NdeI), were cloned to pSeSD1 (29), the *Sulfolobus* expression vector at SphI/SalI sites in a ligation of three DNA fragments, yielding expression plasmids for each mutant *tfb3* gene (pSe^{ptfb3}-*tfb3*N-Zn, pSe^{ptfb3}-*tfb3*M1 and pSe^{ptfb3}-*tfb3*M2, pSe^{ptfb3}-*tfb3*C3S-C25T (Supplementary Table S1)). All mutations on the expression plasmids were verified by DNA sequencing. Importantly, all *tfb3* mutant genes still contained the native promoter of *tfb3*.

Each expression plasmid was introduced into the *S. islandicus* $\Delta tfb3$ mutant by electroporation, yielding strains expressing either N-Zn, or CoilM1, or CoilM2, or C3S-C25T or the WT TFB3. The resulting strains were then used for further experiments.

Western blot analysis

Cell sample of 15 ml culture was pelleted by centrifugation and re-suspended in 150 µl TBST buffer (50 mM Tris-HCL, 100mM NaCl, 0.1% Tween-20, pH7.6). Sonication

of the cell suspension gave total cell extracts, which were fractionated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Protein bands on the gel were transferred onto a nitrocellulose membrane using a Trans-Blot Semi-Dry Transfer Cell (Bio-Rad). For immunoblotting, the membrane was immersed in 5% skim milk blocking agent, then incubated with individual primary antibodies and finally with corresponding second antibodies. TFB3 antisera were raised against *S. solfataricus* TFB3 (18) (kindly provided by Prof. Malcolm F. White) whereas PCNA3 antisera against *S. islandicus* PCNA3 were reported previously. Secondary antibodies were purchased from Thermo Fisher Scientific, and hybridization signals were detected using the ECL western blot substrate (Thermo Fisher Scientific), and visualized by exposure of the membrane to an X-ray film (Agfa HealthCare, Belgium).

Transcriptome analysis

Exponentially growing cultures of E233S1 and $\Delta tfb3$ strains ($OD_{600} = 0.2$) were grown for 6 h in the presence or absence of NQO (6 hpt). Cells in each culture were harvested by centrifugation and used for RNA extraction using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA) by following the instruction of the manufacturer. The quality of total RNA preparations was controlled using NanoDrop 1000 and 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). RNA preparations of high quality were used for construction of RNA-Seq libraries and next generation sequencing in BGI (Shenzhen, China) using Illumina HiSeq™ 4000.

A total of ca.12 millions of high-quality sequence reads were mapped to the reference genome of *S. islandicus* Rey15A, giving genome coverage of 320- to 400-fold for different samples. The gene expression level was calculated by using RPKM method (reads per kb per million reads) (30) and the differentially expressed genes (DEGs) was identified by comparing treated and corresponding untreated samples. The FDR (false discovery rate) was applied to determine the threshold of p value in multiple tests. $FDR \leq 0.001$ and the absolute value of \log_2 ratio ≥ 1 were used to quantify the significance of gene expression difference.

To validate the RNA seq results, seven genes (*upsX*, *upsA*, *cedB*, *SiRe_1957*, *dpo2*, *cdvB1* (*SiRe_1550*) and *SiRe_0187*), including those that were either upregulated or downregulated by NQO were selected for quantitative reverse transcriptase PCR (qRT-PCR) analysis. Correlation values between the result of RNA seq and qPCR was 0.97 for the slope and 0.93 for R-value (Supplementary Figure S2), indicating the two datasets are strongly correlated.

Real-time quantitative PCR analysis

Before cDNA synthesis, residual DNA in each total mRNA preparation (5 μ g) was removed by DNase I treatment. The first strand cDNA synthesis was conducted using Revert Aid First Strand cDNA Synthesis Kit (Thermo Scientific), and the resulting cDNA was diluted and then used as the templates for quantitative PCR analysis. qPCR reaction was conducted with a CFX96 real-time PCR

system (Bio-Rad, Hercules, CA, USA), using Maxima SYBR Green/ROX qPCR Master Mix (Thermo Scientific). Primers employed in qPCR were listed in Supplementary Table S2. Relative amounts of RNAs were calculated using the comparative Ct method by using 16srRNA as the reference gene (31).

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) experiment was conducted as previously described with minor modifications (32). Cell samples for ChIP analysis were first cooled to room temperature, and then, formaldehyde was added to each cooled culture to the final concentration of 1% and incubated at RT for 10 min for crosslink. Glycine was added to 125 mM for another 5 min to quench the cross-link reaction. Cells were then collected by centrifugation, and the resulting cell pellet was washed with PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 2 mM KH₂PO₄), and re-suspended again in 3 ml TBSTT buffer (20 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20, 0.1% Triton X-100, pH 7.5). Cells in the suspension were disrupted by French press, and DNA fragments in the cell extracts were sheared by sonication to yield a size range of 200–1000 bp of genomic DNAs. Immunoprecipitations were performed by incubating 5 μ l antibody or corresponding serum with 100 μ g of cell extract in 3 ml TBSTT at 4°C for overnight, with gentle rotation. Then, 50 μ l DynaBeads of protein G (Thermo Fisher Scientific) was added and incubated for 2 h. Immune complexes were captured by the magnet and washed first with the TBSTT buffer, then with TBSTT containing 500 mM NaCl, and finally with TBSTT containing 0.5% Tween 20 and Triton X-100, followed by wash with TBSTT for three times. The immune complexes was eluted by resuspending the beads in 100 μ l elution buffer (20 mM Tris (pH 7.8), 10 mM ethylenediaminetetraacetic acid and 1% SDS) and heating up to 65°C for 30 min. Eluted samples were first treated with 10 μ g/ml proteinase K for 6 h at 65°C and 10 h at 37°C to remove proteins cross-linked to DNA, and then incubated with RNase A for 2 h at 37°C to remove RNAs in the samples. Finally, the treated samples were extracted with phenol/chloroform/isoamyl alcohol (25:24:1) and chloroform, and DNAs in the samples were precipitated with cold ethanol in the presence of 20 μ g glycogen. The DNA preparations were then used for qPCR analysis.

Cell aggregation analysis

Exponentially growing cultures of *S. islandicus* strains ($OD_{600} = 0.2$) were subjected to different DNA damage treatment. Samples with UV irradiation was taken 3 h post-treatment, while for the treatment with other DNA damaging agents, the samples was taken at 10 hpt. Sampling time in NQO treatment experiments was delayed, relative to that in the UV experiments, because NQO is not toxic; its toxicity to the cell has to be generated by cellular metabolic activities (33). As a result, it takes longer time to reach the maximal level of cell aggregation. Fresh samples were transferred to a glass slide, covered with a coverslip and directly observed under a Nikon Eclipse Ti-E inverted microscope (Nikon, Kobe, Japan) and pictures were taken from a camera connected to the microscope. Quantification data were

obtained from at least 12 fields of view images and 500 single cells for each cell sample and from three independent experiments.

Cell viability analysis

Exponentially growing cultures of *S. islandicus* ($OD_{600} = 0.2$) were treated with NQO with a final concentration of 0–2 μM and incubated for 24 h. Then, 1 ml of culture was taken, and cells were pelleted by centrifugation and resuspended in 1 ml fresh SCV medium. The resulting cell suspensions were serially diluted and plated using the two-layer plating method (24). A total of 100 μl of diluted sample were plated onto gelrite plates in triplicate. Colonies appeared on plates after 7 days of incubation were counted, giving colony formation units (CFUs) per ml culture.

RESULTS

Expression of *tfb3* is highly induced in *Sulfolobus islandicus* upon NQO treatment

To investigate if the expression of *tfb3* in *S. islandicus* could be activated by the treatment of NQO, a chemical that yields bulky adducts on bases of DNA (33,34), the WT *S. islandicus* strain was grown for 9 h in NQO-containing media (hours post-treatment, hpt) during which cell samples were taken at a 3 h interval. Each cell sample was then divided into two portions: one was used for extraction of total RNAs while the other was for preparation of cell extracts for immunoblotting analysis. Cellular levels of *tfb3* mRNAs in the samples were estimated by real-time quantitative reverse transcription PCR (qRT-PCR) as described in ‘Materials and Methods’ section. We found that, while the expression level of *tfb3* in the untreated reference remained almost constant, the expression of the gene was elevated for >40-fold at all three tested time points (Figure 1A). Western analysis of the *S. islandicus* TFB3 using antibodies raised against the *S. solfataricus* TFB3 (a gift from Prof. M. F. White) revealed that a protein band of ca. 20 kDa was specifically recognized, which was present in a large amount in NQO-treated samples but was barely detectable in the untreated samples (Figure 1B).

To estimate the TFB3 protein level after induction, cell extracts were diluted for 8- and 16-fold and the diluted cell extracts were then used for SDS-PAGE and for immunoblotting analysis. Since signals of TFB3 hybridization obtained from 16-fold diluted samples were equivalent to, or stronger than, that in the undiluted cell sample of the reference culture. The NQO-treated cells were estimated to contain >16-fold of TFB3 protein, relative to the untreated cells (Supplementary Figure S3). Therefore, we concluded that *tfb3* gene expression is strongly upregulated in *S. islandicus* cells upon NQO treatment.

TFB3 deficiency by gene deletion reduced cell viability of *S. islandicus* upon NQO treatment

To investigate the function of TFB3, an in-frame deletion mutant of *tfb3* gene was constructed for *S. islandicus*, using the CRISPR-assisted gene deletion procedure (Supplementary Figure S1). The WT and mutant alleles of *tfb3* are

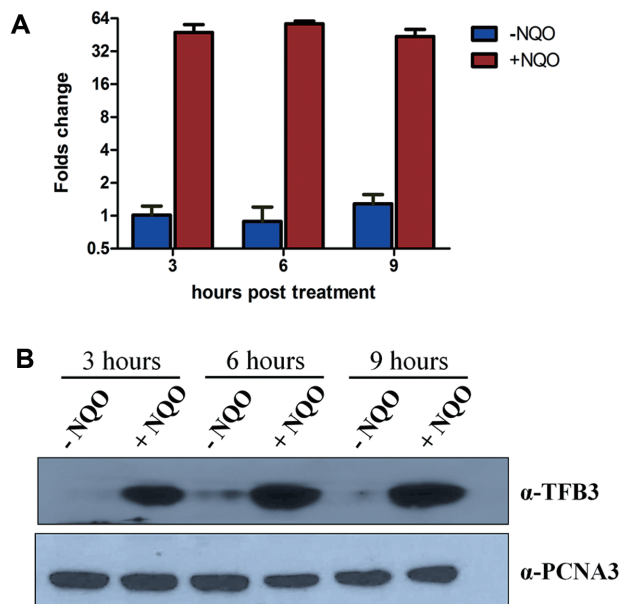


Figure 1. NQO-induced upregulation of *tfb3* in *Sulfolobus islandicus*. (A) qRT-PCR analysis of *tfb3* mRNA in NQO-treated and untreated *S. islandicus* cells. Data were normalized to the level of 16S rRNA. Error bars represent the standard deviation values of three independent replicates. (B) Western blot analysis of TFB3 protein in the cell extracts of NQO-treated and untreated *S. islandicus* using PCNA as a reference. *S. islandicus* E233S1 (WT) strain was grown in the absence (–NQO) or presence of the 2 μM NQO (+NQO) for 9 h (hours post-treatment, hpt) during which cell samples were taken at the indicated time points and used for mRNA extraction (for qPCR) and cell extracts preparation (western analysis).

shown in Figure 2A, and the absence of the *tfb3* gene in the mutant was confirmed by PCR with the primer sets and genomic DNAs of the two strains (Figure 2B). Furthermore, immunoblotting analysis of TFB3 protein in NQO-treated cell extracts using antibodies raised against *S. solfataricus* TFB3 revealed that TFB3 was present in the cell extracts of the WT strain in a large quantity but absent from $\Delta tfb3$ (Figure 2C), consistent with their genotypes.

The mutant was employed for investigation of its resistance to NQO as above described for WT. To do that, the WT strain and $\Delta tfb3$ were grown in the media containing 0, 0.5, 0.75, 1, 1.5 or 2 μM of NQO for 24 h. Investigation of TFB3 expression in cell samples taken from the above cultures by immunoblotting revealed that the gene was strongly activated at all tested concentrations of NQO (Supplementary Figure S4). Growth data of these cultures showed that $\Delta tfb3$ did not exhibit any obvious growth defect either in the absence or presence of NQO, relative to the corresponding WT cultures (Supplementary Figure S5). Nevertheless, determination of their cell viability by plating for CFUs revealed that the survival rate of $\Delta tfb3$ was always lower than that of the WT strain in all tested drug concentrations and the largest difference occurred for the treatment with 0.75–1.0 μM NQO (30–37%, Figure 3). Therefore, $\Delta tfb3$ exhibited a higher sensitivity to the DNA damage agent than the WT strain.

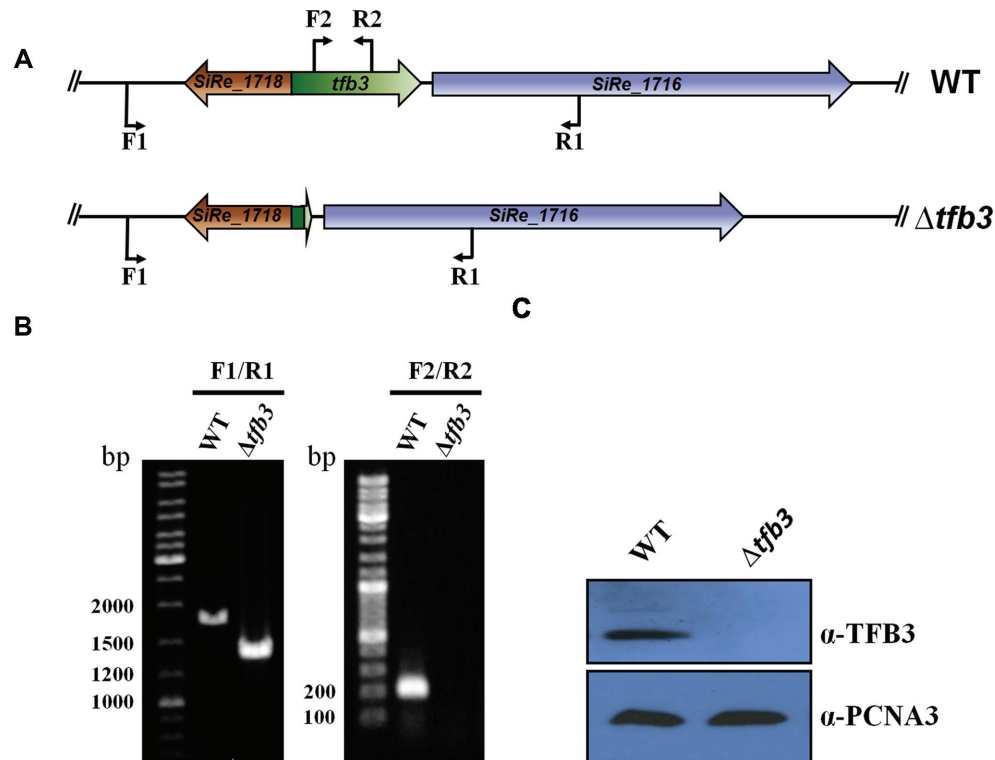


Figure 2. Construction of a $\Delta tfb3$ in-frame deletion mutant of *S. islandicus*. (A) Schematic of PCR identification of the WT *tfb3* gene and $\Delta tfb3$ alleles in *S. islandicus* strains. (B) Verification of the $\Delta tfb3$ genotype by PCR amplification of the *tfb3* gene alleles (PCR). Primer set of F1/R1 was designed to check the gene deletion whereas primer set of F2/R2 was designed to amplify an internal DNA fragment in the *tfb3* gene. (C) Western blot analysis of TFB3 protein in NQO-treated cultures of *S. islandicus*. Immunoblotting was conducted with the TFB3 antibody raised against the *S. solfataricus* TFB3 protein (α -TFB3). WT: *S. islandicus* E233S1, the genetic host; $\Delta tfb3$: *tfb3* deletion mutant constructed with E233S1.

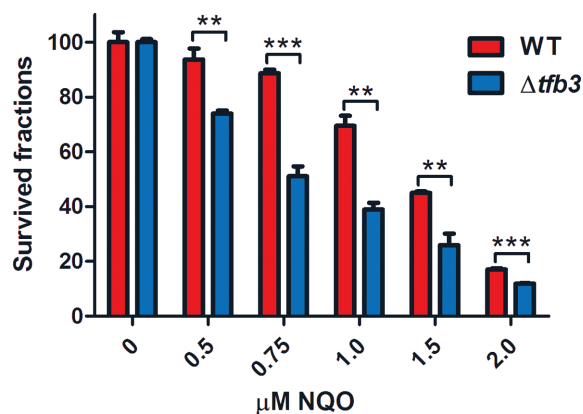


Figure 3. Survival rates of the wild-type *S. islandicus* and $\Delta tfb3$ upon NQO treatment. Cell viability was estimated for *S. islandicus* E233S1 (WT) strain and $\Delta tfb3$ mutant at 24 h post-NQO addition by determination of CFUs. CFU of WT-0 μ M and $\Delta tfb3$ -0 μ M was assigned to 100% for each strain, with which the survived fractions of cells in other cultures were calculated individually.

TFB3 as the activator of a subset of NQO-responsive genes in *S. islandicus*

Then, we investigated how TFB3 deficiency could influence the genome expression in this archaeon by transcriptome analysis using RNA sequencing. Both the deletion mutant and the WT strain were grown in presence or absence of

NQO for 6 h. Cell mass was collected from each culture from which the total mRNAs were extracted, giving RNA preparations for conducting RNA-seq analysis.

First, the transcriptome analysis of the WT strain led to the identification of 313 DEGs upon NQO treatment (>2-fold change), including 139 upregulated genes and 174 downregulated genes (Supplementary Table S3). The same analysis with total RNAs prepared from corresponding $\Delta tfb3$ culture revealed that 78 upregulated genes and 8 downregulated ones were TFB3 dependent (Supplementary Table S3).

Among the 61 genes that showed NQO-upregulation of >4-fold, 51 genes showed TFB3-dependent activation (Supplementary Tables S3 and 4). The most upregulated ones were those in the *ups* operon and all known *ced* genes (*cedA1*, *cedA* and *cedB*). Other regulated genes included those that code for potential DNA transfer related processes, transporter-related or membrane-associated proteins, all of which can be implicated in DNA damage repair. Together, these results indicated that TFB3 regulates a subset of DDR-activated DNA repair genes. A few genes showed TFB3-dependent downregulation and they all code for a hypothetical protein, and the negative regulation could be indirect (see 'Discussion' section). Therefore, we concluded that TFB3 functions as an activator to the transcription of a large subset of NQO-responsive genes in *S. islandicus*.

TFB3 deficiency resulted in the loss of the DNA damage-induced cell aggregation

Since the activation of the expression of *ups* and *ced* genes was both NQO-responsive and TFB3-dependent, this suggested that $\Delta tfb3$ could have lost the capability of cell aggregation. To test that, untreated and NQO-treated cultures of $\Delta tfb3$ and E233S1, the WT reference were grown in SCV or in SCV+NQO for 15 h during which cell samples were taken for microscopic examination. As shown in Figure 4, while no cell aggregates were observed for the WT strain, nor for $\Delta tfb3$ grown in SCV media, >50% of WT cells formed cell aggregates at 12 hpt when grown in SCV+NQO media and the portion of the cells in cell aggregates increased to 82% at 15 hpt. In contrast, no cell aggregates were observed in the corresponding $\Delta tfb3$ culture (Figure 4). These results are consistent with the lack of activation of the expression from the *ups* operon in the mutant as revealed from the RNA seq analysis.

To date, both UV irradiation and NQO treatment are capable of inducing the formation of cell aggregates in different *Sulfolobus* species. Since they both are DNA damage agents, this raised a question if other DNA damage agents could also trigger cell aggregation in this archaeon. To test that, three additional drugs were studied for their effects on *S. islandicus*, including cisplatin that makes cross-link lesions on dsDNA (35), MMS that produces base-alkylating lesion on DNA (36) and HU that induces cell cycle arrests and replication stresses (37). *Sulfolobus islandicus* WT strain and $\Delta tfb3$ were grown in the presence or absence of each drug and grown for 6 h (or 3 h post-UV irradiation). Cell samples were taken and analyzed for the expression of TFB3 protein by immunoblotting analysis, and this revealed that the protein was present in a large amount in the cells treated with cisplatin, or MMS or UV irradiation, as for NQO treatment; however, HU treatment did not influence the *tfb3* expression (Supplementary Figure S6A).

Investigation of cell aggregation by microscopy showed that cell aggregates were observed in *S. islandicus* cultures treated with all these DNA damage agents but absent from the culture of HU treatment as well as from $\Delta tfb3$ cultures treated with each of the tested drugs (Supplementary Figure S6B), consistent with the upregulation of *tfb3* gene expression observed for the former category of drug, but not for the latter one. Together these results indicated that DNA damage treatment triggers the expression of TFB3, which in turn functions as a DNA damage-inducible regulator to activate the expression of a large number of genes involved in DNA damage repair in *S. islandicus*.

TFB3 was specifically associated with the promoters of DDR genes

To yield an insight in to the mode of the TFB3 regulation, we analyzed the association of the transcription factor with the promoters of NQO-responsive genes by ChIP-qPCR. The tested promoters include those of *upsE*, *herA1*, *cedB*, *cdvB1*, with the *cmr- β* promoter (promoter of the subtype III-B CRISPR-Cas *cmr- β* gene operon) and an intragenic region of 16S rRNA as references. Both untreated and NQO-treated samples were analyzed. DNA fragments enriched by ChIP with the TFB3 antisera were quantified by

qPCR, and this revealed that, after NQO treatment, the antisera specifically enriched DNA fragments containing the promoters of *upsE*, *herA1* and *cedB* only in NQO-treated samples and the enrichment was for 17-, 25- and 37-fold, respectively (Figure 5). Essentially no enrichment was observed from the ChIP-qPCR experiments with the same promoters in the untreated samples, and furthermore, the same analysis for the promoters of *cdvB1*, a gene exhibiting TFB3-independent downregulation, and two references, i.e. the *cmr- β* promoter and an internal fragment of 16S rRNA gene, with and without NQO treatment, did not yield any difference between the untreated and NQO-treated samples (Figure 5). These results suggested that TFB3 could form complexes with other transcription factors on the promoters of NQO-responsive genes upon DNA damage treatment and activate gene expression probably by recruiting RNAP to the promoters to form PICs.

Functional characterization of TFB3 conserved domains

Next, we examined what part of TFB3 could be involved in facilitating PIC formation. Unlike its TFB1 and TFB2 paralogs, TFB3 lacks the helix-turn-helix DNA-binding domain present in the C-termini of common TFB proteins. Instead, *Sulfolobus* TFB3 proteins contain a peptide that is predicted to form a coiled-coil (CC) domain (38) (Figure 6A and Supplementary Figure S7B). Nevertheless, the TFB3 protein contains the four cysteine residues that form the Zn ribbon domain in archaeal TFB proteins, whereas the *Sulfolobus* TFB1 protein only has two conserved cysteine residues; those at the first and the last positions were changed as for the *S. solfataricus* TFB3 protein (Figure 6A and Supplementary Figure S7A).

Here, we investigated whether the Zn ribbon and CC domains of *S. islandicus* TFB3 could be essential for activation of gene expression in this archaeon by construction of four *tfb3* mutants and test the function of each mutant protein. These included: (i) N-Zn, a truncated form of TFB3 mutant only containing the N-terminal 49 amino acids of the Zn domain, (ii) CoilM1 and CoilM2—two CC mutants that carry R145A, K146A substitutions and L148A, K149A, L151A substitutions at the conserved CC domains, respectively, and (iii) C3S-C25T, two substitutions in the conserved cysteine residues of the TFB3 Zn ribbon domain, which change the TFB3 Zn ribbon motif to that present in the non-canonical TFB1 proteins (Figure 6A and Supplementary Figure S7). The original promoter of *tfb3* was used for controlling the expression such that their expression was under the same regulation as in the WT strain. After introduction of each TFB3 expression plasmid into $\Delta tfb3$, the expression of the wild-type TFB3 and the four mutant proteins was examined in each transformant by immunoblotting analysis with the TFB3 antibodies. This revealed all these TFB3 proteins were expressed and their expression was subjected to DNA damage activation (Figure 6B). Then, the expression levels of *upsX* and *cedB* upon DNA damage in $\Delta tfb3$ strains carrying a WT *tfb3* gene or one of its mutated derivatives were estimated by qRT-PCR. As shown in Figure 6C, while the plasmid-borne WT *tfb3* gene completely restored the function of the chromosomal *tfb3* gene, none of the tested TFB3 mutant proteins, includ-

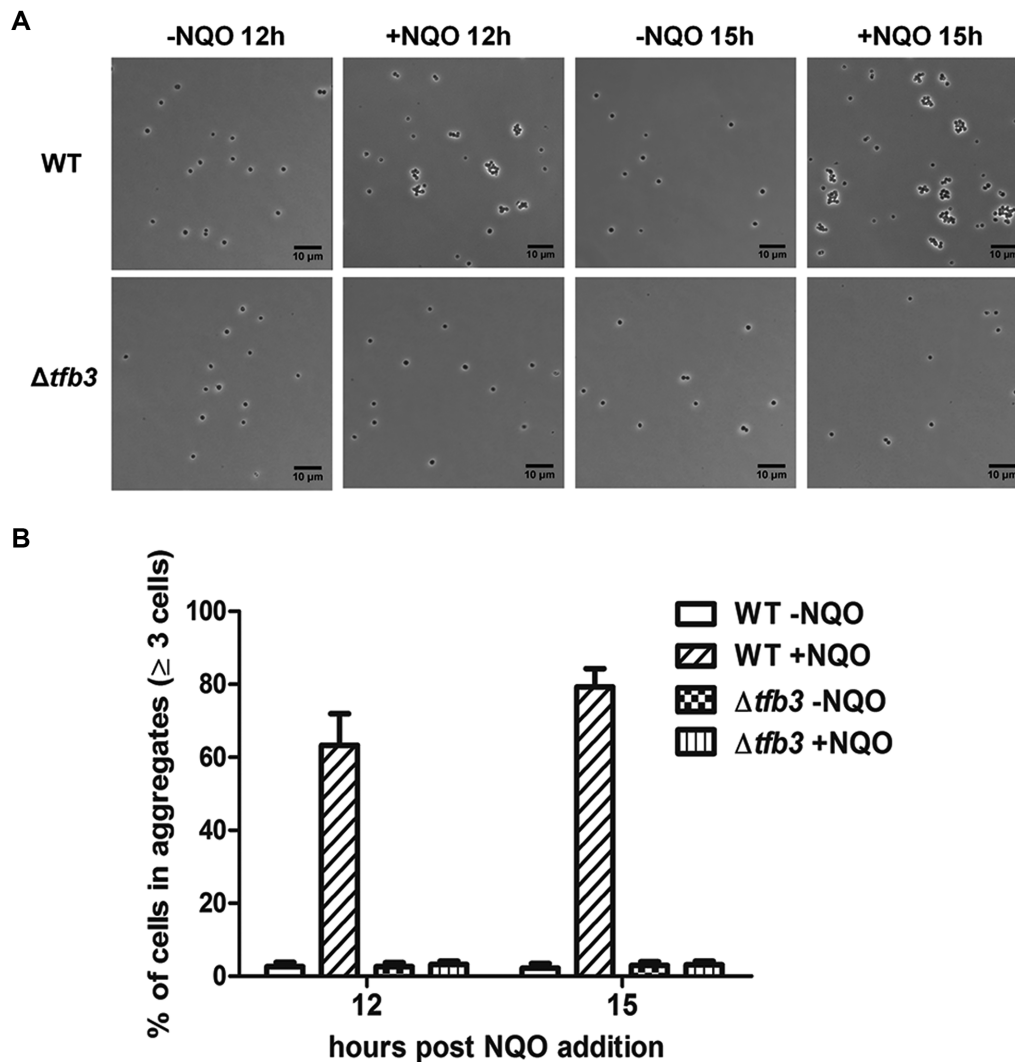


Figure 4. *S. islandicus* $\Delta tfb3$ mutant lost the capability of forming NQO-induced cell aggregation. (A) Microscopy of cells and cell aggregates in samples taken from the cultures of the WT strain (E233S1) and $\Delta tfb3$ mutant. Each strain was grown in the absence (–NQO) or presence (+NQO) of NQO for 24 h, and aliquots of cultures were placed on glass slides, covered with coverslips and directly observed under a phase-contrast microscope. 12 and 15 h: hours after NQO supplementation. (B) Quantification data of cell aggregation in the cell samples shown in panel A. At least 500 cells were analyzed for each cell sample.

ing N-Zn, CoilM1, CoilM2 and C3S-C25T, were capable of activating the expression of these DDR genes. These results indicated that the conserved cysteines in the Zn ribbon and the CC motif are essential for TFB3 function in DDR regulation in this archaeon.

Co-evolution of TFB3 and Ced systems in crenarchaea

To yield an insight into the conservation of TFB3 and its target systems of regulation, we searched for the presence of *tfb3*, the *ups* operon and *ced* genes in archaea using the BLAST search (39). We found that, while the *ups* system was limited to organisms of Sulfolobales as reported previously (40), the truncated version of TFB3 was found in organisms of 16 genera in Crenarchaeota whereas the Ced sys-

tem appeared in 15 genera, and strikingly, 13 of the crenarchaeal genera have both a truncated version of TFB3 protein and the Ced system. All these data were illustrated in Figure 7, using iTOL (Interactive Tree Of Life) (41). Multiple sequence alignment of all identified truncated TFB3 proteins showed that they fell into three distinct classes in which all known TFB3 homologs formed a clade (Supplementary Figure S8). In addition, there also appeared a concurrence of TFB3 and non-canonical TFB1 proteins, the unconventional TFB proteins carrying two cysteine residues in the Zn ribbon domain in Crenarchaeota (18) (Supplementary Figure S7A). Together, these data suggested a coevolution of TFB1, TFB3 and the Ced system and the DDR regulatory circuit of the intercellular DNA transfer in Crenarchaeota, one of the major phyla in the Archaeal domain.

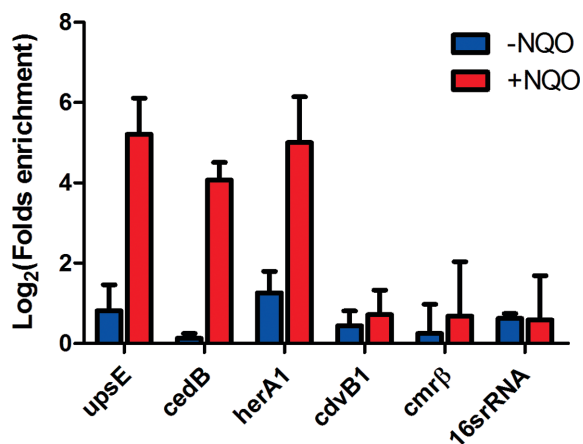


Figure 5. Enrichment of TFB3 association with promoters of upregulated DDR genes. *S. islandicus* E233S1 (WT) was grown in SCV in the presence (+NQO, 2 μ M) or absence of the drug (-NQO) for 3 h. Cell mass was then collected from cultures of E233S1 from which cell extracts were prepared and used for CHIP analysis using the TFB3 antisera. The folds enrichment refers to the CHIP signal from the TFB3 antibody relative to the signal from the sheep serum. The promoter regions of three TFB3-dependent upregulated genes (*upsE*, *cedB* and *herA1* (SiRe.1715)), one TFB3-independent downregulated gene, *cdvB1*, (SiRe.1550 encoding ESCRTIII-1) and *cmrB* operon, together with intragenic region of 16S rRNA gene, were selected for CHIP analysis with the TFB3 antibody. Folds of enrichment were calculated with the data from three independent replicates with the error bars indicating the standard derivation values.

DISCUSSION

DDR in bacteria and eukarya consists of a series of cellular and molecular events, including inhibition of DNA replication, cell cycle arrest and activation of certain DNA damage repair pathways such as nucleotide excision repair, translesion DNA repair and homologous recombination repair (HRR) (42). DDR is initiated by the recognition of a DNA damage signal, usually ssDNA and dsDNA breaks, resulting in the differential expression of a large set of genes that are responsible for the observed DDR events. Nevertheless, bacteria and eukaryotes have evolved distinct mechanisms to control their DDR regulation; while bacteria employ global transcriptional factors to regulate the expression (20,43), eukaryotes have evolved kinase-signaling pathways to control the process (21,22,44,45). In this article, we show that, upon DNA damage, the expression of *tfb3* gene is strongly activated in *S. islandicus*, a crenarchaeal model organism and the expressed protein forms complexes with transcriptional factors on DDR gene promoters and functions as a DNA damage-responsive activator to regulate RNA transcription in this archaeon.

The most interesting group of TFB3-responsive genes identified from the DNA damage study in *S. islandicus* include those in the *ups* operon implicated for UV-responsive cell aggregation of *Sulfolobus* and *ced* genes coding for the intercellular DNA transfer system. Genes of the *ups* operon have been investigated in *S. solfataricus* and *S. acidocaldarius* (40,46) whereas *ced* genes responsible for intercellular DNA transfer have recently been identified in *S. acidocaldarius* and functionally characterized (47). These studies have led to the proposal that the UV-responsive pilus formation (Ups system) and the Ced DNA transfer system (Ced

system) probably function in concert to facilitate DNA exchange for HRR in these archaea (47). Our investigation on DNA damage repair in *S. islandicus* demonstrates that both the Ups and Ced systems are subjected to the regulation by TFB3, a truncated version of the archaeal TFB, further supporting that the two systems cooperate in DNA exchange for HRR.

Important insights have been gained into the activation of DDR genes in this archaeon. First, the *ups* operon in *S. islandicus* is clustered with an NQO-repressed *SiRe.1883* that codes for a putative transcriptional factor (26). However, this gene is absent from the genomes of several *Sulfolobus* species including *S. acidocaldarius* (15), suggesting it may not have a function in the regulation of *ups* operon. Furthermore, the *ced* genes does not appear to be located close to any genes encoding a putative transcriptional factor. Since DNA transfer required both the Ups pili and the Ced proteins, it is more meaningful to have both systems regulated by a common factor and the identification of TFB3 as their common regulator is consistent with this reasoning.

Indeed, several other genes exhibiting TFB3-dependent regulation have also been implicated in DNA transfer-related processes in *S. islandicus*. The encoded proteins include: (i) a PadR family transcriptional regulator (SiRe.1956) and its adjacent gene SiRe.1957; the latter codes for a membrane protein with signal peptide for type IV pili, with the N-terminal part showing similarity to flagellar hook-basal body protein (48), suggestive of a function in Ups pilus assembly or protein export, (ii) an Archaease (SiRe.1319) that is a putative DNA chaperon (Supplementary Table S4), and the encoding gene is clustered with the *cedA* operon (located immediately downstream), implying that the protein may interact with the Ced system in the intercellular DNA transfer, (iii) several putative transporter and membrane proteins that may function in formation of cell aggregation or related process. The conservation of these proteins in Sulfolobales (Supplementary Table S4) suggests that the TFB3-dependent regulation is conserved in these organisms and they all probably play a role in facilitating cell aggregation or/and intercellular DNA transfer upon DNA damage.

Transcriptome data have revealed that TFB3 does not only activate gene expression but it also represses gene expression (Supplementary Table S3), which appeared to be inconsistent with the activator hypothesis for TFB3. The only two genes i.e. SiRe.0266 (membrane protein), SiRe.0267 (sodium:solute symporter) that showed TFB3-dependent repression of >10-fold were organized into an operon and cluster together with a third gene, SiRe.0269 that codes for a small protein of 66 amino acids (26). Interestingly, SiRe.0269 expression is strongly upregulated by TFB3 (45-fold, Supplementary Table S4). Although SiRe.0269 does not show a detectable similarity to any known protein in the current GenBank database in BLAST search, our RNA-seq data suggested that the gene could code for a novel transcription repressor to downregulate the expression of SiRe.0266, SiRe.0267 and thus representing an example of indirect regulation by TFB3 in *S. islandicus*. To this end, albeit as an activator, TFB3 can either mediate activation or repression to DDR genes. To date, TFB3 is

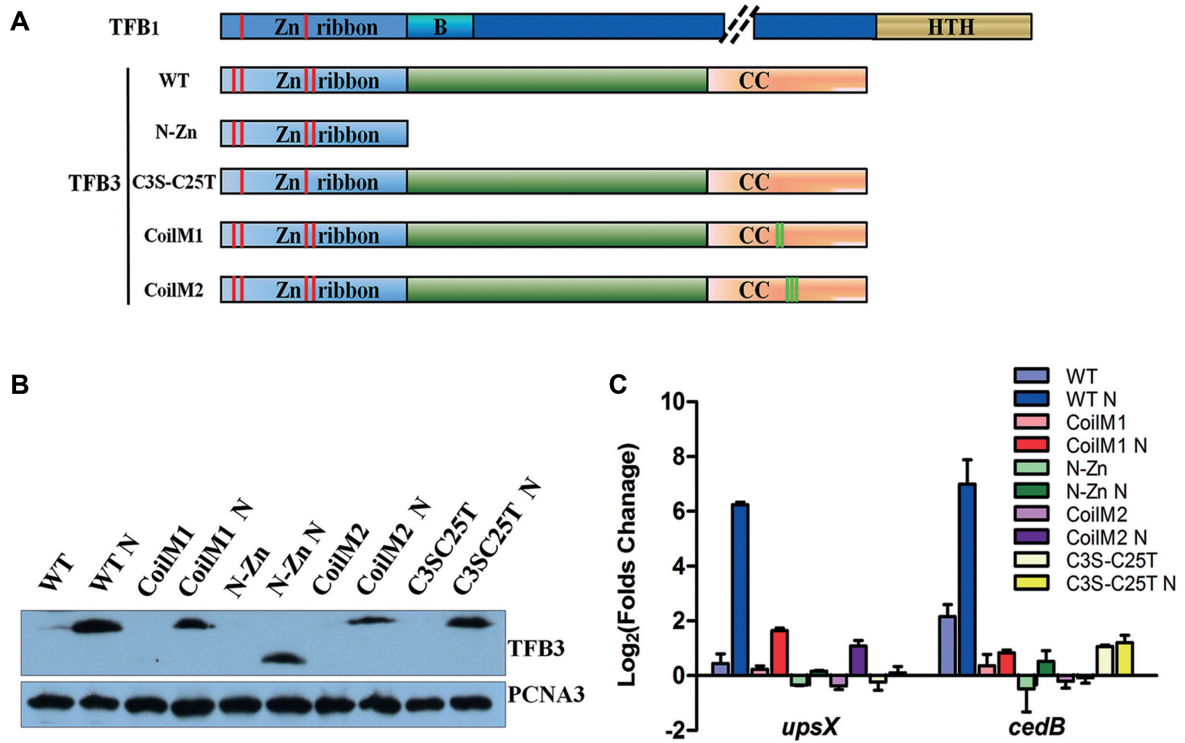


Figure 6. Functional characterization of TFB3 by mutagenesis. (A) Schematic of the WT TFB3 and its mutant derivatives. A CC region was predicted by Coils server (38). CoilM1 and CoilM2 refer to the mutation of R145AK146A and L148AK149AL151A. The first and fourth conserved cysteine in Zn ribbon of TFB3 was replaced with the corresponding ones in TFB1 (SiRe_1555). (B) Western blot analysis of the total cell extracts of the strains carrying different mutated TFB3 after NQO treatment. Sample was taken 3 h after 2 μM NQO treatment. (C) Quantitative analysis of the expression levels of *upsX* and *cedB* in the strains encoding different isoform of TFB3. Sample was taken 3 h after 2 μM NQO treatment.

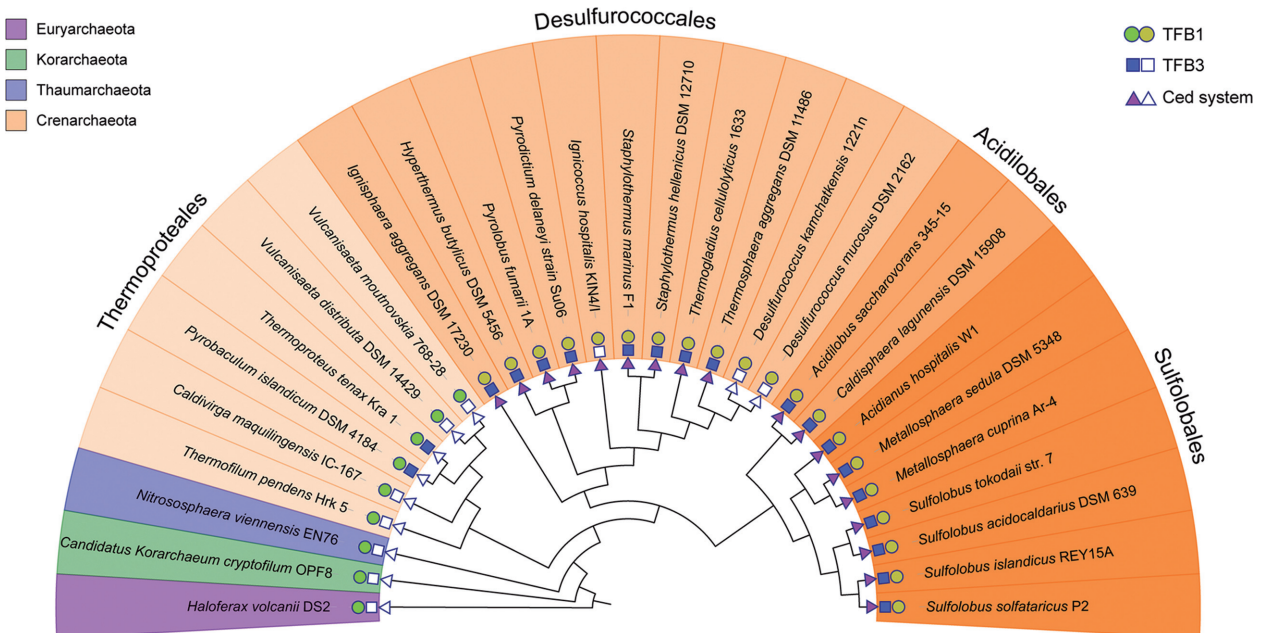


Figure 7. Co-evolution of TFB3 and Ced system in Archaea DNA sequences of 16S rRNA genes of different crenarchaeal species were retrieved from the GenBank database and used for multiple sequence alignment using Cluster X and for construction of phylogenetic tree. The resulting tree was visualized and annotated using iTOL (Interactive Tree Of Life) (41). The presence or absence of a truncated version of TFB (TFB3) is indicated by the filled or empty rectangles, respectively; the presence or absence of the Ced system is shown as the filled or empty triangles individually, whereas the canonical and noncanonical TFB1 proteins are notified as the green and dark yellow circles, respectively.

the only known transcriptional factor that regulates DDR in archaea.

Further insights into the regulation by TFB3 have been gained from ChIP-qPCR analysis of the association of TFB3 with DDR gene promoters and functional analysis of TFB3 mutant proteins. In the latter, mutagenesis of two conserved domains in TFB3 proteins, i.e. the conserved cysteine residues in the Zn ribbon and R145, K146, L148, K149, L151 residues in the C terminal CC motif has revealed that the both domains are essential for the function of the DDR-specific activator. These results are in consistent with those obtained from the study of the Zn ribbon domain of eukaryotic TFIIB functions where the Zn ribbon is essential to the recruitment of RNA Pol II into the PIC (49,50), and that a CC motif in the C terminal of a TFB homolog is essential for the transcriptional activation in mitochondria of *Dictyostelium discoideum* (51). For the former, specific enrichment of TFB3 on the upregulated DDR gene promoter regions suggests that the interaction between TFB3 and DDR gene promoters *in vivo* is specific. Nevertheless, TFB3 does not contain any recognizable DNA-binding motif, suggesting the protein does not interact with DNA sequence specifically. To this end, the specific activation of DDR genes by TFB3 is probably mediated by protein–protein interaction between TFB3 and another transcriptional factor.

Furthermore, it was shown that the *S. solfataricus* TFB3 co-exist with the *S. solfataricus* TFB1, a non-canonical TFB1 and furthermore, the two factors also interact *in vitro* and the interaction enhances RNA transcription activity by an archaeal RNAP (18). The concurrence of TFB3 and the non-canonical TFB1 proteins has now extended to all known orders of organisms in Crenarchaeota except Thermoproteales (Figure 7). These data are consistent with the distribution of the Ced system of DNA transfer (47). Together, this suggests co-evolution of the TFB1 and TFB3 factors with the Ced system in Crenarchaeota. It would be interesting to investigate the mechanisms of activation of DDR genes by the two factors in these crenarchaea.

In addition, several DDR genes exhibit TFB3-independent regulation. First, genes coding for enzymes involved in homologous recombination belong to a class of genes that is commonly regulated by DDR in bacteria and eukaryotes (21,22,45,46). The archaeal proteins include RadA, Mre11, Rad50 that are homologous to the eukaryotic counterparts (52,53) and NurA and HerA, two unique enzymes of the archaeal system and their genes form an operon (54,55). As expected, *radA* and all four genes in the *nurA* operon are upregulated upon NQO treatment (Supplementary Figure S5). Other genes that also show TFB3-independent upregulation include *dpo2*, coding for an enzyme that can bypass DNA lesion on the templates in DNA synthesis in *S. solfataricus* (56), and *cdc6-2* that was shown to inhibit DNA replication in *S. solfataricus* (32) but is dispensable for replication initiation in *S. islandicus* (57). Furthermore, a large group of genes that are downregulated by DNA damage treatment both in the WT strain and in $\Delta tfb3$. These include cell division genes (*cdvA*, *cdvB2*, *cdvB3* and *vps4*) and DNA replication initiator gene (*cdc6-3*) was moderately affected in *tfb3* mutant (Supplementary Table S5). These data suggest that DDR events

like cell-cycle arrest and inhibition of DNA replication and upregulation of homologous recombination are regulated by other transcription factor(s) and the regulation occurs in the TFB3-independent regulation. Together, these findings suggest there are multiple regulatory pathways in the DDR regulation in this archaeon. In conclusion, our work demonstrates that archaea also possess a network of DDR regulation to mediate cell cycle arrest, inhibition of DNA replication and activation of DNA repair pathways and the DDR process in crenarchaea involves novel transcriptional regulators and new DNA repair pathways.

DATA AVAILABILITY

The RNA-seq data have been deposited in the public database with GEO accession: GSE111187.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Prof. Malcolm White for kindly providing us the antibody of TFB3 and our colleagues in the Archaea Centre, University of Copenhagen and those in the Wuhan lab, Huazhong Agricultural University, China for helpful discussions.

FUNDING

National Science Foundation of China [31771380]; Danish Council for Independent Research [DFR-4181-00274]. Funding for open access charge: National Science Foundation of China [31771380]; China Scholarship Council PhD Studentship (to X.F., M.S.).

Conflict of interest statement. None declared.

REFERENCES

1. Werner, F. and Grohmann, D. (2011) Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.*, **9**, 85–98.
2. Browning, D.F. and Busby, S.J. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.*, **14**, 638–650.
3. Gietl, A., Holzmeister, P., Blombach, F., Schulz, S., von Voithenberg, L.V., Lamb, D.C., Werner, F., Tinnefeld, P. and Grohmann, D. (2014) Eukaryotic and archaeal TBP and TFB/TF(II)B follow different promoter DNA bending pathways. *Nucleic Acids Res.*, **42**, 6219–6231.
4. Nagy, J., Grohmann, D., Cheung, A.C., Schulz, S., Smollett, K., Werner, F. and Michaelis, J. (2015) Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nat. Commun.*, **6**, 6161.
5. Gehring, A.M., Walker, J.E. and Santangelo, T.J. (2016) Transcription regulation in archaea. *J. Bacteriol.*, **198**, 1906–1917.
6. Sheppard, C. and Werner, F. (2017) Structure and mechanisms of viral transcription factors in archaea. *Extremophiles*, **21**, 829–838.
7. Jun, S.H., Reichlen, M.J., Tajiri, M. and Murakami, K.S. (2011) Archaeal RNA polymerase and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.*, **46**, 27–40.
8. Contursi, P., Fusco, S., Limauro, D. and Fiorentino, G. (2013) Host and viral transcriptional regulators in *Sulfolobus*: an overview. *Extremophiles*, **17**, 881–895.
9. Gindner, A., Hausner, W. and Thomm, M. (2014) The TrmB family: a versatile group of transcriptional regulators in Archaea. *Extremophiles*, **18**, 925–936.

10. Martinez-Pastor, M., Tonner, P.D., Darnell, C.L. and Schmid, A.K. (2017) Transcriptional regulation in archaea: from individual genes to global regulatory networks. *Annu. Rev. Genet.*, **51**, 143–170.
11. Santangelo, T.J., Cubonova, L., James, C.L. and Reeve, J.N. (2007) TFB1 or TFB2 is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription in vitro. *J. Mol. Biol.*, **367**, 344–357.
12. Hidese, R., Nishikawa, R., Gao, L., Katano, M., Imai, T., Kato, S., Kanai, T., Atomi, H., Imanaka, T. and Fujiwara, S. (2014) Different roles of two transcription factor B proteins in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *Extremophiles*, **18**, 573–588.
13. Micorescu, M., Grünberg, S., Franke, A., Cramer, P., Thomm, M. and Bartlett, M. (2008) Archaeal transcription: function of an alternative transcription factor B from *Pyrococcus furiosus*. *J. Bacteriol.*, **190**, 157–167.
14. She, Q., Singh, R.K., Confalonieri, F., Zivanovic, Y., Allard, G., Awayez, M.J., Chan-Weiher, C.C., Clausen, I.G., Curtis, B.A., De Moors, A. *et al.* (2001) The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 7835–7840.
15. Chen, L., Brugger, K., Skovgaard, M., Redder, P., She, Q., Torarinnsson, E., Greve, B., Awayez, M., Zibat, A., Klenk, H.P. *et al.* (2005) The genome of *Sulfolobus acidocaldarius*, a model organism of the Crenarchaeota. *J. Bacteriol.*, **187**, 4992–4999.
16. Gotz, D., Paytubi, S., Munro, S., Lundgren, M., Bernander, R. and White, M.F. (2007) Responses of hyperthermophilic crenarchaea to UV irradiation. *Genome Biol.*, **8**, R220.
17. Frols, S., Gordon, P.M., Panlilio, M.A., Duggin, I.G., Bell, S.D., Sensen, C.W. and Schleper, C. (2007) Response of the hyperthermophilic archaeon *Sulfolobus solfataricus* to UV damage. *J. Bacteriol.*, **189**, 8708–8718.
18. Paytubi, S. and White, M.F. (2009) The crenarchaeal DNA damage-inducible transcription factor B paralogue TFB3 is a general activator of transcription. *Mol. Microbiol.*, **72**, 1487–1499.
19. Erill, I., Campoy, S. and Barbe, J. (2007) Aeons of distress: an evolutionary perspective on the bacterial SOS response. *FEMS Microbiol. Rev.*, **31**, 637–656.
20. Kreuzer, K.N. (2013) DNA damage responses in prokaryotes: regulating gene expression, modulating growth patterns, and manipulating replication forks. *Cold Spring Harb. Perspect. Biol.*, **5**, a012674.
21. Ciccia, A. and Elledge, S.J. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell*, **40**, 179–204.
22. Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.
23. Peng, N., Han, W., Li, Y., Liang, Y. and She, Q. (2017) Genetic technologies for extremely thermophilic microorganisms of *Sulfolobus*, the only genetically tractable genus of crenarchaea. *Sci. China Life Sci.*, **60**, 370–385.
24. Deng, L., Zhu, H., Chen, Z., Liang, Y.X. and She, Q. (2009) Unmarked gene deletion and host-vector system for the hyperthermophilic crenarchaeon *Sulfolobus islandicus*. *Extremophiles*, **13**, 735–746.
25. Li, Y., Pan, S., Zhang, Y., Ren, M., Feng, M., Peng, N., Chen, L., Liang, Y.X. and She, Q. (2016) Harnessing Type I and Type III CRISPR-Cas systems for genome editing. *Nucleic Acids Res.*, **44**, e34.
26. Guo, L., Brugger, K., Liu, C., Shah, S.A., Zheng, H., Zhu, Y., Wang, S., Lillestol, R.K., Chen, L., Frank, J. *et al.* (2011) Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. *J. Bacteriol.*, **193**, 1672–1680.
27. Peng, W., Feng, M., Feng, X., Liang, Y.X. and She, Q. (2015) An archaeal CRISPR type III-B system exhibiting distinctive RNA targeting features and mediating dual RNA and DNA interference. *Nucleic Acids Res.*, **43**, 406–417.
28. Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K. and Pease, L.R. (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, **77**, 61–68.
29. Peng, N., Deng, L., Mei, Y., Jiang, D., Hu, Y., Awayez, M., Liang, Y. and She, Q. (2012) A synthetic arabinose-inducible promoter confers high levels of recombinant protein expression in hyperthermophilic archaeon *Sulfolobus islandicus*. *Appl. Environ. Microbiol.*, **78**, 5630–5637.
30. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods*, **5**, 621–628.
31. Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative C(T) method. *Nat. Protoc.*, **3**, 1101–1108.
32. Robinson, N.P., Dionne, I., Lundgren, M., Marsh, V.L., Bernander, R. and Bell, S.D. (2004) Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell*, **116**, 25–38.
33. Bailleul, B., Daubersies, P., Galiegue-Zouitina, S. and Loucheux-Lefebvre, M.H. (1989) Molecular basis of 4-nitroquinoline 1-oxide carcinogenesis. *Jpn. J. Cancer Res.*, **80**, 691–697.
34. Han, W., Xu, Y., Feng, X., Liang, Y.X., Huang, L., Shen, Y. and She, Q. (2017) NQO-Induced DNA-Less cell formation is associated with chromatin protein degradation and dependent on A0A1-ATPase in *Sulfolobus*. *Front. Microbiol.*, **8**, 1480.
35. Wong, J.H., Brown, J.A., Suo, Z., Blum, P., Nohmi, T. and Ling, H. (2010) Structural insight into dynamic bypass of the major cisplatin-DNA adduct by Y-family polymerase Dpo4. *EMBO J.*, **29**, 2059–2069.
36. Valenti, A., Napoli, A., Ferrara, M.C., Nadal, M., Rossi, M. and Ciaramella, M. (2006) Selective degradation of reverse gyrase and DNA fragmentation induced by alkylating agent in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.*, **34**, 2098–2108.
37. Liew, L.P., Lim, Z.Y., Cohen, M., Kong, Z., Marjavaara, L., Chabes, A. and Bell, S.D. (2016) Hydroxyurea-mediated cytotoxicity without inhibition of ribonucleotide reductase. *Cell Rep.*, **17**, 1657–1670.
38. Lupas, A., Van Dyke, M. and Stock, J. (1991) Predicting coiled coils from protein sequences. *Science*, **252**, 1162–1164.
39. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
40. Ajon, M., Frols, S., van Wolferen, M., Stoecker, K., Teichmann, D., Driessen, A.J., Grogan, D.W., Albers, S.V. and Schleper, C. (2011) UV-inducible DNA exchange in hyperthermophilic archaea mediated by type IV pili. *Mol. Microbiol.*, **82**, 807–817.
41. Letunic, I. and Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.*, **44**, W242–W245.
42. Giglia-Mari, G., Zotter, A. and Vermeulen, W. (2011) DNA damage response. *Cold Spring Harb. Perspect. Biol.*, **3**, a000745.
43. Baharoglu, Z. and Mazel, D. (2014) SOS, the formidable strategy of bacteria against aggressions. *FEMS Microbiol. Rev.*, **38**, 1126–1145.
44. Sirbu, B.M. and Cortez, D. (2013) DNA damage response: three levels of DNA repair regulation. *Cold Spring Harb. Perspect. Biol.*, **5**, a012724.
45. Blackford, A.N. and Jackson, S.P. (2017) ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. *Mol. Cell*, **66**, 801–817.
46. Frols, S., Ajon, M., Wagner, M., Teichmann, D., Zolghadr, B., Folea, M., Boekema, E.J., Driessen, A.J., Schleper, C. and Albers, S.V. (2008) UV-inducible cellular aggregation of the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by pili formation. *Mol. Microbiol.*, **70**, 938–952.
47. van Wolferen, M., Wagner, A., van der Does, C. and Albers, S.V. (2016) The archaeal Ced system imports DNA. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, 2496–2501.
48. Liu, R. and Ochman, H. (2007) Stepwise formation of the bacterial flagellar system. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 7116–7121.
49. Chen, H.-T. and Hahn, S. (2003) Binding of TFIIB to RNA polymerase II: mapping the binding site for the TFIIB zinc ribbon domain within the preinitiation complex. *Mol. Cell*, **12**, 437–447.
50. Bushnell, D.A., Westover, K.D., Davis, R.E. and Kornberg, R.D. (2004) Structural basis of transcription: an RNA polymerase II-TFIIB cocystal at 4.5 Ångstroms. *Science*, **303**, 983–988.
51. Manna, S., Le, P. and Barth, C. (2013) A unique mitochondrial transcription factor B protein in *Dictyostelium discoideum*. *PLoS One*, **8**, e70614.
52. Seitz, E.M., Brockman, J.P., Sandler, S.J., Clark, A.J. and Kowalczykowski, S.C. (1998) RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev.*, **12**, 1248–1253.
53. Aravind, L., Walker, D.R. and Koonin, E.V. (1999) Conserved domains in DNA repair proteins and evolution of repair systems. *Nucleic Acids Res.*, **27**, 1223–1242.
54. Constantinesco, F., Forterre, P. and Elie, C. (2002) NurA, a novel 5′-3′ nuclease gene linked to rad50 and mre11 homologs of thermophilic Archaea. *EMBO Rep.*, **3**, 537–542.

55. Constantinesco, F., Forterre, P., Koonin, E.V., Aravind, L. and Elie, C. (2004) A bipolar DNA helicase gene, *herA*, clusters with *rad50*, *mre11* and *nurA* genes in thermophilic archaea. *Nucleic Acids Res.*, **32**, 1439–1447.
56. Choi, J.-Y., Eoff, R.L., Pence, M.G., Wang, J., Martin, M.V., Kim, E.-J., Folkmann, L.M. and Guengerich, F.P. (2011) Roles of the four DNA polymerases of the crenarchaeon *Sulfolobus solfataricus* and accessory proteins in DNA replication. *J. Biol. Chem.*, **286**, 31180–31193.
57. Samson, R.Y., Xu, Y., Gadelha, C., Stone, T.A., Faqiri, J.N., Li, D., Qin, N., Pu, F., Liang, Y.X., She, Q. *et al.* (2013) Specificity and function of archaeal DNA replication initiator proteins. *Cell Rep.*, **3**, 485–496.