A novel archaeal regulatory protein, Sta1, activates transcription from viral promoters

Alexandra Kessler\textsuperscript{1,4}, Guennadi Sezonov\textsuperscript{1}, J. Iñaki Guijarro\textsuperscript{2}, Nicole Desnoues\textsuperscript{1}, Thierry Rose\textsuperscript{3}, Muriel Delepierre\textsuperscript{2}, Stephen D. Bell\textsuperscript{5} and David Prangishvili\textsuperscript{1,*}

\textsuperscript{1}Unité de Biologie Moléculaire du Gène chez les Extrémophiles, \textsuperscript{2}Unité de RMN des Biomolécules (CNRS URA 2185), \textsuperscript{3}Unité d’Immunogénétique Cellulaire, Institut Pasteur, 25-28 rue du Dr Roux, 75724 Paris Cedex 15, France, \textsuperscript{4}Department of Microbiology, University of Regensburg, Universitätsstrasse, 31, D-93053 Regensburg, Germany and \textsuperscript{5}MRC Cancer Cell Unit Hutchison/MRC Research Centre, Hills Road, Cambridge CB2 2XZ, UK

Received March 17, 2006; Revised June 30, 2006; Accepted July 3, 2006

ABSTRACT

While studying gene expression of the rudivirus SIRV1 in cells of its host, the hyperthermophilic crenarchaeon Sulfolobus, a novel archaeal transcriptional regulator was isolated. The 14 kDa protein, termed Sulfolobus transcription activator 1, Sta1, is encoded on the host chromosome. Its activating effect on transcription initiation from viral promoters was demonstrated in \textit{in vitro} transcription experiments using a reconstituted host system containing the RNA polymerase, TATA-binding protein (TBP) and transcription factor B (TFB). Most pronounced activation was observed at low concentrations of either of the two transcription factors, TBP or TFB. Sta1 was able to bind viral promoters independently of any component of the host pre-initiation complex. Two binding sites were revealed by footprinting, one located in the core promoter region and the second ~30 bp upstream of it. Comparative modeling, NMR and circular dichroism of Sta1 indicated that the protein contained a winged helix–turn–helix motif, most probably involved in DNA binding. This strategy of the archaeal virus to co-opt a host cell regulator to promote transcription of its genes resembles eukaryal virus–host relationships.

INTRODUCTION

The mechanisms and regulation of gene expression in the Archaea have been studied during the past 25 years [reviewed in (1)]. However, our knowledge on them remains modest in comparison to what is known on transcription in the other two domains of life, the Eukarya and Bacteria. Initial studies revealed that the archaeal basal transcription machinery resembles the core components of the eukaryal RNA polymerase (RNA Pol) II apparatus (2–7). Through the establishment of \textit{in vitro} transcription systems for some archaea (8–13), it became possible to identify the archaeal factors necessary for specific initiation of transcription. Consisting of only the TATA-binding protein (TBP), transcription factor B (TFB), homologous to the eukaryotic TFIIB, and the RNA polymerase, a multi-subunit enzyme, the minimal archaeal transcription pre-initiation complex appears to be a simplified version of the eukaryotic RNA Pol II system. With the ongoing genome sequencing projects many transcription regulators could be identified in archaeal genomes. Surprisingly, many of them were homologs to the members of the bacterial Lrp-like regulator family (14,15). How regulation of an eukaryotic-like system could occur using bacterial-like regulators remains an intriguing question, mainly from an evolutionary point of view. Some of these regulators have been studied in cell-free transcription systems. Except the transcription activators Pfr2 from Methanocaldococcus jannaschii (16), and the homologous Lrp protein Mth from Methanothermococcus thermolithothrophicus (17), these were exclusively repressors: MDR1-repressor of the ABC-transporter-gene from Archaeoglobus fulgidus (18), LrpA from Pyrococcus furiosus (19,20), the negatively autoregulated factor Lrs14 from Sulfolobus solfataricus (21,22), and Prh involved in the heat-shock response of \textit{P. furiosus} (23). However, the physiological functions of most of these regulators are still unclear. It would appear that a majority of \textit{trans}- and \textit{cis}-acting regulatory transcription factors of the Archaea still remain unknown.

In a situation in which efficient genetic tools are not yet available, one possibility to study transcription regulation in hyperthermophilic archaea is offered by diverse crenarchaeal virus–host systems. Although studies on transcription of the Sulfolobus virus SSV1, crucial for the identification of archaeal promoter sequences, were carried out about two decades ago, detailed analysis of transcription of viruses of hyperthermophilic crenarchaea over the replication cycle was performed only recently. \textit{In vivo} transcription studies on the rudiviruses SIRV1 and SIRV2 infecting...
the hyperthermophilic crenarchaeon *Sulfolobus islandicus* demonstrated a rather simple and barely chronological pattern of transcription, with a few cases of temporal regulation (24). SIRV promoters, similar to the host promoters, contain a TATA-box and a TFB responsive element. However, most of them contain an additional virus-specific consensus element. These observations suggested a major role for the host transcription machinery in the transcription of viral genomes, as well as possible involvement of virus-specific transcription factors.

Here, we report on the isolation and characterization of a host-encoded transcription regulator Sta1 involved in the activation of transcription from promoters of the crenarchaeal virus SIRV1.

**MATERIALS AND METHODS**

**Biotinylation of promoter DNAs**

Biotinylated promoter DNA used for the magnetic DNA affinity purification experiments were generated by PCR using biotinylated primers. The promoter regions 56, 134 and 399 were amplified from genomic DNA using the following primer sets: GACTCTGTTCTTGAAGTGGCA and Biotin-ATTTAATGTCCCAAAGTCTATTAGCG for 56, GAAATTTGTGGGCAACAGGAGC and Biotin-AGCAATGACAAATTTAATAGGT for 134, and TACTGTTAAGAAAATGAGATAAC and Biotin-CTTCAATTTCTTTAACAAATATA for 399. Biotinylated T6 promoter was reamplified from the T6 promoter plasmid described previously (13,18) using the primer set TCTGATC-CAACCGTTTGGAGACTC and Biotin-TTAAACTCAGCTTCTAG.

**Magnetic DNA affinity purification of Sta1**

The magnetic affinity purification was carried out as described previously (25) with modifications. For preparation of the affinity beads, 4 mg of streptavidin-coated magnetic beads (Dynabeads; Dynal Biotech) were resuspended in 2 × B & W buffer (10 mM Tris–HCl, pH 7.5, 2 M NaCl and 1 mM EDTA) to a final concentration of 5 mg/ml. The beads were incubated with crude extracts prepared from *S. islandicus* cells infected and non-infected cells (see above). The samples were treated for 1 min with 0.1 U of RNase A, 0.2 with SIRV1/VIII with a multiplicity of infection of 1 and incubated to an OD600 of 0.2.

**Preparation of crude extracts of Sulfolobus cells**

*S. islandicus* REN2H1 cells were grown as described previously (26) to an OD600 of 0.4, pelleted and resuspended in TBS buffer (10 mM Tris–HCl, pH 8.0 and 150 mM NaCl). Cells were lysed through sonication and the soluble protein fraction was collected after centrifugation in a SORVALL SS34 rotor at 170 000 r.p.m. for 20 min at 4°C. For the preparation of crude extracts of virus-infected cells, a growing *S. islandicus* REN2H1 culture was infected at an OD600 of 0.2 with SIRV1/VIII with a multiplicity of infection of 1 and incubated to an OD600 of 0.4.

**Purification of RNA polymerase, TBP and TFB**

RNA polymerase, TBP and TFB were purified as described previously (13,27). Transcription assays and DNase I footprinting were performed using RNA polymerase, TBP and TFB as described previously (18,21).

**Transcription assays and DNase I footprinting**

For transcription assays, PCR products of promoters 56, 134 and 399 were generated from genomic SIRV1 DNA using standard conditions and oligonucleotides. 5547S1F: 5′-GACTCTGTTCTTGAAGTGGCA-3′ and 5679S1R: 5′-TTCAATTTCTTTAACAAATATA for promoter 56; 10964S1R: 5′-AGCAGGAGC-3′ and 11276S1R: 5′-GAAATTTGTGGGCAACAGGAGC-3′ for promoter 134; and 5034S1F: 5′-TCTGTTAAGAAAATGAGATAAC-3′ and 5367S1R: 5′-TTCAATTTCTTTAACAAATATA-3′ for promoter 399. The PCR products were cloned directly into pDrive (Qiagen) by T/A cloning. A plasmid carrying T6 promoter was generated as described previously (13). In vitro transcription reactions were performed using 100 ng of the corresponding plasmid DNA, 0.2 mM NTPs, 10 μg *Sulfolobus* whole-cell extract or 20 ng of TBP and TFB (or as indicated in figure legends), 1 μg RNA polymerase and Sta1 in amounts indicated in the figure legends. The reactions were carried out for 20 min at 70°C in 50 μl transcription buffer (50 mM Tris–HCl, pH 8.0, 75 mM KCl, 25 mM MgCl2 and 1 mM DTT). Reactions were stopped by adding 250 μl NEW buffer (10 mM Tris–HCl, pH 8.0, 750 mM NaCl, 10 mM EDTA, 0.5% SDS and 40 μg/ml glycogen). The in vitro synthesized RNA was isolated by phenol–chloroform extraction followed by ethanol precipitation. Transcription products were detected by primer extension using radiolabeled T7 primer in the case of the T6 promoter template or sequence-specific primers for viral promoter templates as described previously (24).

DNase I footprinting was performed using a 300 bp fragment generated by PCR using the radiolabeled oligonucleotides 10964S1F and 11276S1R (see above). The DNA template was incubated with Sta1 as indicated in the legend to Figure 5 in 50 μl transcription buffer for 10 min at 48°C. Samples were treated for 1 min with 0.1 U of DNase I (Roche). Reactions were stopped by adding 250 μl NEW buffer. DNA fragments were isolated by phenol–chloroform extraction followed by ethanol precipitation. Pellets were resuspended in 20 μl TE buffer. Twenty microliters of 50% formamide loading dye were added and 20 μl of
the denatured samples were analyzed on an 8% denaturing polyacrylamide gel.

**Mass spectrometry**

Protein identification was performed by the Mass spectrometry facility of the MRC Laboratory of Molecular Biology. After SDS–PAGE, protein was in-gel digested with trypsin. Peptides analyzed were recovered on a Voyager-DE STR Biospectrometry Workstation (PerSeptive Biosystems). Peak analysis and database interrogation were performed using the Mascot software package.

**Heterologous expression of SSO0048 and purification of the recombinant protein**

The gene SSO0048 was amplified by PCR from *S. solfataricus* P2 genomic DNA using primers 5′-GGATTCCTCATGTCGAAACCCHATTAA-3′) and 3′-SSO0048 (5′-GGATCCTCGAGTTACAATGGGCTTG-AATTCCT-3′). The PCR product was digested with NdeI and XhoI and ligated to NdeI–XhoI digested pET30a. The sequence of the cloned DNA fragment was shown to be identical to the original *Sulfolobus* sequence. The expression construct was transformed into Rosetta (DE3)pLysS cells. Overexpression of non-tagged Sta1 protein was induced during logarithmic growth of cells by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG) to 1 mM for 4 h. The cell pellet was resuspended in N100 buffer (50 mM Tris–HCl, pH 8.0, 100 mM NaCl and 10 mM β-mercaptoethanol) lysed by sonication and clarified by centrifugation. Sta1 was purified to apparent homogeneity from the crude cell lysate after the removal of heat-denatured cellular proteins by chromatography on a Heparin–Sepharose column. Sta1 was eluted using a linear NaCl gradient. Peak fractions were verified by SDS–PAGE and Coomassie blue staining.

Alternatively, for NMR experiments on Sta1, the protein was produced with a C-terminal hexahistidine-tag (Sta1-h6). This construct contained an 8-residue tail (LEHHHHHH) and a modification of the wild-type protein at position 127 (M instead of K). The gene of Sta1-h6 was cloned in a pET30a vector, and the protein was expressed and purified like recombinant Sta1, with only the Heparin–Sepharose step being replaced by an affinity chromatography using an Ni-NTA column. As assessed by biochemical assays and circular dichroism (CD) in the far-UV region, the tag does not influence the structure or the activity of Sta1.

**Circular dichroism**

CD in the far-UV region was performed on an Aviv 215 spectropolarimeter (Aviv Biomedical Inc., Lakewood, NJ). The concentration of Sta1 prepared in 10 mM sodium acetate, pH 5.5 (buffer NA), ranged between 20 and 100 μM. It was determined from the molar extinction coefficient of the protein calculated as described previously (28) CD spectra were recorded at 20°C between 180 and 260 nm with a step of 0.5 nm, a bandwidth of 1 nm and an optical path of 0.02 cm. The integration time was 4 or 1 s for points between 180 and 200 nm or 200.5 and 260 nm, respectively. Three scans were averaged. The CD spectrum of the protein was deconvoluted in terms of secondary structure content using the CONTIN (29) algorithm implemented in CDPro (30).

**Analytical centrifugation**

Sedimentation/diffusion equilibrium experiments were run at 20°C on an XL-I or XL-A ultracentrifuge (Beckman Coulter Inc.) of the Plateforme de Biophysique (Institut Pasteur). The ultracentrifuges were equipped with an AN-60 ti four hole rotor. Homodimerization of Sta1 was analyzed using rates between 14 and 22 kr.p.m. with samples of Sta1 (4.2, 8.4 or 42 μM) prepared in buffer NA supplemented with 150 mM NaCl.

Binding of Sta1 to a 30 bp DNA oligonucleotide, called Reg2, was followed with spinning rates of 12–20 kr.p.m. using samples obtained in a buffer containing 20 mM Tris–HCl, 150 mM NaCl and 1 mM Na EDTA (pH 8.0). HPLC-purified single-stranded oligonucleotides (5′-AATT-TATTAATTAAAAAGATATTTATA-3′ and its complementary strand) were purchased from Proligo (Sigma-Aldrich). Oligonucleotides were mixed at an equinolar ratio in running buffer and annealed by incubation at 75°C for 10 min followed by a slow (2 h) return to room temperature. Experiments were run with 20 μM of Sta1 (protein only experiment), 5 μM DNA (DNA only experiment) or 15 μM DNA/30 μM Sta1 (binding experiment).

**NMR**

NMR experiments were acquired on an Inova 600 (Varian Inc., Palo Alto, CA) spectrometer with a 14.1 Tesla magnetic field. The spectrometer was equipped with a cryoprobe. Spectra were recorded, processed and analyzed using Vnmr 6.1C (Varian), NMRPipe (31) and NMRView 5.2 (32). Purified Sta1-h6 was dialyzed against 20 mM NH₂HCO₃ and freeze-dried. The lyophilized protein was dissolved in 20 mM CD₃COONa, pH 5.5 (uncorrected meter reading) prepared with 15 or 100% D₂O, for experiments in H₂O or D₂O, respectively. Experiments were performed at 37°C with a protein concentration of 0.3 mM. Homonuclear ¹H NOESY (nuclear Overhauser effect spectroscopy) spectra (33) were acquired with a 100 (H₂O) or 80 (D₂O) ms mixing time. The spectral width was 11 p.p.m., with 32 or 64 accumulations per free-induction decay and 400 (H₂O) or 256 (D₂O) complex data points in the indirect dimension.

**Comparative modeling**

A BLAST (34) search of the PDB with the sequence of Sta1 produced a single hit with a low E-value (0.007). The hit corresponded to the protein Mj233 from *M.jannaschii* (PDB code 1KU9). The structure of Mj233 was used as a template to obtain a model of Sta1 using Modeller v6.2 (35). The geometrical quality of the model was assessed using Procheck 3.5.4 (36).
RESULTS

Purification of DNA-binding proteins

SIRV1 promoters have been previously identified experimentally (24). In initial in vitro transcription studies on them, using a reconstituted system of *S. solfataricus*, in many cases only weak transcription initiation could be observed in comparison to strong transcription from the T6 promoter from the *Sulfolobus shibatae* virus 1. SSV1, known to be efficiently transcribed by the *Sulfolobus* transcription machinery. For example, transcription initiation from the SIRV1 56 gene promoter could be improved using whole-cell extracts of non-infected cells of *S. islandicus* (Figure 1). This observation suggested that viral promoters may need additional factors, present in the whole-cell extract, to turn on transcription of their genes.

Promoters of three genes of SIRV1, 56 and 399 (unknown function) and 134 (encoding for the structural protein), were chosen for a search of proteins involved in the regulation of transcription of the viral genome, as the previously obtained transcription map of SIRV1 had indicated that the expression of the three genes could be under transcriptional control (24).

The promoters were amplified from the viral DNA by PCR using biotinylated primers and applied in magnetic DNA affinity purification experiments (see Materials and Methods), using whole-cell extracts prepared from non-infected as well as virus-infected host cells. An unspecific competitor, pUC18 DNA, was added in high excess. The T6 promoter of SSV1, which was previously shown to be efficiently transcribed in a reconstituted transcription system of *Sulfolobus* (1,13,27), served as a control. In the conditions of the experiment, no protein was observed to bind to the T6 promoter (Figure 2). In contrast, a 14 kDa protein was bound specifically to all three SIRV promoters (Figure 2). The same result was obtained using crude extracts prepared from both virus-infected and non-infected cells, suggesting that the 14 kDa protein was encoded by the *Sulfolobus* host. Through elution from the beads, the protein could be purified to homogeneity, as judged by SDS–PAGE (Supplementary Figure S1).

Effect of the 14 kDa DNA-binding protein on viral transcription

In order to get insights into its function, the highly purified 14 kDa protein was studied using a cell-free transcription system of the host. The system consisted of recombinant TBP and TFB as well as highly purified RNA polymerase from *Sulfolobus* (13). As DNA templates, we used the same viral promoters that were used for affinity purification. The T6 promoter again served as a control. Although the 14 kDa protein had no effect on transcription initiation on the T6 promoter, a stimulation of transcription was observed for the viral gene promoters 56 and 134 (Figure 3A). Owing to its origin and activating effect, we term the protein *Sulfolobus* transcription activator, Sta1.

Identification and heterologous expression of the *sta1* gene, activating effect of the recombinant protein

For identification of the gene encoding Sta1, the protein was identified by using MALDI–TOF mass spectrometry. It was identified as a *S. islandicus* homolog of the gene SSO0048 of *S. solfataricus*, a species closely related to *S. islandicus*. The putative protein encoded by this gene, owing to its predicted helix–turn–helix motif in the annotation of *S. solfataricus* genome sequence (37), was presumed to be a transcription factor with homology to the *S. solfataricus* Lrs14 transcription regulator (18,22).

The SSO0048 gene of *S. solfataricus* was cloned and expressed in *Escherichia coli* in native form, as well as with a C-terminal His-tag. The recombinant protein in both forms, Sta1 and Sta1-h6, was purified to apparent homogeneity (Supplementary Figure S2).

The activity of the recombinant protein was inspected by in vitro transcription experiments using the SIRV1 gene promoter 134 as a template. To ensure that the in vitro transcription start site was identical to the one in vivo, the primer extension product of the in vitro transcription reaction was

![Figure 1](image1.png)

Figure 1. Transcription activation in the presence of whole-cell extracts from *S. islandicus* REN2H1. In vitro transcription from SSV promoter T6 and the SIRV promoter 56 using either (A) a reconstituted transcription system consisting of 20 ng recombinant TFB and TBP and 200 ng of the RNA polymerase from *S. solfataricus* or (B) whole-cell extracts from *S. islandicus* REN2H1. (A) Lane 1, T6 promoter; lane 2, promoter 56 of SIRV1. (B) Ten micrograms of whole-cell extracts from *S. islandicus* REN2H1 were added per reaction. Lane 1, T6 promoter; lane 2, promoter 56.

![Figure 2](image2.png)

Figure 2. DNA affinity purification of the 14 kDa DNA-binding protein from whole-cell extracts of *S. islandicus* REN2H1. Lane 1, size markers; lanes 2, 3 and 4, proteins purified by their binding to promoters 56, 134 and 399, respectively, from non-infected cell extracts; lanes 6, 7 and 8, the same as lanes 2, 3 and 4, correspondingly, but from SIRV1-infected cell extracts; and lane 5, control experiment with T6 promoter and non-infected cell extract. Proteins were silver stained.
analyzed together with its sequence reaction. The recombinant Sta1 was further tested in different amounts in the reconstituted transcription system with the promoter 134 as DNA template. The results shown in Figure 3B demonstrate that the recombinant protein at a concentration of 0.4 μg/ml has the same stimulating effect on transcription as the one purified from Sulfolobus cells. Increasing the concentration of the recombinant Sta1 > 0.4 μg/ml no stronger stimulation was observed (Figure 3B).

**Structural analysis of Sta1**

A model of the putative helix–turn–helix region of Sta1 was obtained by comparative modeling and validated experimentally using NMR and CD (Figure 4). The protein MJ233 from *M. jannaschii*, which is the closest homolog of Sta1 with known structure, served as a template (Materials and Methods). MJ233 forms a homodimer in which each monomer contains a winged helix–turn–helix (wHTH) motif in its N-terminal region and two C-terminal α-helices involved in the dimerization interface (38). The wHTH module consists of a two- or three-stranded antiparallel β-sheet and three α-helices.

1D spectra (data not shown) and 2D NOESY 1H NMR experiments of Sta1-h6 showed several characteristics indicating that the protein was rich in α-helices and contained β-sheets. In order to test the model, we assigned several signals of the NOESY spectra of Sta1-h6 acquired in D2O and H2O (Supplementary Figure S3). As β-sheets produce well-resolved downfield-shifted NH and Hα signals, and aromatic proton signals generally show very good dispersion, we focused on the antiparallel β-sheet predicted by the model, which contained two tyrosine residues (Y93 and Y95) in the second strand. We identified two tyrosine spin systems with downfield-shifted NH and Hα resonances, which indicated that the corresponding aromatic residues were located in a β-sheet. Remarkably, we found several NOEs implicating these tyrosine residues that were in accordance with the topology of the antiparallel β-sheet of the model. A careful analysis of the NOESY spectra allowed us to unambiguously assign several long-range NOEs in agreement with the model between residues 51 and 95 (Figure 4B). Hence, our results showed that Sta1 had indeed a wHTH motif that resembled to that of MJ233. Of note, analysis of the NOESY spectra indicated that Sta1 did not contain any extra β-sheet.

Sedimentation–diffusion equilibrium experiments performed by centrifugation indicated that, similar to MJ233 and generally, prokaryotic transcription regulators Sta1 showed a dimer–monomer equilibrium. The dissociation constant (KD) of this equilibrium was 5 μM at 20°C. Far-UV CD spectra of Sta1 were recorded at concentrations higher than the KD to observe the dimer’s CD. Once normalized, the latter was concentration independent between 20 and 100 μM and was dominated by contributions of α-helical structures (Supplementary Figure S4). As estimated by deconvolution of the CD spectrum, Sta1 contained 39% of α-helices and 16% of β-sheets. Taking into account that the content of α-helices of the wHTH motif (residues 30–97) was ~25%, the rest of the molecule (residues 1–29 and 98–129) should contain ~20 residues in α-helices.

**Functional analysis of Sta1**

For functional studies of Sta1, the 134 promoter was chosen, as it revealed the strongest level of activation (Figure 3A). In order to identify Sta1-binding sites, we employed DNase I footprinting assays on the promoter 134. Two non-sensitive regions were clearly visible (Figure 5). One binding site was shown to be located in the core promoter region, the second ~30 nt upstream of it. About 20 bp were protected in both regions at low concentrations of the protein (Figure 5). Increasing the protein concentration resulted in extension of
the protected areas, suggesting multiple binding. The results of the footprinting confirmed the initial observation that Sta1 was able to bind to viral promoters in the absence of other components of the transcription pre-initiation complex (Figure 2).

Sta1 showed a monomer–dimer equilibrium. To assess, which species binds DNA, we performed sedimentation/equilibrium experiments of Sta1 at 20°C in the presence or absence of a 30 bp oligonucleotide (Reg2), corresponding to a fragment of the protected region 2 in the footprinting experiment (Figure 5). Experiments were performed at protein concentrations above the Sta1 dimerization K_D. Sedimentation/equilibrium profiles (Supplementary Figure S5) of Reg2 alone were well-fitted by a single species with a mass corresponding to double-stranded DNA (18.4 kDa). Fitting of the Sta1 profiles showed an equilibrium of monomeric and dimeric forms of the protein, with only the dimeric form being substantially populated at 20 μM, the concentration used in the experiment. Importantly, the binding experiment was well described by the association of a dimer of protein with double-stranded DNA, with a dissociation constant of 10 ± 2 × 10^{-6} M. Hence, these results indicate that Sta1 binds to DNA as a dimer.

Considering that Sta1 binds DNA as a homodimer, its binding sites are expected to comprise inverted repeat sequences, as is the case for many transcription regulators. Imperfect palindromes detected in both the 20-bp-long protected regions are highlighted by arrows in Figure 5.

The apparent proximity of one of the binding sites of Sta1 with the TATA-box and the BRE element suggested that the activating effect of Sta1 might influence binding by the general transcription factors, TBP and/or TFB. Thus, the effect of Sta1 was studied in in vitro transcription experiments with varying concentrations of TBP and TFB. As expected, in absence of TBP, no transcription initiation could occur (Figure 6A). In the presence of low amounts of TBP, however, the activating effect of Sta1 was observed. Quantification of band intensities by phosphorimaging revealed that the strongest activation effect was observed using low amounts of TBP. Indeed, in the presence of 1 ng of TBP, transcription of promoter 134 is about five times stronger if Sta1 is added and the effect is even more dramatic (nearly 10-fold difference) in the presence of 5 ng of TBP (Figure 6A). The activating effect of Sta1 was also observed in analogous titration experiments in which the concentration of TFB was varied (Figure 6B).

The activating effect of Sta1 was also observed in analogous titration experiments in which the concentration of TFB was varied (Figure 6B). The effect of Sta1 was shown to be specific for SIRV1 promoters: transcription efficiency of the Sulfolobus promoter T6 was not affected in experiments with varying concentrations of either TBP or TFB (Figure 6B). These results are in line with the failure to pull-down Sta1 using the immobilized T6 promoter.

### DISCUSSION

In order to isolate proteins involved in transcription of genes of the Sulfolobus virus SIRV1, DNA affinity purification experiments were conducted. With the help of three different viral promoters immobilized on magnetic beads, we isolated a 14 kDa DNA-binding protein from crude extracts of host cells. All three promoters bound to the same protein, which turned out to be encoded on the chromosome of the Sulfolobus host. When we used crude extracts from virus-infected host cells in the purification experiments, the same protein was bound to all three promoters. The protein was identified as the product of the S.islandicus homolog of the gene SSO0048 from S.solfataricus, and was named Sta1.

Analysis of Sta1, either purified from host cell extracts or recombinant, showed an activating effect on transcription from two viral promoters in in vitro transcription experiments.

Sta1 is the first archaeal transcription regulator isolated by pull-down assays (25) with archaeal promoters directly from cells. All other archaeal transcription regulators have been identified by in silico analysis. Our results open possibilities...
for the identification of unknown proteins involved in the regulation of archaeal gene expression.

Binding sites for Sta1 were determined by DNase I footprinting. Two distinct protected regions of ~20 bp could be clearly identified. One of them was located in the core promoter region and the other was ~30 nt upstream of it. The analysis of the protected sequences allowed us to identify the imperfect 16 bp inverted repeat which is also present in the promoter region of the SIRV1 gene (Figure 7). The alignment of the identified and putative Sta1-binding sites let us to design the consensus site as ATNT-N8-A/TNAT (Figure 7). Location of a binding site in immediate proximity of the TATA-box and the BRE element suggested that the activating effect of Sta1 could be associated with the TBP and/or TFB, for example, by enhancing their recruitment or by stabilization of their binding. To confirm the possibility of a TBP/TFB-dependent effect, we performed in vitro transcription experiments under suboptimal concentrations of TBP and TFB. The results demonstrated that in the presence of low amounts of one of the two factors, either TBP or TFB, Sta1 is necessary for transcription initiation from the assayed viral promoter, especially in the case of low amounts of TBP.

In general, mechanisms of transcription activation in archaea are poorly understood. Sta1 is the first transcription activator from the Crenarchaeota for which an activating effect has been directly demonstrated in in vitro studies. The current knowledge on the molecular basis of transcription activation in the entire archaeal domain of life is limited to results gleaned from analysis of the recently described factor Ptr2 from M. janaschii and its Lrp ortholog from M. thermolithotrophicus (16,17). Activation by Ptr2 is generated by recruitment of the TATA-binding protein to the promoter, and conveys its stimulatory effect, in contrast to Sta1, from two upstream-located binding sites (16,17). Significantly, Sta1 is only distantly related to Ptr2 and appears to operate by distinct mechanisms from Ptr2, facilitating transcription at limiting TBP and TFB, suggesting that it belongs to a novel class of archaeal transcriptional activators.

Genomic analysis implies that Sta1 could represent a group of archaeal-specific transcription regulators. Conserved domains search (39) identified an archaea-specific domain in Sta1, annotated as predicted transcription regulator (CDD 12 688). Clear homologs of Sta1, containing 120–130 amino acids, are present on all three sequenced genomes of Sulfolobus species, at least in five copies in each of them, based on an E-value threshold of 0.01. However, given the small size of the protein, additional homologs may be present but not identified as significant. Applying the same threshold,
The work was supported by grants from the Deutsche Forschungsgemeinschaft (PR 663/1–2), and the EMBO. The 600 MHz spectrometer was funded by the Région Ile de France and the Institut Pasteur. Funding to pay the Open Access publication charges for this article was provided by the Institut Pasteur, Paris, France.

Conflict of interest statement. None declared.

REFERENCES


SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank Patrick Forterre and Simonetta Gribaldo for stimulating discussions, Sew Youn Peak-Chew of the MRC Laboratory of Molecular Biology for Mass Spectrometry. The work was supported by grants from the Deutsche


