ORIGINAL PAPER

Harold J. Schreier · Kelly A. Robinson-Bidle
Alla M. Romashko · Gira V. Patel

Heterologous expression in the Archaea: transcription from *Pyrococcus furiosus gdh* and *mlrA* promoters in *Haloferax volcanii*

Abstract Multicopy plasmids containing the promoter regions for *gdh* and *mlrA* genes from *Pyrococcus furiosus* were propagated in *Haloferax volcanii*. High-level expression was detected from *gdh* promoter sequences, with transcription initiating at the same start-site as that found in *P. furiosus*. For *mlrA*, several transcripts were detected, with one initiating at the *P. furiosus* start-site; removal or disruption of the likely *P. furiosus boxA* element resulted in the disappearance of this transcript, indicating that these sequences were utilized by the *H. volcanii* RNA polymerase for initiation. This is the first demonstration of the utilization of promoters from a hyperthermophilic archaeon in a mesophilic haloarchaeon and provides further evidence for the unity of transcription processes in the domain Archaea.

Introduction

In recent years, transcriptional processes in the Archaea have been the focus of a considerable amount of research (Reeve et al. 1997; Thomm 1996). Archaeal RNA polymerase (RNAP) is highly homologous to eukaryotic RNA polymerase II (RNAP II) in both major subunit and transcription factor composition. The two largest archaeal subunits, B and A, correspond to the RpoB and RpoA subunits of RNAP II, respectively, and homologs to eukaryotic TFIIB, TFIIS, TFIID, and TATA-binding protein have been found in all members of the Archaea (Brown and Doolittle 1997; Reeve et al. 1997). The similarity to RNAP II is reinforced by the finding that both yeast and human TATA-binding proteins are able to substitute for native transcription factors in cell-free archaeal transcription systems (Thomm 1996; Wettach et al. 1995).

In addition to the structural resemblance, archaeal RNAP has been shown to bind to and initiate from sequences upstream of transcription start-sites that are quite similar to the TATA box-like elements utilized by eukaryotic RNAP II (Baumann et al. 1995; Thomm 1996). Called *boxA* elements, these sequences are centered approximately 27 bp upstream of the transcription start-point (tsp) and contain an AT-rich hexanucleotide. They have been defined from both in vivo and in vitro studies of promoter regions from one member of the Crenarchaeota, *Sulfolobus* (Reiter et al. 1990), and several representatives of the Euryarchaeaota, including *Methanococcus* (Gohl et al. 1992; Hausner et al. 1991), *Pyrococcus* (Hethke et al. 1996), *Haloferax* (Palmer and Daniels 1995), and *Halobacterium* (Danner and Soppa 1996).

Because the *boxA* core hexanucleotide is conserved among members of the Archaea, it is not surprising that cell-free systems derived from one archaeon are not only able to recognize promoters from the same organism but are also capable of transcribing promoters of different origin. For instance, an extract from *Methanococcus thermolithotrophicus* was able to direct transcription from promoters of histone and 7S genes from *Methanothermus fervidus* (Koller et al. 1992; Thomm et al. 1992), and a *Pyrococcus furiosus* cell-free system was capable of initiating transcription of promoters of genes from *Methanococcus* and *Methanofermus* species (Thomm 1996).

On the other hand, although a cell-free extract from *Sulfolobus shibatae* was able to initiate transcription of certain plasmid- and phage-encoded genes from

H.J. Schreier (✉) · K.A. Robinson-Bidle[1] · A.M. Romashko · G.V. Patel
Center of Marine Biotechnology, The University of Maryland Biotechnology Institute, Columbus Center, Suite 236, 701 E. Pratt Street, Baltimore, MD 21202, USA
Tel. +1-410-234-8874; Fax +1-410-234-8896
e-mail: Schreier@umbi.umd.edu

H.J. Schreier · K.A. Robinson-Bidle[1]
Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD 21228, USA

Current address:
[1] Marine Biology Research Division, Scripps Institution of Oceanography, University of California, La Jolla, CA 92093-0202, USA

Halobacterium halobium (also known as *H. salinarium*), transcription of a chromosomal locus could not be detected (Hudepohl et al. 1991). Such a result could be explained by the absence of particular factors in the crude *S. shibatae* extract necessary to achieve specific *H. halobium* gene expression. However, the sulfur-dependent thermophilic *S. shibatae* and the mesophilic and halophilic *H. halobium* have not only adapted to very different environments but are also distantly related as members of different branches of the Archaea, the Crenarchaeota and Euryarchaeota, respectively (Brown and Doolittle 1997). Thus, internal biochemical and biophysical parameters may be sufficiently different between the two organisms to account for the inability of one to recognize promoter signals of certain genes from the other. Given the extensive range of diversity among the Archaea as well as the intricate nature of polymerase–promoter interactions, it is likely that the requirements for these interactions are more complex than what is specified by the *boxA* hexanucleotide. Such appears to be the case for tRNA genes from *Methanococcus* (Hausner et al. 1991) and *Haloferax* (Palmer and Daniels 1995), which have strict requirements for sequences adjacent to *boxA* for maximal expression. Similarly, the presence of alternative TBPs and TFIIBs have been reported in the halophilic Archaea (Reeve et al. 1997), and promoters recognized by these transcription factors would require additional control regions specific for their utilization.

The development of plasmids and transformation protocols (Robb et al. 1995) have made it possible to examine archaeal transcriptional control in vivo, and a few studies have used these approaches to express genes in the same or related hosts (Danner and Soppa 1996; Palmer and Daniels 1995). However, very few in vivo analyses have examined heterologous expression among Archaea from very different environments. In the present study, we utilize a vector developed for *Haloferax volcanii* (Palmer and Daniels 1995) to examine whether the *gdh* and *mlrA* promoters from *P. furiosus* may be expressed in the heterologous host. The former, which is responsible for glutamate dehydrogenase, is highly expressed in *P. furiosus* in an apparently unregulated manner (DiRuggiero and Robb 1996). The latter is poorly expressed compared to *gdh* and is regulated in *P. furiosus* approximately five- to tenfold by the presence of maltose in the growth medium (Robinson et al. 1994; Robinson and Schreier 1994). We found that both hyperthermophile promoters were capable of being recognized by the halophilic host and that expression occurred from the same tsp in both organisms, although there were differences between levels of expression as well as the use of alternative *boxA* elements.

Materials and methods

Strains and growth conditions

Haloferax volcanii strain WFD11 (DSM 5716) was provided by Dr. C. Daniels. Cultures were grown at 37°C with vigorous shaking in liquid media or plated on agar plates as described previously (Robb et al. 1995). When necessary, mevinolin was added to liquid or solid media at a concentration of 8 and 16 μg/ml, respectively.

Pyrococcus furiosus (DSM 3638) was cultured at 100°C as described previously (Robb et al. 1992). Filter-sterilized maltose was added to autoclaved media to a final concentration of 10 mM. Cultures were harvested approximately 16 h after inoculation and were immediately chilled on ice. Cells were centrifuged at $10\,000 \times g$ for 20 min and washed with imidazole buffer [1 M imidazole-HCl, pH 7.1, 100 mM dithiothreitol (DTT), and 10 mM Na_2EDTA] and were placed at −70°C if not used immediately.

Escherichia coli strains DH5α [F′/*endA1 hsdR17* ($r_K^- m_K^+$) *supE44 thi-1 recA1 gyrA* (Nal^r) *relA1* Δ(*lacIZYA-argF*)*169 deoR* (φ80*dlac*Δ (*lacZ*)*M15*] and JM110 [F′ *traD36 supE44 thi proA$^+$B$^+$/rpsL* (Str^r) *thr leu lacY galK galT ara FhuA dam dcm lacI^q* Δ(*lac-proAB*) Δ (*lacZ*)*M15*] were grown in Luria broth at 37°C. Cultures containing plasmids were grown in Luria broth supplemented with ampicillin at 50 μg/ml.

Plasmids

Plasmids used in this study are derivatives of the mevinolin-resistant (Mev^r) multicopy shuttle vector pWL222 (Palmer and Daniels 1995), and promoter-containing sequences harbored within each plasmid are shown in Fig. 1. To clone *P. furiosus mlrA* and *gdh* promoter regions, *Hin*dIII and *Xba*I restriction sites were introduced 5′ and 3′ of each promoter, respectively, by the polymerase chain reaction (PCR), using as templates plasmid pKAR4, a pBluescript KS (Stratagene, La Jolla, CA, USA) derivative harboring a 3.2-kb *Hin*dIII fragment containing the *mlrA* promoter region (Robinson et al. 1994), and L8.1 (a gift from J. DiRuggiero), a λ phage containing the *gdh* promoter region. Following PCR amplification, preparations were digested with *Hin*dIII and *Xba*I, and fragments were isolated and then ligated into plasmid pWL222 linearized with *Hin*dIII/*Xba*I (Sambrook et al. 1989). Insertion of DNA fragments between these sites replaced the 48-bp *H. volcanii* tRNA^{Lys} promoter element harbored in pWL222. Ligation mixtures were then transformed into *E. coli* strain DH5α, selecting ampicillin-resistant colonies as described previously (Sambrook et al. 1989). Plasmids containing inserts were screened either directly by restriction analysis or via PCR using the PROEXI primer (see following), which complements downstream tRNA^{ProM} sequences (Palmer and Daniels 1994), and a second primer internal to the cloned fragment. After purification, the presence of cloned inserts in all plasmid constructs were confirmed by dideoxy DNA sequencing (Sanger et al. 1977) using the Sequenase Version 2.0 Sequencing Kit (Amersham, Arlington Heights, IL, USA).

To introduce plasmid DNA into *H. volcanii*, plasmids were first passed through *E. coli* strain JM110 as described by Chung et al. (1989). Ampicillin-resistant transformants were selected, and plasmid DNA was extracted from a single transformant for each derivative. Plasmids were then
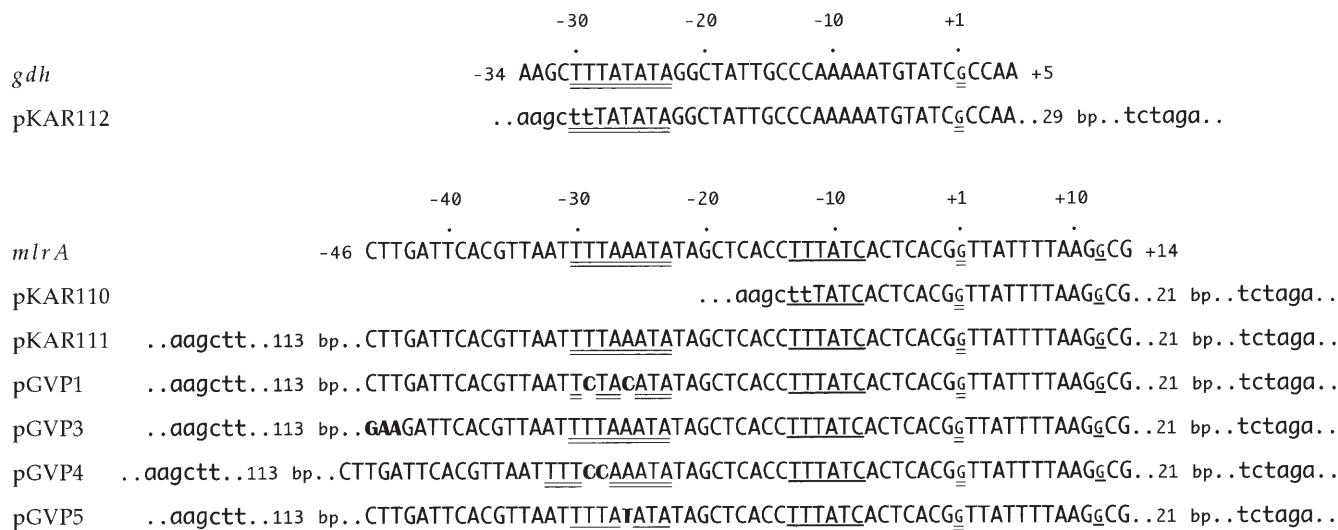
```
                         -30        -20         -10       +1
                          .          .           .        .
 gdh                 -34 AAGCTTTATATAGGCTATTGCCCAAAAATGTATCgCCAA +5
 pKAR112                ..aagcttTATATAGGCTATTGCCCAAAAATGTATCgCCAA..29 bp..tctaga..


                    -40        -30        -20        -10        +1        +10
                     .          .          .          .          .          .
 mlrA          -46 CTTGATTCACGTTAATTTTAAATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG +14
 pKAR110                           ...aagcttTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
 pKAR111  ..aagctt..113 bp..CTTGATTCACGTTAATTTTAAATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
 pGVP1    ..aagctt..113 bp..CTTGATTCACGTTAATTCTACATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
 pGVP3    ..aagctt..113 bp..GAAGATTCACGTTAATTTTAAATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
 pGVP4    ..aagctt..113 bp..CTTGATTCACGTTAATTTTCCAAATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
 pGVP5    ..aagctt..113 bp..CTTGATTCACGTTAATTTTATATATAGCTCACCTTTATCACTCACGgTTATTTTAAGgCG..21 bp..tctaga..
```

Fig. 1. Nucleotide sequence of *Pyrococcus furiosus gdh* and *mlrA* promoter regions contained in plasmids used in this study. *Numbering* refers to position relative to the *P. furiosus* transcription start-point (tsp) for each gene. Likely *P. furiosus* and *Haloferax volcanii* promoters are *doubly* and *singly underlined*, respectively, with their accompanying initiation sites indicated by *small capitals*. (The *H. volcanii* promoter is indicated for transcript *b*; see text for details.) Each plasmid was created by cloning fragments into plasmid pWL222 at *Hin*dIII and *Xba*I sites (*lower case*) as described in Materials and methods. *Bold-faced positions* indicate base changes created by substitutions or insertions

used to transform *H. volcanii* strain WFD11 by the polyethyleneglycol procedure described by Cline et al. (1995), selecting for Mev[r] transformants. Colonies appeared 7–10 days after plating. To confirm that Mev[r] resulted from the presence of a pWL222 derivative, DNA was isolated as described by Ng et al. (1995), and the presence of tRNA[ProM] sequences carried on the pWL222 derivative in each transformant was demonstrated by southern analysis (Sambrook et al. 1989). In addition, verification that cloned sequences were not altered when propagated in *H. volcanii* was done by sequencing the promoter-tRNA[ProM] regions of plasmids isolated from *H. volcanii* after passing through *E. coli* strain DH5α.

RNA isolation

RNA was isolated from *H. volcanii* as follows. Strains were grown in 10-ml cultures at 37°C with shaking to an OD_{550} of 1.0. Cultures were transferred to 15-ml polyethylene tubes, centrifuged for 10 min at $3500 \times g$ at 4°C, and pellets were suspended in 5 ml RNAzol B (Tel-Test, Friendswood, TX, USA). After suspension, 0.5 ml chloroform was added, mixed thoroughly, and samples were incubated on ice for 5 min. Suspensions were then centrifuged for 15 min at $12\,000 \times g$ and 4°C. The aqueous phases were transferred to diethylpyrocarbonate-treated tubes, and an equal volume of isopropanol was added and placed at 4°C for 15 min, followed by centrifugation for 15 min at $12\,000 \times g$ and 4°C. After washing with 75% ethanol and centrifugation for 8 min at $7500 \times g$, pellets were air-dried and resuspended in H_2O to give a concentration of approximately 5 mg/ml. RNA was isolated from *P. furiosus* cultures as described previously (DiRuggiero and Robb 1995).

Primer extension analysis

Primer extension reactions were done using end-labeled primers as follows. Primer (30 pmol) was added to 5 μl [γ-^{32}P]ATP (10 mCi/ml), 1 μl 10× polynucleotide kinase buffer (700 mM Tris-HCl, pH 7.6, 100 mM $MgCl_2$, 50 mM DTT), 0.5 μl T4 polynucleotide kinase (4 U) (Promega, Madison, WI, USA), 0.5 μl 400 mM DTT, and 1 μl H_2O, and the mix was incubated for 45 min at 37°C, followed by 10 min at 65°C. Labeled primer was purified from unincorporated label using NENsorb 20 nucleic acid purification cartridges (DuPont, Boston, MA, USA) according to the manufacturer's instructions. After purification and lyophilization, the primer was suspended in 20 μl H_2O. The PROEXI primer (5′-CCCAAAGCGAGAATCATACCAC-3′) (Palmer and Daniels 1994) was used to detect expression from pWL222 and its derivatives in *H. volcanii*. For primer extension studies using RNA isolated from *P. furiosus*, *mlrA* and *gdh* start-points were determined using primers MLRA7 (5′-GCTCCTCAAACCACT-3′) (Robinson and Schreier 1994) and GDHPE (5′-AATAACAATTTCATA-3′), respectively. Expression from the plasmid-encoded Mev[r] locus (3-hydroxy-3-methylgultaryl-CoA reductase) was done using the MEVPE primer (5′-TTCGTGGAG-GCGGAGG-3′).

Extension reactions were done as described by Dennis and Chow (1995). After coprecipitating RNA and probe, pellets were resuspended in 10 μl KTE buffer (20 mM Tris-HCl, pH 8.5, 80 mM KCl, 0.5 mM Na_2EDTA) and incubated for 10 min at 65°C, followed by incubation at 50°C for 1 h. Extensions were catalyzed by the addition of 10 μl MNM extension mix (10 mM $MgCl_2$, 1 mM each dNTP, 10 mM DTT) and 2 μl (400 U) murine Moloney leukemia virus re-

verse transcriptase (Amersham) and incubated for 45 min at 37°C. After precipitation with 3 μl 4 M NaOAc and 70 μl 100% ethanol at −80°C for 20 min, reactions were centrifuged for 15 min at 4°C, and pellets were washed in 70% ethanol and dried at room temperature. Pellets were suspended in formamide loading dye and electrophoresis was done using an 8% denaturing polyacrylamide sequencing gel (Sambrook et al. 1989). After electrophoresis, the gel was dried and exposed to XAR-5 autoradiography film.

Mutagenesis

Mutagenesis was done using the Chameleon double-stranded site-directed mutagenesis kit (Stratagene) following the manufacturer's procedures. For these constructs, mutations were introduced into *mlrA* promoter sequences within the *Hin*dIII-*Xba*I fragment of plasmid pKAR111a, a pCRII (Invitrogen, Carlsbad, CA, USA) derivative, using mutagenic oligonucleotides. After screening for plasmids containing alterations, *Hin*dIII-*Xba*I fragments were then inserted into plasmid pWL222 as described. Mutations either within the putative *mlrA boxA* element (pGVP1, pGVP4, and pGVP5) or upstream of the element (pGVP3) are shown in Fig. 1. The presence of each mutation was confirmed by sequencing the final constructs.

Materials

All reagents were of the highest purity available. Mevinolin was obtained from Dr. A. Alberts, Merck Sharpe & Dohme Research Laboratories. Restriction endonucleases and DNA modification enzymes were purchased from Boehringer-Mannheim (Indianapolis, IN, USA), Promega, New England Biolabs (Beverly, MA, USA), or Amersham. DNA oligonucleotides were synthesized by the Center of Marine Biotechnology's Bioanalytical Sequencing Laboratory.

Results

The *Haloferax volcanii* plasmid

To examine expression from *P. furiosus* promoters in *H. volcanii*, multicopy plasmid pWL222 (~17 copies/cell) described by Palmer and Daniels (1995) was used. This plasmid has been used to dissect the *H. volcanii* tRNA^{Lys} promoter, which directs expression of a yeast tRNA^{ProM} gene (Palmer and Daniels 1994, 1995). Using reverse transcriptase primer extension analysis to detect transcription of tRNA^{ProM} sequences (see Materials and methods), we found that initiation from the *H. volcanii* tRNA^{Lys} promoter in pWL222 occurred at a G residue 20 bp downstream from the *H. volcanii boxA* element (not shown), consistent with previous results (Palmer and Daniels 1995). Importantly, no extension product could be detected using the PROEXI primer when we examined RNA prepared from plasmid-less strain WFD11 (not shown).

Expression from the *Pyrococcus furiosus gdh* promoter in *H. volcanii*

To determine whether the *H. volcanii* transcription machinery is capable of directing expression from the *P. furiosus gdh* promoter, transcript analysis was done using RNA extracted from strain WFD11(pKAR112). Plasmid pKAR112 contains *gdh* promoter sequences (−35 to +35, relative to the tsp) fused to tRNA^{ProM} sequences (see Fig. 1). As shown in Fig. 2, lane 2, a transcript was detected in strain WFD11(pKAR112) that originated from within *gdh* sequences. Transcription initiation occurred at a G residue, 22–24 bases downstream from likely *boxA* sequences (shown in Fig. 1). Initiation at the same site was found when we examined RNA extracted from *P. furiosus* (Fig. 2, lane 1) and was in agreement with results obtained by others (DiRuggiero and Robb 1996). When we compared transcription from the very highly expressed plasmid-encoded Mev^r determinant (not shown) to *gdh* expression, we estimated that *gdh*-derived transcripts were approximately fivefold lower than the level observed for Mev^r transcripts This estimate was based on the length of time required to visualize equivalent levels of Mev^r and *gdh* primer extension products. Thus, *H. volcanii* RNAP appears to efficiently utilize the *P. furiosus gdh* promoter. Furthermore, because *P. furiosus gdh* promoter sequences in plasmid pKAR112 only include the DNA region to −35, our results demonstrated that *gdh* expression in *H. volcanii* did not require specific sequences upstream of −35, relative to the tsp.
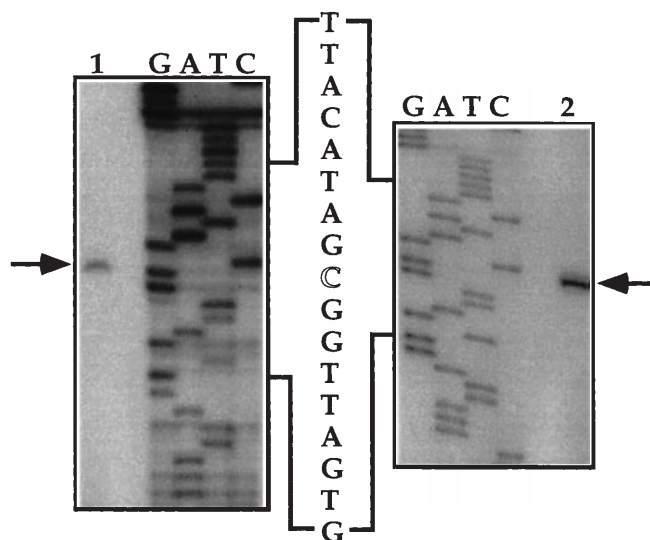

Fig. 2. Transcript analysis of the *P. furiosus gdh* promoter in *P. furiosus* and *H. volcanii*. *Lane 1*: the GDHPE primer was used for primer extension and DNA sequencing reactions (L8.1 as template) with total RNA (~5 μg) isolated from *P. furiosus* as described in Materials and methods. *Lane 2*: primer extension and DNA sequencing reactions (pKAR112 as template) used the PROEXI primer and total RNA (~20 μg) from *H. volcanii* strain WFD11(pKAR112) as described. Exposure time for the lane 2 primer extension reaction was 65 h at −70°C. The tsp, which is shown by *arrows*, is indicated as an *outlined* base in the relevant antisense nucleotide sequence (extending in the 5′- to 3′-direction from the *bottom* to the *top* of the figure)
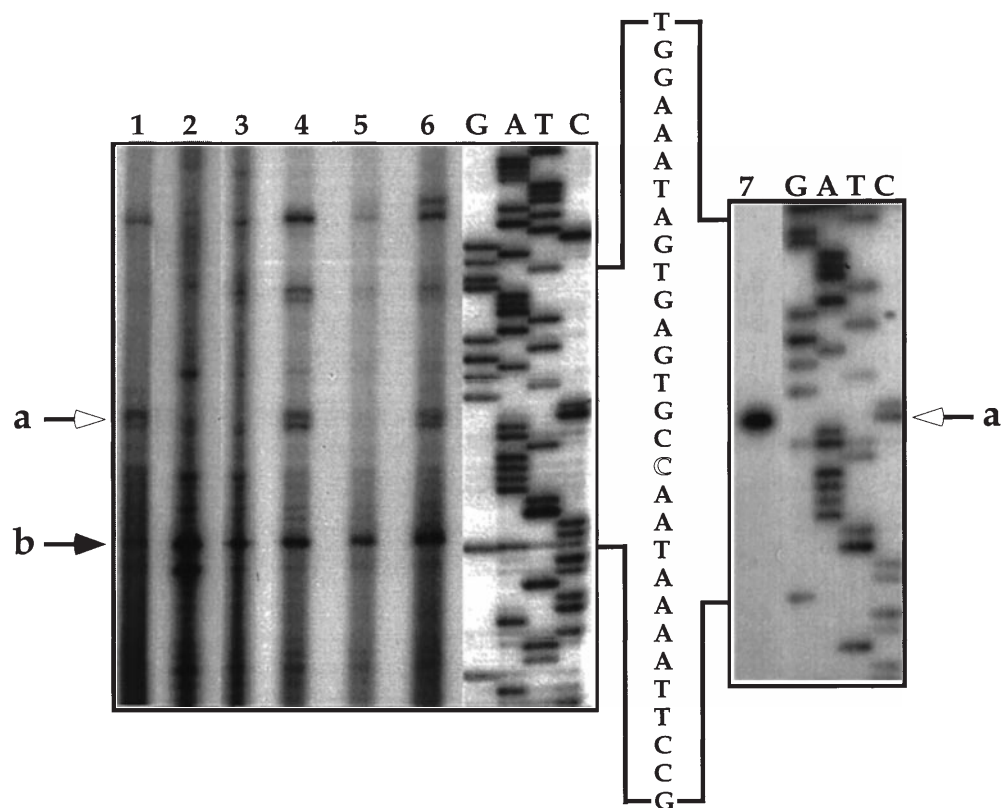
Fig. 3. Transcript analysis of the *P. furiosus mlrA* promoter in *P. furiosus* and *H. volcanii*. Primer extension reactions using the PROEXI primer and total RNA (~50 µg) from *H. volcanii* strain WFD11 containing plasmids pKAR111 (*lane 1*), pKAR110 (*lane 2*), pGVP1 (*lane 3*), pGVP3 (*lane 4*), pGVP4 (*lane 5*), and pGVP5 (*lane 6*) along with sequencing reactions (pKAR111 as template) used the same primer as described in Materials and methods. The MLRA7 primer was used for primer extension (*lane 7*) and accompanying DNA sequencing reactions (pKAR4 as template) with total RNA (~10 µg) from a *P. furiosus* culture grown in the presence of maltose as described. Results for lanes 1–6 are a composite of several exposure times ranging from 10 to 14 days at −70°C. *Arrows* indicate start-points for transcript *a*, which is indicated as an *outlined base* in the antisense strand sequence (extending in the 5′- to 3′-direction from the *bottom* to the *top* of the figure), and transcript *b*

Expression from the *P. furiosus mlrA* promoter

Because *gdh* transcription in *P. furiosus* occurs in an apparently unregulated high-level manner (DiRuggiero and Robb 1996), we asked whether *H. volcanii* is capable of expressing the regulated *mlrA* promoter that is conditionally expressed in *P. furiosus*. Like *gdh*, transcription of *mlrA* in *P. furiosus* begins at a G residue 22–24 bp downstream from a potential *box A* element (Fig. 3, lane 7) (Robinson and Schreier 1994). Strain WFD11(pKAR111) contains the *mlrA* promoter region from positions −159 to +35, relative to the *P. furiosus* tsp, fused to tRNA^ProM sequences. When we looked for tRNA^ProM transcripts, we found several that appeared to initiate at different sites (Fig. 3, lane 1). Two transcripts differing in size by one base could have arisen from the binding of RNAP at or near the same promoter utilized in *P. furiosus*. Designated transcript *a* in Fig. 3, the larger transcript appeared to have initiated from the same tsp found in *P. furiosus*, suggesting that *mlrA* promoter sequences can be recognized by the *H. volcanii* transcription machinery. The shorter transcript may have been produced posttranscriptionally or resulted from the binding of RNAP to overlapping *boxA* sequences. While we presently cannot rule out either possibility, the fact that

the relative levels of these transcripts to each other varied from experiment to experiment (not shown) as well as the results from mutation analysis (see following) makes it likely that the shorter transcript is derived from the larger one. A third transcript, designated transcript *b* in Fig. 3, initiated 11 bp downstream from the *P. furiosus* tsp and was present at a level as high as or higher than transcript *a*. Finally, a fourth transcript that was longer than the other three was detected that appeared to have been the result of read-through from upstream *mlrA* sequences. We note that the levels of all transcripts were very low and, based on a visual comparison of their levels of expression to those of *gdh*-derived transcripts, we estimate them to be from 10- to 20 fold less than the level observed for the *gdh* transcript.

To begin studying the origins of transcripts *a* and *b*, we examined their synthesis in WFD11 cells carrying plasmid pKAR110, which harbors *mlrA* sequences from −14 to +35, relative to the *P. furiosus* tsp. Using the PROEXI primer as probe, several extension products were detected in RNA isolated from strain WFD11(pKAR110) (Fig. 3, lane 2). However, the transcript pattern was significantly different from that observed for strain WFD11(pKAR111). While transcript *b* was still evident, its level was approximately fivefold greater than that observed in

WFD11(pKAR111). In addition, a second shorter transcript appeared whose start-point was 2 bp downstream from the tsp for transcript *b*. More importantly, both transcript *a* and its accompanying shorter transcript were no longer detectable, indicating that synthesis of these transcripts required *mlrA* sequences upstream of −14. These results suggested that transcript *b* was not the result of the processing of transcript *a* or other transcripts derived from upstream sequences. Rather, synthesis of transcript *b* was dependent on a likely *boxA* element situated 18–19 bp upstream of the transcript's initiation site, within *mlrA* sequences contained on plasmid pKAR110 (see Fig. 1).

If transcript *a* was derived from the same *boxA* used by *P. furiosus* instead of being processed from a larger mRNA, then alteration of likely *boxA* sequences should affect its synthesis in *H. volcanii*. To explore this possibility we constructed plasmids having alterations within or outside the putative *P. furiosus boxA* element and examined the effect on *mlrA*-directed expression in *H. volcanii*. Plasmids pGVP1 and pGVP4 contain modifications within *mlrA boxA* sequences (Fig. 1), with the former possessing T to C and A to C substitutions within the element (−29 and −26, respectively) and the latter containing a CC insertion at −28, relative to transcript *a*'s tsp. If used as a promoter in *H. volcanii*, these changes should severely alter the *boxA* element in a manner that would make it difficult for RNAP to recognize these sequences for binding or initiation. As can be seen in Fig. 3 (lanes 3 and 5), transcript *a* could not be detected in both strains WFD11(pGVP1) and WFD11(pGVP4), while transcript *b* production was unaffected. Thus, synthesis of transcript *a* required the particular DNA sequence in the −29 through −26 region. It should be noted that the shorter transcript accompanying transcript *a* in strain WFD11(pKAR111) was also not detected in these strains, which is consistent with the notion that synthesis of both transcripts requires the involvement of the same or overlapping sequences.

Two strains with plasmids that did not show any significant change in the production of transcript *a* and its accom-

panying shortened transcript were WFD11(pGVP3) and WFD11(pGVP5) (Fig. 3, lanes 4 and 6, respectively). Plasmid pGVP3 carries a GAA in place of the CTT at positions −48 to −46 (see Fig. 1). These sequences are within a region shared by promoters of at least three maltose-regulated genes in *P. furiosus* (Schreier, unpublished), suggesting that they may play a role in regulation. The absence of any effect on transcript *a* production in strain WFD11(pGVP3) indicates that this region is not important for its expression in *H. volcanii*. For plasmid pGVP5, the inability of the transversion mutation at position −26 to significantly influence RNAP's recognition of these sequences may be explained by the fact that the alteration maintains the characteristic A + T composition that is common to the central core of *boxA* elements (see Discussion).

Discussion

Studies using both in vitro and in vivo systems have shown that archaeal transcription signals appear to be relatively uniform among members of the Archaea (Thomm 1996). That is, the promoter *boxA* element conforms to the general consensus 5′-T/CTTAAN-3′ (Fig. 4) and does not vary substantially from this sequence in those Crenarchaea and Euryarchaea that have been examined. However, in vivo studies examining whether promoters from one archaeon may be recognized by the RNAP of an archaeon from different kingdoms or from very different environments have not been done. Aravalli and Garrett (1997) demonstrated maintenance of plasmids from the crenarchaeon *Sulfolobus* in the euryarchaeon *P. furiosus*. Furthermore, they showed that expression of a plasmid-borne alcohol dehydrogenase gene from *S. solfataricus* could be obtained in *P. furiosus*. However, Aravalli and Garrett (1997) did not analyze transcription start-sites for plasmid-encoded genes, and it is not known whether initiation of these genes in both organisms occurred from the same promoters. In the present study we

Fig. 4. Comparison of *gdh* and *mlrA* promoter regions to haloarchaeal consensus sequences. The consensus for *H. volcanii* and *H. salinarium* promoter elements was derived by Palmer and Daniels (1995) and Danner and Soppa (1996), respectively. Bases within the *H. volcanii boxA* core are *underlined*; those that do not conform to the *H. salinarium* (*halobium*) consensus are indicated by *lowercase letters*

```
GcTTTATAtagGCT   -19...  ...+1 C        gdh Pf

AaTTTTAAAtatag   -19...  ...+1 G        mlrA Pf (transcript a)

AtTTTAAAtaTaGC   -18...  ...+2 T        mlrA Pf (transcript a₋₁)

AaTTTTAtAtatag   -19...  ...+1 G        mlrA Pf (pGVP5)

AagcTTTAtCactC   -19...  ...+1 G        mlrA Hv (pKAR110; transcript b)

caccTTTAtCactC   -19...  ...+1 G        mlrA Hv (pKAR111; transcript b)

GtcATTTTACCcaC   -19...  ...+1 G        tRNA^Lys Hv

      T TA                             H. volcanii consensus
      T  AN
      C AT

    A GTTTTTAACCGGC  -19               H. salinarium consensus
    G  AAAAG T CT
```

have shown that a halophilic, mesophilic archaeon is capable of recognizing the signals of two promoters from a hyperthermophilic sulfur-dependent archaeon. To our knowledge, this is the first demonstration of in vivo expression of promoters from one archaeon in an archaeon from a very different environment. The ability of *H. volcanii* to recognize *P. furiosus gdh* and *mlrA* promoters extends the range of transcription signal uniformity among the Archaea.

If the very different environments from which *H. volcanii* and *P. furiosus* are found is considered, our finding that the *gdh* promoter can be recognized in the haloarchaeon suggests that promoter sequence may be the primary determinant for RNAP interaction. This promoter contains a consensus *boxA* between -29 and -23 that is highly homologus to the general archaeal *boxA*, with four of six or six of six matches, depending on the hexanucleotide used to define the element (see Fig. 4). Interestingly, the region between -32 and -19 shares 10 of 14 bp of the *H. salinarium* consensus established via saturation mutagenesis of the *fdx* promoter (Danner and Soppa 1996) (Fig. 4). It is likely that the high-level, apparently constitutive expression observed by others for *gdh* in *P. furiosus* (DiRuggiero and Robb 1996) may also be attributed to the same features. We note, however, that other factors may be involved in the ability for the *H. volcanii* RNAP to recognize the *P. furiosus gdh* promoter. In the present study, the region directly upstream of -32 was fused to *H. volcanii* sequences provided by plasmid pWL222, and we cannot rule out the possiblity that these sequences played a role in the ability for RNAP to recognize the *gdh* promoter as well as the level of expression. These sequences have the potential to affect DNA conformation near the *gdh boxA* because they are haloarchaeal and contain a 60%–65% G + C composition (Moore and McCarthy 1969) compared *to P. furiosus* DNA, which is typically 38% G + C (Fiala and Stetter 1986). It has been shown that the structure of GC-rich DNA near certain promoters in Haloarchaea plays a role in the strength of those promoters (Yang et al. 1996). Whether the presence of *P. furiosus* DNA upstream of -32 would have influenced *gdh* expression in *H. volcanii* is presently unknown.

For the *mlrA* promoter, several factors appear to be involved in determining where RNAP initiates transcription as well as the relative frequency of initiation. Like the *gdh* promoter, the DNA sequence within the region most likely used as the *mlrA* promoter in *P. furiosus* is highly homologous to the *H. volcanii boxA* consensus (Fig. 4). A *boxA* hexanucleotide is found between positions -29 and -24 and another is identifiable between -28 and -23. The DNA region encompassing both *boxA* sequences is homologous to the *H. salinarium* consensus, with the former sharing 8 of 14 and the latter 10 of 14 (see Fig. 4). However, unlike the *gdh* promoter, which yielded one unique transcript in *H. volcanii*, several transcripts were produced from the *mlrA* promoter region. Two transcripts, which differed in size by one base, were found to require nucleotide sequences upstream of -23, relative to the *P. furiosus* start-point, as removal or alteration of these sequences resulted

in the elimination of both transcripts. The longer transcript, transcript ***a***, initiated at the same site that was found for *mlrA* mRNA from *P. furiosus*. The production of these two transcripts in *H. volcanii* may be explained by the binding of the basal transcription complex to either of the two *boxA* sequences. However, changing the A to T at position -25 (plasmid pGVP5), which resulted in moving the putative -28 to -23 *boxA* away from consensus by one base (see Fig. 4), did not significantly influence expression of either transcript. Thus, *H. volcanii* RNAP may not have a preference for either *boxA* element, binding to both with the same apparent efficiency. Because only one start site for *mlrA* is found in *P. furiosus*, it is conceivable that RNAP in that organism is directed to only one of the *boxA* sequences by the presence of accessory factors. A requirement for such factors is consistent with the fact that *mlrA* is a regulated promoter in *P. furiosus*.

Our preliminary experiments using extracts obtained from maltose-grown cultures have provided evidence for the presence of several proteins having an affinity for this region (Romashko, Patel, and Schreier, unpublished). Such factors may not be found in *H. volcanii* or are not present under the growth conditions used for analysis. Alternatively, the conformation of the DNA structure around the *mlrA* promoter may be sufficiently different in *H. volcanii* from that found in *P. furiosus* to affect how RNAP binds to these sequences. While this does not appear to be a factor in *gdh* expression, it is likely that DNA conformation is sufficiently different in *H. volcanii* from that found in *P. furiosus*. Plasmid topology, and by extension that of the chromosome, is known to be different between mesophilic and hyperthermophilic Archaea, with the former being negatively supercoiled and the latter close to the relaxed state (Charbonnier and Forterre 1994). Thus, the difference in supercoiling may have an effect on how the *mlrA* promoter region is recognized by the *H. volcanii* RNAP. Finally, we cannot rule out the possibility that the shortened transcript was derived from the larger via a posttranscriptional processing event. Such an event would be consistent with the low-level expression for transcript ***a*** and our finding that this transcript and its shortened companion are both short-lived under the conditions used for growth and mRNA isolation (Romashko and Schreier, unpublished).

The low-level expression observed for *mlrA* transcripts in *H. volcanii* was surprising given the high homology of the *mlrA* promoter region to the consensus of *H. volcanii* and haloarchaeal promoters in general. Unlike the vector construct containing the *gdh* promoter, vector sequences were placed more than 100 bp upstream of the *P. furiosus mlrA boxA* sequence. As noted for the *gdh* construct, it is conceivable that the low G + C character of the *P. furiosus* DNA within this region may have played a role in restricting RNAP initiation frequency. Preliminary studies, however, have indicated that the same transcript pattern was obtained when vector sequences were placed adjacent to -34 as well as at -284, relative to the *P. furiosus mlrA* start-site (Romashko and Schreier, unpublished). Alternatively, the low-level expression may result from competition of the *H. volcanii* transcription system for other promoters

in the *mlrA* region. One of these promoters is responsible for a major transcript that initiated 11 bp downstream from the *P. furiosus* start-point. Expression of this transcript, referred to as transcript ***b***, was likely the result of binding of RNAP to *boxA* sequences located 9–10 bp downstream from *mlrA boxA* sequences (Figs. 1 and 4). This alternative *boxA* shares 5 of 6 bp of the *H. volcanii* consensus and 6 of 14 bp of the *H. salinarium* consensus (Fig. 4). When vector sequences were placed adjacent to this alternative *boxA* (as in plasmid pKAR110), a significant increase in transcript ***b*** expression was noted. This increase may result from the high G + C content of vector sequences as well as the alteration of the *boxA* itself, changing the C to an A at position −32, bringing the promoter one position closer to consensus (Fig. 4). Whether removal of the alternative *boxA* would influence the level of transcript ***a*** expression is presently unknown. Interestingly, transcripts originating from the transcript ***b*** start-site have not been detected in *P. furiosus* under any condition, indicating that the alternative *H. volcanii boxA* is not recognized by the *P. furiosus* transcription system.

Our finding that *H. volcanii* is capable of recognizing promoters from *P. furiosus*, an archaeon that is not only from a very different environment but is also distantly related phylogenetically, suggests that *H. volcanii* may be useful for carrying out similar studies on promoters from other archaeal systems. Indeed, given the broad range of transcriptional unity in the Archaea, it is imaginable that other archaeal hosts having well-characterized transformation and vector systems may be useful for similar studies of promoters from less well developed systems. Such studies will help elucidate both the similarities and differences between members of the Archaea and should provide some important information in understanding the factors involved in transcription of both unregulated and regulated promoters.

References

Aravalli RN, Garrett RA (1997) Shuttle vectors for hyperthermophilic archaea. Extremophiles 1:183–191

Baumann P, Qureshi SA, Jackson SP (1995) Transcription: new insights from studies on Archaea. Trends Genet 11:279–283

Brown JR, Doolittle WF (1997) Archaea and the prokaryote-to-eukaryote transition. Microbiol Mol Rev 61:456–502

Charbonnier F, Forterre P (1994) Comparison of plasmid DNA topology among mesophilic and thermophilic eubacteria and archaebacteria. J Bacteriol 176:1251–1259

Chung CT, Niemela SL, Miller RH (1989) One step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci USA 86:2172–2175

Cline SW, Pfeifer F, Doolittle WF (1995) Transformation of halophilic archaea. In: Robb FT, Sowers KR, DasSarma S, Place AR, Fleischmann EM, Schreier HJ (eds) Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY, pp 197–204

Danner S, Soppa J (1996) Characterization of the distal promoter element of halobacteria *in vivo* using saturation mutagenesis and selection. Mol Microbiol 19:1265–1276

Dennis PP, Chow J (1995) Analysis of RNA transcripts in Archaea. In: Robb FT, Sowers KR, DasSarma S, Place AR, Fleischmann EM, Schreier HJ (eds) Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY, pp 163–177

DiRuggiero J, Robb FT (1995) RNA extraction from sulfur-utilizing thermophilic archaebacteria. In: Robb FT, Sowers KA, DasSarma S, Place AR, Fleischmann EM, Schreier HJ (eds) Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY, pp 97–100

DiRuggiero J, Robb FT (1996) Enzymes of central nitrogen metabolism from hyperthermophiles: characterization, thermostability, and genetics. Adv Protein Chem 48:311–339

Fiala G, Stetter KO (1986) *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch Microbiol 145:56–61

Gohl HP, Hausner W, Thomm M (1992) Cell-free transcription of he *nifH1* gene of *Methanococcus thermolithotrophicus* indicates that promoters of archaeal *nif* genes share basic features with the methanogen consensus promoter. Mol Gen Genet 231:286–295

Hausner W, Frey G, Thomm M (1991) Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA^{VAL} gene of *Methanococcus vannielii*. J Mol Biol 222:495–508

Hethke C, Geerling ACM, Hausner W, de Vos WM, Thomm M (1996) A cell-free transcription system for the hyperthermophilic archaeon *Pyrococcus furiosus*. Nucleic Acids Res 24:2369–2376

Hudepohl U, Gropp F, Horne M, Zillib W (1991) Heterologous *in vitro* transcription from two archaebacterial promoters. FEBS Lett 285:257–259

Koller G, Reeve J, Frey G, Thomm M (1992) Transription in vitro and in vivo of the 7S RNA gene associated with the ribosomal RNA operon in the hyperthermophilic archaeon *Methanothermus fervidus*. FEMS Microbiol Lett 98:95–102

Moore R, McCarthy B (1969) Characterization of the deoxyribonucleic acid of various halophilic bacteria. J Bacteriol 99:248–254

Ng W-L, Yang C-F, Halladay JT, Arora P, DasSarma S (1995) Isolation of genomic and plasmid DNAs from *Halobacterium halobium*. In: Robb FT, Sowers KA, DasSarma S, Place AR, Fleischmann EM, Schreier HJ (eds) Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY, pp 179–184

Palmer JR, Daniels CJ (1994) A transcriptional reporter for in vivo promoter analysis in the archaeon *Haloferax volcanii*. J Bacteriol 60:3867–3869

Palmer JR, Daniels CJ (1995) In vivo definition of an archaeal promoter. J Bacteriol 177:1844–1849

Reeve JN, Sandman K, Daniels CJ (1997) Archaeal histones, nucleosomes, and transcription initiation. Cell 89:999–1002

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

Robb FT, Park J-B, Adams MWW (1992) Characterization of an extremely thermostable glutamate dehydrogenase: a key enzyme in the primary metabolism of the hyperthermophilic archaebacterium, *Pyrococcus furiosus*. Biochem Biophys Acta 1120:267–272

Robb FT, Place AR, Sowers KR, Schreier HJ, DasSarma S, Fleishmann EM (eds) (1995) Archaea: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY

Robinson KA, Schreier HJ (1994) Isolation, sequence, and characterization of the maltose-regulated *mlrA* gene from the hyperthermophilic archaeum *Pyrococcus furiosus*. Gene 151:173–176

Robinson KA, Robb FT, Schreier HJ (1994) Isolation of maltose-regulated genes from the hyperthermophilic archaeum *Pyrococcus furiosus* by subtractive hybridization. Gene 148:137–141

Sambrook J, Fritsch F, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

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

Thomm M, Sandmann K, Frey G, Koller G, Reeve J (1992) Transcription of the histone encoding *hmfB* from the hyperthermophilic archaeon *Methanothermus fervidus*. J Bacteriol 174:3508–3513

Wettach J, Gohl HP, Tschochner H, Thomm M (1995) Functional interaction of yeast and human TATA-binding proteins with an archaeal RNA polymerase and promoter. Proc Natl Acad Sci USA 92:472–476

Yang C-F, Kim J-M, Molinari E, DasSarma S (1996) Genetic and topological analyses of the bop promoter of *Halobacterium halobium*: stimulation by DNA supercoiling and non-B-DNA structure. J Bacteriol 178:840–845