

Characterization of Two Heat Shock Genes from *Haloferax volcanii*: a Model System for Transcription Regulation in the *Archaea*

YEN-PING KUO,¹ DOROTHEA K. THOMPSON,¹ ANDREW ST. JEAN,² ROBERT L. CHARLEBOIS,²
AND CHARLES J. DANIELS^{1*}

Department of Microbiology, The Ohio State University, Columbus, Ohio 43210,¹ and Department of Biology, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada²

Received 1 April 1997/Accepted 7 August 1997

The expression of two heat-responsive *cct* (chaperonin-containing Tcp-1) genes from the archaeon *Haloferax volcanii* was investigated at the transcription level. The *cct1* and *cct2* genes, which encode proteins of 560 and 557 amino acids, respectively, were identified on cosmid clones of an *H. volcanii* genomic library and subsequently sequenced. The deduced amino acid sequences of these genes exhibited a high degree of similarity to other archaeal and eucaryal *cct* family members. Expression of the *cct* genes was characterized in detail for the purpose of developing a model for studying transcription regulation in the domain *Archaea*. Northern (RNA) analysis demonstrated that the *cct* mRNAs were maximally induced after heat shock from 37 to 55°C and showed significant heat inducibility after 30 min at 60°C. Transcription of *cct* mRNAs was also stimulated in response to dilute salt concentrations. Transcriptional analysis of *cct* promoter regions coupled to a yeast tRNA reporter gene demonstrated that 5' flanking sequences up to position –233 (*cct1*) and position –170 (*cct2*) were sufficient for promoting heat-induced transcription. Transcript analysis indicated that both basal transcription and stress-induced transcription of the *H. volcanii* *cct* genes were directed by a conserved archaeal consensus TATA motif (5'-TTTATA-3') centered at –25 relative to the mapped initiation site. Comparison of the *cct* promoter regions also revealed a striking degree of sequence conservation immediately 5' and 3' of the TATA element.

All living organisms adapt to adverse environmental conditions by evolving specific molecular responses. In particular, the universally conserved heat shock response occurs when cells are exposed to elevated temperatures, resulting in the rapid and transient overproduction of a limited class of proteins called the heat shock proteins (HSPs). The HSPs produced in the domains *Bacteria* and *Eucarya* are highly conserved both in structure and function, and their induction is generally regulated at the transcription initiation level (recently reviewed in reference 31).

Among the HSPs, the Cct family is a recently identified group of molecular chaperonins that are distinct from members of the well-studied Hsp60/GroEL family. So far, members of the Cct family have been identified only in the domains *Archaea* and *Eucarya*, where they are known variously as T-complex polypeptide, or TCP (15); chaperonin-containing Tcp-1, or Cct (25, 62); TCP-1 ring complex, or TriC (13); thermosome complex (59); and thermophilic factor-55, or TF55 (54). Protein members of the Cct and Hsp60/GroEL families display similar double-ringed toroidal structures as their active forms. Differences between the two families include the absence of an HSP10/GroES-like accessory protein necessary for Cct function (25) and the hetero-oligomeric structure of Cct complexes, which contain either 2 (58) or 8 to 10 different subunits (61). Among the members of the *Archaea*, members of the Cct protein family have been documented in many hyperthermophiles, including *Sulfolobus shibatae* (18,

54), *Sulfolobus solfataricus* (24, 32), *Pyrodicticum occultum* (42, 43), *Desulfurococcus* sp. strain SY (19), and *Thermoplasma acidophilum* (58–60) and in the methanogen *Methanopyrus kandleri* (1). The *S. shibatae* TF55 complex was the first archaeal Cct family member to be reported (54) and has been the most intensively studied to date (for a review, see reference 53). The wide distribution of these proteins among the *Archaea*, as well as the observation that expression of these proteins is regulated at the transcription level (53), make these genes an attractive model system for examining the mechanisms of gene regulation in the domain *Archaea*.

Compared to the domains *Bacteria* and *Eucarya*, very little is known about the molecular mechanisms controlling gene expression in the domain *Archaea*. What is known has been largely inferred from the presence of promoter sequences and transcription machinery components homologous to the eucaryal counterparts. These include an RNA polymerase II (pol II) TATA-like sequence within the archaeal core promoter element (47, 52), a complex RNA polymerase containing 8 to 12 subunits exhibiting sequence similarity and immunological cross-reactivity with subunits of the eucaryal RNA polymerases (23, 63), a TATA-binding protein homolog (33, 38, 45, 48), and transcription factors TFIIB (10, 38, 46) and TFIIS (20, 28). While such similarities suggest that the archaeal and eucaryal transcription systems are fundamentally related, they say very little about the underlying mechanisms responsible for gene regulation in members of the domain *Archaea*. To address this problem, we chose to examine the heat shock response as a potential model system for studying gene regulation in the halophilic archaeon *Haloferax volcanii*. This work was facilitated by the availability of a complete set of overlapping cosmid clones representing 96% of the *H. volcanii* DS2 genome

* Corresponding author. Mailing address: Department of Microbiology, The Ohio State University, 484 West 12th Ave., Columbus OH 43210. Phone: (614) 292-4599. Fax: (614) 292-8120. E-mail: daniels.7@osu.edu.

(4) as well as a preliminary study reporting that several of these cosmids carried heat-responsive genes (55).

Here, we report the cloning, sequencing, and transcript analysis of two *cct* family member genes in halophilic archaea. Using an in vivo transcription reporter assay, we also show that 5' flanking sequences up to position -233 (*cct1*) and position -170 (*cct2*) are sufficient for promoting normal heat-induced transcription.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. *Escherichia coli* ED8767 cells carrying cosmids A199 and 268 were grown in Luria-Bertani (LB) medium supplemented with kanamycin (30 µg per ml). *E. coli* strains DH5α-F', XL-1 Blue MRF', and JM110 were routinely cultured in LB medium supplemented with ampicillin (100 µg per ml) during propagation of pWL-based plasmids (26) or ampicillin (100 µg per ml), X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (0.002%), and IPTG (isopropyl-β-D-thiogalactopyranoside) (40 µM) during selection for pUC18 and pUC19 plasmids. Construction of the promoter-reporter expression module in the *H. volcanii*-*E. coli* pWL201 shuttle plasmid has been previously described (36). *Haloferax volcanii* WFD11 (3) was grown in complex medium as previously described (35), which was supplemented with 20 µM lovastatin (a gift from Merck and Co., Inc.), when necessary, to ensure maintenance of the pWL-based expression plasmid.

Cloning, restriction analysis, and DNA sequencing. Cosmid A199, which contains 34 kbp of the *H. volcanii* genome, was previously shown to carry a heat-responsive locus (55). Using Northern blotting, we determined that two *Mlu*I fragments (4.0 kb and 2.3 kb in size) of cosmid A199 hybridized to a 1.68-kb heat shock-specific mRNA. These *Mlu*I fragments were purified from agarose gel slices with the Prep-A-Gene DNA purification system (Bio-Rad), blunt ended with the Klenow fragment (Gibco BRL) by a standard protocol (49), and cloned into the *Sma*I site of pUC19. The two pUC19 clones, designated HS5 (2.3-kb fragment) and HS21 (4.0-kb fragment), were analyzed by restriction digestion, and overlapping subclones were constructed for DNA sequence analysis. DNA sequencing was carried out by the dideoxy chain termination method (50) with Sequenase T7 polymerase (U.S. Biochemicals) and the 7-deaza-dGTP sequencing kit (U.S. Biochemicals). Sequencing products were subjected to 6% denaturing (8.3 M urea) polyacrylamide gel electrophoresis and visualized by X-ray film autoradiography.

The heat shock gene of cosmid 268 was identified in the random sequence analysis of this cosmid. Cosmid 268 was randomly sheared, and the ends were repaired according to the protocol of S. J. Surzycki (University of Indiana) as communicated to us by G. Burger (Université de Montréal). A Hospitak "Up-Mist" medication nebulizer was used to shear 3 to 5 µg of DNA in 500 µl of 25% glycerol. Nebulization was carried out with compressed nitrogen at 5 lb/in² (35 kPa) for 20 s. This procedure produced DNA fragments averaging 1,000 bp in size. The DNA was then ethanol precipitated and the ends were repaired with T7 DNA pol II (New England Biolabs), Klenow polymerase, and deoxynucleoside triphosphates. After a second ethanol precipitation, the DNA was treated with T4 polynucleotide kinase and ligated into the *Sma*I site of pUC18 (Pharmacia Biotech). The resulting ligation was transformed into *E. coli* XL-1 Blue MRF' (Stratagene), and the cells were plated onto medium containing ampicillin, X-Gal, and IPTG. Once grown, 768 colonies were picked into sterile, lidded microtiter plates containing LB medium with ampicillin and incubated overnight. Glycerol was then added to the microtiter plates to a final concentration of 15%, and the cells were frozen at -80°C. Colony blots of the library were prepared with a 96-prong device to stamp frozen cells from the microtiter plates onto sheets (22 by 22 cm) of GeneScreen Plus nylon membrane (Dupont NEN) laid on ampicillin-containing medium in square dishes. After incubation overnight, the membranes were processed by a sequential lysing (10% sodium dodecyl sulfate [SDS]), denaturation (0.5 M NaOH, 1.5 M NaCl), neutralization (1.5 M NaCl, 0.5 M Tris [pH 7.4]), and rinsing (2× SSC [1× SSC contains 150 mM NaCl and 15 mM trisodium citrate]) procedure (5 min for each step) followed by autoclaving of the membranes for 1 min on the fast exhaust setting. Immediately before use, the membranes were scrubbed in a washing solution (0.5% SDS, 1 mM EDTA, 5× SSC) to remove cellular debris. The colony blots were screened by hybridization with gel-isolated *Mlu*I restriction fragments purified with GeneClean (Bio 101). Individual clones were grown, and their plasmid DNA was extracted with the Nucleobond AX PC-20 column system (Macherey-Nagel). Sequencing reactions for both ends of each positive clone were prepared with the PRISM dye-terminator cycle sequencing kit (Applied Biosystems) with standard M13 forward and reverse primers (provided by R. K. Singh of the National Research Council of Canada). Reactions were run on an ABI 373A automated sequencer. Sequence trace data were assembled and edited with the Staden package, version 1995.0 (12). Cct sequences were aligned by the Clustal method of the DNASTAR program.

Transcription induction. To investigate the expression of the stress-responsive *cct* genes, transcription induction studies were performed by subjecting *H. volcanii* cultures ($A_{560} = 0.6$ to 0.8) growing at 37°C to elevated temperatures or to reduced salt concentrations and then immediately isolating total RNA with the

RNeasy system (Qiagen, Inc.), which is recommended for the purification of RNA molecules with sizes of greater than 200 bp. RNA samples were stored in 75% ethanol at -70°C until ready for analysis. To determine the maximum induction temperature for *cct* gene expression, a temperature course was performed in which *H. volcanii* cultures were incubated at 45, 50, 55, 60, 65, or 70°C for 45 min. Once the optimal temperature was determined, a time course for induction was used in which cell aliquots of 5 ml were removed from an *H. volcanii* culture at 0, 15, 30, 45, 60, 75, and 90 min at 60°C. To study the effect of reduced salt concentrations on *cct* gene expression, 60 ml of an *H. volcanii* culture was divided into four 50-ml culture tubes. Cells were isolated by centrifugation at $2,000 \times g$; resuspended in 15 ml of complex media containing NaCl concentrations of 2.2 (100%), 1.8 (80%), 1.3 (60%), or 0.9 (40%) M; and incubated at 37°C with shaking for 1 h before total RNA was isolated.

Prior to formaldehyde agarose gel electrophoresis, RNA was isolated by centrifugation at 14,000 rpm for 10 min, resuspended in RNase-free distilled water, and quantitated spectrophotometrically at an A_{260} . Approximately 5 to 10 µg of total RNA was resuspended in 5 µl of RNase-free distilled water and then incubated at 55°C for 15 min after addition of 19.4 µl of sample treatment buffer (20 mM MOPS [morpholinepropanesulfonic acid; pH 7.0], 50 mM sodium acetate, 10 mM EDTA, 6.4% formaldehyde, and 48% deionized formamide). Following incubation, 5 µl of loading buffer (1 mM EDTA [pH 8.0], 0.25% bromophenol blue, 0.25% xylene cyanol, and 50% glycerol) was added to each RNA sample. Formaldehyde agarose gel electrophoresis and RNA capillary transfer to Zeta-probe nylon membrane were performed according to standard protocols (51). Transcripts were sized by comparing their migration to a 0.24- to 9.5-kb RNA ladder (Gibco BRL).

The synthetic oligonucleotide HSPE (5'-CATCTTGTCCATCCCTTTGGG GCC-3') was end labeled at its 5' terminus with [γ -³²P]ATP and 5 U of T4 polynucleotide kinase (49) (Gibco BRL) and used in the transcription induction study to probe for the *cct1* and *cct2* transcripts. This probe is complementary to a sequence element that is conserved in both *cct* mRNAs. Therefore, hybridization signals observed with this probe are likely to represent the expression of both *cct1* and *cct2*. All Northern blot hybridizations were performed in 0.25 M Na₂HPO₄ (pH 7.2) and 7% SDS. The radioactive HSPE probe was allowed to bind to its complementary RNA sequence at 50°C overnight. Blots were washed twice (10 min each) in a solution containing 2× SSC and 0.5% SDS, followed by a single wash for 5 min in a solution containing 0.2× SSC and 0.5% SDS. Hybridization signals were quantitated with the InstantImager (Packard Instrument Company).

Promoter fusion and transcript analyses. For in vivo promoter fusion studies, sequences flanking the *cct1* and *cct2* N-terminus coding regions were amplified by PCR with Vent DNA polymerase (New England Biolabs). Amplification reactions were carried out in a 100-µl volume containing Vent DNA polymerase (2 U), 100 to 200 ng of template DNA, 1.0 µM primers, 200 µM deoxynucleoside triphosphates, and 10 µl of 10× Vent buffer (New England Biolabs) plus 1 ml of bovine serum albumin concentrate (100×). The *cct1* gene was amplified with the primers 5'-Δ-233 (5'-CGGCAAGCTTCGACAGCAACTGAGAC-3') and 3' REVcct1 (5'-AGGCTCTAGACTGGCTCATAGTCATCGC-3'), and the *cct2* gene was amplified with the primers 5'-Δ-177 (5'-AGAGAAGCTTGACGACC GGGCACAAT-3') and 3' REVcct2 (5'-ATAATCTAGACTGGCTCATAG CATTCC-3'). Forward primers were complementary to template strand sequences and had 5' termini corresponding to positions -233 (*cct1*) and -170 (*cct2*) relative to the transcription start. The reverse primers were complementary to regions of the nontemplate strand that included the first two or three amino acid codons of *cct1* and *cct2*, respectively. PCR amplification was carried out in 30 cycles of denaturation (94°C, 1 min), annealing (50°C, 1 min), and extension (72°C, 5 min). For cloning purposes, *Hind*III and *Xba*I endonuclease sites were introduced at the 5' and 3' termini, respectively, of the amplified DNAs.

The *cct1* and *cct2* 5' flanking sequences were cloned into the plasmid pWL201 expression vector (39) as *Hind*III-*Xba*I fragments. In the case of *cct1*, whose 5' flanking region contained a *Hind*III endonuclease restriction site, the putative core promoter region was first cloned into the *Hind*III and *Xba*I sites of pWL201, and then the remaining upstream sequence was ligated into the expression construct as a *Hind*III-*Hind*III fragment. Promoter fusions were verified by DNA sequencing as described above, and these clones were designated pWLcct1-233 or pWLcct2-170.

Promoter fusion constructs were transformed into competent *H. volcanii* cells as described previously (37). *H. volcanii* cells containing the appropriate promoter-reporter construct were grown at 37°C until the cell density reached mid-log phase ($A_{550} = 0.5$ to 0.7). Cultures were then transferred from 37°C to a 60°C shaking water bath to induce the heat shock response. Aliquots of 1.5 ml were removed just prior to heat shock (zero time point) and 45 min after incubation at 60°C. Total RNA was isolated immediately with the TRIzol reagent (Gibco BRL) as described by the manufacturer and stored in 75% ethanol at -70°C.

Northern analysis was used to detect the in vivo levels of tRNA^{Pro} transcript. RNA stored in 75% ethanol was centrifuged (7,500 × g, 5 min at room temperature), dried briefly under vacuum, and resuspended in 20 µl of RNA loading solution (7 M urea, 10% [wt/vol] glycerol, 0.05% xylene cyanol, 0.05% bromophenol blue). RNA species were resolved by electrophoresis through a 6% denaturing (8.3 M urea) polyacrylamide gel and transferred to Zeta-probe nylon

membrane with an electrophoretic blotter (Idea Scientific). RNA was fixed to the membrane by baking under vacuum (80°C, 30 min). The exon 1-specific tRNA^{Pro} oligonucleotide PROEX1 (5'-CCCAAAGCGAGAATCATACC AC-3') and the oligonucleotide LEU3E (5'-GGGGACGAGATTCGAATC GCGAACCCTACG-3') were labeled with [γ -³²P]ATP at their 5' termini as described above. LEU3E, which is specific for chromosomally encoded leucine tRNA, was used as an internal control for RNA recovery. Radioactive probes were allowed to hybridize to their complementary RNA sequences in hybridization solution at 65°C (15 to 20 h). The membrane was washed twice in 2× SSC and 0.5% SDS for 15 min each at room temperature. The values for the tRNA^{Pro}-derived signals were normalized by calculating the ratio of tRNA^{Pro} total counts over the LEU3E total counts. The change in transcript level was determined by dividing the normalized values obtained under heat shock conditions with those obtained under nonshock conditions.

To map the *cct1* and *cct2* transcription start sites, the antisense synthetic oligonucleotide PROEX1, which is complementary to sequences in exon 1 of tRNA^{Pro}, was used as the primer for cDNA synthesis and the accompanying DNA sequencing reactions. The oligonucleotide was labeled with [γ -³²P]ATP at its 5' terminus as described above. Approximately 10 pmol of the labeled primer was coprecipitated with 10 to 15 μ g of *H. volcanii* total RNA isolated from cells carrying the promoter fusion plasmid pWLcct1-233 or pWLcct2-170 and resuspended in hybridization buffer (0.3 M NaCl, 2 mM EDTA, 10 mM TrisCl [pH 7.5]). This mixture was incubated at 80°C for 4 min, and annealing was performed at 50°C for 1 h. The primer-RNA hybrid was precipitated and subjected to reverse transcription with Superscript II (Gibco BRL) reverse transcriptase according to the manufacturer's instructions. RNAs in the final reaction mixture were degraded by digestion with RNase A (0.5 μ g per μ l) at 37°C for 10 min. DNA sequence reactions were also performed with pWLcct1-233 or pWLcct2-170 plasmid DNA with the PROEX1 oligonucleotide as the primer. The cDNA products and DNA sequence reactions were analyzed by 6% denaturing (8.3 M urea) polyacrylamide gel electrophoresis.

Nucleotide sequence accession number. The sequences of the *cct1* and *cct2* genes have been assigned the GenBank accession no. AF010470 and AF010469, respectively.

RESULTS

Cloning and sequencing of the *H. volcanii* *cct* genes. Using a minimal set of overlapping cosmid clones (4) and radiolabeled cDNAs derived from *H. volcanii* DS2 total RNA, Trieselmann and Charlebois (55) identified seven strong heat-responsive loci that mapped to the chromosome (cosmids A199, H11, 10D2, 268, 452, 456, and 531). *Mlu*I restriction digestion of cosmid A199 yielded a 4.0-kb fragment and a 2.3-kb fragment, both of which hybridized to a 1.68-kb RNA transcript that was specifically induced upon heat shock at 60°C (data not shown). The same *Mlu*I fragments were shown previously to contain transcriptionally active regions responsive to thermal stress (55). The 2.3- and 4.0-kb *Mlu*I fragments were cloned into the *Sma*I site of pUC19 and designated HS5 and HS21, respectively. Partial sequencing of the pUC19 clones revealed an open reading frame that encoded an amino acid sequence sharing high similarity to Cct family members, such as thermophilic factor 55 (TF55) from *S. shibatae* and the cytosolic eucaryal T-complex polypeptide 1 (TCP-1), a hetero-oligomeric chaperonin. Therefore, we designated the open reading frame *cct1*. Overlapping subclones of HS5 and HS21 were then constructed for sequencing of the complete *cct1* gene from *H. volcanii*. A second halophilic archaeal *cct* gene (called *cct2*) was discovered during the sequencing of the previously identified heat-responsive locus (55) of cosmid 268.

DNA sequence analysis of the monocistronic *cct1* and *cct2* genes showed open reading frames of 560 and 557 amino acids, respectively, giving a predicted molecular mass of 59 kDa for both proteins. Additionally, the polypeptides encoded by *cct1* and *cct2* are highly acidic, as indicated by their calculated pIs of 3.9 and 4.1, respectively.

Protein sequence comparisons. Pairwise sequence comparisons between the *H. volcanii* Cct proteins and other Cct family members showed a high degree of similarity to the TF55 α and β subunits (41 to 46%) of *S. shibatae* and to human TCP-1 (30 to 35%). Figure 1 presents an amino acid sequence alignment of the two *H. volcanii* Cct proteins together with the eucaryal

TCP-1 protein from humans and the two TF55 subunits from *S. shibatae*. Extensive sequence similarity between these *H. volcanii* proteins and other Cct members occurs throughout their primary sequences, including the putative ATP-binding site (the GDGTT motif at amino acids 96 to 100) and a structural element adjacent to the ATP-binding site with the hexapeptide sequence TITNDG (residues 62 to 67) (22). The VVTNDG motif of *H. volcanii* Cct1 does not perfectly match the TITXDG consensus sequence of eucaryal Ccts; however, the replacement of the first threonine and/or isoleucine with valine in this motif occurs in other archaeal Ccts, such as the thermosome of *M. kandleri* (1). Both halophilic proteins also contain short Gly-Gly-Met repeats at their carboxy termini. Although this motif is absent in other archaeal and eucaryal Ccts, such C-terminal repeats are characteristic of the distantly related Hsp60/GroEL proteins (15, 44). Since deletion of this repeat region did not appear to affect the biological activity of *E. coli* GroEL in vitro (34), its functional role is presently unknown. However, it has been suggested that the repeat might be involved in membrane anchoring of the protein in vivo (6, 57).

Transcription induction study. To determine if the *H. volcanii* *cct* genes are controlled at the level of transcription, we examined the effects of heat and salt stress on the levels of *cct* mRNA by using Northern analysis (Fig. 2). To examine the effects of temperature on induction, *H. volcanii* cells were grown at 37°C and then shifted to higher temperatures for 45 min each. As shown in Fig. 2A, the peak for transcription induction occurred at 55 to 60°C and represented a 12- to 13-fold increase in transcript synthesis (Fig. 2A, lanes 4 and 5). Transcript levels began to decline at 65°C (Fig. 2A, lane 6), indicating that *cct* mRNAs exhibit a rapid and transient heat shock induction pattern typical of heat shock proteins. Since the HSPE oligonucleotide used as a probe in these experiments is complementary to sequences in both the *cct1* and *cct2* mRNAs, any observed increase in transcription during stress was likely to represent the response from both genes. A time course induction performed at 60°C is presented in Fig. 2B. The results show that synthesis of the *H. volcanii* *cct* transcripts increased dramatically after 30 min and continued to rise after 75 min. In addition to the heat shock response, it had been observed previously that three proteins were induced in *H. volcanii* when these cells were subjected to salt dilution stress (11). Two of these proteins (91 and 79 kDa) appeared to comigrate with two proteins observed to be induced during heat shock (11). Consequently, we wanted to know if *cct* genes also responded to salt dilution. For this purpose, *H. volcanii* cells were grown to an A_{550} of 1.0, centrifuged, and resuspended in complex media containing reduced NaCl concentrations and then incubated in a 37°C shaking incubator for 60 min. Total RNA from these cells was analyzed by Northern blot hybridization as in the temperature and time course studies. As shown in Fig. 2C, transcripts hybridizing to the *cct* probe were induced 2.7-fold (lane 3) and 4.6-fold (lane 4) in response to 60 and 40% of the normal NaCl level, respectively. Although the degree of induction by salt dilution was not as dramatic as that observed under heat shock conditions, these genes appeared to be induced in response to decreased salt in the external environment.

Transcription reporter assays for *cct* gene expression and transcript mapping. Using a previously reported plasmid-based expression module for *H. volcanii* (36), we identified regions of the *cct1* and *cct2* genes necessary for heat shock-induced transcription. Fusions were made by coupling the 5' flanking regulatory regions of *cct1* and *cct2* to the yeast tRNA^{Pro} reporter gene on the *H. volcanii*-*E. coli* pWL201

Hv Cct 1	M	S	-	-	-	Q	R	M	Q	Q	Q	P	M	I	I	L	G	E	D	S	Q	R	T	S	G	Q	D	A	Q	S	M	N	I	T	A	G	K	A	V	A	E	A	V	R	T	T	L	G	46					
Hv Cct 2	M	S	-	-	-	Q	R	M	Q	Q	Q	P	M	I	I	L	G	E	D	A	Q	R	V	K	D	R	D	A	Q	E	Y	N	I	R	A	A	R	A	V	A	E	A	V	R	S	T	L	G	46					
Ss A	M	A	S	-	-	-	-	-	-	-	-	-	-	P	V	L	L	L	K	E	G	T	S	R	T	G	R	D	A	L	R	N	N	I	L	A	A	K	T	L	A	E	M	L	R	S	S	L	G	40				
Ss B	M	A	T	A	T	V	A	T	P	E	G	I	P	V	T	I	L	K	E	G	S	S	R	T	Y	G	K	E	A	L	R	A	N	I	A	A	V	K	A	I	E	E	A	L	K	S	T	Y	G	50				
Hm Tcp1	M	E	G	P	-	-	-	-	-	-	-	-	-	-	-	L	S	V	F	G	D	R	S	T	G	E	T	I	R	S	Q	N	V	M	A	A	S	I	A	N	I	V	K	S	S	L	G	37						
Hv Cct 1	P	K	G	M	D	K	M	L	V	D	S	G	G	Q	V	V	V	T	N	D	G	V	T	I	L	K	E	M	D	I	D	H	P	A	A	N	M	I	V	E	V	S	E	T	Q	E	D	E	V	G	96			
Hv Cct 2	P	K	G	M	D	K	M	L	V	D	S	M	G	D	V	T	I	T	N	D	G	V	T	I	L	K	E	M	D	I	D	N	P	T	A	E	M	I	V	E	V	A	E	T	Q	E	D	E	V	G	96			
Ss A	P	K	G	L	D	K	M	L	T	D	S	F	G	D	V	T	I	T	N	D	G	A	T	I	V	K	D	M	E	I	Q	H	P	A	A	K	L	L	V	E	A	A	K	A	Q	D	A	E	V	G	90			
Ss B	P	R	G	M	D	K	M	L	V	D	S	L	G	D	T	T	I	T	N	D	G	A	T	I	L	D	K	M	D	L	Q	H	P	T	G	K	L	L	V	Q	I	A	K	Q	D	E	E	T	A	100				
Hm Tcp1	P	V	G	L	D	K	M	L	V	D	D	I	G	D	V	T	I	T	N	D	G	A	T	I	L	K	L	L	E	V	E	H	P	A	A	K	V	L	C	E	L	A	D	L	Q	D	K	E	V	G	87			
Hv Cct 1	D	G	T	T	T	A	V	I	N	A	G	E	L	L	D	Q	A	E	D	L	L	D	S	D	V	H	A	T	T	I	A	Q	G	Y	R	Q	A	A	E	K	A	K	E	V	L	E	D	N	A	I	146			
Hv Cct 2	D	G	T	T	T	A	V	I	N	A	G	E	L	L	K	N	A	E	D	L	L	E	Q	D	I	H	P	T	A	I	T	R	G	F	N	L	A	S	E	K	A	R	E	I	D	P	I	A	E	146				
Ss A	D	G	T	T	S	A	V	I	L	A	G	E	L	L	E	K	A	E	S	L	L	D	Q	N	I	H	P	T	I	I	E	G	Y	K	K	A	Y	T	K	A	L	E	L	P	Q	L	G	T	140					
Ss B	D	G	T	T	T	A	V	I	L	A	G	E	L	L	K	A	E	D	L	L	V	K	E	I	H	P	T	I	I	V	S	G	Y	K	K	A	E	E	I	A	L	K	T	I	Q	D	I	A	Q	150				
Hm Tcp1	D	G	T	T	S	V	I	I	A	A	E	L	L	K	N	A	E	D	L	L	V	K	E	I	H	P	T	S	V	I	S	G	Y	R	L	A	C	K	E	A	V	R	Y	I	N	E	N	L	I	137				
Hv Cct 1	E	V	T	E	D	D	-	-	-	-	-	-	-	R	E	T	L	T	K	I	A	A	T	A	M	T	G	K	G	A	E	S	A	K	D	L	-	-	L	S	E	L	V	D	A	V	L	A	I	188				
Hv Cct 2	R	V	D	P	D	D	-	-	-	-	-	-	-	E	E	L	L	K	K	V	A	E	T	S	M	T	G	K	S	S	E	L	N	K	E	L	-	-	L	A	D	L	V	I	R	A	V	R	Q	V	188			
Ss A	R	I	D	I	R	D	L	N	S	V	A	R	D	T	L	R	K	I	A	F	T	T	L	A	S	K	F	T	A	E	G	A	E	L	N	K	I	I	D	M	V	I	D	A	I	V	N	V	A	190				
Ss B	P	V	S	I	N	D	-	-	-	-	-	-	T	D	V	L	R	K	V	A	L	T	S	L	G	S	K	A	V	A	G	A	R	E	Y	-	-	L	A	D	L	V	V	K	A	V	A	Q	V	A	192			
Hm Tcp1	V	N	T	D	E	-	-	-	-	-	-	-	L	G	R	D	C	L	I	N	A	A	K	T	S	M	S	S	K	I	T	I	G	I	N	G	D	F	-	-	F	A	N	M	V	V	D	A	V	L	A	I	K	180
Hv Cct 1	D	D	D	G	-	-	-	-	-	-	-	-	I	D	T	N	N	V	S	I	E	K	V	V	G	G	T	I	D	N	S	E	L	V	E	G	V	I	V	D	K	E	R	V	D	E	N	M	P	Y	A	V	E	233
Hv Cct 2	V	E	A	N	D	G	S	H	V	V	D	L	E	N	V	S	I	E	T	Q	T	G	R	S	S	E	S	E	S	E	L	V	L	T	G	A	V	I	D	K	E	D	P	V	H	D	M	P	V	Q	F	D	238	
Ss A	E	P	L	P	N	G	G	Y	N	V	S	L	D	L	I	K	I	D	K	K	G	G	S	I	E	S	V	L	V	K	G	L	V	L	D	K	E	V	V	H	P	G	M	P	R	R	V	T	240					
Ss B	E	-	L	R	G	D	K	W	Y	V	D	L	D	N	V	Q	I	V	K	K	H	G	S	I	N	D	T	Q	L	V	Y	G	I	V	V	D	K	E	V	V	H	P	G	M	P	K	R	I	E	241				
Hm Tcp1	Y	T	D	I	R	G	Q	P	R	Y	P	V	N	S	V	N	I	L	K	A	H	G	R	S	Q	M	E	S	M	L	I	S	G	Y	A	L	N	C	V	V	G	S	Q	G	M	P	K	R	I	V	230			
Hv Cct 1	D	A	N	I	A	I	L	D	D	A	L	E	V	R	E	T	E	I	D	A	E	V	N	V	T	D	P	D	L	Q	Q	F	L	D	Q	E	E	K	Q	L	K	E	M	V	D	Q	L	V	E	283				
Hv Cct 2	A	A	D	V	L	L	L	N	E	P	V	E	E	T	D	I	D	T	N	V	S	T	E	S	P	D	Q	L	Q	K	F	L	D	Q	E	E	A	Q	L	K	K	V	D	Q	L	V	D	288						
Ss A	K	A	K	A	I	A	V	L	D	A	L	E	V	E	K	P	E	I	S	A	K	I	S	I	S	P	E	Q	I	K	A	F	L	D	E	E	S	K	Y	L	K	D	M	V	D	K	L	A	S	290				
Ss B	N	A	K	I	A	L	D	A	A	S	L	E	V	E	K	P	E	L	D	A	E	I	R	I	N	D	P	T	Q	M	H	K	F	L	E	E	E	N	I	L	K	E	K	V	D	K	I	A	S	291				
Hm Tcp1	N	A	K	I	A	C	L	D	F	S	L	Q	K	T	K	M	K	L	G	V	Q	V	V	I	T	D	P	E	K	L	D	Q	I	R	Q	R	E	S	D	I	T	K	E	R	I	Q	K	I	L	A	280			
Hv Cct 1	V	G	A	D	A	V	F	V	G	D	G	I	D	D	M	A	Q	H	Y	L	A	K	E	G	I	L	A	V	R	R	A	K	S	S	D	L	K	R	L	A	R	A	T	G	L	R	V	V	S	333				
Hv Cct 2	S	G	A	D	V	V	F	C	Q	K	G	I	D	D	L	A	Q	H	Y	L	A	K	Q	G	I	L	A	V	R	R	T	K	K	S	D	I	R	F	L	K	N	I	T	G	A	A	V	V	S	338				
Ss A	I	G	A	N	V	I	C	Q	K	G	I	D	D	I	A	Q	H	Y	L	A	K	G	I	L	A	V	R	R	V	K	R	S	D	I	E	K	L	E	K	A	L	G	A	R	I	I	S	340						
Ss B	T	G	A	N	V	I	C	Q	K	G	I	D	D	E	V	A	Q	H	Y	L	A	K	G	I	L	A	V	R	R	A	K	K	S	D	L	E	K	L	A	R	A	T	G	L	R	V	I	S	N	341				
Hm Tcp1	T	G	A	N	V	I	L	T	T	G	G	I	D	D	M	C	L	K	V	F	V	E	A	G	A	M	A	V	R	R	V	L	K	R	D	L	K	R	T	A	K	A	S	G	A	T	I	L	S	T	330			
Hv Cct 1	L	D	D	I	E	A	D	D	-	-	-	-	-	L	G	F	A	G	S	V	G	Q	K	D	V	G	G	D	E	R	I	F	V	E	D	V	-	E	D	A	K	S	V	T	L	I	L	R	G	376				
Hv Cct 2	L	D	S	I	E	A	A	V	-	-	-	-	-	L	-	-	G	R	A	S	V	R	R	D	E	A	G	E	L	F	Y	V	E	G	I	G	D	V	H	G	V	T	L	L	L	R	G	379						
Ss A	I	K	D	A	T	P	D	D	-	-	-	-	-	L	G	Y	A	E	L	V	E	E	R	V	G	N	D	K	M	V	F	I	E	G	-	A	K	N	L	K	A	V	N	I	L	L	R	G	383					
Ss B	I	D	E	L	T	S	Q	D	-	-	-	-	-	L	G	Y	A	E	L	V	E	E	R	K	V	G	N	D	K	M	V	F	V	E	G	-	A	K	N	P	K	S	V	S	I	L	I	R	G	384				
Hm Tcp1	L	A	N	L	E	G	E	E	T	F	E	A	A	M	L	G	Q	A	E	E	V	V	Q	E	R	I	C	D	E	L	I	L	I	K	N	-	T	K	A	R	T	S	A	S	I	I	L	R	G	379				
Hv Cct 1	G	T	E	H	V	V	D	E	L	E	R	A	I	E	D	S	L	G	V	V	R	T	T	L	E	D	G	K	V	L	P	G	G	A	P	E	T	E	L	S	L	Q	L	R	E	F	A	D	S	426				
Hv Cct 2	S	T	D	H	V	V	D	E	L	E	R	G	V	Q	D	A	L	D	V	V	R	S	T	V	A	D	G	R	V	L	P	G	G	A	I	E	V	E	L	A	S	R	L	R	N	Y	A	D	S	429				
Ss A	S	N	D	M	A	L	D	E	A	E	R	S	I	N	D	A	L	H	A	L	R	N	I	T	L	E	P	V	I	L	P	G	G	A	I	E	L	E	L	A	M	K	L	R	E	Y	A	R	S	433				
Ss B	G	L	E	R	V	V	D	E	T	E	R	A	L	R	D	A	L	G	T	V	A	D	V	I	R	D	G	R	A	V	A	G	G	A	V	E	I	E	I	A	K	R	L	R	K	Y	A	P	Q	434				
Hm Tcp1	A	N	D	F	M	C	D	E	M	E	R	S	L	H	D	A	L	C	V	V	K	R	V	L	E	S	K	S	V	V	L	P	G	G	A	V	E	A	A	L	S	I	Y	L	E	N	Y	A	T	S	429			
Hv Cct 1	V	G	G	R	E	Q	L	A	V</																																													

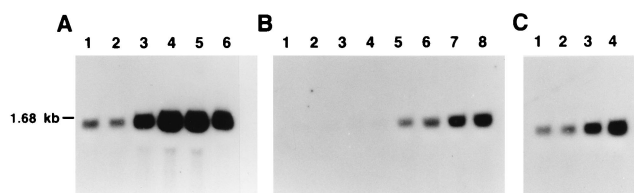


FIG. 2. Transcription induction of the *H. volcanii* *cct* genes. Total RNA isolated from *H. volcanii* cells under nonstress and stress conditions was analyzed by Northern hybridization. (A) Temperature course for *cct1* mRNA induction. *H. volcanii* cells were grown to mid-log phase at 37°C (lane 1) and then incubated at 45, 50, 55, 60, or 65°C (lanes 2 to 6, respectively) for 45 min. (B) Time course for *cct1* mRNA induction. *H. volcanii* cells were grown at 37°C (lane 1) and then incubated for 5, 10, 15, 30, 45, 60, or 75 min (lanes 2 to 8, respectively) at 60°C. (C) Effect of salt shock on *cct1* mRNA expression. *H. volcanii* cells were exposed to complex media that contained NaCl concentrations of 2.2 (100%), 1.8 (80%), 1.3 (60%), or 0.9 (40%) M (lanes 1 to 4, respectively) for 60 min. The size of the major hybridizing species is indicated.

specifically stimulated in response to heat shock, with induction folds of 9.4 and 13.6, respectively. The minor hybridization signals observed above the tRNAProM-derived signals were most likely due to inefficient termination by the expression module terminator under heat shock conditions. Inclusion of sequences further upstream of position -233 of *cct1* did not affect the heat shock response of this gene (data not shown). Constructs containing sequences beyond -170 of *cct2* were not examined.

A comparison of the *cct1* and *cct2* 5' flanking regions revealed that these genes shared significant sequence conservation in the regions that extended to -40 bp from their presumed translation start sites (Fig. 4A). This region included the sequence 5'-TTTATA-3' (positions -23 to -28), which matches the consensus TATA motif (5'-T/CTTAT/AA-3') of archaeal promoters (17, 63) and resembles the TATA box of eucaryal pol II core promoters (47, 52).

To determine whether the 5'-TTTATA-3' sequence motif functions as a TATA element for *cct* transcription, we performed primer extension analysis to identify the 5' termini of the *cct1* and *cct2* transcripts. Total RNA was isolated from *H. volcanii* cultures carrying the pWLcct1-233 and pWLcct2-170 reporter plasmids grown under normal conditions or subjected to heat shock at 60°C for 60 min. As shown in Fig. 4B, discrete primer extension products were observed in lanes 1 and 2, indicating that transcription of the *cct1* gene initiated from the

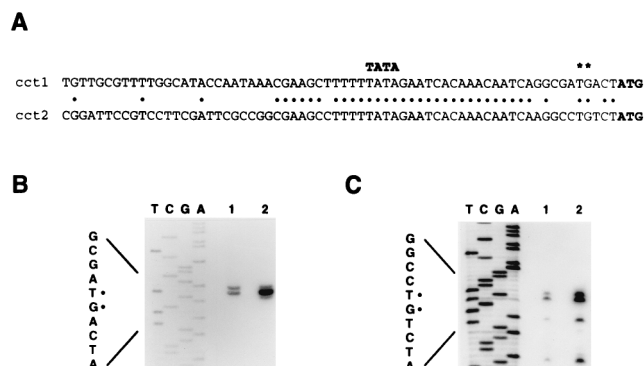


FIG. 4. Sequence comparison of the *H. volcanii* *cct* promoter regions and 5' transcript mapping. (A) Nucleotide sequence alignment of the *cct1* and *cct2* promoter regions. Nucleotides conserved between the two 5' flanking regions are indicated. The putative TATA element, the transcription start sites (marked by asterisks), and the translational start codon (in boldface) are also presented. (B) Mapping of the *H. volcanii* *cct1* transcription start site by primer extension. Dideoxy-terminated DNA sequencing reaction products (lanes T, C, G, and A) along with the primer extension reaction products (lanes 1 and 2) are shown. Lane 1, RNA isolated from cells grown under normal conditions; lane 2, RNA from cells exposed to heat shock at 60°C for 45 min. (C) Mapping of the *H. volcanii* *cct2* transcription start site by primer extension. Lane 1, RNA isolated from cells grown under normal conditions; lane 2, RNA isolated from cells exposed to heat shock at 60°C for 45 min. Initiation sites are marked with dots in panels B and C (asterisks in panel A).

TG dinucleotide, with a preference for G, located approximately 25 bp downstream of the TATA element. The same initiation site was used under both nonstress and stress conditions. Similarly, both basal and heat-activated transcription of *cct2* initiated from the same relative TG residues (Fig. 4C). Primer extension analysis with a primer specific for the *cct2* mRNA and total RNA isolated from *H. volcanii* cells lacking the reporter plasmids indicated that the same transcription start site was used by the native gene (data not shown).

DISCUSSION

For the purpose of understanding the molecular mechanisms governing archaeal gene regulation, we began an investigation of transcription signals by using the inducible *cct* genes as models. In this study, we report the sequencing and transcription analysis of two *cct* family member genes from the halophilic archaeon *H. volcanii*. Comparative sequence analysis of the deduced Cct proteins indicated high similarity to each other (60.9%) and to *S. shibatae* TF55 α and β subunits (43 to 48%), and human TCP-1 (33 to 37%).

Recent studies of core promoter architecture and transcription initiation machinery indicate that the archaeal and eucaryal transcription systems are homologous at a fundamental level, while both differ significantly from the transcription apparatus controlling bacterial gene expression (reviewed in references 21 and 27). In agreement with the recent descriptions of eucaryal-like features of archaeal transcription, we identified an AT-rich TATA-like motif (5'-TTTATA-3') reminiscent of eucaryal pol II promoters approximately 25 nucleotides upstream from the initiation sites for *cct1* and *cct2*. Primer extension analysis indicated that both *cct1* and *cct2* transcripts initiated approximately 25 bp downstream from an archaeal consensus TATA element (17) under both basal and stress-activated transcription (Fig. 4). This is consistent with the use of a consensus TATA element. As expected, the putative *cct* promoter elements bear no sequence resemblance to the -35 and -10 consensus sequences of *E. coli* σ^{32} promoters (9), and

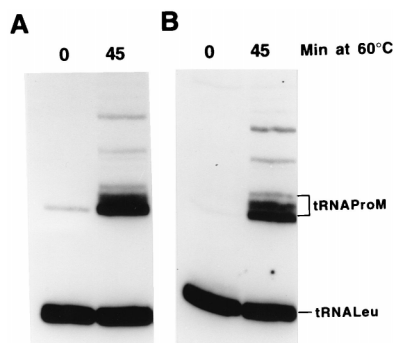


FIG. 3. Heat-induced transcription from *cct1* and *cct2* promoter fusions. Northern hybridizations were performed with the reporter-specific probe PRO-EX1. Cellular RNAs isolated from cells carrying pWLcct1-233 (A) and pWLcct2-170 (B) were isolated under nonshock (zero time point) and heat shock (45 min at 60°C) conditions. Hybridization signals corresponding to tRNAProM transcripts are indicated. Chromosomally encoded leucine tRNA was used as an internal control to evaluate RNA recoveries.

we did not detect the heat shock consensus element (5'-nGAAnnTTCn-3') responsible for the activation of eucaryal heat shock genes (40, 41) in the *cct* promoter regions. These observations suggest that the recognition signals, and possibly the underlying regulatory mechanism, controlling heat shock gene expression in the domain *Archaea* are significantly different from those described for the domains *Bacteria* and *Eucarya*.

In the study reported here, transcriptional analysis of *cct* promoter regions fused to a yeast tRNA^{Pro} reporter gene demonstrated that 5' flanking *cct1* sequences up to position -233 and 5' flanking *cct2* sequences up to position -170 efficiently promote heat-induced transcription. In a recent study, Kagawa and coworkers (19) identified regions of homology upstream from the N terminus of the genes encoding the α and β subunits of *S. shibatae* TF55. Although an AT-rich pseudopalindromic sequence was detected at around position -90 relative to the transcription initiation site, most of the sequence conservation of these putative promoters centered in the 5'-TTTATA-3' TATA element and flanking regions (18). The tetranucleotide TATA element identified upstream of the *H. volcanii* *cct* coding regions is identical to the promoter element reported for TF55. In addition, nucleotide sequence alignment of the *cct1* and *cct2* promoter regions showed that sequences immediately flanking the TATA elements of these genes were also conserved (Fig. 4A). These shared promoter elements encompass the TATA element and adjacent sequences and may constitute the heat shock promoter of the *H. volcanii* *cct* genes. Consistent with this proposal, Palmer and Daniels (37) demonstrated that a purine-rich sequence located immediately 5' of the tRNA^{Lys} TATA element was essential for transcriptional activity in vivo.

Other HSP-encoding genes have been documented in the archaea (7, 8, 16, 30). An examination of the 5' sequences flanking the inferred Hsp70 coding region in *Haloarcula marismortui* (16) revealed a potential TATA element, 5'-TTTTAACC-3', located approximately 30 bases upstream from the presumed translational start codon (the transcription start site was not mapped [16]). This potential promoter lacks the central TATA tetranucleotide present in the TF55 and *cct* TATA elements and the conserved sequences flanking these motifs. However, it is not known whether the halobacterial *hsp70* is actually heat responsive, since studies investigating the expression of this gene were not described (16). Comparative sequence analysis of putative TATA elements located upstream of other archaeal heat shock genes, such as the heat-inducible *dnaK* and *dnaJ* genes from *Methanosarcina mazei* S6 (7), revealed typically AT-rich regions with promoter sequences different from those of the *H. volcanii* *cct* TATA elements. These differences in core promoter sequence suggest that members of the domain *Archaea* may utilize different classes of heat shock promoters, which might reflect differences in transcription regulation.

Our finding that expression of the halophilic *cct* genes is induced in response to both heat shock and salt dilution suggests that the *H. volcanii* Cct protein complex might be involved in cellular adaptation to general environmental stresses. As molecular chaperonins, TCP-1-containing complexes are involved in mitotic spindle formation (62) and have been shown to mediate the folding of various protein substrates, including actin, tubulin, neurofilament, firefly luciferase, and hepatitis B virus capsid (5, 14, 29, 56). In vitro functional studies of archaeal Ccts from *S. shibatae* (54), *S. solfataricus* (24), *P. occultum* (42), and *T. acidophilum* (60) have also shown that these complexes, which are composed of two different polypeptide subunits, possess ATPase activity and per-

form such chaperonin functions as peptide binding and protein folding (reviewed in reference 2). The presence of two divergent Cct-encoding genes in *H. volcanii* suggests that the putative halophile Cct complex is also heteromeric in structure.

ACKNOWLEDGMENTS

This work was supported by a grant from the Department of Energy, DE-FG02-91ER20041, to C.J.D. and by a grant from the Natural Sciences and Engineering Research Council to R.L.C.. R.L.C. and C.J.D. are associates of the Canadian Institute for Advanced Research.

The first two authors, Y.-P.K. and D.K.T., made equal contributions to the completion of this work.

REFERENCES

- Andra, S., G. Frey, M. Nitsch, W. Baumeister, and K. O. Stetter. 1996. Purification and structural characterization of the thermosome from the hyperthermophilic archaeum *Methanopyrus kandleri*. FEBS Lett. **379**:127-131.
- Baross, J. A., and J. F. Holden. 1996. Overview of hyperthermophiles and their heat shock proteins. Adv. Protein Chem. **48**:1-34.
- Charlebois, R. L., W. L. Lam, S. W. Cline, and W. F. Doolittle. 1987. Characterization of pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an archaeobacterium. Proc. Natl. Acad. Sci. USA **84**:8530-8534.
- Charlebois, R. L., L. C. Schalkwyk, J. D. Hofman, and W. F. Doolittle. 1991. Detailed physical map and set of overlapping clones covering the genome of the archaeobacterium *Haloferax volcanii* DS2. J. Mol. Biol. **222**:509-524.
- Chen, X., D. S. Sullivan, and T. C. Huffaker. 1994. Two yeast genes with similarity to TCP-1 are required for microtubule and actin function in vivo. Proc. Natl. Acad. Sci. USA **91**:9111-9115.
- Chitnis, P. R., and N. Nelson. 1991. Molecular cloning of the genes encoding two chaperone proteins of the cyanobacterium *Synechocystis* sp. PCC 6803. J. Biol. Chem. **266**:58-65.
- Clarens, M., A. J. Macario, and E. Conway de Macario. 1995. The archaeal *dnaK-dnaJ* gene cluster: organization and expression in the methanogen *Methanosarcina mazei*. J. Mol. Biol. **250**:191-201.
- Conway de Macario, E., C. B. Dugan, and A. J. Macario. 1994. Identification of a *grpE* heat-shock gene homolog in the archaeon *Methanosarcina mazei*. J. Mol. Biol. **240**:95-101.
- Cowing, D. W., J. C. Bardwell, E. A. Craig, C. Woolford, R. W. Hendrix, and C. A. Gross. 1985. Consensus sequence for *Escherichia coli* heat shock gene promoters. Proc. Natl. Acad. Sci. USA **82**:2679-2683.
- Creti, R., P. Londei, and P. Cammarano. 1993. Complete nucleotide sequence of an archaeal (*Pyrococcus woesei*) gene encoding a homolog of eukaryotic transcription factor IIB (TFIIB). Nucleic Acids Res. **21**:2942.
- Daniels, C. J., A. H. Z. McKee, and W. F. Doolittle. 1984. Archaeobacterial heat-shock proteins. EMBO J. **3**:745-749.
- Dear, S., and R. Staden. 1991. A sequence assembly and editing program for efficient management of large sequencing projects. Nucleic Acids Res. **19**:3907-3911.
- Frydman, J., and F.-U. Hartl. 1994. Molecular chaperone functions of hsp70 and hsp60 in protein folding, p. 251-283. In R. I. Morimoto, A. Tissières, and C. Georgopoulos (ed.), The biology of heat shock proteins and molecular chaperones. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Frydman, J., E. Nimmesgern, H. Erdjument-Bromage, J. S. Wall, P. Tempst, and F.-U. Hartl. 1992. Function in protein folding of TricA, a cytosolic ring complex containing TCP-1 and structurally related subunits. EMBO J. **11**:4767-4778.
- Gupta, R. S. 1995. Evolution of the chaperonin families (Hsp60, Hsp10 and Tcp-1) of proteins and the origin of eukaryotic cells. Mol. Microbiol. **15**:1-11.
- Gupta, R. S., and B. Singh. 1992. Cloning of the HSP70 gene from *Halobacterium marismortui*: relatedness of archaeobacterial HSP70 to its eubacterial homologs and a model for the evolution of the HSP70 gene. J. Bacteriol. **174**:4594-4605.
- Hain, J., D. Reiter, U. Hüdepohl, and W. Zillig. 1992. Elements of an archaeal promoter defined by mutational analysis. Nucleic Acids Res. **20**:5423-5428.
- Kagawa, H. K., J. Osipiuk, N. Maltsev, R. Overbeek, E. Quate-Randall, A. Joachimski, and J. D. Trent. 1995. The 60 kDa heat shock proteins in the hyperthermophilic archaeon *Sulfolobus shibatae*. J. Mol. Biol. **253**:712-725.
- Kagawa, Y., T. Ohta, Y. Abe, H. Endo, M. Yohda, N. Kato, I. Endo, T. Hamamoto, M. Ichida, T. Hoaki, and T. Maruyama. 1995. Gene of heat shock protein of sulfur-dependent archaeal hyperthermophile *Desulfurococcus*. Biochem. Biophys. Res. Commun. **214**:730-736.
- Kaine, B. P., I. J. Mehr, and C. R. Woese. 1994. The sequence, and its evolutionary implications, of a *Thermococcus celer* protein associated with transcription. Proc. Natl. Acad. Sci. USA **91**:3854-3860.

21. Keeling, P. J., and W. F. Doolittle. 1995. Archaea: narrowing the gap between prokaryotes and eukaryotes. *Proc. Natl. Acad. Sci. USA* **92**:5761–5764.
22. Kim, S., K. R. Willison, and A. L. Horwich. 1994. Cytosolic chaperonin subunits have a conserved ATPase domain but diverged polypeptide-binding domains. *Trends Biochem. Sci.* **19**:543–548.
23. Klenk, H. P., P. Palm, F. Lottspeich, and W. Zillig. 1992. Component H of the DNA-dependent RNA polymerases of Archaea is homologous to a subunit shared by the three eucaryal nuclear RNA polymerases. *Proc. Natl. Acad. Sci. USA* **89**:407–410.
24. Knapp, S., I. Schmidt-Krey, H. Hebert, T. Bergman, H. Jornvall, and R. Ladenstein. 1994. The molecular chaperonin TF55 from the thermophilic archaeon *Sulfolobus solfataricus*. A biochemical and structural characterization. *J. Mol. Biol.* **242**:397–407.
25. Kubota, H., G. Hynes, and K. Willison. 1995. The chaperonin containing t-complex polypeptide 1 (TCP-1). Multisubunit machinery assisting in protein folding and assembly in the eukaryotic cytosol. *Eur. J. Biochem.* **230**:3–16.
26. Lam, W. L., and W. F. Doolittle. 1989. Shuttle vectors for the archaeobacterium *Haloferax volcanii*. *Proc. Natl. Acad. Sci. USA* **86**:5478–5482.
27. 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.
28. Langer, D., and W. Zillig. 1993. Putative *tflls* gene of *Sulfolobus acidocaldarius* encoding an archaeal transcription elongation factor is situated directly downstream of the gene for a small subunit of DNA-dependent RNA polymerase. *Nucleic Acids Res.* **21**:2251.
29. Lingappa, J. R., R. L. Martin, M. L. Wong, D. Ganem, W. J. Welch, and V. R. Lingappa. 1994. A eukaryotic cytosolic chaperonin is associated with a high molecular weight intermediate in the assembly of hepatitis B virus capsid, a multimeric particle. *J. Cell Biol.* **125**:99–111.
30. Macario, A. J., C. B. Dugan, and E. Conway de Macario. 1991. A *dnaK* homolog in the archaeobacterium *Methanosarcina mazei* S6. *Gene* **108**:133–137.
31. Mager, W. H., and A. J. De Kruijff. 1995. Stress-induced transcriptional activation. *Microbiol. Rev.* **59**:506–531.
32. Marco, S., D. Urena, J. L. Carrascosa, T. Waldmann, J. Peters, R. Hegerl, G. Pfeifer, H. Sack-Kongehl, and W. Baumeister. 1994. The molecular chaperone TF55. Assessment of symmetry. *FEBS Lett.* **341**:152–155.
33. Marsh, T. L., C. I. Reich, R. B. Whitelock, and G. J. Olsen. 1994. Transcription factor IID in the Archaea: sequences in the *Thermococcus celer* genome would encode a product closely related to the TATA-binding protein of eukaryotes. *Proc. Natl. Acad. Sci. USA* **91**:4180–4184.
34. McLennan, N. F., A. S. Girshovich, N. M. Lissin, Y. Charters, and M. Masters. 1993. The strongly conserved carboxyl-terminus glycine-methionine motif of the *Escherichia coli* GroEL chaperonin is dispensable. *Mol. Microbiol.* **7**:49–58.
35. Nieuwlandt, D. T., and C. J. Daniels. 1990. An expression vector for the archaeobacterium *Haloferax volcanii*. *J. Bacteriol.* **172**:7104–7110.
36. Palmer, J. R., and C. J. Daniels. 1994. A transcriptional reporter for in vivo promoter analysis in the archaeon *Haloferax volcanii*. *Appl. Environ. Microbiol.* **60**:3867–3869.
37. Palmer, J. R., and C. J. Daniels. 1995. In vivo definition of an archaeal promoter. *J. Bacteriol.* **177**:1844–1849.
38. Palmer, J. R., and C. J. Daniels. Unpublished data.
39. Palmer, J. R., D. T. Nieuwlandt, and C. J. Daniels. 1994. Expression of a yeast intron-containing tRNA in the archaeon *Haloferax volcanii*. *J. Bacteriol.* **176**:3820–3823.
40. Pelham, H. R. B. 1982. A regulatory upstream promoter element in the *Drosophila* Hsp 70 heat-shock gene. *Cell* **30**:517–528.
41. Perisic, O., H. Xiao, and J. T. Lis. 1989. Stable binding of *Drosophila* heat shock factor to head-to-head and tail-to-tail repeats of a conserved 5 bp recognition unit. *Cell* **59**:797–806.
42. Phipps, B. M., A. Hoffmann, K. O. Stetter, and W. Baumeister. 1991. A novel ATPase complex selectively accumulated upon heat shock is a major cellular component of thermophilic archaeobacteria. *EMBO J.* **10**:1711–1722.
43. Phipps, B. M., D. Typke, R. Hegerl, S. Volker, A. Hoffmann, K. O. Stetter, and W. Baumeister. 1993. Structure of a molecular chaperone from a thermophilic archaeobacterium. *Nature (London)* **361**:475–477.
44. Picketts, D. J., C. S. Mayanil, and R. S. Gupta. 1989. Molecular cloning of a Chinese hamster mitochondrial protein related to the “chaperonin” family of bacterial and plant proteins. *J. Biol. Chem.* **264**:12001–12008.
45. Qureshi, S. A., P. Baumann, T. Rowlands, B. Khoo, and S. P. Jackson. 1995. Cloning and functional analysis of the TATA binding protein from *Sulfolobus shibatae*. *Nucleic Acids Res.* **23**:1775–1781.
46. Qureshi, S. A., B. Khoo, P. Baumann, and S. P. Jackson. 1995. Molecular cloning of the transcription factor TFIIB homolog from *Sulfolobus shibatae*. *Proc. Natl. Acad. Sci. USA* **92**:6077–6081.
47. Reiter, W. D., P. Palm, and W. Zillig. 1988. Analysis of transcription in the archaeobacterium *Sulfolobus* indicates that archaeobacterial promoters are homologous to eukaryotic pol II promoters. *Nucleic Acids Res.* **16**:1–19.
48. 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.
49. 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.
50. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
51. Selden, R. F. 1987. Analysis of RNA by Northern hybridization, p. 4.9.1–4.9.7. In R. B. Frederick, M. Ausubel, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.), *Current protocols in molecular biology*. Greene Publishing Associates and Wiley-Interscience, New York, N.Y.
52. Thomm, M., and G. Wich. 1988. An archaeobacterial promoter element for stable RNA genes with homology to the TATA box of higher eukaryotes. *Nucleic Acids Res.* **16**:151–163.
53. Trent, J. D. 1996. A review of acquired thermotolerance, heat-shock proteins, and molecular chaperones in archaea. *FEMS Microbiol. Rev.* **18**:249–258.
54. Trent, J. D., E. Nimmesgern, J. S. Wall, F.-U. Hartl, and A. L. Horwich. 1991. A molecular chaperone from a thermophilic archaeobacterium is related to the eukaryotic protein t-complex polypeptide-1. *Nature (London)* **354**:490–493.
55. Trieselmann, B. A., and R. L. Charlebois. 1992. Transcriptionally active regions in the genome of the archaeobacterium *Haloferax volcanii*. *J. Bacteriol.* **173**:30–34.
56. Vinh, D. B., and D. G. Drubin. 1994. A yeast TCP-1-like protein is required for actin function in vivo. *Proc. Natl. Acad. Sci. USA* **91**:9116–9120.
57. Vodkin, M. H., and J. C. Williams. 1988. A heat shock operon in *Coxiella burnetii* produces a major antigen homologous to a protein in both mycobacteria and *Escherichia coli*. *J. Bacteriol.* **170**:1227–1234.
58. Waldmann, T., A. Lupas, J. Kellermann, J. Peters, and W. Baumeister. 1995. Primary structure of the thermosome from *Thermoplasma acidophilum*. *Biol. Chem. Hoppe-Seyler* **376**:119–126.
59. Waldmann, T., E. Nimmesgern, M. Nitsch, J. Peters, G. Pfeifer, S. Muller, J. Kellermann, A. Engel, F.-U. Hartl, and W. Baumeister. 1995. The thermosome of *Thermoplasma acidophilum* and its relationship to the eukaryotic chaperonin TRiC. *Eur. J. Biochem.* **227**:848–856.
60. Waldmann, T., M. Nitsch, M. Klumpp, and W. Baumeister. 1995. Expression of an archaeal chaperonin in *E. coli*: formation of homo- (alpha, beta) and hetero-oligomeric (alpha+beta) thermosome complexes. *FEBS Lett.* **376**:67–73.
61. Willison, K. R., and H. Kubota. 1994. The structure, function, and genetics of the chaperonin containing TCP-1 (Cct) in eukaryotic cytosol, p. 299–312. In R. I. Morimoto, A. Tissieres, and C. Georgopoulos (ed.), *The biology of heat shock proteins and molecular chaperones*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
62. Yaffe, M. B., G. W. Farr, D. Miklos, A. L. Horwich, M. L. Sternlicht, and H. Sternlicht. 1992. Tcp1 complex is a molecular chaperone in tubulin biogenesis. *Nature (London)* **358**:245–248.
63. Zillig, W., P. Palm, H.-P. Klenk, D. Langer, U. Hüdepohl, J. Hain, M. Lanzendorfer, and I. Holz. 1993. Transcription in the Archaea, p. 367–391. In M. Kates, D. Kushner, and A. Matheson (ed.), *The biochemistry of archaea (Archaeobacteria)*, vol. 26. Elsevier, New York, N.Y.