

# Transcriptional Regulation in the Hyperthermophilic Archaeon *Pyrococcus furiosus*: Coordinated Expression of Divergently Oriented Genes in Response to $\beta$ -Linked Glucose Polymers

WILFRIED G. B. VOORHORST,<sup>1</sup> YANNICK GUEGUEN,<sup>1</sup> ANS C. M. GEERLING,<sup>1</sup> GERTI SCHUT,<sup>1</sup>  
ISABELL DAHLKE,<sup>2</sup> MICHAEL THOMM,<sup>2</sup> JOHN VAN DER OOST,<sup>1\*</sup> AND WILLEM M. DE VOS<sup>1</sup>

Laboratory of Microbiology, Department of Biomolecular Sciences, Wageningen Agricultural University,  
NL-6703 CT Wageningen, The Netherlands,<sup>1</sup> and Institut für Allgemeine Mikrobiologie  
der Christian-Albrechts-Universität zu Kiel, D-24118 Kiel, Germany<sup>2</sup>

Received 12 February 1999/Accepted 19 April 1999

**The genetic organization, expression, and regulation of the *celB* locus of the hyperthermophilic archaeon *Pyrococcus furiosus* were analyzed. This locus includes the *celB* gene, which codes for an intracellular  $\beta$ -glucosidase, and a divergently orientated gene cluster, *adhA-adhB-lamA*, which codes for two alcohol dehydrogenases and an extracellular  $\beta$ -1,3-endoglucanase that is transcribed as a polycistronic messenger (the *lamA* operon). During growth of *P. furiosus* on either the  $\beta$ -1,4-linked glucose dimer cellobiose or the  $\beta$ -1,3-linked glucose polymer laminarin, the activities of both  $\beta$ -glucosidase and endoglucanase were increased at least fivefold compared with levels during growth on maltose or pyruvate. Northern blot analysis revealed an enhanced transcription of both the *celB* gene and the *lamA* operon in the presence of these glucose-containing substrates. The *in vivo* and *in vitro* transcription initiation sites of both the *celB* gene and the *lamA* operon were identified 25 nucleotides downstream of conserved TATA box motifs. A number of repeating sequences have been recognized in the *celB-adhA* intergenic region, some of which might be part of a transcriptional regulator-binding site.**

The clustering of archaeal genes into operons as well as the size of the archaeal genome resembles the organization of the bacterial chromosome (1, 41, 47). However, the process of gene expression is quite different in the domains *Bacteria* and *Archaea*. Unlike that of bacteria, the transcription initiation machinery of archaea is very similar to that of eucarya (1, 2, 15, 26, 39). The archaeal RNA polymerase is structurally related to the RNA polymerases II of members of the domain *Eucarya* (26). In addition, archaeal transcription requires two eucaryal-like transcription factors, i.e., the TATA-binding protein (TBP) and transcription factor IIB (TFIIB; the archaeal homolog is called TFB) (11, 15, 16, 18, 45). The binding site of TBP in many archaeal promoters, like that of the eucaryal RNA polymerase II promoters, is a TATA box located approximately 25 to 30 nucleotides upstream of the transcription initiation site (14, 30, 31).

In contrast to the growing understanding of the basic machinery of transcription initiation in archaea, our understanding of the mechanisms of the regulation of this process is slight (2). The few studies that have addressed control of archaeal transcription have been limited to describing modulations of gene expression. Such effects have been reported for several hyperthermophilic crenarchaeotes and their viruses (47), as well as for some methanogens (5, 6, 28, 29). The molecular analysis of transcription regulation in hyperthermophilic euryarchaeota is limited to a report of two metabolic enzymes of *Pyrococcus furiosus* that appear to be upregulated at the transcriptional level by the  $\alpha$ -linked glucose disaccharide maltose (32). Recently, more detailed analyses of transcriptional activators from *Haloferax* spp. and *Methanobacterium thermoautotrophicum*, which are involved in synthesis of the gas vesicle

and the molybdenum-containing formylmethanofuran dehydrogenase, respectively, have been reported (17, 23).

Here we present a molecular analysis of transcription regulation in *P. furiosus*, which is able to grow on a wide range of substrates, such as proteins and carbohydrates (9, 13). Because of this apparent flexible metabolism, it was anticipated that the hydrolytic enzymes involved in polymer utilization are subjected to some form of regulation. In this study, we have analyzed the regulation of enzymes that are involved in the growth of *P. furiosus* on  $\beta$ -linked glucose polymers.

## MATERIALS AND METHODS

**Sequence analysis of the *celB* locus of *P. furiosus*.** To complete the nucleotide sequence of the *P. furiosus celB* locus (44), we sequenced the pyrococcal DNA in pLUW500 covering the 3.5-kb *Pst*I-*Bam*HI fragment, including the *adhA*, *adhB*, and *lamA* genes (Fig. 1). Nucleotide sequencing was carried out on a model 373A automated DNA sequencer (Applied Biosystems) with a Prism Ready Reaction Dye/Deoxy Terminator cycle sequencing kit or on a LiCor 4000L sequencer with a Thermo Sequenase fluorescence-labelled-primer cycle sequencing kit with 7-deaza-dGTP (Amersham) and infrared-labelled oligonucleotides (MWG-Biotech).

Computer analyses of nucleotide and deduced amino acid sequences were carried out with the PC/GENE program (version 5.01; IntelliGenetics Inc.) and the Genetics Computer Group package (version 7.0) at the CAOS/CAMM center of the University of Nijmegen (Nijmegen, The Netherlands) and with the multiple-sequence program CLUSTAL W (version 1.6) available on the internet (BCM Search Launcher; Human Genome Center, Baylor College of Medicine, Houston, Tex.).

**Growth and induction conditions.** *P. furiosus* DSM3638 was cultured at 98°C in synthetic seawater as previously described (20) with pyruvate (40 mM), cellobiose (10 mM), maltose (10 mM), or laminarin (2 g/liter) as the growth substrate. Cells were grown for several generations on these substrates to allow adaptation before any analysis was performed. Growth of *P. furiosus* was monitored by spectrophotometrically analyzing increases in optical density (OD) and by determining hydrogen production with a Packard gas chromatograph. Under these conditions late exponential growth phase corresponds with an OD at 450 nm (OD<sub>450</sub>) of 0.8 to 0.9. For the induction with cellobiose, cells were grown on pyruvate to an OD<sub>450</sub> of 0.7 to 0.8 and supplemented with cellobiose (10 mM). Samples were taken before and after induction and immediately cooled on ice prior to further processing.

**Preparation of cell extracts and enzyme activities.** Cells grown to late exponential growth phase were harvested by centrifugation, washed with fresh me-

\* Corresponding author. Mailing address: Laboratory of Microbiology, Wageningen Agricultural University, Hesselink van Suchtelenweg 4, 6703 CT Wageningen, The Netherlands. Phone: 31 317 483100. Fax: 31 317 483829. E-mail: john.vanderoost@algemeen.micr.wau.nl.

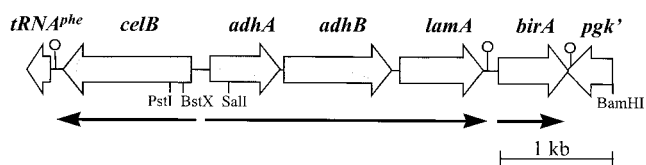


FIG. 1. Genetic organization of the *celB* locus. The locations and orientations of the genes (open and solid arrows) are indicated. Relevant restriction sites used for cloning and in vitro transcription analysis are shown, as are putative transcription termination signals ( $\nabla$ ).

dium, and subsequently resuspended in citrate buffer as described previously (44). Following sonication to disrupt the cells, the cell debris was pelleted and the resulting supernatant was used as the cell extract for activity measurements and immunological analysis. The  $\beta$ -glucosidase activity was determined by hydrolysis of  $\beta$ -D-glucopyranoside-*p*-nitrophenyl (Boehringer Mannheim GmbH, Mannheim, Germany) (20). Protein concentration was measured with Bradford reagents (Bio-Rad Laboratories), with bovine serum albumin as the standard.

**PAGE and Western blotting.** Electrophoretic analysis of protein samples was performed by sodium dodecyl sulfate–11% polyacrylamide gel electrophoresis (SDS-PAGE) (25). Protein samples for SDS-PAGE were prepared by heating them for 5 min at 100°C in an equal volume of sample buffer (0.1 M Tris-HCl, 5% SDS, 0.9% 2-mercaptoethanol, 20% glycerol [pH 6.8]). Proteins separated by SDS-PAGE were transferred to a nitrocellulose membrane by semidry blotting (Bio-Rad Laboratories). Immunological detection was performed as previously described with antiserum raised against the AdhA and LamA proteins purified from *Escherichia coli* (13).

**Isolation of total mRNA, Northern blot analysis, and primer extension.** Cells were harvested by centrifugation, and RNA was isolated with guanidinium isothiocyanate and  $\beta$ -mercaptoethanol as previously described (44). For Northern blot analysis, 15  $\mu$ g of RNA was separated on a formaldehyde–1% agarose gel. Following gel electrophoresis the RNA was transferred by capillary blotting to a Hybond-N<sup>+</sup> membrane (34). Gene-specific probes were obtained after appropriate digestion of DNA from plasmid pLUW500 or pLUW501 (44), purified with GeneClean (Bio 101, La Jolla, Calif.), and labelled by nick translation (34). During the hybridizations Southern blots were included to verify the specificities of the probes. Primer extension experiments on the isolated or synthesized RNA templates were done as previously described (16). The following oligonucleotides were used for the indicated templates: 5'-CCA AGA ATA TCC AAA CAT GAA G-3' (for *celB*) and 5'-GGC AAT CTT CTC TAA CCT ATC AAC-3' (for *adhA*).

**Cell-free transcription system of *P. furiosus*.** Cell-free transcription reactions with partially purified transcription factors and highly purified RNA polymerase (Superdex fraction) were essentially carried out as previously described (16). Cell-free transcription assays of *celB* and *adhA* were performed at 70°C at an optimal potassium chloride concentration of 300 mM (rather than 250 mM for transcription of *gdh*) [16].

**Nucleotide sequence accession number.** The nucleotide sequence of the *celB* locus (Fig. 1) has been submitted to the GenBank and EMBL data banks and given the accession no. AF013169.

## RESULTS

### Genetic organization of the *celB* gene and the *lamA* operon.

Analysis of the upstream flanking region of the previously isolated *P. furiosus*  $\beta$ -glucosidase-encoding *celB* gene (44) revealed a cluster of three genes. This gene cluster, hereinafter called the *lamA* operon (see below), encodes two alcohol dehydrogenases (the short-chain AdhA and the iron-dependent AdhB), as well as a  $\beta$ -1,3-endoglucanase (LamA, member of family 16 of the glycosyl hydrolases) (Fig. 1). The functional expression of these three enzymes in *E. coli* and their subsequent purification and characterization have been described elsewhere (13, 43). Downstream of the *lamA* operon is the *birA* gene (Fig. 1), the predicted translation product of which is homologous to the biotin ligase of *Bacillus subtilis* (46). Downstream of the *birA* gene, on the opposite strand, a (partial) gene has been identified (*pgk'*) (Fig. 1), the product of which has a high degree of similarity with 2-phosphoglycerate kinase from methanogenic archaea (4, 27) (Fig. 1).

The *celB* gene and the *lamA* operon are separated by a 175-bp intergenic region that is AT rich (71%) in comparison with the average AT content (62%) determined for the pyro-

coccal genome (7) (Fig. 1 and see below). The *adhA* and *adhB* genes and the *adhB* and *lamA* genes are separated by only 10 and 12 nucleotides, respectively. These intergenic sequences are purine rich and most likely function as ribosome-binding sites. Immediately downstream of the *lamA* gene a pyrimidine-rich stretch of residues (12 pyrimidines of 15 residues) resembling a typical archaeal termination sequence is located (3, 7, 40). A 41-bp intergenic region separates the *lamA* operon from the predicted initiation codon of the *birA* gene, and the 3' end of the *birA* gene overlaps by 12 nucleotides the *pgk'* gene, which is located in the opposite orientation. On the coding strands, both genes are followed by stretches of pyrimidine residues (mainly thymidines) with homology to archaeal termination sequences.

**Coregulation of CelB, AdhA, and LamA.** Cells of *P. furiosus* grown on cellobiose (glucose- $\beta$ -1,4-glucose) as the sole carbon and energy source were found to contain high levels of activity of intracellular, cellobiose-hydrolyzing  $\beta$ -glucosidase (CelB; 12 to 18 U/mg), confirming previous results (20). In contrast, in pyrococcal cells grown on pyruvate, 10-fold-lower CelB activity was detected (1.5 to 2.1 U/mg). The same *P. furiosus* cultures were used to detect the products of the *lamA* operon. Low but significant activities of the alcohol dehydrogenase AdhA and the endoglucanase LamA could be detected in cell extracts and in the medium of a cellobiose culture, respectively. Because no accurate activity measurements could be performed, the presence of AdhA and LamA was analyzed by immunological detection with specific antisera as well. Western blotting clearly indicated that the production of AdhA and LamA was stimulated when pyrococcal cells were grown on cellobiose or laminarin, unlike when cells were grown on maltose (glucose- $\alpha$ -1,4-glucose) or pyruvate (Fig. 2). These results indicate that *P. furiosus* enzymes encoded by the *lamA* operon and the *celB* gene are coregulated in response to the  $\beta$ -linked sugars cellobiose and laminarin.

For the identification of compounds that could induce CelB activity, cells of *P. furiosus* were grown on pyruvate, maltose, cellobiose, laminarin, or peptides and subsequently analyzed for the presence of CelB activity and CelB protein. High CelB activity was measured in cell extracts of cells grown on cellobiose or laminarin. In contrast, background levels of CelB were found in cells grown on pyruvate, maltose, or peptides. No

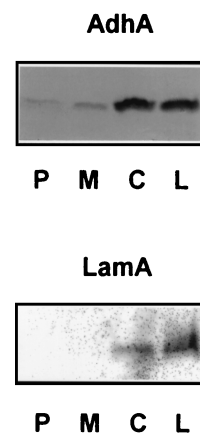


FIG. 2. Western blot analysis of AdhA and LamA of *P. furiosus* when it was grown on different substrates. Shown are results with cell extracts (AdhA) and concentrated culture supernatants (LamA) of *P. furiosus* cells after cells were grown on pyruvate (P), maltose (M), cellobiose (C), and laminarin (L), separated by SDS-PAGE, blotted on nitrocellulose membranes, and detected with polyclonal antibodies raised against AdhA and LamA, respectively.

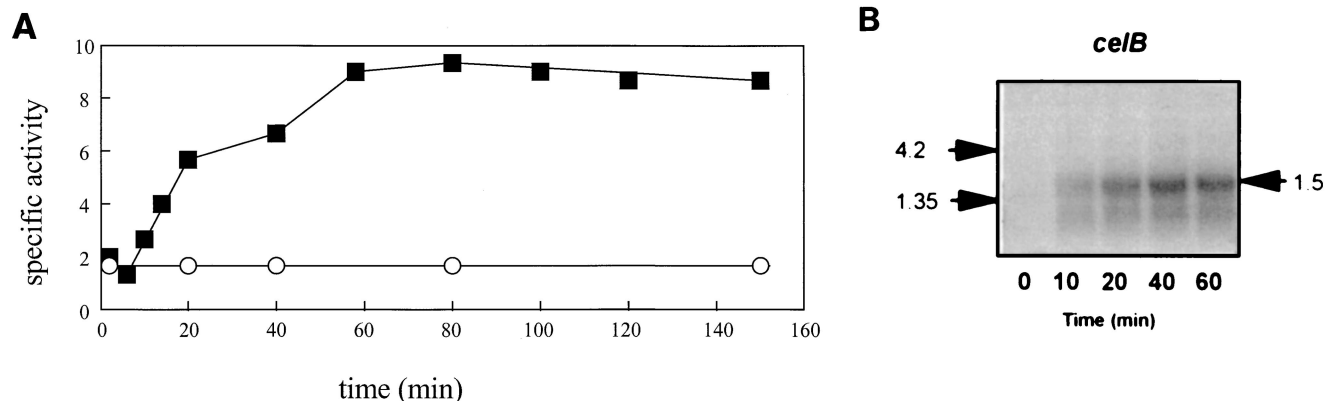


FIG. 3. Kinetics of the induction of  $\beta$ -glucosidase activity and *celB* transcription by cellobiose. (A) Specific  $\beta$ -glucosidase activity (in units per milligram) in *P. furiosus* cells grown on pyruvate with (■) or without (○) cellobiose. (B) Northern blot analysis of total RNA extracted from *P. furiosus* cells grown on pyruvate following addition of cellobiose. Hybridization was performed with a  $^{32}\text{P}$ -labelled probe specific for *celB*. A control experiment was performed with the constitutively expressed *P. furiosus por* gene, which encodes pyruvate ferredoxin oxidoreductase, indicating that equal amounts of RNA were applied in all lanes (data not shown and reference 42). The size of the *celB* transcript (1.5 kb) was estimated by using a RNA molecular size standard (Life Technologies) as indicated (4.2 or 1.35 kb).

induction was observed in pyruvate-grown *P. furiosus* cells upon addition of either glucose or isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG).

**Regulation of expression of the *celB* gene.** The effect of the  $\beta$ -linked glucose polymers on *celB* gene expression was studied in more detail in an experiment in which cellobiose was added to a *P. furiosus* culture growing on pyruvate at different stages of the growth curve. Specific  $\beta$ -glucosidase activity was found to increase most rapidly when cellobiose was added in the mid-exponential phase. After addition of cellobiose, the specific  $\beta$ -glucosidase activity rapidly increased within 1 h from the background level up to 9 U/mg (Fig. 3A).

To determine whether the observed induction of CelB takes place at the transcriptional level, Northern blot analysis was performed on samples from an induction experiment in which pyruvate-grown cells were supplemented with cellobiose at mid-exponential growth phase. A hybridizing band with the expected size of 1.5 kb was obtained with a *celB*-specific probe when RNA was extracted from induced cells but not from noninduced, pyruvate-grown cells (Fig. 3B). In the induced cells (with a doubling time of approximately 50 min), the *celB* transcript was detectable within 10 min after induction with cellobiose and the intensity reached a maximum after 20 min (Fig. 3B).

**Regulation of transcription of the *lamA* operon in *P. furiosus*.** The presence of CelB, AdhA, and LamA in *P. furiosus* cells grown on different substrates suggested a coregulation of the *celB* gene and the *lamA* operon. To analyze whether the transcription of the *lamA* operon is indeed induced by cellobiose, we performed Northern blot analysis with total RNA extracted from *P. furiosus* cells grown on pyruvate, cellobiose, or maltose. Northern blot hybridization with a probe specific for the *adhA* transcript resulted in a hybridizing band with RNA extracted from cells grown on cellobiose but not with RNA from cells grown on maltose or pyruvate (Fig. 4). Subsequent washing of the filter and rehybridization with a probe specific for *adhB* or *lamA* resulted in the same hybridization pattern as with the *adhA*-specific probe (Fig. 4). The hybridization signals were similar for all probes (2.5 to 3.0 kb), indicating that the gene cluster *adhA-adhB-lamA* is transcribed as a polycistronic messenger; the calculated size of the *lamA* operon, from the *adhA* transcription start site (see below) to the *lamA* stop codon, is 2,803 bp (Fig. 1). Although the same RNA was used, the transcript of the *lamA* operon appeared

less abundant and less discrete than the transcript of the *celB* gene, probably due to instability of the longer *lamA* messenger (see also below). A probe specific for the *birA* gene revealed a hybridizing band of approximately 0.7 kb on the Northern blot (not shown). It is concluded that transcription of the *lamA* operon appears to be terminated downstream of *lamA*, probably at the aforementioned thymidine-rich region.

**Transcription initiation.** In addition to the previously determined *in vivo* transcription start site of *celB* (44), the transcription initiation site of the adjacent *lamA* operon was identified by primer extension on total RNA extracted from cellobiose-grown *P. furiosus*. The transcription start site of the *lamA* transcript could be identified 10 nucleotides upstream of the *adhA* translation start site (ATG) (Fig. 5A). At a distance of 25 nucleotides, the transcription start site was preceded by a hexanucleotide sequence that resembles the archaeal TATA box sequence: ATTATA (Fig. 6) (3, 7, 40).

In a cell-free transcription system for *Pyrococcus* that consists of TBP, TFB, and RNA polymerase (15), efficient transcription initiation from the promoter of the *P. furiosus* glutamate dehydrogenase gene (*gdh*) has been accomplished (8, 16). This system has been applied to analyze transcription initiation of the divergently orientated *lamA* operon (Fig. 5A) and *celB* gene (Fig. 5B). Primer extension with the *in vitro*-produced transcripts of the *celB* gene and the *lamA* operon showed that

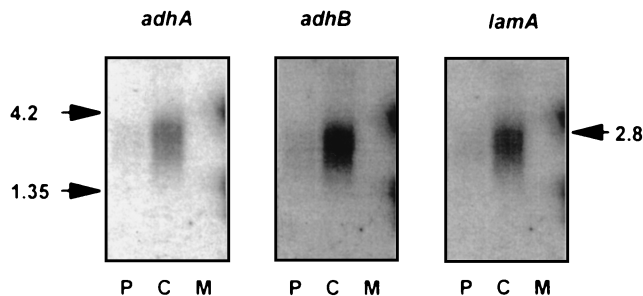


FIG. 4. Detection of the transcript of the *lamA* operon in RNAs extracted from cells grown on pyruvate (P), cellobiose (C), and maltose (M). Equal amounts of RNA were used for agarose gel electrophoresis. The Northern blots were hybridized with  $^{32}\text{P}$ -labelled gene-specific probes as indicated. The size of the transcript (2.8 kb) was estimated by using a RNA molecular size standard (Life Technologies) as indicated (4.2 or 1.35 kb).

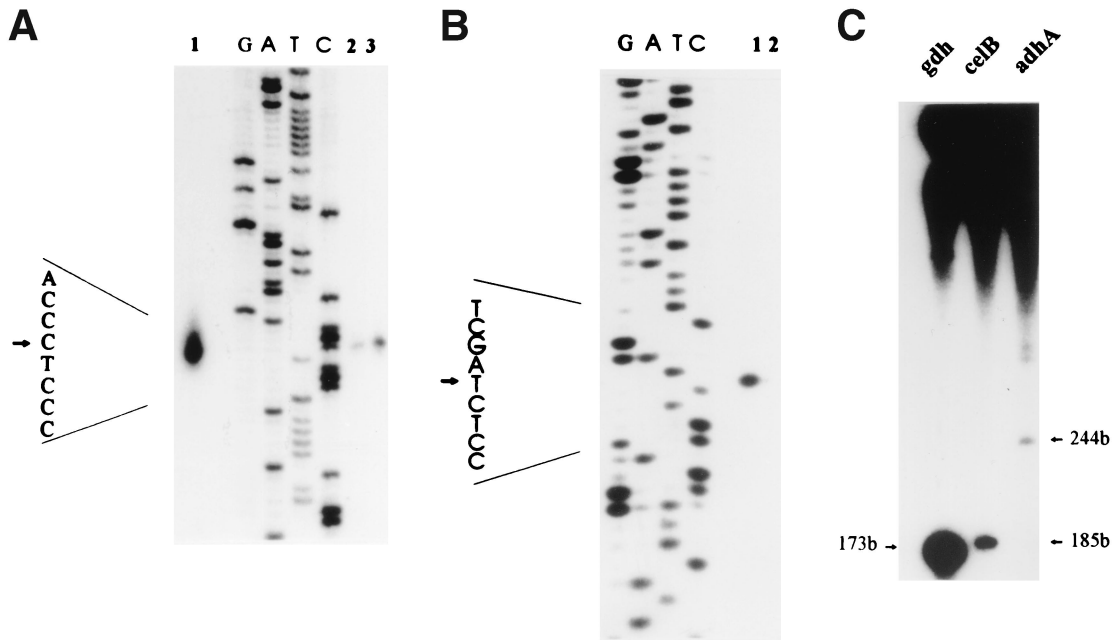


FIG. 5. In vivo and in vitro transcriptional analyses of the *celB* locus. (A) Analysis of the transcription initiation site at the *adhA* promoter by primer extension. Primer-extended products of transcripts produced in vivo (lane 1) and in vitro (lane 2), together with the sequence ladder (lanes G, A, T, and C) generated with the same primer on the noncoding strand of *adhA*, were loaded and separated on polyacrylamide gel. Annealing reactions were performed at 60°C (lane 2) and 50°C (lane 3). (B) In vitro analysis of the transcription initiation site at the *celB* promoter by primer extension. Primer-extended products were loaded next to the sequence ladder (lanes G, A, T, and C) generated with the same primer on the noncoding strand of *celB*. Annealing reactions were performed at 50°C (lane 1) and 60°C (lane 2). (C) Comparison of the in vitro transcription efficiencies of the *gdh*, *celB*, and *adhA* genes. RNA polymerase, the Superdex fraction of a TFB, and the heparin-Sepharose fraction of a TFA were assayed for specific activity in cell-free transcription system reactions as described in the work of Hethke et al. (16). Labeled runoff transcripts of the *gdh* (16) and, after optimization, of the *celB* and *adhA* genes were separated on an 8% polyacrylamide-urea gel and identified by autoradiography. Only one-fifth of the product generated with the *gdh* template was loaded. The arrows label the runoff transcripts and show their lengths in bases.

initiation in vitro occurred at the same site as that determined for initiation in vivo (Fig. 5B) (44). To compare the in vitro activities of these regulated promoters with that of the strong constitutively expressed *gdh* promoter, RNA products transcribed from linearized templates were analyzed. When plasmid pLUW504 (44) was cleaved in the *celB* coding region with *Bst*XI and in the *adhA* coding region with *Sal*I (Fig. 1), the

expected runoff transcripts of 185 and 244 nucleotides, respectively, were synthesized (Fig. 5C, lanes 2 and 3). Quantitation of the runoff products revealed that the activity of the *celB* promoter was only 1% and that the activity of the *adhA* promoter was only 0.2% of that of the *gdh* promoter obtained with the template pLUW479 (Fig. 5C, lane 1) (*gdh* runoff transcript, 173 nucleotides [16]).

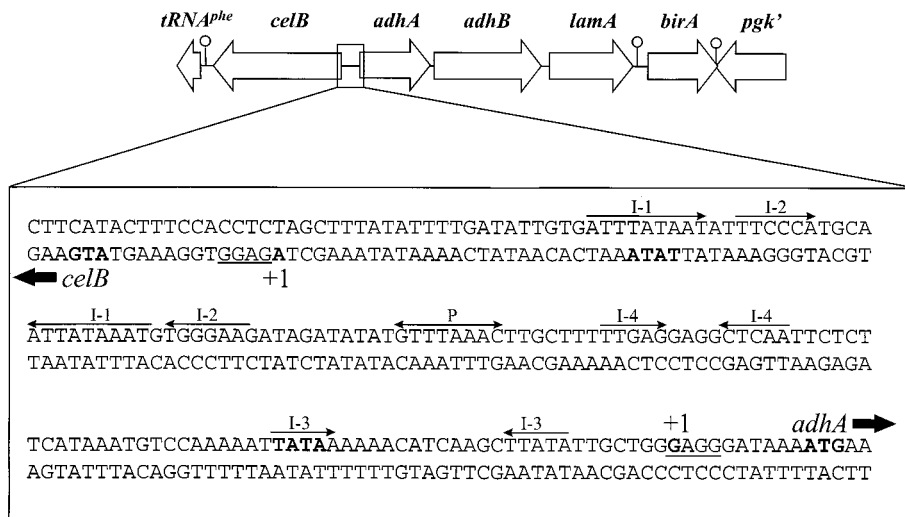


FIG. 6. Intergenic region between *celB* and *adhA*. The TATA box, transcription initiation sites (+1), and translation start sites (boldface letters and gray shading) are indicated; putative ribosome-binding sites are underlined. Many repeating sequences are present; for clarity, only inverted repeats (I-1 to I-4) and the palindromic sequence (P) are shown; inverted repeat 4 resembles a recognition site for a bacterial transcription regulator (FNR), as discussed in the text.

## DISCUSSION

A molecular analysis of the substrate-dependent transcriptional regulation of four genes that constitute the *celB* locus in the hyperthermophilic archaeon *Pyrococcus furiosus* is presented. In an in vitro experiment, we recently demonstrated that two glycosyl hydrolases, the *celB*-encoded  $\beta$ -glucosidase and the *lamA*-encoded  $\beta$ -1,3-endoglucanase LamA, degrade the laminarin polysaccharide completely to glucose in a synergistic manner (13). The physiological model is that extracellular LamA is involved in the partial hydrolysis of laminarin to oligomers, which are imported via an unknown transporter and subsequently degraded to glucose by intracellular CelB.

In this study, we analyzed substrate-dependent fluctuations of CelB and LamA activities. Pyrococcal cultures grown on pyruvate or maltose have background levels of CelB activity, and LamA is not detectable by Western blot analysis. In contrast, at least fivefold-higher activities of CelB and higher amounts of LamA were detected in cultures grown on cellobiose or laminarin than in cultures grown on maltose, pyruvate, or peptides (Fig. 2 and 3). Induction studies showed a rapid increase of CelB activity levels upon addition of the  $\beta$ -linked glucose polymer cellobiose or laminarin to a culture grown on pyruvate (Fig. 3). Northern blot analysis revealed that the observed control occurs at the transcriptional level. Activation of *celB* transcription, within 10 min, resulted in the observed rapid induction response of CelB activity (Fig. 3B). It is concluded that the *celB* gene is transcribed as a monocistronic mRNA, probably indicating the functionality of a typical T(C)-rich archaeal termination sequence immediately downstream of its coding region (CCATTTCATTTTTTCTTTGTTTTTT [44]). Likewise, the *lamA* transcript could be detected only in total RNA extracted from cells grown on cellobiose or laminarin, indicating coregulation of the divergently orientated transcripts. The effector is probably a  $\beta$ -linked glucose disaccharide (cellobiose or laminaribiose) or some unknown derivative (see below).

Analysis of the activity as well as immunological detection indicates coregulation of AdhA and CelB or LamA (Fig. 2). Moreover, Northern blot analysis showed that the *lamA* operon is transcribed as a polycistronic mRNA (Fig. 4). Unlike with LamA and CelB, however, it is less obvious to imagine a physiological link between the additional gene products, the short-chain alcohol dehydrogenase AdhA and the iron-dependent alcohol dehydrogenase AdhB, and laminarin or cellobiose metabolism. A putative function may be the reduction of aldehydes to alcohol, as an electron sink during sugar fermentation; however, only traces of alcohols (ethanol) are detectable among the fermentation products after cultivation of *P. furiosus* on maltose and cellobiose (19a, 21). Alternatively, the dehydrogenases may be involved in the oxidative or reductive conversion of saccharides to glucose, which is most likely required for the processing via the  $\beta$ -glucosidase and, subsequently, the glycolytic enzymes. In an experiment with a range of primary and secondary alcohols (ethanol to decanol), the purified AdhA oxidized 2-pentanol or 2-hexanol with the highest activity (40 U/mg) (42); sugars such as cellobiose were oxidized but at lower rates (2 to 3 U/mg) (44a). Hence, these preliminary analyses do not rule out the possibility that oxidized or reduced polyols (saccharides) are the physiological substrates for AdhA. However, the physiological implications of the two dehydrogenases being coexpressed with the glycosyl hydrolases will probably remain unclear until the actual physiological substrate(s) for the LamA-CelB system is elucidated.

In vitro initiation of transcription of the *celB* gene and the *lamA* operon occurred at the same sites as those identified in

vivo (Fig. 5A and B) (44). The efficiencies of transcription initiation of the *celB* and *lamA* transcripts in the cell-free system were significantly lower than that of the *P. furiosus* *gdh* gene (Fig. 5C) (16). This suggests suboptimal conditions of the in vitro experiment which may have been caused by, for example, the secondary structure of the intergenic promoter region, which contains a number of repeats (see below and Fig. 6). One possibility is that these lower efficiencies of transcription initiation reflect the in vivo situation, implying the requirement of a transcription activator to counteract the putative obstruction; alternatively, it cannot be ruled out that the impeded efficiency is due to an artifact of the described in vitro transcription experiment with these particular DNA fragments.

Both of the promoters of the *celB* gene and the *lamA* operon contain a TATA box sequence (ATTATA) that closely resembles the *P. furiosus* consensus [(T/A)TTATA] (41), and both TATA boxes are located 25 nucleotides upstream of the promoters' transcription initiation sites (Fig. 6). These promoter elements, equivalent to the eucaryal TATA box, have been shown to interact with the TBP and are involved in directing transcription (15, 33). Based on DNase I footprinting experiments with the *P. furiosus* *gdh* promoter, it has recently been demonstrated that TBP results in a footprint between positions -20 and -34, centered around the TATA box. In addition, TFB bound cooperatively to TBP generates an extended footprinting pattern ranging from positions -19 to -42 (15). The sequences surrounding the TATA box of both the *celB* gene and the *lamA* operon showed a relatively high degree of conservation (Fig. 6), possibly indicating a role in modulating transcription efficiency.

The AT-rich *celB-adhA* intergenic region was further analyzed for potential *cis*-acting elements. Interestingly, we found a palindromic sequence (GTTTAAAC) located exactly in between the two TATA elements. Moreover, four inverted repeats of at least 5 residues were detected (Fig. 6). Inverted repeat 4 (**TTGAGXXXXCTCAA**) has a striking similarity (indicated in bold) to the consensus binding site for the FNR (fumarate nitrate regulation) family of bacterial transcription regulators (**TTGATXXXXATCAA**) (37). The location of this putative binding site (54 bp from the *adhA* and 86 bp from the *celB* transcription initiation sites) is in agreement with that found for transcription regulators in bacterial systems (12). As with the *E. coli* *lac* operon, potential inducers of the anticipated transcription regulator might be either a substrate (or a derivative of the substrate), namely, cellobiose, laminaribiose, or a reporter molecule such as cAMP. In the last case, the regulator might resemble CRP (cAMP receptor protein), a regulator that is structurally related to FNR (consensus binding site, **GTTGAXXXXXTCAAC** [37]). However, no archaeal FNR or CRP homologs have yet been reported. Hence, the significance of the identified repeats with respect to transcriptional regulation remains to be established.

The fast-growing sequence databases contain a large number of genes that code for archaeal homologs of bacterial transcription regulators (4, 10, 22, 24, 35, 36). In addition, a number of palindromic binding sites of transcriptional repressors have been reported to be functional in halophilic and methanogenic archaea, although the corresponding repressors have not yet been identified (19, 38). To the best of our knowledge, the actions of only three archaeal transcription regulators have recently been characterized in considerable detail: (i) the transcription activation of gas vesicle synthesis in *Haloferax* spp. by a eucaryal-like leucine-zipper-type regulator (23), (ii) the molybdenum-activated expression of formylmethanofuran dehydrogenase in *M. thermoautotrophicum* (17), and (iii) the negative autoregulation by a homolog of the bacterial leucine-

responsive regulatory protein (Lrp) in *P. furiosus* (2a). Future studies will aim at elucidating the molecular details of the transcriptional regulation described herein, specifically, the actual nucleotide sequence of the regulator's binding site and the identification of this transcriptional regulator.

#### ACKNOWLEDGMENTS

Part of this work was supported by contracts BIOT-CT93-0274 and BIOT-CT96-0488 of the European Union.

#### REFERENCES

- Baumann, P., S. A. Qureshi, and S. P. Jackson. 1995. Transcription: new insights from studies on Archaea. *Trends Genet.* **11**:279–283.
- Bell, S. D., and S. P. Jackson. 1998. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. *Trends Microbiol.* **6**:222–228.
- Brinkman, A. B., I. Dahlke, J. E. Tuininga, T. Lammers, V. Dumay, E. de Heus, J. H. G. Lebbink, M. Thomm, W. M. de Vos, and J. van der Oost. Submitted for publication.
- Brown, J. W., C. J. Daniels, and J. N. Reeve. 1989. Gene structure, organization, and expression in archaeobacteria. *Crit. Rev. Microbiol.* **16**:287–338.
- Bult, C. J., O. White, G. J. Olsen, L. X. Zhou, R. D. Fleischmann, G. G. Sutton, J. A. Blake, L. M. FitzGerald, R. A. Clayton, J. D. Gocayne, A. R. Kerlavage, B. A. Dougherty, J. F. Tomb, M. D. Adams, C. I. Reich, R. Overbeek, E. F. Kirkness, K. G. Weinstock, J. M. Merrick, A. Glodek, J. L. Scott, N. S. M. Geoghagen, J. F. Weidman, J. L. Fuhrmann, D. Nguyen, T. R. Utterback, J. M. Kelley, J. D. Peterson, P. W. Sadow, M. C. Hanna, M. D. Cotton, K. M. Roberts, M. A. Hurst, B. P. Kaine, M. Borodovsky, H. P. Klenk, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* **273**:1058–1073.
- Cohen-Kupiec, R., C. Blank, and J. A. Leigh. 1997. Transcriptional regulation in Archaea: *in vivo* demonstration of a repressor binding site in a methanogen. *Proc. Natl. Acad. Sci. USA* **94**:1316–1320.
- Cohen-Kupiec, R., C. J. Marx, and J. A. Leigh. 1999. Function and regulation of *glnA* in the methanogenic archaeon *Methanococcus maripaludis*. *J. Bacteriol.* **181**:256–261.
- Dalgaard, J. Z., and R. A. Garrett. 1993. Archaeal hyperthermophile genes, p. 535–563. In M. Kates, D. J. Kushner, and A. T. Matheson (ed.), *The biochemistry of Archaea (Archaeobacteria)*. Elsevier, Amsterdam, The Netherlands.
- Eggen, R. I. L., A. C. M. Geerling, K. Waldkoetter, G. Antranikian, and W. M. De Vos. 1993. The glutamate dehydrogenase-encoding gene of the hyperthermophilic archaeon *Pyrococcus furiosus*: Sequence, transcription and analysis of the deduced amino acid sequence. *Gene* **132**:143–148.
- Fiala, G., and K. O. Stetter. 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaeobacteria growing optimally at 100°C. *Arch. Microbiol.* **161**:168–175.
- Fitz-Gibbon, S., A. J. Choi, J. H. Miller, K. O. Stetter, M. I. Simon, R. Swanson, and U. J. Kim. 1997. A fosmid-based genomic map and identification of 474 genes of the hyperthermophilic archaeon *Pyrobaculum aerophilum*. *Extremophiles* **1**:36–51.
- Frey, G., M. Thomm, B. Brudigam, H. P. Gohl, and W. Hausner. 1990. An archaeobacterial cell-free transcription system. The expression of tRNA genes from *Methanococcus vannielii* is mediated by transcription factor. *Nucleic Acids Res.* **18**:1361–1367.
- Gralla, J. D., and J. Collado-Vides. 1996. Organization and function of transcription regulatory elements, p. 1232–1262. In F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli and Salmonella: cellular and molecular biology*, 2nd ed. ASM Press, Washington, D.C.
- Gueguen, Y., W. G. B. Voorhorst, J. van der Oost, and W. M. de Vos. 1997. Molecular and biochemical characterization of an endo-beta-1,3-glucanase of the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Biol. Chem.* **272**:31258–31264.
- Hausner, W., G. Frey, and M. Thomm. 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.
- Hausner, W., J. Wettach, C. Hethke, and M. Thomm. 1996. Two transcription factors related with the eucaryal transcription factors TATA-binding protein and transcription factor IIB direct promoter recognition by an archaeal RNA polymerase. *J. Biol. Chem.* **271**:30144–30148.
- Hethke, C., A. C. M. Geerling, W. Hausner, W. M. de Vos, and M. Thomm. 1996. A cell-free transcription system for the hyperthermophilic archaeon *Pyrococcus furiosus*. *Nucleic Acids Res.* **24**:2369–2376.
- Hochheimer, A., R. Hedderich, and R. K. Thauer. 1999. The DNA binding protein Tfx from *Methanobacterium thermoautotrophicum*: structure, DNA binding properties and transcriptional regulation. *Mol. Microbiol.* **31**:641–650.
- Hüdepohl, U., W. D. Reiter, and W. Zillig. 1990. *In vitro* transcription of two rRNA genes of the archaeobacterium *Sulfolobus* sp. B 12 indicates a factor requirement for specific initiation. *Proc. Natl. Acad. Sci. USA* **87**:5851–5855.
- Ken, R., and N. R. Hackett. 1991. *Halobacterium halobium* strains lysogenic from phage ΦH contain a protein resembling coliphage repressors. *J. Bacteriol.* **173**:955–960.
- Kengen, S. W. Personal communication.
- Kengen, S. W., E. J. Luesink, A. J. Stams, and A. J. Zehnder. 1993. Purification and characterization of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur. J. Biochem.* **213**:305–312.
- Kengen, S. W. M., and A. J. M. Stams. 1994. Formation of L-alanine as a reduced end product in carbohydrate fermentation by the hyperthermophilic archaeon *Pyrococcus furiosus*. *Arch. Microbiol.* **161**:168–175.
- Klenk, H. P., R. A. Clayton, J. F. Tomb, O. White, K. E. Nelson, K. A. Ketchum, R. J. Dodson, M. Gwinn, E. K. Hickey, J. D. Peterson, D. L. Richardson, A. R. Kerlavage, D. E. Graham, N. C. Kyrpides, R. D. Fleischmann, J. Quackenbush, N. H. Lee, G. G. Sutton, S. Gill, E. F. Kirkness, B. A. Dougherty, K. McKenney, M. D. Adams, B. Loftus, S. Peterson, C. I. Reich, L. K. McNeil, J. H. Badger, A. Glodek, L. X. Zhou, R. Overbeek, J. D. Gocayne, J. F. Weidman, L. McDonald, T. Utterback, M. D. Cotton, T. Spriggs, P. Artiach, B. P. Kaine, S. M. Sykes, P. W. Sadow, K. P. d'Andrea, C. Bowman, C. Fujii, S. A. Garland, T. M. Mason, G. J. Olsen, C. M. Fraser, H. O. Smith, C. R. Woese, and J. C. Venter. 1997. The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**:364–370.
- Krüger, K., T. Hermann, V. Armbruster, and F. Pfeifer. 1998. The transcriptional activator GvpE for the halobacterial gas vesicle genes resembles a basic region leucine-zipper regulatory protein. *J. Mol. Biol.* **279**:761–771.
- Kyrpides, N. C., and C. A. Ouzounis. 1995. The eubacterial transcriptional activator Lrp is present in the archaeon *Pyrococcus furiosus*. *Trends Biochem. Sci.* **20**:140–141.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–685.
- Langer, D., J. Hain, P. Thuriaux, and W. Zillig. 1995. Transcription in Archaea: similarity to that in Eucarya. *Proc. Natl. Acad. Sci. USA* **92**:5768–5772.
- Lehmacher, A., and R. Hensel. 1994. Cloning, sequencing and expression of the gene encoding 2-phosphoglycerate kinase from *Methanothermobacter ferredoxinus*. *Mol. Gen. Genet.* **242**:163–168.
- Morgan, R. M., T. D. Pihl, J. Nölling, and J. N. Reeve. 1997. Hydrogen regulation of growth, growth yields, and methane gene transcription in *Methanobacterium thermoautotrophicum* ΔH. *J. Bacteriol.* **179**:889–898.
- Nölling, J., and J. N. Reeve. 1997. Growth- and substrate-dependent transcription of the formate dehydrogenase (*fdhCAB*) operon in *Methanobacterium thermoformicum* Z-245. *J. Bacteriol.* **179**:899–908.
- Palmer, J. R., and C. J. Daniels. 1995. *In vivo* definition of an archaeal promoter. *J. Bacteriol.* **177**:1844–1849.
- Reiter, W. D., U. Hüdepohl, and W. Zillig. 1990. Mutational analysis of an archaeobacterial promoter: essential role of a TATA box transcription efficiency and start-site selection. *Proc. Natl. Acad. Sci. USA* **87**:9509–9513.
- Robinson, K. A., and H. J. Schreier. 1994. Isolation, sequence and characterization of the maltose-regulated *mhrA* gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Gene* **151**:173–176.
- Rowlands, T., P. Baumann, and S. P. Jackson. 1994. The TATA-binding protein: a general transcription factor in eukaryotes and archaeobacteria. *Science* **264**:1326–1329.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Sensen, C. W., H. P. Klenk, R. K. Singh, G. Allard, C. C. Chan, Q. Y. Liu, S. L. Penny, F. Young, M. E. Schenk, T. Gaasterland, W. F. Doolittle, M. A. Ragan, and R. L. Charlebois. 1996. Organizational characteristics and informational content of an archaeal genome: 156 kb of sequence from *Sulfolobus solfataricus* P2. *Mol. Microbiol.* **22**:175–191.
- Smith, D. R., L. A. Doucette-Stamm, C. Deloughery, H. Lee, J. Dubois, T. Aldredge, R. Bashirzadeh, D. Blakely, R. Cook, K. Gilbert, D. Harrison, L. Hoang, P. Keagle, W. Lumm, B. Pothier, D. Qiu, R. Spadafora, R. Vicaire, Y. Wang, J. Wierzbowski, R. Gibson, N. Jiwani, A. Caruso, D. Bush, H. Safer, D. Patwell, S. Prabhakar, S. McDougall, G. Shimer, A. Goyal, S. Pietrovski, G. M. Church, C. J. Daniels, J. Mao, P. Rice, J. Nölling, and J. N. Reeve. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: functional analysis and comparative genomics. *J. Bacteriol.* **179**:7135–7155.
- Spiro, S., and J. R. Guest. 1990. Adaptive responses to oxygen-limitation in *Escherichia coli*. *Trends Biochem.* **16**:310–314.
- Stolt, P., and W. Zillig. 1992. *In vivo* studies on the effects of immunity genes on early lytic transcription in the *Halobacterium salinarum* phage Φ H. *Mol. Gen. Genet.* **235**:197–204.
- Thomm, M. 1996. Archaeal transcription factors and their role in transcription initiation. *FEMS Microbiol. Rev.* **18**:159–171.
- Thomm, M., W. Hausner, and C. Hethke. 1994. Transcription factors and

- termination of transcription in *Methanococcus*. Syst. Appl. Microbiol. **16**: 648–655.
41. **van der Oost, J., M. Ciaramella, M. Moracci, F. M. Pisani, M. Rossi, and W. M. de Vos.** 1998. Molecular biology of hyperthermophilic Archaea. Adv. Biochem. Eng. Biotechnol. **61**:87–115.
42. **van der Oost, J., G. Schut, S. W. M. Kengen, W. R. Hagen, M. Thomm, and W. M. de Vos.** 1998. The ferredoxin-dependent conversion of glyceraldehyde-3-phosphate in the hyperthermophilic archaeon *Pyrococcus furiosus* represents a novel site of glycolytic regulation. J. Biol. Chem. **273**:28149–28154.
43. **Voorhorst, W. G. B.** 1998. Molecular characterization of hydrolytic enzymes from hyperthermophilic Archaea. Ph.D. thesis. Wageningen Agricultural University, Wageningen, The Netherlands.
44. **Voorhorst, W. G. B., R. I. L. Eggen, E. J. Luesink, and W. M. de Vos.** 1995. Characterization of the *celB* gene coding for beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus* and its expression and site-directed mutation in *Escherichia coli*. J. Bacteriol. **177**:7105–7111.
- 44a. **Voorhorst, W. G. B., and J. van der Oost.** Unpublished data.
45. **Wettach, J., H. P. Gohl, H. Tschochner, and M. Thomm.** 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.
46. **Wilson, K. P., L. M. Shewchuk, R. G. Brennan, A. J. Otsuka, and B. W. Matthews.** 1992. *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. Proc. Natl. Acad. Sci. USA **89**:9257–9261.
47. **Zillig, W., P. Palm, H. P. Klenk, D. Langer, U. Hüdepohl, J. Hain, M. Lanzendorfer, and I. Holz.** 1993. Transcription in Archaea, p. 367–391. In M. Kates, D. J. Kushner, and A. T. Matheson (ed.), The biochemistry of Archaea (Archaeobacteria). Elsevier, Amsterdam, The Netherlands.