Abstract

Since the inception of this research program, the project has focused on two central questions: What is the relationship between the "eukaryal-like" transcription machinery of archaeal cells and its counterparts in eukaryal cells? And, how does the archaeal cell control gene expression using its mosaic of eukaryal core transcription machinery and its bacterial-like transcription regulatory proteins? During the grant period we have addressed these questions using a variety of in vivo approaches and have sought to specifically define the roles of the multiple TATA binding protein (TBP) and TFIIB-like (TFB) proteins in controlling gene expression in Haloferax volcanii. H. volcanii was initially chosen as a model for the Archaea based on the availability of suitable genetic tools; however, later studies showed that all haloarchaea possessed multiple tbp and tfb genes, which led to the proposal that multiple TBP and TFB proteins may function in a manner similar to alternative sigma factors in bacterial cells. In vivo transcription and promoter analysis established a clear relationship between the promoter requirements of haloarchaean genes and those of the eukaryal RNA polymerase II promoter. Studies on heat shock gene promoters, and the demonstration that specific tfb genes were induced by heat shock, provided the first indication that TFB proteins may direct expression of specific gene families. The construction of strains lacking tbp or tfb genes, coupled with the finding that many of these genes are differentially expressed under varying growth conditions, provided further support for this model. Genetic tools were also developed that led to the construction of insertion and deletion mutants, and a novel gene expression scheme was designed that allowed the controlled expression of these genes in vivo. More recent studies have used a whole genome array to examine the expression of these genes and we have established a linkage between the expression of specific tfb genes and the regulation of nitrogen metabolism and other global cellular responses.

Summary:

Since the inception of this research program, the project has focused on two central questions: What is the relationship between the "eukaryal-like" transcription machinery of archaeal cells and its counterparts in eukaryal cells? And, how does the archaeal cell control gene expression using its mosaic of eukaryal core transcription machinery and its bacterial-like transcription regulatory proteins? In the initial phase we understood that archaeal cells possessed a core RNA polymerase (RNAP) that was closely related in subunit complexity and antigenic reactivity to the eukaryal RNA polymerase II (RNAPII) enzyme and that these cells had at a subset of the eukaryal general transcription factor proteins. Basic questions about promoter structure and regulation were just beginning to be analyzed using newly established in vitro transcription systems for the methanogens, but no experiments had been performed using in vivo approaches. This was the goal of the initial project; namely, to define the DNA sequence requirements for transcription in vivo.

The development of genetic tools for Haloferax volcanii, a halophilic archaeon, in the 1980's provided the first opportunity to examine the requirements for transcription in vivo. In our initial study we developed a plasmid-based in vivo transcription reporter system, using a yeast tRNA gene as the reporter, and examined the promoter requirements for the expression of an H. volcanii tRNA gene, tRNAlys. Results from this study clearly identified sequences in the -25 region, the TATA element, as an essential component of the promoter, and also implicated regions adjacent to the TATA element. The required sequence region upstream of the TATA element was later shown to be the BRE element, a second essential component of the promoter. These results established that the promoter structure in archaeal cells was indeed similar to that described for the RNAPII promoter of eukaryal cells. These results were also consistent with studies from eukaryal systems where it was shown that the TATA region was the binding site for the eukaryal general transcription factor TATA binding protein (TBP) and the BRE element was the binding site for the TFIIB protein.
In the next two grant periods we extended these studies to examine the expression of heat shock genes (cct genes) in *H. volcanii* as a model for regulated gene expression. During this time we also acquired sequence data (pre-genome sequencing) establishing that *H. volcanii* possessed multiple TBP and TFB encoding genes. The occurrence of multiple general transcription factor proteins raised several interesting questions. For example, are the individual *tbp* and *tfb* genes differentially expressed? And if so, do they function to direct transcription of specific genes in a manner analogous to the actions of multiple sigma factors in bacterial cells? Studies on heat shock regulation showed that the chaperone-encoding *cct* genes were controlled at the level of transcription initiation and that the heat shock regulatory elements were located within, or were closely associated with the BRE-TATA regions. The absence of other potential regulatory elements, and the co-localization of the regulatory regions to the core promoter, suggested that regulation might involve the use of alternative TBP or TFBS. A second study was undertaken to determine if the *tbp* and *tfb* genes exhibited regulated expression. Using northern analysis we detected differences in the mRNA levels for the individual *tbp* (4 genes) and *tfb* (6 genes) at differing stages of the growth curve. The most striking differences were between the populations in balanced growth and stationary phase. These trends were also confirmed in western analysis where polyclonal antibodies raised against representative TBP and TFB proteins were used to detect the different proteins. Expression levels of these genes were generally unchanged during heat shock; however, the *tfb* gene showed a strong induction, suggesting that it may play a role in directing the heat shock transcription response. An in vivo analysis of the *tfb* promoter region using the transcription reporter system developed earlier, established that the *tfb* gene was regulated by heat shock and promoter mutagenesis studies led to the discovery that this gene was also controlled in part by an antitermination-like mechanism. In addition to the BRE-TATA regions, the *tfb* gene possessed a strong termination signal in its leader region. Removal of this termination element led to high levels of constitutive expression. These studies established that the *cct* genes were controlled at the level of transcription and that there were strong indicators to support the model that the alternative TBP and TFB proteins may indeed function in regulated gene expression, perhaps functioning in a manner analogous to alternative sigma factors. The *tfb* terminator provided an opportunity to examine the in vivo requirements for transcription termination, a question that had not been addressed. Using a modified form of the tRNA reporter construct we were able to establish that a T-tract of five consecutive T residues functioned as a strong termination signal. Further testing of T-tract sequences revealed a consensus sequence for effective in vivo transcription termination: 5'-TT(T/C)(T/C)TT-3. The analysis of the termination element is ongoing and is now being combined with the results from array studies (described below).

During this time we also began an informal collaboration with Dr. Vito Delvecchio, from the University of Scranton, who had begun a DOE funded project to complete the genome sequence of *H. volcanii* strain DS2 using a directed sequence approach. We assisted in the assembly and gene prediction of sequences from the large plasmids of this organism. A significant amount of sequence data was obtained from this project that was helpful to our studies and to halophile research community in general; unfortunately, the sequence project did not led to a completed genome sequence. The excitement of these preliminary results, combined with a growing number of researchers using *H. volcanii* as a model archaeal system, led to the initiation of a new project to sequence the genome of *H. volcanii*. Dr. Eisen at Tigr directed the new NSF funded project. We again had the opportunity to participate in the analysis of the sequence, and as data were made available, we were able to identify additional genes associated with the transcription apparatus as well as many candidate regulatory proteins. The influx of new data greatly expanded the scope of our initial research goals; for example, these new data established that there were eleven, not six *tfb* genes, and genes encoding two eukaryal general transcription factor proteins, TFIIA (and the related protein RpoM) and TFE were also evident. In this same timeframe there were significant advances in the genetic tools available for *H. volcanii*. We constructed an *H. volcanii pyrF* deletion strain and plasmids carrying the complementing gene. This was a significant advance since the resulting uracil auxotroph could then used to construct deletion strains using a pop-in/pop-out scheme with 5-fluoroorotic acid (5FOA) as the counter selection. The 5-FOA counter selection approach is widely used in yeast cell strain construction.
The \textit{pyrF} deletion and \textit{pyrF} complement gene also allowed us to construct insertion mutants by homologous recombination. Mutant strains generated by this approach were stable if the cells were maintained under selective growth conditions, for example, non-uracil containing medium. The later grant periods built on the newly acquired sequence data and the advanced genetic tools to construct mutant strains lacking transcription related genes. We proposed to use the \textit{pyrF}-system to examine the essentiality of the individual \textit{tbp} and \textit{tfb} genes, and other eukaryal-like transcription related proteins. These later genes included those encoding the archaeal histone, TFIIIS (RpoM), TFE and RNAP subunits that were shown to be non-essential in eukaryal cells. The resulting mutants would be examined for growth defects by examining their growth under a variety of nutrient and stress conditions, and a proteome-based approach would be used to examine the effect of their absence on global gene expression.

We were successful in some aspects of our proposed goals, but we were unable to bring many of the experiments to a level of completeness that would merit publication at that time. Our bottleneck was our inability to achieve high levels of separation of proteins by 2D gels despite our earlier successes in obtaining general separation. Other labs were experiencing similar issues. The extremely low pI of the halophile proteins presented a significant technical challenge. Haloarchaeal proteins have a median pI of 4.3 and there are few buffer systems that allow separation of proteins in this range. Preparation of proteins was also complicated by the presence of 2 M KCl in cell extracts, which is the normal physiological concentration of salt in these cells. During this time period four reports, including two from the same lab, described "new" methods for 2D separation of haloarchaeal proteins. To help in this area we have recently begun a collaboration with Dr. Maupin-Furlow, University of Florida, who is using 2D gel separation and mass spectrometry to characterize \textit{H. volcanii} proteins (see below). While these conditions slowed our progress, we continued with the analysis of \textit{H. volcanii} genome sequence data (this also led to the development of new database tools) and construction of mutant strains. The availability of a draft genome sequence provided an opportunity to approach these questions by measuring transcript levels. We were able to use the data from the DOE funded study to obtain funds from NSF for a SERG award to construct a whole-genome tiled array for \textit{H. volcanii}. This tool has provided an alternative route to examine the roles of the transcription factors. Results of the ongoing studies and the results of the array analyses are summarized below. The publications of these findings will acknowledge both funding sources.

We have continued to make progress in constructing mutant strains and we have developed a system for inducible gene expression in \textit{H. volcanii} that will be of general utility. We have successfully constructed individual mutant strains with gene disruptions in each of the four \textit{tbp} genes. The \textit{tbp}1 and \textit{tbp}2 mutants exhibited a discernable growth lag when grown in defined medium; however, based on these results, we conclude that none of the individual \textit{tbp} genes are essential. We obtained insertion mutants for six of the then identified nine \textit{tfb} genes; these included: \textit{tfb}2, \textit{tfb}4, \textit{tfb}5, \textit{tfb}7, \textit{tfb}8 and \textit{tfb}9 genes. The \textit{tfb}5 mutant strain also exhibited a growth lag. The pop-in/pop-out strategy has been used to construct strains where the \textit{tfb}1 and \textit{tfb}2 genes have been deleted and a strain lacking the \textit{rpoM} (a TFIIIS related protein) gene. The \textit{tfb}1 and \textit{tfb}2 deletion strains are particularly interesting since expression of the \textit{tfb}2 gene is correlated with the heat shock response and \textit{tfb}1 expression increases during nitrogen limitation and growth in stationary phase (see below). We have also made progress in developing new genetic tools. We have constructed a double deletion strain lacking both the \textit{pyrF} and \textit{hisD} genes, and have prepared plasmids with the respective complementing genes. The double deletion strain now provides the flexibility to use a second selection. This is helpful when selecting mutants in pop-in/pop-out schemes; the selection for 5-FOA resistance and histidine prototrophy greatly limits the population of recombinants. The most exciting development is the identification of a positive regulator, HutR, which acts as an inducer of the histidine degradation operon (hutHIGU) when the co-inducer urocanate is present. We have prepared a \textit{hutR} deletion strain, which is the recipient for recombinant plasmids carrying the \textit{hutR} gene and associate \textit{hut} operon promoter region (hutR-module). We have successfully shown that the hutR-module can direct the regulated expression of genes located on the chromosome when introduced by homologous recombination and when the module and a target gene are present on an \textit{H. volcanii} plasmid (pWL500). We are
currently developing plasmids for the re-introduction of the tfb1 and tfb2 genes into their respective deletion strains to assess their roles and to facilitate the testing of mutant forms of these proteins to support our studies on the transcription system.

The availability of the DNA array has greatly extended our ability to measure the expression patterns of the transcription related proteins and to discern patterns of coordinate regulation. We have examined transcriptome for cells growing under a variety of conditions; these include completely defined and complex media, stationary phase, varying salt conditions and starvation for nitrogen and carbon. From these experiments were have established that six of the tfb genes, tfb1-tfb6, are expressed at high levels. The remains tfb genes are expressed at detectable levels under some growth conditions, but the magnitude of their expression is below our significance threshold of the 75th percentile, and their expression does not appear to be regulated by any of the physiological variations we have tested. Each of the four tbp genes was expressed at significant levels despite the genetic data suggesting no individual tbp gene was essential. These genes were not induced under the conditions examined; however, their expression decreased in response to a variety of stresses and their pattern of expression was similar to that observed for the core transcription machinery. This was consistent with the notion that the TBP proteins have a general role and perhaps are redundant; the occurrence of only one tbp gene in some haloarchaea also favors this model. In contrasted to this, the expression patterns of the tfb genes exhibited significantly higher and specific changes during different growth conditions. The most notable correlations were the high expression of tfb1 and tfb5 during nitrogen limitation and stationary phase and the induction tfb4 during growth in the presence of high salt concentrations. Changes in the expression of the tfb1 and tfb5 genes was also correlated with changes in the expression of genes associate with nitrogen metabolism and membrane transport systems. Similarly, changes in the expression levels of genes associated with central metabolism were observed during stationary phase growth, again suggesting a specific role for these transcription factors in directing global gene expression. These initial observations strongly support the model where a change in the expression of tfb genes affects global gene expression patterns of the cell. We have not yet determined whether their control is direct or indirect; numerous transcription regulator genes are also affected and a cascade of regulation is also possible. More experiments are needed to address the precise actions of these GTFs; however, the specific association of tfb gene expression with specific cellular responses, for example the relationship of tfb1 expression to nitrogen metabolism, is physiologically similar to the action of alternative sigma factors in bacterial cells. In this regard, the haloarchaea appear to be more versatile, and have possibly evolved more complex regulatory systems than those observed for the other currently described Archaea.

Traineeship: Students and postdoctoral fellows.

This research program has provided partial support for seven Ph.D. students, four have completed their degree and three are ongoing, and three M.Sc. students. They are: Ph.D Students completed: Ms. Yen Ping Kuo, Ms. Dorthea Thompson, Mr. David Armbruster and Mr. William Ray; Current Ph.D Students: Ms. Anice Sabag-Daigle, Ms. Shoko Morimoto, and Mr. Rick Nist; M.Sc. Students completed: Ms. Aparna Donti, Mr. Brian Gabel, and Ms. Kara Bette. The work performed by these students has been presented at national and international meetings; abstract titles are provided below. Dr. John Palmer, postdoctoral fellow, also received partial support from this program.

Publications and presentations at national and international meeting.

Fifteen publications acknowledge support from this program and six additional manuscripts are in preparation. Work from this project has presented at both national and international meetings. Titles of these presentations are presented below.
Student presentations:


Archaeal tRNA Genes are Transcribed by a Eucaryal-like Transcription System. J.R. Palmer*, D.K. Thompson, and C.J. Daniels. 15th International tRNA Workshop, June 1995


Invited presentations by PI at national and international meetings.


Genetics in the Halophilic Archaea. ASM General Meeting, May 2003.


Multiple TBP and TFB genes in the Archaea: extending the RNA polymerase II transcription system. Society for Industrial Microbiology, July, 2000


Multiple Transcription Factors in the Halophilic Archaea Gordon Research Conference on Archaea, July 1996.

Transcription in the Archaea and the Orgins of the Eucaryal Transcription Pandigm Divisional Lecturer; American Society for Microbiology, May 1996.

Control of tRNA Gene Expression in the Archaea. 15th International tRNA Workshop, Madison WI. June 1995.

Publications related to this grant.


Publications in preparation:


