DISCOVERY OF NEW REGULATORY TRANSCRIPTION FACTORS IN

PYROCOCCUS FURIOSUS IN METAL STRESS AND COLD SHOCK

by

GUANQUN YUAN

(Under the Direction of Robert A. Scott)

ABSTRACT

All living cells need to adapt to stress conditions in their environment. Under stress, the cells' primary response is at the level of transcription. This response may involve the binding of regulatory transcription factors to specific DNA sequences near the promoters or transcription initiation sites. Experiments have been done to retrieve the regulatory transcription factors in the model prokaryote, *Pyrococcus furiosus*, under metal (Fe) deficiency and cold shock. A putative regulatory transcription factor PF0250p has been found and binding was verified.

INDEX WORDS: Archaea, *Pyrococcus furiosus* (*Pf*), transcription, regulatory transcription factors (rTFs), metal stress, cold shock, protein capture assay, electrophoretic mobility shift assay (EMSA), fluorescence DNase footprinting, PF0250p

DISCOVERY OF NEW REGULATORY TRANSCRIPTION FACTORS IN

PYROCOCCUS FURIOSUS IN METAL STRESS AND COLD SHOCK

by

GUANQUN YUAN

B.S., Fudan University, Shanghai, China 2003

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2007

DISCOVERY OF NEW REGULATORY TRANSCRIPTION FACTORS IN

PYROCOCCUS FURIOSUS IN METAL STRESS AND COLD SHOCK

by

GUANQUN YUAN

Major Professor:	Robert A. Scott
Committee:	I. Jonathan Amster
	Jeffrey L. Urbauer

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

For survival, all cells must be able to adapt in response to changes in their environment [1]. Environmental stress may involve depletion of an essential nutrient, the presence of a toxic species, or suboptimal conditions (temperature, pH, salinity, etc.). Under stress, the cell's primary response is at the level of transcription, in which expression levels are altered for a set of key genes. One mechanism used to affect this response involves the binding of regulatory transcription factors (rTFs) to specific DNA sequences (operators) near the promoters or transcription initiation sites of the target genes. My research goal is to discover new rTFs used in response to specific stresses (e.g., cold shock and iron depletion). This will be accomplished by using genomic and proteomic technologies that allow discovery and characterization of the rTF, its operator, and the genes regulated by the rTF-operator combination. As a test case, I have conducted this research on a model prokaryote, the hyperthermophilic marine archaeon *Pyrococcus furiosus* (*Pf*).

1.1. Archaeal Transcription and Regulation

1.1.1. Archaea

All cellular organisms can be divided into three domains: Eucarya, Archaea, and Bacteria [2]. Figure 1.1 shows a phylogenetic tree of cellular organisms.
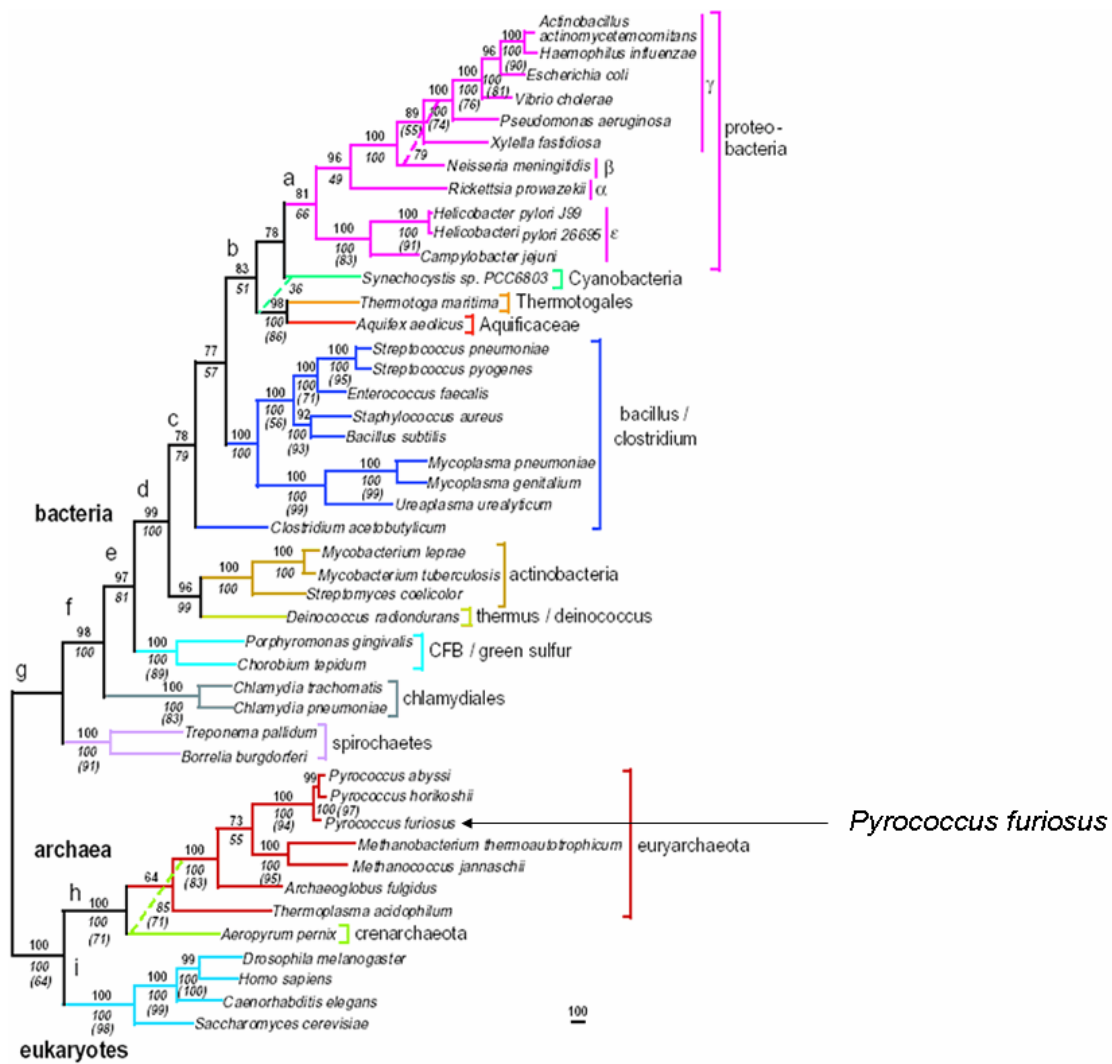
Figure 1.1. The phylogenetic tree of cellular organisms [2, 3]. The position of the *Pf* genome is indicated by the arrow.

2

Many of the prokaryotic organisms that are found in extreme environments (hot, cold, extremes of pH or salinity) belong to the archaeal domain. Such extreme conditions exceed the optimal conditions for growth and reproduction for the majority of organisms. Archaea are similar to bacteria in many aspects of structure and metabolism. However, the genetic transcription and translation machinery of archaea is more similar to that of eukaryotes, rather than that of bacteria [4]. On the other hand, the examples of transcriptional regulation that have been characterized in archaea show similarity to bacterial regulatory systems. This makes the archaeal domain a good testing ground for discovery of transcriptional regulatory mechanisms.

1.1.2. Transcription in Archaea

Shared by all three domains of life, transcription is the process whereby a sequence is transcribed into a complementary RNA sequence by a group of enzymes denoted as DNA-dependent RNA polymerases. Simply put, it is the transfer of genetic information from DNA into RNA.

In the past two decades, significant advances have been achieved in the understanding of archaeal transcription. Like bacteria, archaea contain only one RNA polymerase, but the transcription processes are significantly different. Bacterial promoters consist of two consensus sequences located at -10 and -35 bp upstream of the transcription start site, and a family of proteins known as sigma factors recognize the two sequences and help initiate the RNA polymerase transcription process. Though archaea only contain one RNA polymerase like

bacteria, the subunits in the archaeal RNA polymerase are homologous to those of eukaryotes in both structure and function. Indeed, the mechanism of archaeal transcription more closely resembles that of the eucaryal RNAP II system [4-7]. Like eucarya, the archaeal RNA polymerase is not capable of recognizing promoter sequences on its own. It must rely on additional general transcription factors, proteins that recruit RNA polymerase during initiation of transcription. Two basal transcription factors，TATA-box-binding protein (TBP) and transcription factor B (TFB), are indispensable in the process of transcription [8-10].

For both eucarya and archaea, an AT-rich sequence (often TATA, and referred to as a TATA box) acts as the promoter and is recognized by the TATA box-binding protein (TBP). The position of the TATA box is about -26 bp relative to the transcription start site in archaeal genes. Subsequently, the TFIIB-related transcription factor B (TFB) binds to the TBP–DNA complex. Then RNA polymerase (RNAP) binds to the TFB-TBP-promoter complex and initiates transcription. The adjacent purine-rich BRE (transcription factor B recognition element, just upstream of the TATA box) directs efficient pre-initiation complex assembly, which helps the sequence-specific interactions with TFB and determines the orientation of the transcription complex [4, 11, 12]. Figure 1.2 shows the mechanism of transcription in archaea compared to the mechanism of transcription in bacteria.
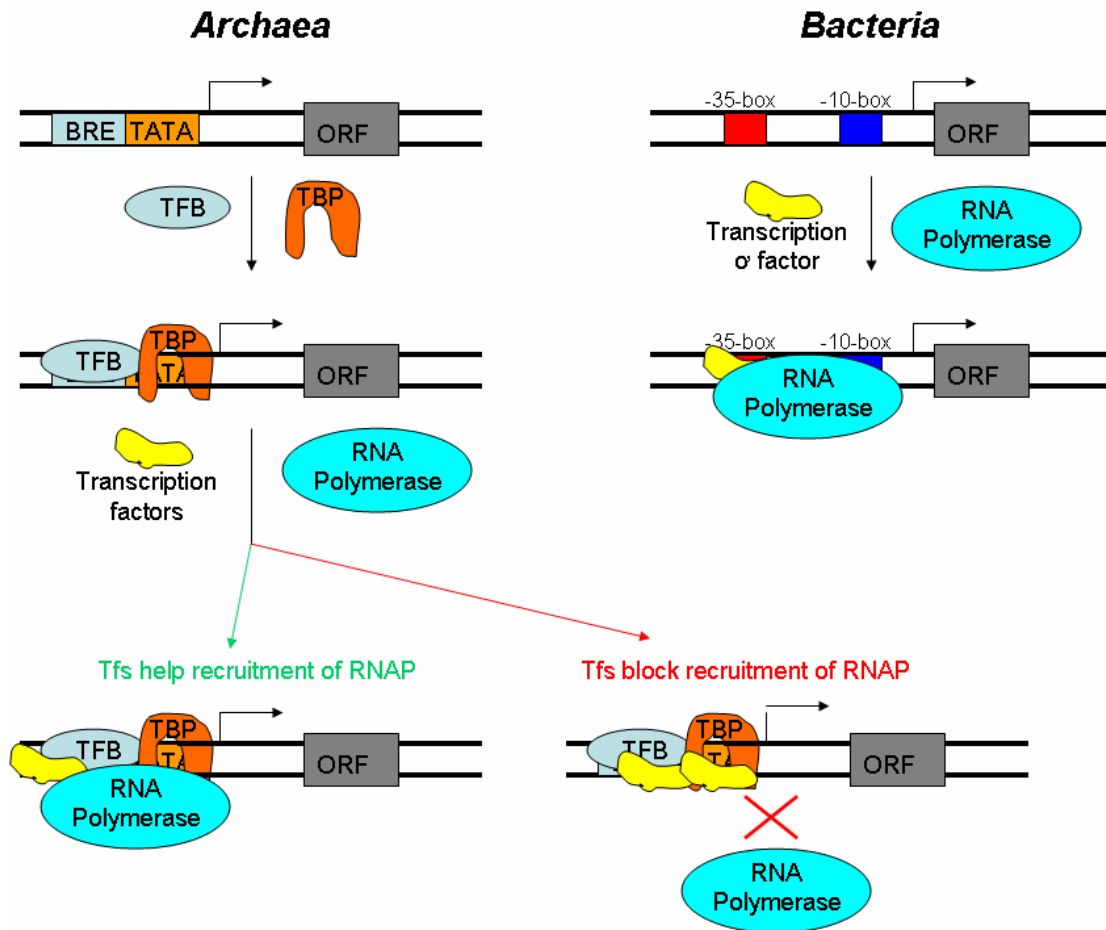
Figure 1.2. The mechanism of basal transcription in archaea, compared with that of bacteria. In archaea, the general transcription factor TBP binds to the TATA box and TFB binds to BRE and TBP-TATA. Subsequently, they recruit the RNA polymerase (RNAP) and initiate transcription. Regulatory transcription factors can either help or block the recruitment of the RNAP. The positioning of the factors in the above diagram is based on a review of transcription in archaea [13, 14]. Bacteria contain two consensus promoter sequences located at -10 and -35 bp upstream of the transcription start site. Sigma factors, a family of proteins, recognize the two sequences and help initiate the RNAP transcription process.

5

1.1.3. Transcriptional Regulation in Archaea

Regulatory transcription factors (rTFs) can either increase the rate of transcription (called activators) or decrease the rate of transcription (called repressors) for a given gene. To date, only a few transcriptional regulatory systems have been reported in archaea [4, 15, 16]. Most of the reported rTFs act as transcriptional repressors. Studies showed two possible transcriptional repression mechanisms in archaea. In one mechanism, rTFs can bind to or near the TATA/BRE promoter sequence, inhibiting TBP/TFB binding to the promoter [13, 17, 18]. The other repression mechanism involves rTF binding near or at the transcription start site, inhibiting the recruitment of RNAP [11, 19, 20].

Bell summarized the archaeal transcriptional regulators and their modes of action. This summary is shown in Table 1.1 [13].

The Lrp (Leucine-responsive Regulatory Protein) family contains the best studied archaeal regulatory transcription factors. An Lrp-like protein Lrs14 of *Sulfolobus solfataricus*, a hyperthermophilic archaeon, is an autoregulator. Gel shift and DNase I footprint experiments verified that this Lrs14 protein specifically binds to multiple DNA sequences in its own promoter, overlapping the TATA box [18]. An LrpA regulatory transcription factor from *Pyrococcus furiosus* also shows negative autoregulation (the gene is repressed by its own gene product) [21].

Table 1.1. Summary of Identified transcription factors in Archaea [13].

Name	Species	Activator or repressor	Ligand	Mechanism
GvpE	Halophilic archaea	Activator	Unknown	Unknown
LrpA	P. furiosus	Repressor	Unknown	Blocks RNAP recruitment
Lrs14	S. solfataricus	Repressor	Unknown	Blocks TBP and TFB recruitment
LysM	S. solfataricus	Possible activator	Lysine	Unknown
Mdr1	A. fulgidus	Repressor	Metal ions	Blocks RNAP recruitment
NrpR	M. maripaludis	Repressor	2 oxo glutarate	Possibly blocks RNAP
PhrA	P. furiosus	Repressor	Unknown	Blocks RNAP recruitment
Ptr2	M. jannaschii	Activator	Unknown	Facilitates TBP binding
Ss-LrpB	S. solfataricus	Unknown	Unknown	Unknown
TrmB	T. litoralis	Repressor	Sugars	Probably TBP and TFB recruitment

Metal-dependent repressor 1 (MDR1) from *Archaeoglobus fulgidus* is regulated by metal ion availability [22]. In vitro experiments indicate that the MDR1 protein specifically binds to three operator elements in its own promoter in a metal-dependent manner. MDR1 negatively regulates transcription of its own gene by preventing the recruitment of RNA polymerase.

There is still much to do to understand transcriptional regulation in archaea. Many questions remain unanswered: if the regulatory systems of archaea resemble those in bacteria, why does it have eukaryotic-like transcriptional machinery? What transcriptional regulatory mechanisms do archaea use to survive under extreme environmental stress?

1.2. Discovery of Stress-related rTFs in *Pf.*

1.2.1. The Purpose of the Study

We have developed an approach called STRES (Survey of Transcriptional Response to Environmental Stress) to discover rTFs and their operators in *Pf*, a model prokaryotic organism. STRES involves technologies including microarray expression profiling, bioinformatics, DNA-affinity protein capture, and protein-DNA biophysical methods to verify rTF binding sequences (operators) and to investigate transcriptional regulatory mechanisms. I used STRES to study the transcriptional regulatory response of *Pf* to cold shock and to iron depletion.

The *Pf* gene expression profile of cold shock and iron depletion has been published as a genome-wide microarray study [23, 24]. The main purpose of

8

using these microarray assays is to show differential expression levels between two growth conditions (before and after cold shock, with or without iron metal) in order to tell what genes are activated or repressed. But the microarray expression profile alone is not enough to determine the transcriptional regulatory mechanism. As a result, in the STRES approach, we integrated the DNA-affinity protein capture and biophysical methods to identify the transcription factors and discover their binding sites, and bioinformatics to investigate rTF binding sites genome-wide.

1.2.2. *Pyrococcus furiosus (Pf)*

Found most often at the vents of the seafloor near volcanoes, *Pf* is a hyperthermophilic species of Archaea that grows at temperatures between 70 ºC and 103 ºC and between pH 5 and 9. It grows optimally near 100 ºC and at a pH near 7. It consumes carbohydrates or peptides as its carbon and energy sources [25]. The complete sequence of the *Pf* genome was reported in 2001 [26, 27]; the latest annotation shows that the whole genome contains 2125 genes and is approximately 1.9 Mb in size.

Choosing *Pf* as a model system has several advantages [28]: (1) the *Pf* genome size is small; (2) *Pf* can use a wide variety of both simple and complex carbohydrates as carbon sources and either peptides or ammonia as a nitrogen source; (3) *Pf* has a well-regulated metabolism which occurs mainly at the transcriptional level; (4) the similarities of transcription initiation components

between archaea and eukaryotes [29] make *Pf* an excellent model to study what could correspondingly be applied to eukaryotic biology.

1.2.3. Metalloregulation

Many cells require the presence of metal ions in enzymatic reactions or as components in proteins [30]. Maintaining an optimum level of essential metal ions in the cell is very important. Metal homeostasis involves the uptake and storage of the beneficial metal [31], the digestion or excretion of excessive amounts of metal [32], and detoxification of toxic metals [33]. Often these functions in the cell are controlled at the level of transcription by metalloregulatory proteins [34].

As mentioned above, the only well studied metal-related transcriptional regulator in archaea is metal-dependent repressor 1 (MDR1) from *Archaeoglobus fulgidus*, regulated by metal ion availability [14]. Several metal-related transcriptional regulators have been studied in bacteria. The major regulator of Fe homeostasis in bacteria is the ferric uptake regulator (Fur) protein [35, 36].

1.2.4. Cold Shock

Environmental temperature change is one of the most often performed stress experiments, applied to bacteria, plants, and animals [37]. A sudden temperature drop stops the synthesis of most of the proteins in the organism, but a small number are induced and referred to as cold shock proteins. The hypothetical function of producing cold shock proteins is to maintain the normal function of the organism to the extent possible [38]. Many proteins are well-characterized as

10

cold shock proteins: CspA of *Escherichia coli* [39], initiation factor 2 (IF2) [40], ribosomal binding factor A (RbfA) [41], and DEAD-box RNA helicase [42]. The cold shock proteins in bacteria have been reviewed recently [43].

1.3 Chemostatic Method for Growth of *Pf*

Cells of *Pf* were grown under different conditions to prepare total RNA for the microarray experiments and cell extract for the DNA-affinity protein capture experiments. The Adams lab has well established technology developed using microarrays for expression profiling in *Pf*. They have developed a chemostat (called the "Gertitron" after its inventor, Gerti Schut) for growing *Pf* under dynamically controlled conditions.

Pf cells were grown with maltose as the primary carbon source in a 20-liter custom fermentor, using the "standard" medium as reported [44]. For microarray profiling, there are at least two types of growth conditions compared, condition A versus condition B. Condition A is the control growth, which in our case is standard maltose growth, while condition B is the stress growth, e.g. cold shock or Fe limitation.

1.3.1 Significance of Microarray Expression Profiling

Microarray technology is a powerful high-throughput method to determine the level of transcriptional regulation, in our case in *Pf*. Each ORF (open reading frames) in *Pf* is replicated and ligated onto a specific spot on a chip. mRNA is isolated from the cell extracts of *Pf*, grown under two separate conditions (e.g.,

11

with sufficient metal or depleted of metal). The mRNA is reverse-transcribed into cDNA which is labeled with two different fluorescent tags, one for each growth condition. The fluorescent-tagged cDNA from the two growth conditions is mixed and hybridized with the DNA on the chip. If a particular ORF is expressed only in one condition, the place where it is ligated on the chip will only show one fluorescent emission. If the ORF is expressed in both conditions it will show both of the fluorescence "colors". Intensities of the two fluorescence emissions of each spot on the chip are examined to determine the amount of mRNA resulting from each ORF under each condition. This technique makes possible rapid determination and comparison of the regulation of all ORFs within *Pf* under different growth conditions [23, 45].

If a certain gene is highly up-regulated or down-regulated in one condition versus another, some rTF may have been involved. In transcriptional regulation, the rTF binds to its operator near the promoter sequence upstream of the gene, which either helps the transcription (up-regulated) or hinders the transcription (down-regulated). The microarray data reveal both up- and down-regulated genes. From this list, a specific gene is chosen. The promoter sequence of the gene is cloned and used as bait to capture specific binding proteins, among which potential rTFs may be found. This is the idea of the DNA-affinity protein capture assay (Chapter 2).

1.3.2. Chemostatic Growth and Microarray Results for Cold Shock

Microarray expression profiles of the *Pf* response to cold shock have been reported [24]. Two types of growth experiments were performed, termed adapted and cold shock growths. For the adapted growth experiments, cultures were grown at 95 and 72 °C from start to finish. In the cold shock growth experiments, cultures were started at 95 °C for 2 h. As the cell density reached 3×10^7 cells/mL, they were rapidly cooled to 72 °C over a period of 15 min, by pumping the cells through a glass cooling coil at a constant temperature of 20 °C. Then cells were harvested 1, 2, and 5 h after the shock, and cells were lysed [23, 24].

In the microarray experiment for the adapted growth at 95 °C and 72 °C, 245 ORFs exhibited significant up- or down-regulation (2.5-fold or higher change in expression). For cells subjected to 1, 2, or 5 h of cold shock and cells adapted to 72 °C, there were 49, 35, 30 and 59 ORFs up-regulated, respectively [24]. Table 1.2 summarizes some data from DNA microarray analyses of mRNA levels from cold shock experiments. Genes PF0029, PF0030 and PF0031 belong to the same operon and are up-regulated after 1 h cold shock with about 4- to 5-fold increase, which is significant among all the 1 h up-regulated genes. PF0031 is the first gene of the operon, so I picked it as the target gene. Genes PF0324, PF0325, and PF0326 are significantly up-regulated after 5 h of cold shock and genes PF1479 and PF1480 are significantly up-regulated after 1, 2, and 5 h of cold shock. Both these operons are transcribed in the forward direction, so I also picked genes PF0324 and PF1479 (first gene of each operon) for later protein capture experiments.

1.3.3. Chemostatic Growth and Microarray Results for Iron Depletion

Microarray expression profiles testing the response of *Pf* to iron deficiency have been carried out by Angeli Menon in the Adams' laboratory [46]. "Control" *Pf* medium (with maltose) contains a final concentration of 7.4 µM Fe [23, 24, 46]. Fe-limited medium (Fe < 0.8 µM) was obtained by omitting Fe (and using acid-washed glassware and high-purity water). Some experiments were also done by adding bathophenanthroline sulfonate (BPS) to the growth medium to chelate extra iron. Cells were grown at 95 °C under standard maltose conditions in the absence of S^0, and the cell extract was collected and used for the protein capture experiment (Chapter 3). Table 1.3 summarizes data from DNA microarray expression profiles from three Fe-sufficient and three Fe-limited batch cultures. 28 ORFs were up-regulated >2.5 fold and 30 ORFs were down-regulated by >2.5-fold by Fe limitation. Again I picked three genes that showed either large up-regulation or down-regulation for later protein capture experiments. Gene PF0723 is significantly up-regulated (7.5-fold) after iron depletion. Gene PF0857 shows 4.7-fold up-regulation. On the other hand, gene PF0340 is significantly down-regulated by 4.8-fold.

Table 1.2. Relative RNA levels from DNA microarray expression profiles of cold shock experiments [24]. Numbers are $\log_2$ of relative expression levels (given in parentheses).

ORF Number	Description	1 hr [a]	2 hr [b]	5 hr [c]
PF0029	[Hypothetical protein]	2.2±0.6 (4.6)	1.8±0.8 (3.5)	1.4±0.6 (2.7)
PF0030	[Conserved hypothetical protein]	2.3±0.7 (4.8)		1.9±0.6 (3.7)
PF0031	[Threonine synthase]	2.3±0.7 (4.9)		
PF0324	[Conserved hypothetical protein]		2.4±0.9 (5.2)	4.6±2.4 (23.8)
PF0325	[Conserved hypothetical protein]		2.3±0.6 (4.8)	4.4±1.1 (21.3)
PF0326	[Conserved hypothetical protein]			2.3±0.9 (4.9)
PF1479	[Oxidoreductase, Fe-S subunit]	2.3±0.8 (4.9)	2.5±0.5 (5.6)	
PF1480	[Formaldehyde:Fd oxidoreductase wor5]	2.3±0.4 (5.0)	2.4±0.4 (5.1)	2.1±0.7 (4.2)

a   the cell extract collected after 1hr cold shock
b   the cell extract collected after 2hr cold shock
c   the cell extract collected after 5hr cold shock

Table 1.3. DNA microarray analyses of relative RNA levels from iron depletion experiments showing ORFs whose expression are up-regulated (red) and down-regulated (green) [46].

Gene Mate ORF number	GeneMate Name or General Function	Fold
PF0723	**the putative high-affinity iron transporter**	**7.5**
PF0857	**ferrous iron transport protein b**	**4.7**
PF1285 & PF1286	the putative FeS cluster assembly/repair proteins	**5.0 & 6.1**
PF1657-PF1666	two amino acid biosynthetic clusters	up to **5.8**
PF0935-PF0942	potentially involved in histidine biosynthesis	up to **4.5**
PF0725-PF0728	conserved hypothetical Proteins	up to **3.5**
PF1282	rubredoxin	**5.5**
PF0346	aldehyde:ferredoxin oxidoreductase	**3.0**
PF1893	transcriptional regulatory protein, asnC family	**3.0**
PF2051	arsr family transcriptional regulator	**2.2**
PF0340	**putative HTH transcription regulator**	**4.8**
PF0339	methyltransferase	**4.4**

CHAPTER 2

DNA-AFFINITY PROTEIN CAPTURE EXPERIMENTS TO CATCH

POTENTIAL TRANSCRIPTIONAL REGULATORS

The DNA-affinity protein capture protocol was developed as a collaboration with former group members, originally described in Yu Chen's MS thesis [47]. This experiment utilizes DynaBead-bound (immobilized) dsDNA containing the DNA sequence upstream of a selected gene. After mixing with cell extract, a series of incubation, washing, heparin challenge, and recovery steps are used to remove non- and weakly binding proteins. Proteins retained by the DNA are visualized by SDS-PAGE and identified by standard proteomics MS identification. The basic procedure involves in-gel tryptic digestion, MALDI-TOF, and data analysis using peptide mapping software like MSfit and Mascot. Finally the proteins can be prioritized for further experiments [28].

2.1. Materials and Methods

2.1.1. Materials

2.1.1.1. Buffers

Table 2.1 lists the buffer necessary for the DNA-affinity protein capture experiments.

Table 2.1. Buffers for the DNA-affinity protein capture experiments

Buffer Name	Components
A. 2xB&W buffer, 50mL	10mM Tris
	1mM EDTA
	2.0M NaCl
	adjust the pH to ≈7.0 by HCl
B. Pull-down assay washing buffer, 50mL	50mM Tris,
	1mM EDTA
	100mM KCl
	adjust the pH to ≈7.0 by HCl
	5% Glycerol
	1mM DDT,
	0.1% TritonX100
C.Heparin challenge buffer, 50mL	0.1mg/mL Sodium Heparin in the washing buffer B above.
	ph≈7.0
(The above buffers need to be filtered with Whatman filter)	
D.4xLaemmli buffer(Protein SDS gel electrophoresis sample mixing buffer), 10mL	156mM Tris-HCl, pH≈6.8
	50% glycerol
	8% SDS
	0.1% Bromophenol blue
	store at 4℃.

2.1.1.2. DNA Oligonucleotides

All unlabeled or biotinylated DNA primers were purchased from Integrated DNA Technologies, Inc. The PF1601 promoter was obtained from former group member Dr. Meiyao Wang.

2.1.1.3. *Pf* Cell Extract Preparation

15 L of *Pf* cell culture were grown in anaerobic conditions as discussed in Chapter 1, control condition vs. cold shock conditions and control condition vs. iron depletion conditions. The concentration of protein in the cell extract was determined by DC Protein Assay Kit II (BioRad).

2.1.1.4. Other Reagents and Materials

Table 2.2 summarizes all the chemicals and other materials used in the experiments and their sources.

2.1.2. Preparation of Promoter DNA by PCR

2.1.2.1. Design of DNA Probe

The characteristics of Archaeal transcription are similar to eukaryotic class II transcription, with TATA elements about 26 bp upstream of transcription start sites. However, for most of the genes in the *Pf* genome, we only know the translation start sites. So in order to design a promoter sequence that is assured to contain any regulatory transcription factor binding site, we often design the

promoter DNA from a few hundred bp upstream of the translation start to about one hundred bp downstream from the translation start site. The promoter DNA is designed with one biotinylated primer to allow attachment to Steptavidin-coated DynaBeads.

2.1.2.2. PCR Protocol

The PCR protocol was obtained from Sigma. The components are listed in Table 2.3.

2.1.2.3. Gradient PCR

The best annealing temperature varies with different primer pairs. In order to identify the optimal PCR conditions for each pair of primers, gradient PCR was usually performed.

2.1.2.4. DNA Electrophoresis

The PCR product was checked by gel electrophoresis on a 1.0-1.2 % agarose gel, run at a constant 75 V for 40 min and stained with ethidium bromide.

2.1.2.5. dsDNA Purification

PCR products were not pure enough for further experiments, since they contain PCR primers and template genomic DNA. PCR products were purified with a QIAquick PCR Purification Kit to remove extra biotinylated primers. If additional purification was necessary, the QIAquick Gel Purification Kit was used.

Table 2.2. Summary of all the chemicals used

Name of the Chemical	Company
a-cyano-4-hydroxycinnamic acid	Sigma
Ammonium bicarbonate	Fisher Scientific Inc.
Bromophenol blue	Fisher Scientific Inc.
Coomassie Brilliant Blue (CBB)	Fisher Scientific Inc.
Dithiothreitol (DTT)	Fisher Scientific Inc.
dNTP	Sigma
Dynabeads M280 Streptavidin	DYNAL Biotech.
Gel Purification kit	Qiagen
Gelcode Blue staining	Pierce Biotechnology Inc.
Hitrap Chelating HP column	Amersham Biosciences
IPTG	Alexis Biochemicals
Kanamycin	Sigma
MPC magnet	DYNAL Biotech.
PCR Purification kit	Qiagen
Plasmid Midi kit	Qiagen
Silver nitrate	Fisher Scientific Inc.
Sodium heparin	Fisher Scientific Inc.
Taq DNA polymerase and buffer	Fisher Scientific Inc.
Taq DNA polymerase and buffer	Qiagen
Tris base	Fisher Scientific Inc.
Trypsin	Promega

Table 2.3. The components of the PCR reaction

Components	Volume
MgCl$_2$	6µl (25µm/l)
Template	2µl(10ng/µl)
10X buffer	10µl
dNTP	2µl (200µmol)
Primer A	1µl (40pmol)
Primer B	1µl (40pmol)
Taq DNA polymerase	1µl
H$_2$O	77µl
Total	100µl

2.1.2.6. DNA Quantification

Purified DNA was quantified using a Hoefer DyNA Quant 200 fluorometer,

2.1.3. DNA-affinity Protein Capture [47, 48]

Figure 2.1 shows a schematic flow chart of the DNA-affinity protein capture approach [48]. The procedure was slightly modified based on the protocol obtained from former group member Dr. Meiyao Wang [48]. In order to immobilize the DNA, I washed 100 µL beads with 100 µL 1x B&W Buffer three times and then added 100 µL DNA solution to the beads (the DNA concentration was about 50 ng/µL), and incubated at room temperature for 20 min. I removed the DNA solution and washed the beads one more time with 100 µL 1x B&W buffer. I added 100 µL cell extract (the concentration was about 3 mg/mL) to the immobilized DNA and incubated at 55 ºC for 30 min. After incubation, I washed the beads with 50 µL washing buffer three times and 50 µL heparin challenge buffer. To retrieve proteins from the beads, I resuspended 40 µL beads in 1x Laemmli buffer and incubated the beads at 55 ºC for 10 min. Finally I boiled the suspensions for 7 min and loaded all the samples onto a 4-20% Tris/HCl gel for electrophoresis. The gel was run at 200 V for 60 min. The running buffer was tris/glycine/SDS. After gel electrophoresis, the gel was stained using either silver staining or Commassie brilliant blue staining. Finally the proteins were identified by mass spectrometry following in-gel tryptic digestion.
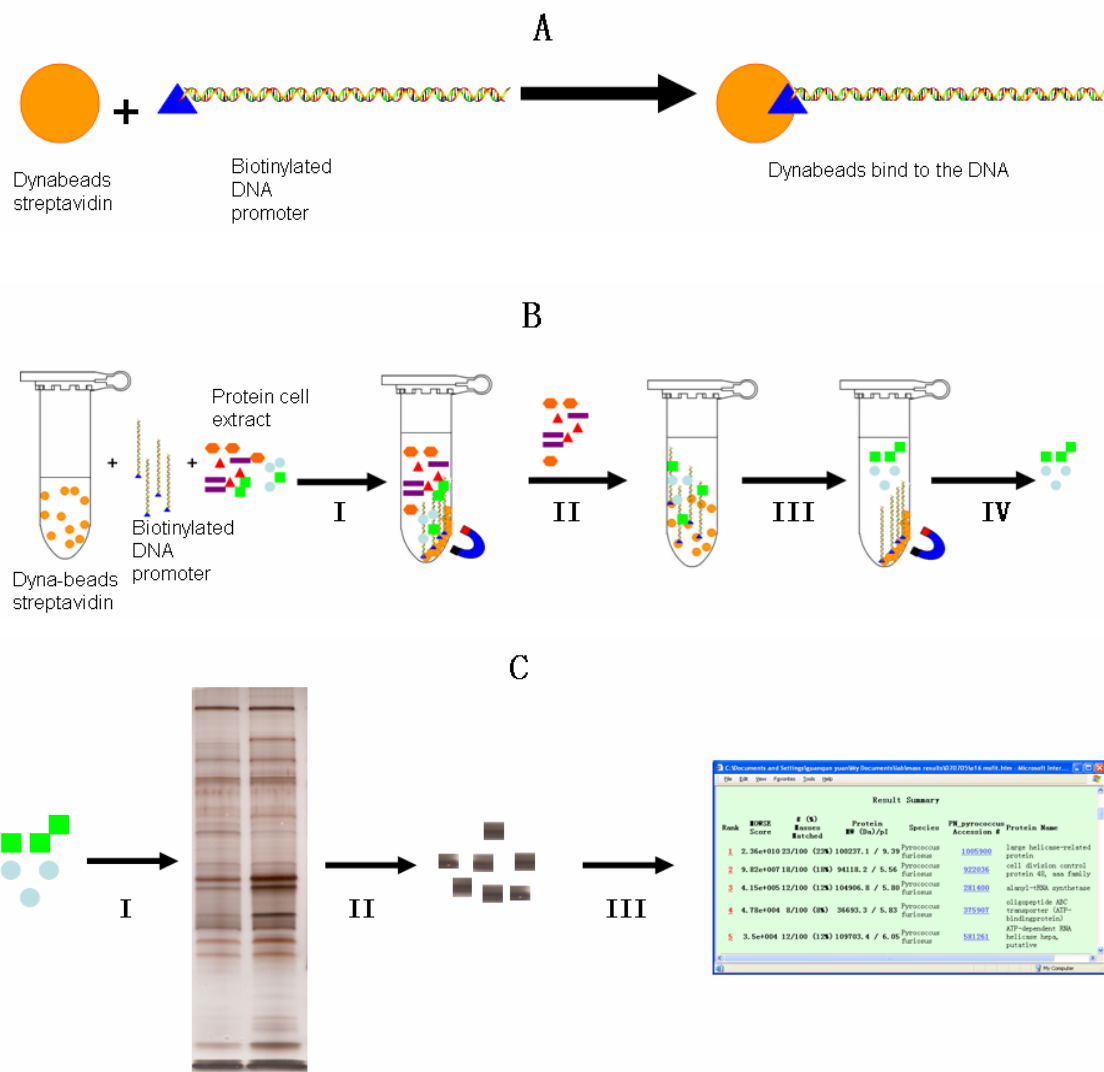
Figure 2.1. Schematic flow chart of DNA-affinity protein capture to identify potential transcription regulators [47, 48].

A. Streptavidin-coated Dynabeads bind to the biotinylated DNA promoter.

B. (I) The bead-bound immobilized DNA is incubated with protein cell extract. (II) DNA binding proteins form a protein/DNA complex with the immobilized DNA. Then a magnet is applied to separate the beads together with the protein/DNA complex from the non DNA-binding proteins. (III) The protein/DNA complex is then dissociated and the proteins that bind to the DNA are released. (IV) The magnet is applied again to separate the bead-bound DNA from the proteins in the solution. The purpose is to collect the eluted protein for future experiments.

C. (I) The protein collected from the previous step is loaded onto the SDS-PAGE gel to separate the proteins by mass. (II) Selected bands are cut from the gel and subjected to in-gel tryptic digestion. (III) The proteins are then identified by mass spectrometry.

24

2.1.4. In-gel Digestion and Mass Spectrometry

2.1.4.1. In-gel Digestion [47, 48]

After a selected gel band was cut out and chopped up, it was rehydrated with 50 µL 50 mM NH_4HCO_3 for 10 min and dehydrated in 80 µL 50% acetonitrile in 50 mM NH_4HCO_3 for 15 min. I repeated the rehydration/dehydration cycle twice more. The dried gel spots were vacuum-centrifuged for about 10 min in order to further dry the gel. Then the gel was resuspended in 10 ng/µL trypsin in 50 mM NH_4HCO_3 solution and left on ice for 30-40 min to absorb trypsin. The gel spots were then incubated at 37 °C for 4.5 h to overnight. In order to extract the peptides, I washed the gel spots with 10-15 µL 25 mM NH_4HCO_3, then 10-15 µL 75% acetonitrile, 0.5% TFA in water, twice. I collected all the supernatants and concentrated the solution to 4-5 µL by vacuum centrifugation. Finally, I added 1 µL 5% TFA to make a final volume of 5-6 µL and a final TFA concentration between 0.1-1%.

2.1.4.2. Mass Spectrometry Analysis

The tryptic peptides were then loaded onto the plate ready for MALDI-TOF analysis by using ZipTip$_{\mu\text{-}C18}$ (Millipore) tips following the manufacture's instructions. I washed the tip three times with acetonitrile followed by three times with 0.1% TFA. To apply the protein to the tip, I pipetted the tip up and down in the tryptic peptides solution, and then washed the tip three times with 0.1% TFA. Finally, I eluted the peptides from the Ziptip with 2 µL freshly made saturated

matrix solution (α-cyano-4-hydroxycinnamic acid in 50% acetonitrile, 0.1% TFA) and spotted the solution directly onto a MALDI plate. The plate was sent to the Chemical and Biological Sciences Mass Spectrometry Facility, UGA, for MALDI-TOF analysis. The mass spectra were calibrated using trypsin autolysis products as internal standards.

MALDI-TOF results in a list of tryptic peptide masses that are analyzed using peptide mass mapping (PMM). Each protein produces a unique list of peptide fragments that can be predicted from known protease cleavage specifities. PMM is a method for identifying proteins by comparing observed mass (m/z) with predicted masses of digested proteins listed in a database. The list of masses was searched in the NCBI genome database of archaeal species using the MS-Fit program (as made available at amster.chem.uga.edu) or the MASCOT software available on the web (http://www.matrixscience.com/).

2.2 DNA-Affinity Protein Capture Experiments to Catch Potential Transcriptional Regulators of Cold Shock Response

Based on the cold-shock microarray data (Chapter 1), I chose three genes that were strongly up-regulated to use in DNA-affinity protein capture experiments. Gene PF0031 is significantly up-regulated 1 h after cold shock. Gene PF0324 is significantly up-regulated 5 h after cold shock. Gene PF1479 is significantly up-regulated at all three time intervals (1, 2, and 5 h after cold shock).

2.2.1 DNA-Affinity Protein Capture Experiments with Promoter of PF1479

2.2.1.1 Promoter of PF1479

Pf genomic DNA was used as template to PCR amplify the region from genome coordinates 1382719 to 1383019 which spans -200 to +100 relative to the translation start of PF1479 (Figure 2.2). Figure 2.3 shows the gradient PCR results for amplification of the PF1479 promoter.

2.2.1.2 DNA-Affinity Protein Capture Results

The PF1479 promoter DNA was immobilized and used for DNA-affinity protein capture, as described. Two separate experiments using either cell extract from standard growth conditions or extract from cells harvested 5 h after cold shock were compared by SDS PAGE separation of the retained proteins. If a gel band appeared with higher intensity in one condition rather than the other, this band was cut out and subjected to in-gel digestion and MS identification. Six bands showed higher intensity in 5 h cell extract and were identified in this way (Figure 2.4). These proteins were candidates for potential promoter-binding proteins that respond to cold shock. Band 3 was identified to be PF0250p, which has a MASCOT score of 94, a very significant score (a score above 60 is considered significant).

Another useful comparison to identify sequence-specific DNA-binding proteins is to apply the same cell extract to immobilized promoter DNA and to immobilize nonpromoter (control) DNA. DNA containing the promoter of PF1601 was used

27

as a control (since this gene is not regulated under cold shock). Two gel bands showed higher intensity when captured with PF1479 promoter and were identified (Figure 2.5). All the DNA-binding proteins captured by the PF1479 promoter and identified by MS are summarized in Table 2.4.

2.2.2. DNA-Affinity Protein Capture Experiments with Promoter of PF0031

2.2.2.1. Promoter of PF0031

Like PF1479, *Pf* genomic DNA was used as template to PCR amplify the region from 39447 to 39126, which spans -200 to +121 relative to the translation start site of PF0031 (Figure 2.6).

2.2.2.2. DNA-Affinity Protein Capture Results

Figure 2.7 compares SDS PAGE analysis of proteins captured from different cell extracts (this time, standard growth conditions compared with extract of cells harvested 1 h after cold shock) using immobilized PF0031 promoter DNA. The intensities of both lanes looked similar, so no candidate proteins were selected.

Figure 2.8 compares proteins captured from the 1 h cell extract on either PF0031 promoter DNA or on the PF1601 promoter control DNA. I found only one band with higher intensity and this band was identified as PF0250p with a MASCOT score of 119, well beyond the threshold (Table 2.5).

```
5    GAGTCCCAGT   AGACATTTGT   ATAAAGGTAT   GGATATCAAG

     CATGTCAAGA   GATAGGGAGA   ATTAAAAAGG   CAGAATTCTT

     GAAAAAGTTG   AGGTTTAAAA   AATATGTTAA   AAACAACACT

     TTTTATGTTT   AGAATTATAT   TTGAAGGTTT   CAATAATCTT

     TATATGAACT   AACTCCCAAC   TATATTCCTG   GTGATATAAT

     ATGAGCGAAG   AAGTTCAAGA   AAGGATATGG   ATTCTAATAA

     CTCCAGACAA   ATGTAGTGGA   TGTAGGCTAT   GTGAAGTCAC

     CTGTTCTCTG   GAACATGAGG  3
```

Upper primer: biotinylated-GAGTCCCAGT AGACATTTGT

Lower primer: CCTCATGTTC CAGAGAACAG

Figure 2.2. The *Pf* genomic DNA sequence, -200 to +100 relative to the translation start site of PF1479. The bottom two sequences are the upper and lower primer sequences. The red ATG is the translation start site.
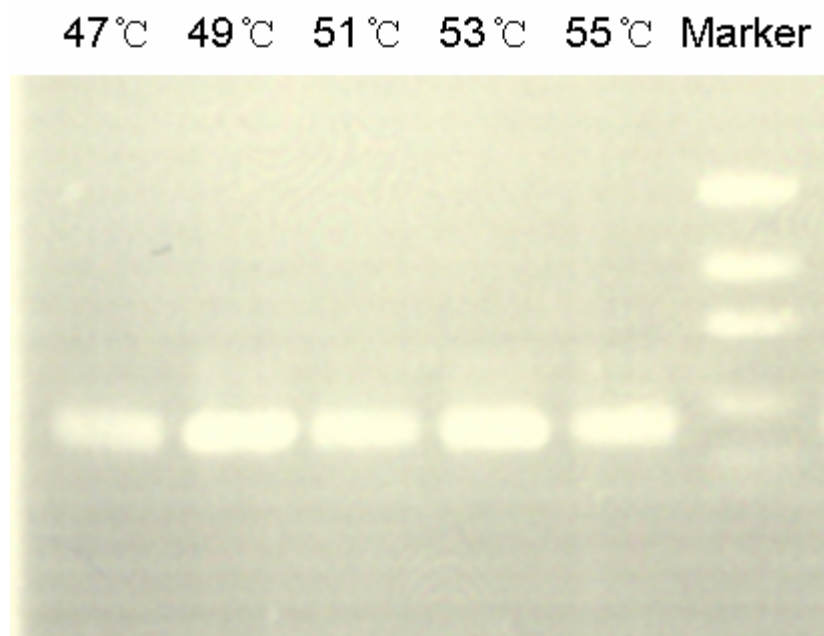
Figure 2.3. Gradient PCR results for PF1479 promoter DNA with annealing temperature 47, 49, 51, 53, and 55 °C.
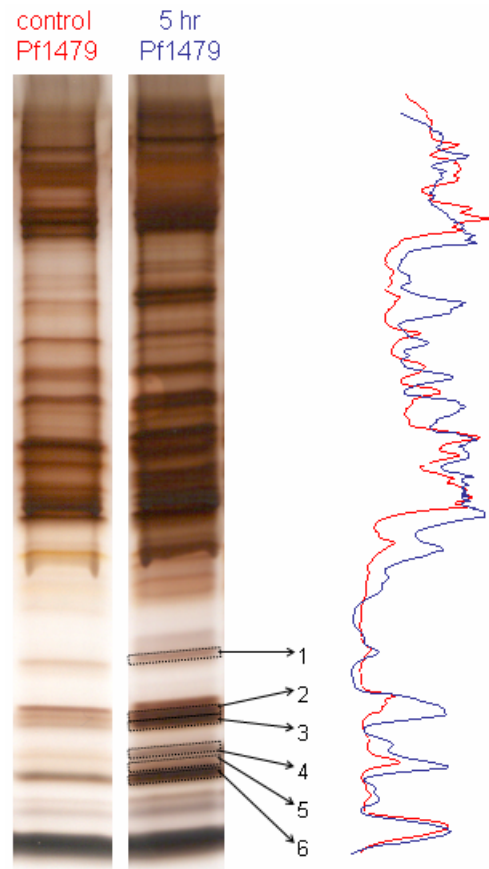
Figure 2.4. 1D SDS PAGE separation of protein mixtures dissociated from PF1479 promoter DNA. PF1479 promoter was used to capture proteins in control cell extract (left lane), and in extract of cells harvested 5 h after cold shock (right lane). The gel bands (numbered 1-6) which showed higher intensity in the right lane were subjected to MS identification. The gel was silver-stained.
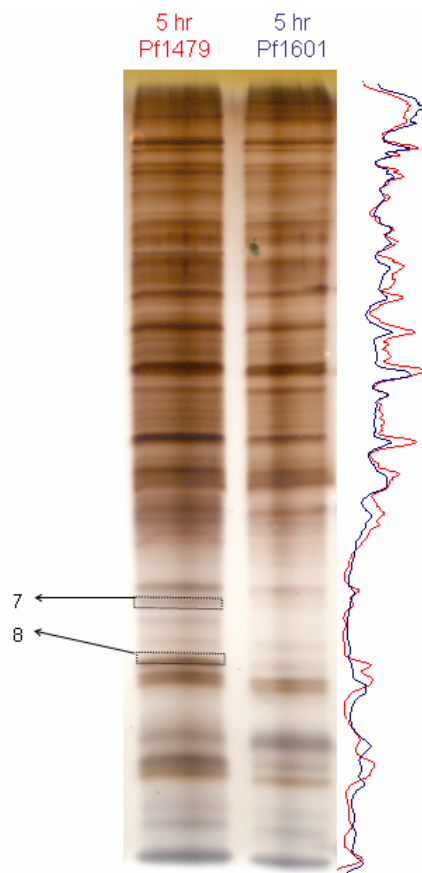
Figure 2.5. 1D SDS PAGE separation of proteins captured from the extract of cells harvested 5 h after cold shock. PF1479 promoter (left lane) or PF1601 promoter (right lane, control) were used to capture proteins from this cell extract. The gel bands numbered 7, 8, which showed higher intensity on PF1479 (left lane) were subjected to MS identification. The gel was silver-stained.

Table 2.4. Protein identification from promoter PF1479 (gel band numbers as in Figures 2.4 and 2.5).

Gel Band	ORF#	Protein Annotation	Score	Protein MW/pI	Sequence %
1	PF1794	inosine-5'-monophosphate dehydrogenase related protein I	54	21349.3 / 9.21	50%
2	PF0248	hypothetical protein	64	17715.7 / 5.80	52%
3	PF0250	transcriptional regulatory protein, Irp family	93	17984.4 / 9.20	56%
4	PF2019	hypothetical protein	74	14215 / 4.87	26%
5	PF1476	hypothetical protein	57	19065 / 9.01	52%
6	PF0396	hypothetical protein	57	21112 / 9.16	37%
7	PF1794	inosine-5'-monophosphate dehydrogenase related protein I	45	21349.3 / 9.21	42%
8	PF1734	transcriptional regulatory protein, asnC family	68	17075.0 / 7.78	42%
	PF0113	hypothetical protein	69	17792.9 / 6.86	46%

Note:
1. The protein annotations are based on the information from TIGR , The Institute of Genomic Research.
2. The score is the Mascot score calculated on the website :www.matrixscience.com
A score over 60 indicates a significant match.
3. Sequence % is the sequence coverage of matched peptides from the identified protein.

5`GATCTTTCCCATCTCGTCAAAAATCTAAAATTATTAGGCCAGAATAACAAT

GTTAATCGGAATAACAGTTATACCTAAGGAGAAAAGTTTATATACCCATATTC

AACCTTACAATGCTCTACCTCAATAATACAGAAAAGAAACCATTGTAAGGC

TTAAAAAAAGAAAAATAAGCACTCAAATTTAAAAATAGAGAAAT ATG GTGTC

ATATATGGATAACGAGACAGATCCCGAAGGATCCCCATTATTGAGGGCTAT

GAAGTTAGAGGAGAAACTGAACATCTCCAAGATATACTTAAAACTTGAAAT

GCACAATCCTAC`3

Upper primer: biotinylated- GATCTTTCCCATCTCGTCAA

Lower primer: GTAGGATTGTGCATTTCAAG

Figure 2.6. The *Pf* genome sequence in the region -200 to +121 relative to the translation start site of PF0031. The bottom two sequences are the upper and lower primer sequences. The red ATG is the translation start site.
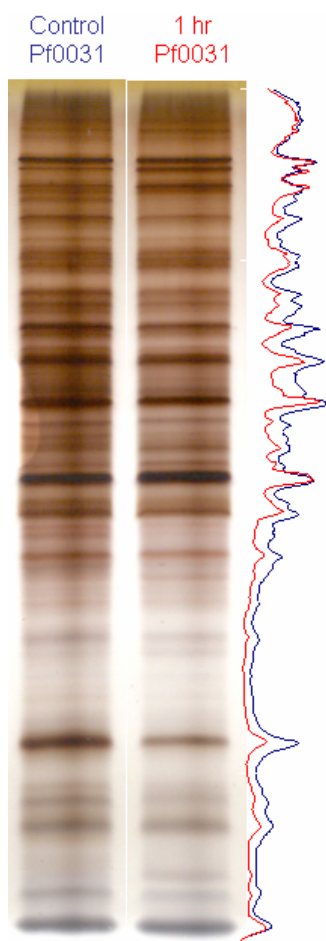
Figure 2.7. 1D SDS PAGE separation of protein mixtures captured from cell extract of standard growth conditions (left) and from extract of cells harvested 1 h after cold shock (right) using immobilized PF0031 promoter DNA. No obvious differences were observed. The gel was stained with silver staining method.
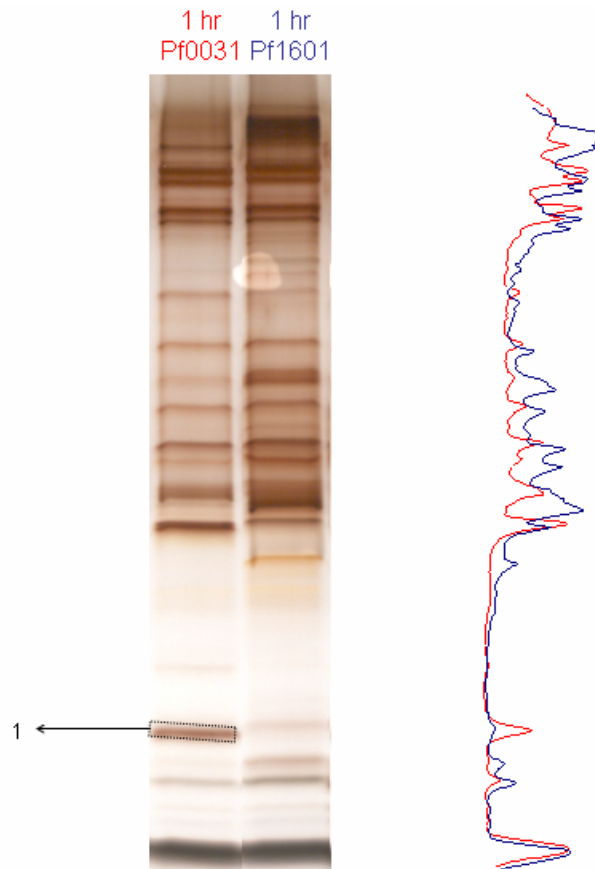
Figure 2.8. 1D SDS PAGE separation of proteins captured from extract of cells harvested 1 h after cold shock. The left lane is the result of using the PF0031 promoter while the right lane is the result of using the PF1601 promoter (control DNA). The gel band numbered 1, which showed higher intensity on PF0031, was subjected to MS identification. The gel was stained with silver staining method.

Table 2.5. Protein identification from promoter PF0031 (gel band labeled in Figure 2.8)

Gel Band	ORF#	Protein Annotation	Score	Protein MW/pI	Sequence %
1	PF0250	transcriptional regulatory protein, lrp family	119	17984.4 / 9.20	62%

Note:
1. The protein annotations are based on the information from TIGR , The Institute of Genomic Research.
2. The score is the Mascot score calculated on the website :www.matrixscience.com
A score over 60 indicates a significant match.
3. Sequence % is the sequence coverage of matched peptides from the identified protein.

2.2.3. DNA-Affinity Protein Capture Experiments with Promoter of PF0324

2.2.3.1. Promoter of PF0324

Pf genomic DNA was used as template to PCR amplify the region from 338566 to 338881, which spans -207 to +108 relative to the translation start of PF0324 (Figure 2.9).

2.2.3.2. DNA-Affinity Protein Capture Results

Figure 2.10 compares SDS PAGE analysis of proteins captured from different cell extracts (standard growth conditions compared with cell extract harvested 5 h after cold shock) using immobilized PF0324 promoter DNA. Four proteins were identified (Table 2.6).

Figure 2.11 compares proteins captured from the 5 h cell extract on either PF0324 promoter DNA or on the PF1601 control DNA. The intensities of both lanes looked similar; therefore no candidate proteins were selected.

5`CGGTTTCTTTTCTGTTCCTTCTAGCATCTCTAGTTTTAAATAATGGGTTG

GGGGTTGTAATGTCTCTTGTTCTTCTTGGTCTAGCCATATTCCAATGGGTC

ACAAAAAAGTAAGCATCTATTACTTCAAGAAACTTGCAATTACCTATATACC

TGGAGGCTTCTACTCACTATTAGGTCTCAACACTGAGGCCTAAAAAGTTT

GGAGATGGAAAAAGTGAATAAATTTGGAATGGCAATGTTGGCTTTGTTGG

CAGCATTTGGCCTTATTCTTGGGGCCGGAGCCAACTTTAGAGACTACAAC

GCAAGCAGAAGC`3

Upper primer: biotinylated- CGGTTTCTTTTCTGTTCCTT

Lower primer: GCTTCTGCTTGCGTTGTAGT

Figure 2.9. The *Pf* genome sequence in the region -207 to +108 relative to the translation start site of PF0324. The bottom two sequences are the upper and lower primer sequence. The red ATG is the translation start site.
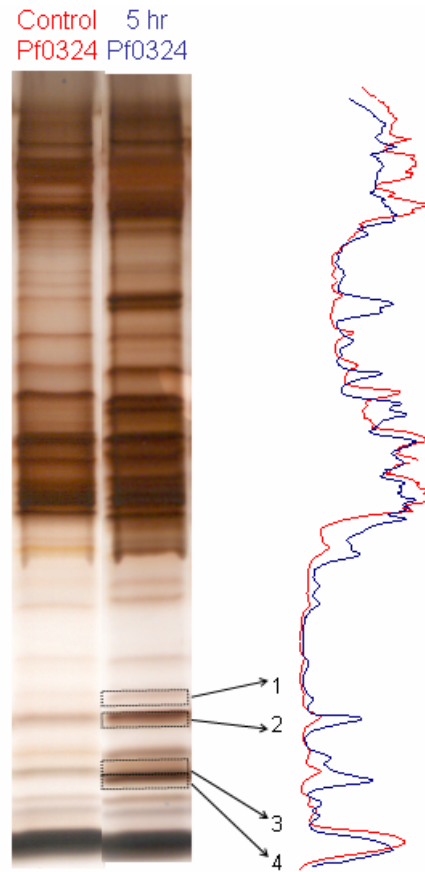
Figure 2.10. 1D SDS PAGE separation of protein mixtures captured from cell extract of standard growth conditions (left) and from extract of cells harvested 5 h after cold shock (right) using immobilized PF0324 promoter DNA. The gel bands (numbered 1-4) that showed higher intensity on 5 h cold shock than on the control cell extract were subjected to MS identification. The gel was silver-stained.
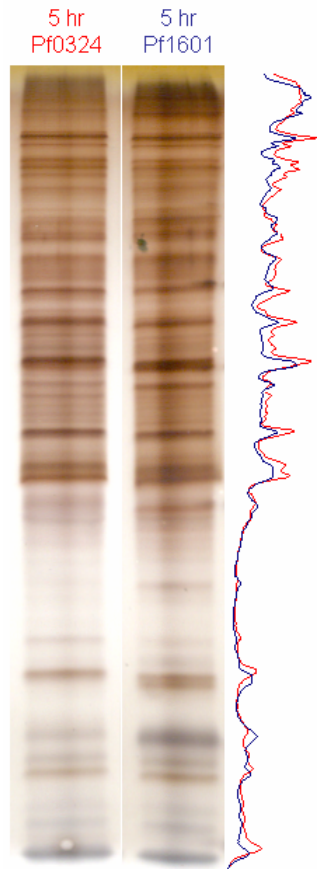
Figure 2.11. 1D SDS PAGE separation of proteins captured from extract of cells harvested 5 h after cold shock. The left lane is the result of using the PF0324 promoter while the right lane is the result of using the PF1601 promoter (control DNA). No obvious difference was observed. The gel was stained with silver staining method.

Table 2.6. Protein identification from promoter PF0324

Gel Band	ORF#	Protein Annotation	Score	Protein MW/pI	Sequence %
1	PF2053	transcriptional regulatory protein, asnC family	54	17204 / 9.50	39%
2	PF0248	hypothetical protein	60	17715.7 / 5.80	54%
3	PF0574	hypothetical protein	56	15622 / 6.07	39%
4	PF0396	hypothetical protein	56	21112 / 9.16	34%

Note:

1. The protein annotations are based on the information from TIGR , The Institute of Genomic Research.

2. The score is the Mascot score calculated on the website :www.matrixscience.com

A score over 60 indicates a significant match.

3. Sequence % is the sequence coverage of matched peptides from the identified protein.

2.3. DNA-Affinity Protein Capture Experiments to Catch Potential Transcriptional Regulators Responding to Iron Depletion

As shown in Chapter 1, based on the microarray data of iron depletion, gene PF0723 is significantly up-regulated (7.5-fold) after iron depletion. Gene PF0857 shows 4.7-fold up-regulation. On the other hand, gene PF0340 is significantly down-regulated (4.8-fold).

2.3.1. Promoters of PF0340, PF0723 and PF0857

Pf genomic DNA was used as template to PCR amplify the three genes. The region for PF0340 is from 353859 to 353661 which spans -100 to +98 relative to the translation start site (Figure 2.12). For PF0723, the region from 721143 to 721418 was PCR amplified and was -177 to +98 relative to the translation start site (Figure 2.13). The PCR amplification for PF0857 covered the region from 833882 to 833690, -188 to +4 relative to the translation start site (Figure 2.14).

2.3.2. DNA-Affinity Protein Capture Experiments of Promoter PF0723 and Iron-Depleted Cell Extract

The PF0723 promoter DNA was immobilized and used for DNA-affinity protein capture, as described. Two separate experiments using either cell extract from standard growth conditions or extract from cells harvested from the iron-depleted condition (described in Chapter 1) were compared by SDS PAGE separation of the retained proteins. Four bands showed higher intensity in iron-depleted cell extract and were identified by MS (Figure 2.15). These proteins

were candidates for potential promoter-binding proteins that respond to iron depletion. Band 4 was identified to be PF0250p.

The same experiments were performed using three promoters PF0340, PF0723 and PF0857 with both the control cell extracts and the iron-depleted cell extract. This time the iron-depleted cell extract comes from a growth in which the iron chelate BPS was added to cells with no added iron to make sure the iron concentration was low. Figure 2.16 shows the SDS PAGE analysis of the DNA-affinity protein capture experiments. Because the gel was not stained properly, it was hard to find any difference, but two bands were still identified. All the DNA-binding proteins captured by the three promoters and identified by MS are summarized in Table 2.7.

2.4. Summary

The protein capture experiments were very effective in identifying potential DNA binding proteins. Among those protein identified, PF0250 bound specifically to promoter PF0031, PF1479 and PF0723 and the MASCOT scores were significant. Further experiments to find out the role of this protein in transcription regulation under stress were explored (Chapter 3).

Other potentially interesting targets for future work include annotated transcriptional regulators PF1734 from Table 2.4 and PF2053 from Table 2.6.

3`GGCAGAAGGATAAATCTCGTAAAGATAATGAAGTATCCTTCTCCTTAGCT

CGCTTCGGTGTAATGACCGTAAGATGTGTGGGTCGATTATTGGTGGCACT

GGAGCATCCCTCCATTATTGGCATTGAATTTTATCCCACCTATATTATAATTG

TTTGGATAATTAATAAATGTTTCGGAAATGTAATTGGGAATC`5

Upper primer: biotinylated-GATTCCCAATTACATTTCCG

Lower primer: GGCAGAAGGATAAATCTCGT

Figure 2.12. The *Pf* genome sequence in the region -100 to +98 relative to the translation start site of PF0324. The bottom two sequences are the upper and lower primer sequence. The red CAC (reverse complement of GTG) is the translation start site.

45

5`GGATTAAGTTCTGTAGATGTTCGATATTCGAATTTTAAATATTTGACATGT

GAAATTTGTAGGTATTTTAAAAGAAACCTTTGACTGAGTGGCTCAAAATTT

CGAATGTGAAATAGTAGCGTTAATAAACTACCCTAATTATTTTAGAATGTCC

TAATAAATTTAGGAGGTGCAAAAATGATAGGACAGTTTCTAATTACATTTAG

AGAAGCATTAGAAGCAGCGATAATTGTTGCTATTATAATTGCCTACCTTAAA

CGTACTGATAGGGAAGAG`3

Upper primer: biotinylated-GGATTAAGTTCTGTAGATGTTCG

Lower primer: CTCTTCCCTATCAGTACGTTTA

Figure 2.13. The *Pf* genome sequence in the region -177 to +98 relative to the translation start site of PF0723. The bottom two sequences are the upper and lower primer sequence. The red ATG is the translation start site.

3`GCACGATTACTCCCTCCTTACAATAATCTTACTTGCAATTCCCCAGCCAA

GAGCTATTTTTGATGAACCTATCCCAATGATTACAGGTCCAGGCTTTCTAC

TTTTTATCATTCTAACCCTAACACCTGGGGTTATACCCATTGAGAGCAATTT

GCCTCTA GCATTAACTCCACTCTGGATGTCAACTACAAC`5

Upper primer: biotinylated-GTTGTAGTTGACATCCAGAGTG

Lower primer: GCACGATTACTCCCTCCTTA

Figure 2.14. The *Pf* genome sequence in the region -188 to +4 relative to the translation start site of PF0857. The bottom two sequences are the upper and lower primer sequence. The red CAC is the translation start site.
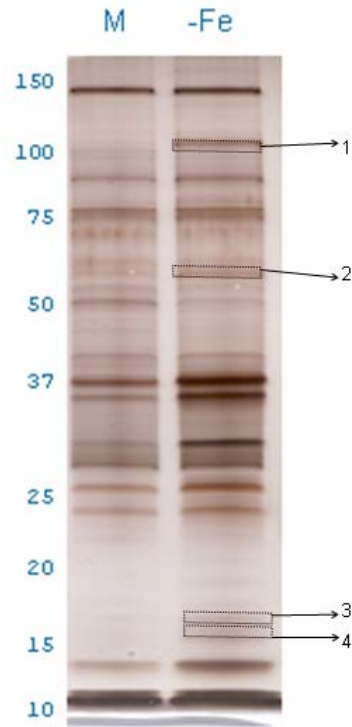
Figure 2.15. 1D SDS PAGE separation of protein mixtures captured from cell extract of standard growth conditions (left) and from extract of cells harvested from iron-depleted growth condition (right) using immobilized PF0723 promoter DNA. The gel bands (numbered 1-4) which showed higher intensity on iron-depleted cell extract than on the control cell extract were subjected to MS identification. The gel was stained with silver staining method.
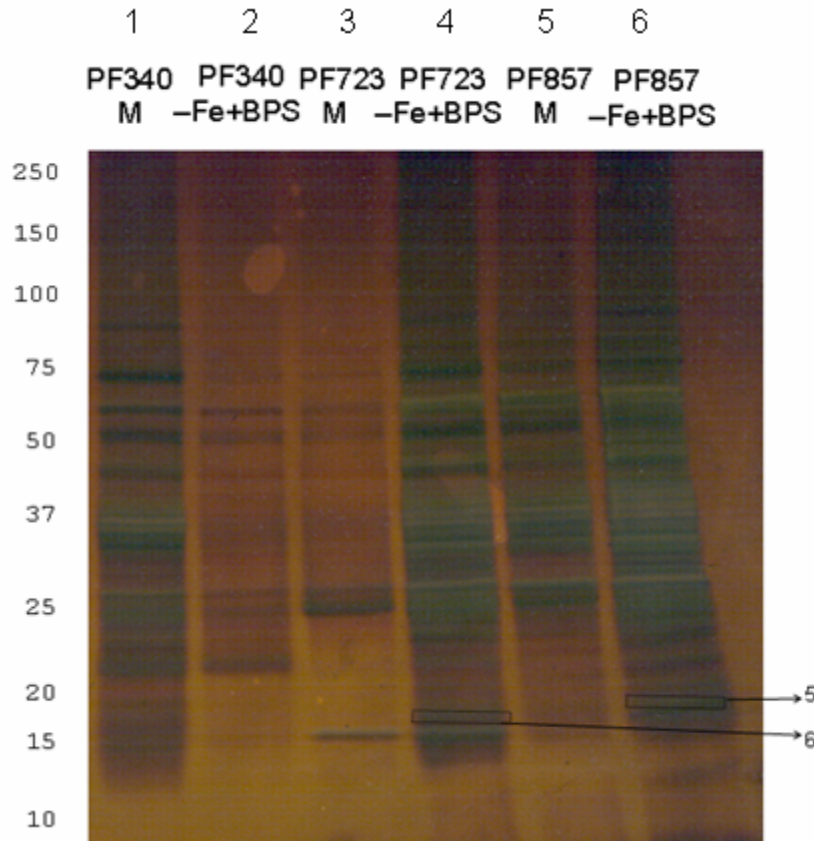
Figure 2.16. 1D SDS PAGE separation of protein mixtures from the DNA-affinity protein capture experiment for iron-depleted, BPS-added (-Fe+BPS) condition using immobilized PF0340, PF0723, and PF0857 promoter DNA.

Lanes 1 and 2 show the retained proteins captured by PF0340 in control cell extract (Lane 1) and from extract of cells harvested from -Fe+BPS condition (Lane 2). Lanes 3 and 4 show the proteins retained by PF0723 promoter from control cell extract (Lane 3) and from extract of cells harvested from -Fe+BPS condition (Lane 4). Lanes 5 and 6 show the retained proteins captured by PF0857 promoter from control cell extract (Lane 5) and from extract of cells harvested from -Fe+BPS condition (Lane 6). The gel bands (numbered 5-6) with

49

higher intensity in -Fe+BPS cell extract were subjected to MS identification. The gel was stained with silver staining method.

Table 2.7. Protein identification from iron stress protein capture experiments.

Gel Band	ORF#	Protein Annotation	Score	Protein MW/pI	Sequence %
1	PF1051	large helicase-related protein	91	100236.26/9.89	32%
2	PF1901	oxygen-independent coproporphyrinogen IIIoxidase	55	74833.7/9.19	26%
3	PF1560	transcription termination-antitermination factornusa	47	16454.5/9.49	32%
4	PF0250	transcriptional regulatory protein, Irp family	58	17984.4 / 9.20	54%
5	PF1281	superoxide reductase	87	14324.3/5.69	94%
6	PF1281	superoxide reductase	80	14324.3/5.69	86%

Note:

1. The protein annotations are based on the information from TIGR , The Institute of Genomic Research.

2. The score is the Mascot score calculated on the website :www.matrixscience.com
A score over 60 indicates a significant match.

3. Sequence % is the sequence coverage of matched peptides from the identified protein.

CHAPTER 3

CHARACTERIZATION OF PUTATIVE REGULATORY TRANSCRIPTION FACTOR, PF0250p

3.1. Bioinformatics Analysis of Identified Promoter DNA Binding Protein PF0250p

The sequence of PF0250 was aligned and compared with all the sequences in NCBI database. The conserved domains in the Conserved Domain Database use the web-based NCBI BLAST program. (http://www.ncbi.nlm.nih.gov/). Figure 3.1 shows the alignment result. PF0250p is homologous to the Lrp protein with about 97.4% aligned. PF0250p is also homologous to HTH_ARSR, Arsenical Resistance Operon Repressor and similar prokaryotic, metal regulated homodimeric repressors with about 94.9% alignment. The ArsR subfamily is characterized as the helix-turn-helix bacterial transcriptional regulatory proteins. PF0250p is homologous to HTH_ASNC (100% alignment), the helix_turn_helix AsnC type of transcriptional regulator, and HTH_8 (97.6% alignment), a bacterial regulatory protein of the Fis family. AsnC is characterized as an autogenously regulated activator of asparagine synthetase A transcription in *Escherichia coli*. Many alignments also show the characteristics of a helix-turn-helix (HTH) DNA-binding domain. Figure 3.2 shows the predicted secondary structure of PF0250p, verifying the helix-turn-helix structure.
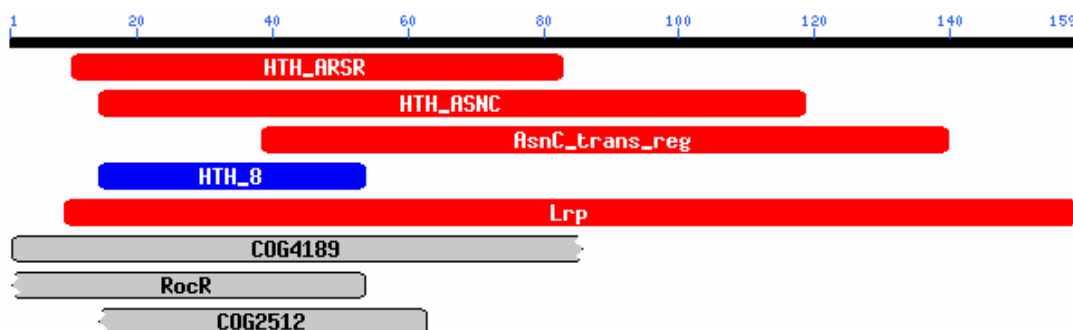
Figure 3.1. The result of a conserved domain BLAST search based on PF0250p. The top black line shows the length of sequence (PF0250p); the bars below indicate the conserved domains found.

1. CD00090, HTH_ARSR, Arsenical Resistance Operon Repressor and similar prokaryotic, metal regulated homodimeric repressors.

2. smart00344, HTH_ASNC, helix_turn_helix ASNC type; AsnC: an autogenously regulated activator of asparagine synthetase A transcription in Escherichia coli

3. pfam01037, AsnC_trans_reg, AsnC family. The AsnC family is a family of similar bacterial transcription regulatory proteins.

4. pfam02954, HTH_8, Bacterial regulatory protein, Fis family.

5. COG1522, Lrp, Transcriptional regulators [Transcription]

6. COG4189, Predicted transcriptional regulator [Transcription]

7. COG3829, RocR, Transcriptional regulator containing PAS, AAA-type ATPase, and DNA-binding domains

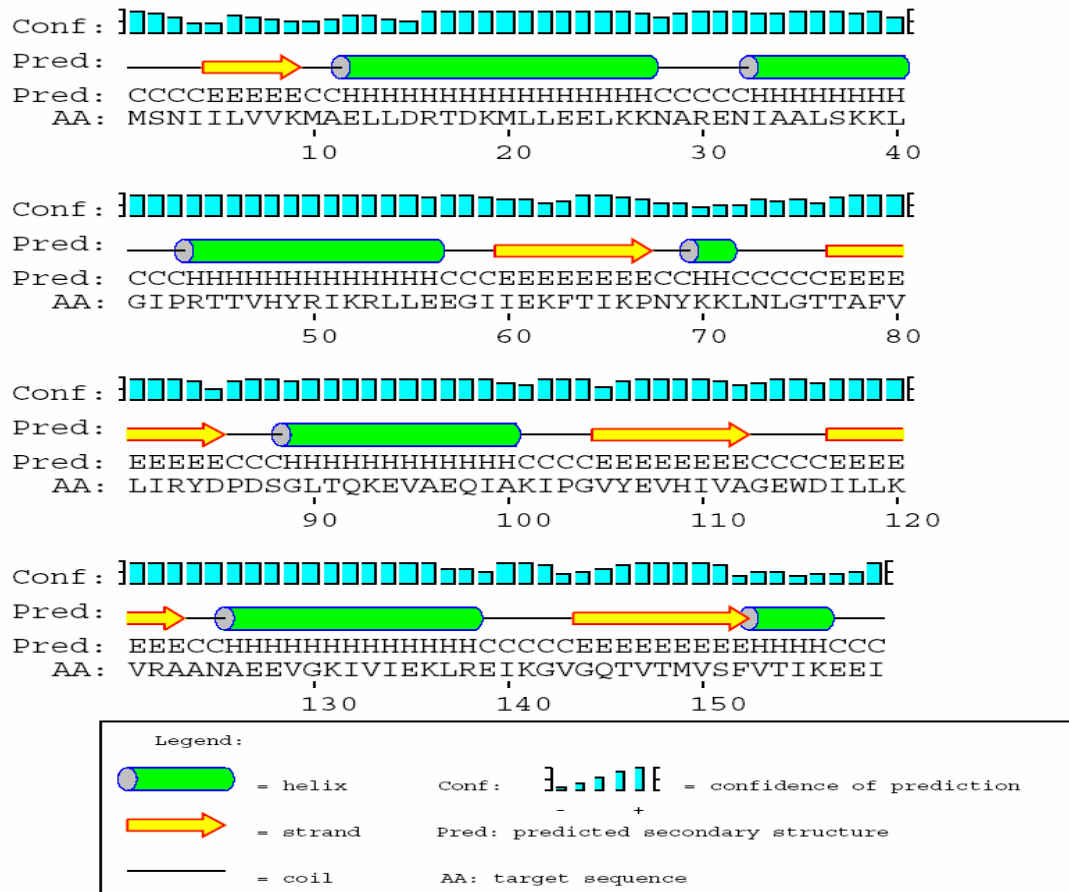8. COG2512, Uncharacterized membrane-associated protein/domain

```
Conf: ]▯▯▯▯▫▫▯▯▯▯▫▫▯▯▯▯▫▫▯▯▯▯▯▯▯▯▯▯▯▯▯▫▫▯▯▯▯▯▯▯▯▯▯▯▫▯E
Pred: ———————▶      ◯▬▬▬▬▬▬▬▬▬▬▬   ◯▬▬▬▬▬▬▬
Pred: CCCCEEEEECCHHHHHHHHHHHHHHHHCCCCCHHHHHHHHH
  AA: MSNIILVVKMAELLDRTDKMLLEELKKNARENIAALSKKL
              |         |         |         |
             10        20        30        40

Conf: ]▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▫▫▯▯▯▯▯▫▫▫▫▫▯▯▯▯▫▫▯▯▯▯E
Pred: ———◯▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬   ▶   ◯▬▬   ▶
Pred: CCCHHHHHHHHHHHHHHHCCCEEEEEEEEEECCHHCCCCCEEEE
  AA: GIPRTTVHYRIKRLLEEGIIEKFTIKPNYKKLNLGTTAFV
              |         |         |         |
             50        60        70        80

Conf: ]▯▯▯▯▫▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▫▫▯▯▯▫▫▯▯▯▯▯▫▫▫▯▯▯▯E
Pred: ▶    ◯▬▬▬▬▬▬▬▬    ▶    ▶
Pred: EEEEECCCHHHHHHHHHHHHHHCCCCEEEEEEEEECCCCEEE
  AA: LIRYDPDSGLTQKEVAEQIAKIPGVYEVHIVAGEWDILLK
              |         |         |         |
             90       100       110       120

Conf: ]▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▯▫▫▫▯▯▯▫▫▫▯▯▯▯▯▯▫▫▫▫▫▯E
Pred: ▶   ◯▬▬▬▬▬▬▬▬▬▬▬   ▶   ▶◯▬▬▬
Pred: EEECCHHHHHHHHHHHHHHHHCCCCCEEEEEEEEEEEHHHHCCC
  AA: VRAANAEEVGKIVIEKLREIKGVGQTVTMVSFVTIKEEI
              |         |         |
             130       140       150
```

Legend:
◯▬▬▬ = helix   Conf:]▫▫▯▯▯E = confidence of prediction
▶ = strand   Pred: predicted secondary structure
——— = coil   AA: target sequence

Figure 3.2. Predicted secondary structure of PF0250p, showing the helix-turn-helix structure (residues 12-40).

3.2 Recombinant His-tagged Protein Expression & Purification

3.2.1. PF0250p Solubility Prediction

The molecular weight of PF0250p is predicted to be 18.88 kDa, including the engineered hexahistidine tag (His-Tag) (the MW of PF0250p without the His-Tag is 17.95 kDa). Figure 3.3 shows the protein sequence and the solubility prediction results. The protein is predicted to be 59.2% insoluble according to online solubility prediction.

3.2.2. Cloning of PF0250

I obtained the clone of PF0250 in the pET24dBAM vector from Dr. Adams group, UGA (Figure 3.4). pET24dBam is modified from Novagen's pET24d by insertion of GGCTCATCACCATCACCATCAC between the NcoI and BamHI sites, generating the His-Tag sequence. I sequence-verified the insertion and transformed the plasmid into BL21 Codon Plus *E. coli* cells for expression.

PF0250 Protein Sequence:

MSNIILVVKMAELLDRTDKMLLEELKKNARENIAALSKKLGIPRTTVHYRIKR

LLEEGIIEKFTIKPNYKKLNLGTTAFVLIRYDPDSGLTQKEVAEQIAKIPGVYEV

HIVAGEWDILLKVRAANAEEVGKIVIEKLREIKGVGQTVTMVSFVTIKEEI

The input protein sequence has a **59.2 percent** chance of **insolubility** when overexpressed in *E. coli*.

The amino acid composition of the input sequence is:

8 Arginine (R)
6 Asparagine (N)
5 Aspartic Acid (D)
17 Glutamic Acid (E)
9 Glycine (G)
18 Lysine (K)
4 Proline (P)
4 Serine (S)

Figure 3.3. PF0250p protein sequence (without the engineered His-Tag) and the solubility prediction results from http://www.biotech.ou.edu/. The protein is predicted to be 59.2% insoluble.

pET24dBam

*Nco*I   *Bam*HI

         MetAlaHisHisHisHisHisHis_   _GlySer

TAC   CATGGCTCATCACCATCACCATCACG   GATCC

ATGGTAC   CGAGTAGTGGTAGTGGTAGTGCCTAG   G
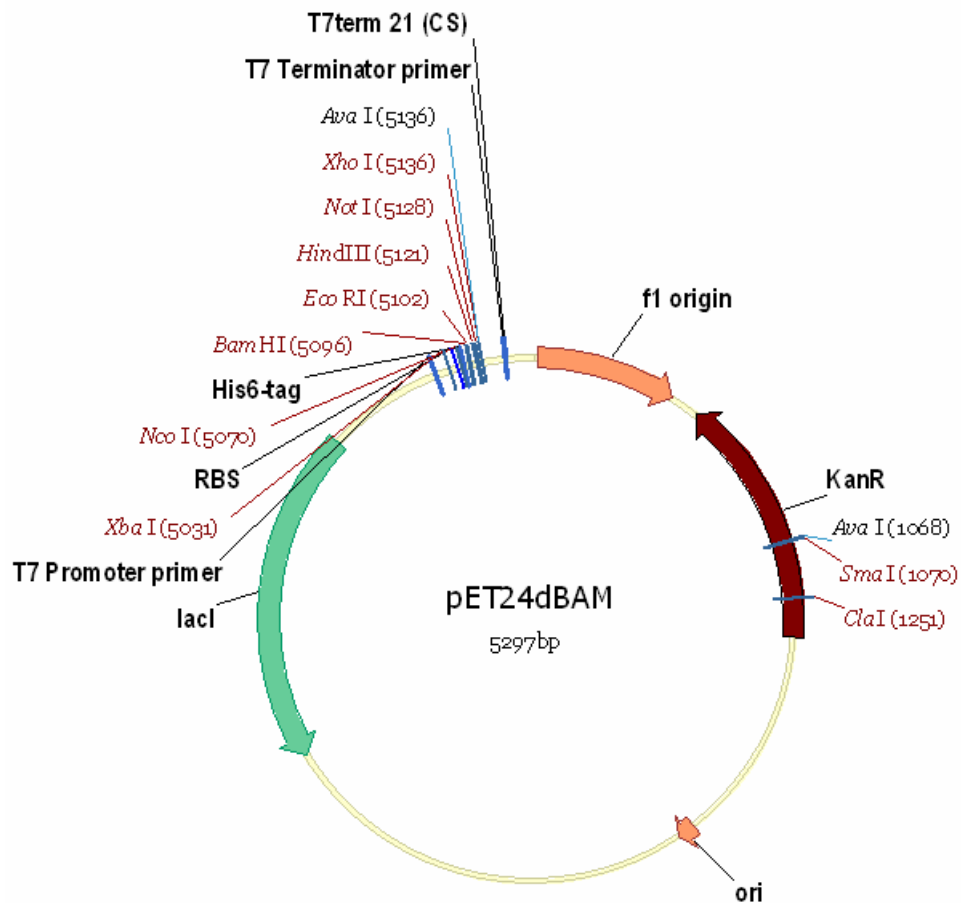
Met-Ala-(His)$_6$-Gly-Ser-GENE

Figure 3.4. pET24dBam expression vector uses the T7 promoter and contains

the insert (top) to create the N-terminal His-Tag construct of the inserted gene.

The plasmid is resistant to kanamycin antibiotics.

3.2.3. Expression and Purification of PF0250p

3.2.3.1. Expression Method

I inoculated 5 mL of kanamycin-containing LB media from a plate with the BL21 Codon Plus expression strain transformed with PF0250-inserted pET24dBAM. The cells were grown in 5 mL culture for 8 h at 37 °C, then used to inoculate a 50 mL culture, which was grown at 37 °C for another 8 h. 1 L of LB media was inoculated with this 50 mL culture. From the time of inoculation, I began taking samples for OD_{600} readings and gel identification every 0.5 h to monitor culture growth. Once the OD_{600} value of the culture reached 0.8, I induced the 1 L culture with IPTG immediately. The final concentration of IPTG was 0.4 mM. I took OD_{600} readings each hour after induction and allowed the culture to grow a total of 6 h at 37 °C. SDS-PAGE analysis was used on the each hour collected samples to determine the extent of expression.

To harvest the cells, I split the 1 L culture into two 500 mL centrifuge bottles and the cells were collected by centrifuging at 5,000xg for 15 min at 4 °C; the cell pellet was frozen and stored at -20 °C until further purification.

3.2.3.2. Purification Method

The cell pellet was thawed and mixed with 20 mL Lyses buffer (Table 3.1). I then sonicated the solution on ice and centrifuged the mixture to collect the supernatant for FPLC analysis and kept the pellet for future use.

A HiTrap Chelating HP 1 mL column, charged with NiSO$_4$ was used to purify the His-Tagged PF0250p protein. Cell extract supernatant was loaded with binding buffer (Table 3.1), then the elution buffer (Table 3.1) was applied as a gradient. I collected all the fractions and used SDS-PAGE to identify which fractions contained PF0250p. Dialysis was used to desalt protein fractions [49].

3.2.4. Expression and Purification of PF0250p

Figures 3.5 and 3.6 show the growth curve and the corresponding SDS-PAGE demonstrating the induced expression of PF0250p. HiTrap purification revealed very little protein in the soluble fraction. As predicted, PF0250p is insoluble and most of the expressed protein was in the pellet after cell lysis. Figure 3.7 shows SDS-PAGE analysis of PF0250p in the gradient eluting fractions; there was too little pure PF0250p for further experiments.

Another Scott group member, Sungguan Hong, introduced a method to enhance the recovery of PF0250p from the pellet, involving 6M urea denaturation, followed by dilution for renaturation (refolding) (see table 3.2 for the buffer compositions). This procedure was successful in solubilizing more PF0250p and Sungguan kindly gave me some of his purified protein, which I will denote as PF0250$_a$. Figure 3.8 shows the SDS-PAGE analysis of the proteins in the gradient-eluting fractions purified from the urea-renaturation method.

I also changed the growth conditions for PF0250p expression to reduce the chance of forming inclusion bodies after IPTG induction. This involved changing induction conditions to reduce the production of expressed protein. After initial

growth at 37 °C, induction with 0.1 mM IPTG and subsequent growth at 18 °C overnight was tried. Finally, after the purification steps I was able to collect PF0250p, which I denote PF0250$_b$. Figure 3.9 shows the SDS-PAGE analysis of the proteins in the gradient fractions using the new growth condition. The concentration of pooled protein PF0250$_b$ was 0.20 mg/mL.

When I concentrated PF0250$_b$, it precipitated. After centrifugation, I collected the supernatant in which the concentration of protein was very low. I denote this soluble PF0250p as PF0250$_c$. The concentration of protein PF0250$_c$ was 0.05 mg/mL.

3.3. Electrophoretic Mobility Shift Assay (EMSA) to Confirm the Binding of PF0250p to Promoters

3.3.1. Introduction

Electrophoretic mobility shift assay (EMSA) is also called gel shift assay. It is a widely used technique to find protein-DNA interactions. This kind of experiment can be performed to verify that the target protein binds to a given DNA or RNA sequence. EMSA usually involves gel electrophoresis of a protein-DNA mixture using various protein/DNA ratios on a polyacrylamide gel or agarose gel. The principle is based on the different migration of different sizes of molecules moving through the gel.
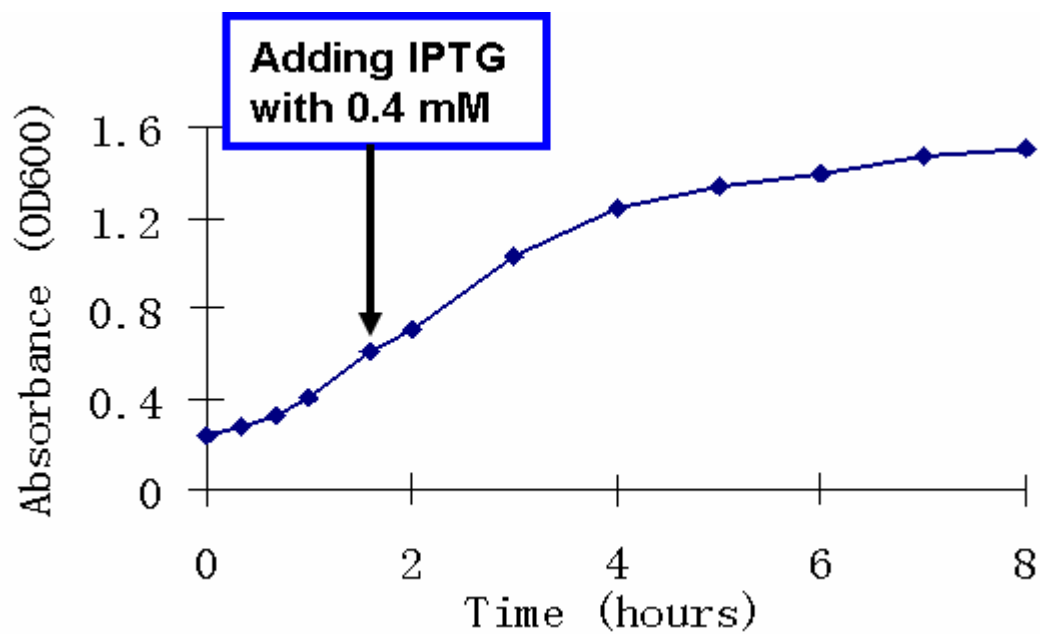
Figure 3.5. The growth curve during expression of PF0250p (OD_{600} vs. time). 0.4 mM IPTG was added at OD_{600} = 0.612, time= 1.6 h.
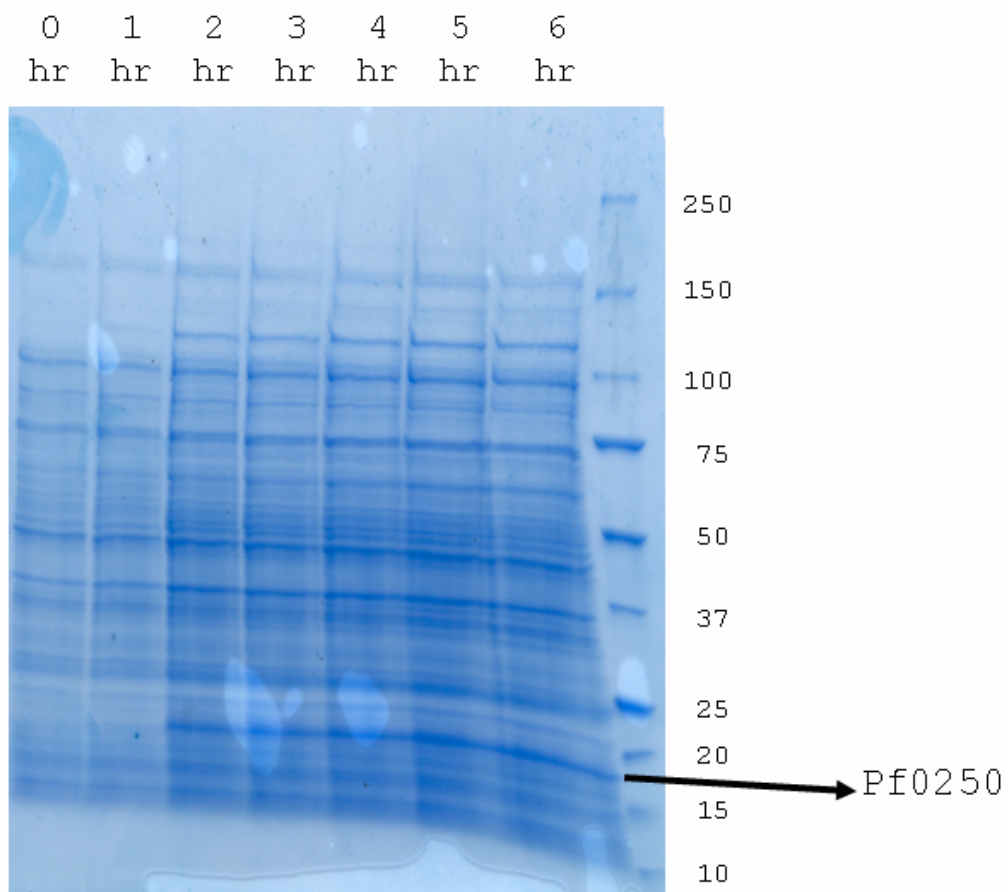
Figure 3.6. The SDS-PAGE of samples from expression culture. The band at ca.

18-19 kDa appears in the 2-h sample (after IPTG induction at 1.6 h). This band

was then identified as PF0250p by in-gel digestion and mass spectrometry.

62

Table 3.1. Makeup of the buffers used for HiTrap purification of His-Tagged proteins.

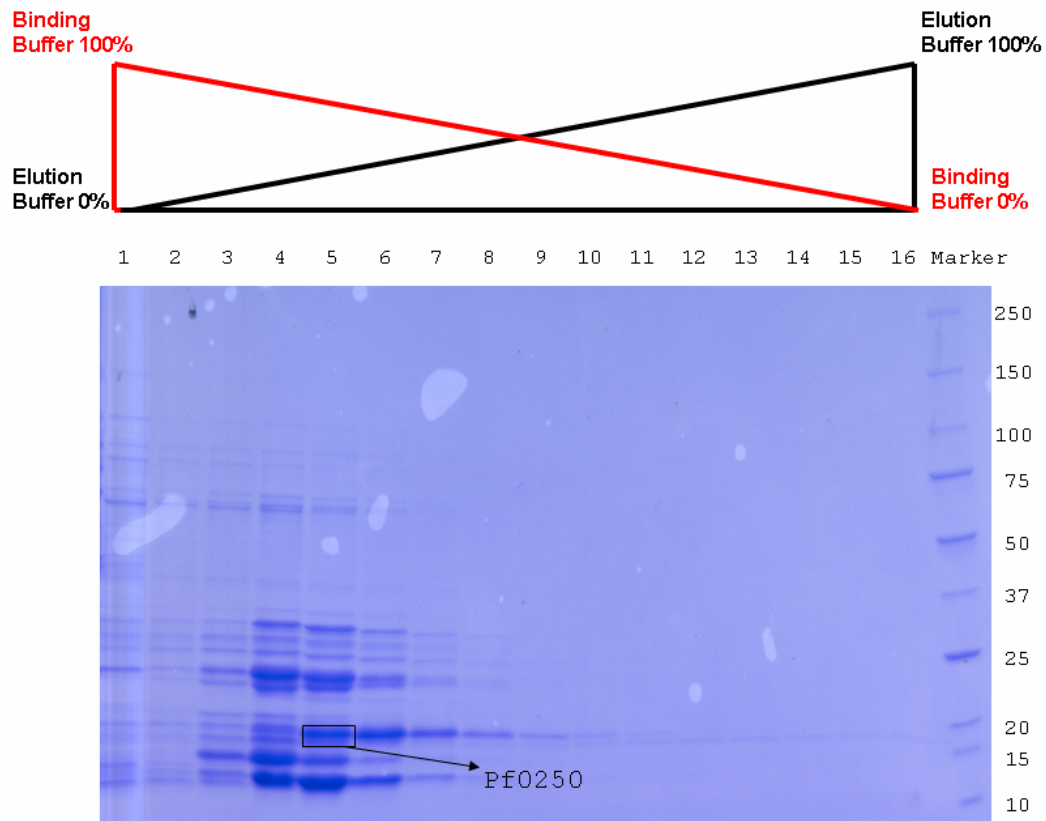| Lyses buffer |
|---|
| 20 mM Na Phosphate (pH 7.4)
0.5 M NaCl
1 mM PMSF
10 mM β-mercaptoethanol
filter 0.45 µm (or 0.2 µm) |
| **Binding Buffer** |
| 20 mM Na Phosphate (pH 7.4)
0.5 M NaCl
filter 0.45 µm (or 0.2 µm) |
| **Eluting Buffer** |
| 0.02 M Na Phosphate (pH 7.4)
0.5 M NaCl
0.5 M imidazole
filter 0.45 µm (or 0.2 µm) |

Figure 3.7. SDS-PAGE analysis of proteins in the gradient-eluting fractions from the HiTrap purification column. The indicated band was excised and identified as PF0250p by in-gel digestion and mass spectrometry. The expression level was low and PF0250p was difficult to separate from other proteins.

Table 3.2. Makeup of the buffers used by Sungguan Hong for HiTrap purification of PF0250$_a$.

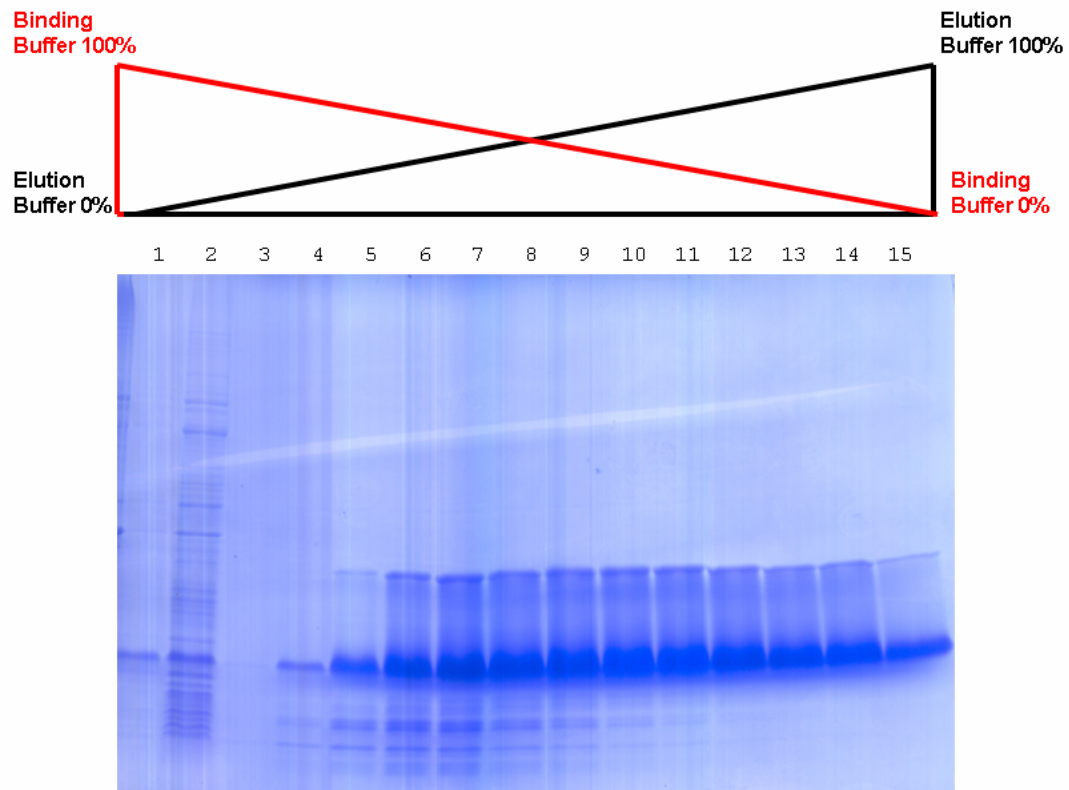| *Lyses buffer* |
|---|
| 20 mM Na Phosphate (pH 7.4)
0.5 M NaCl
1 mM PMSF
10 mM β-mercaptoethanol
6M urea
filter 0.45 µm (or 0.2 µm) |
| *Binding Buffer* |
| 20 mM Na Phosphate (pH 7.4)
0.5 M NaCl
6M urea
filter 0.45 µm (or 0.2 µm) |
| *Eluting Buffer* |
| 0.02 M Na Phosphate (pH 7.4)
0.5 M NaCl
0.5 M imidazole
6M urea
filter 0.45 µm (or 0.2 µm) |

Figure 3.8. SDS PAGE analysis of the proteins in the gradient fractions from the PF0250p purification carried out by Sungguan Hong. All buffers contain 6M urea. Pooled fractions from lanes 8 - 15 were desalted and denoted PF0250$_a$.
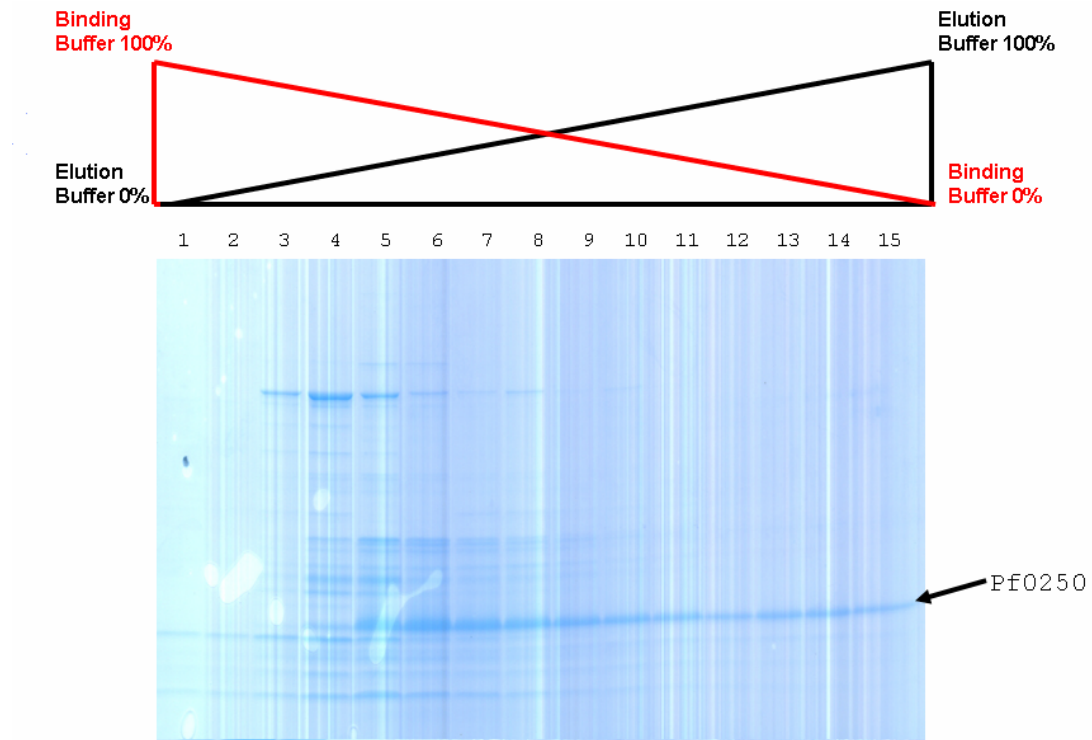
Figure 3.9. SDS PAGE analysis of the proteins in the gradient fractions my second purification attempt (0.1 M IPTG, followed by 18 °C growth). Pooled fractions from lanes 9-15 were desalted and denoted PF0250$_b$.

3.3.2. Method and Materials

A standard EMSA protocol was developed by Gina Lipscomb, another member of the Scott group. The results presented in Chapter 2 demonstrated that the promoter DNA of PF1479, PF0723, and PF0031 all captured PF0250p in the protein capture assay. In contrast, the PF0340 promoter did not capture the protein. I used these four promoters as templates in EMSA assays. PF0250p was added to the reaction as a gradient against the same concentration of promoters. The protein used was $PF0250_a$ (from Sungguan Hong) and $PF0250_b$ purified in the lower expression induction. The protein and DNA were mixed in 10 µL EMSA reaction buffer (20 mM HEPES, pH7.5, 0.1 M KCl, 5% glycerol, 1 mM DTT, 1 mM EDTA). The protein-DNA mixtures were incubated at 55 °C for 30 min, after which the 10 µL mixtures were loaded onto 5% polyacrylamide TBE gel (Biorad). The electrophoresis was run at room temperature under constant voltage of 200 V for 30-40 min. The TBE gel was stained in 40 mL 0.01% SYPRO Green for 10 min [50, 51].

3.3.3. EMSA Results

Figure 3.10 shows the EMSA results of the PF0723 promoter mixed with both proteins $PF0250_a$ and $PF0250_b$ and Figure 3.11 shows the same experiment with the PF1479 promoter. The concentration of DNA was 20 nM, the same in each well. The proteins were added to make a gradient of protein-DNA ratio from 0 to 72. In both experiments, with increasing concentration of protein $PF0250_a$ added, the concentration of free DNA stays the same, suggesting that this protein

doesn't bind to either promoter DNA. On the other hand, the concentration of free DNA promoter gradually deceases with increasing concentration of $PF0250_b$, suggesting that this protein does remove free promoter DNA from the "free DNA" gel band. However, no bands shifted to slower migration (up on the gel) can be observed in either case, although a lot of stain seems to be in the wells of the higher protein/DNA ratio lanes. This may be because of the insolubility of the protein; PF0250p does not enter the gel, even though it binds to the promoter DNA.

Figure 3.12 shows the EMSA results of the PF0031 promoter with $PF0250_b$. The concentration of the DNA promoter was 10 nM, the same in each well. The protein-DNA ratio ranged from 0 to 72. Once again, with increasing concentration of $PF0250_b$, the concentration of free DNA promoter decreases.

Figure 3.13 shows the EMSA results of the PF0340 promoter with $PF0250_b$. No obvious binding is observed, validating the protein capture experiments.

Moreover, I also tested the binding affinity between $PF0250_c$ and the PF1479 promoter. Figure 3.14 uses 10 nM DNA, the same in each well. The lack of sufficient protein allowed protein-DNA ratios only up to 16. Once again, more protein results in less free DNA, which is completely absent at a protein-DNA ratio of 16.

3.3.4. Summary

The EMSA experiments highlighted the different DNA binding affinity of PF0250p purified by different methods. PF0250p prepared by using urea to de-

and renature the protein (PF0250$_a$) doesn't bind to any of the promoter DNA we tried, suggesting that this purification method altered the properties of PF0250p. On the other hand, the soluble PF0250p (PF0250$_b$ and PF0250$_c$) show interaction of the protein with the PF0031, PF0723 and PF1479 promoters, but clear gel-shifted bands were not detected. We suggest that this is due to the insolubility of the protein, which retains its bound DNA in the loading wells of the EMSA gel.

3.4. Fluorescence DNase Footprinting to Confirm the Binding of PF0250p to Promoters

3.4.1. Introduction

Like EMSA, DNase I footprinting is also a technique that detects DNA-protein interaction. It can provide additional information about the location of binding on the DNA for sequence-specific DNA-binding proteins. It is based on the idea that a protein bound to DNA will protect the DNA from enzymatic (nuclease) cleavage. [52]
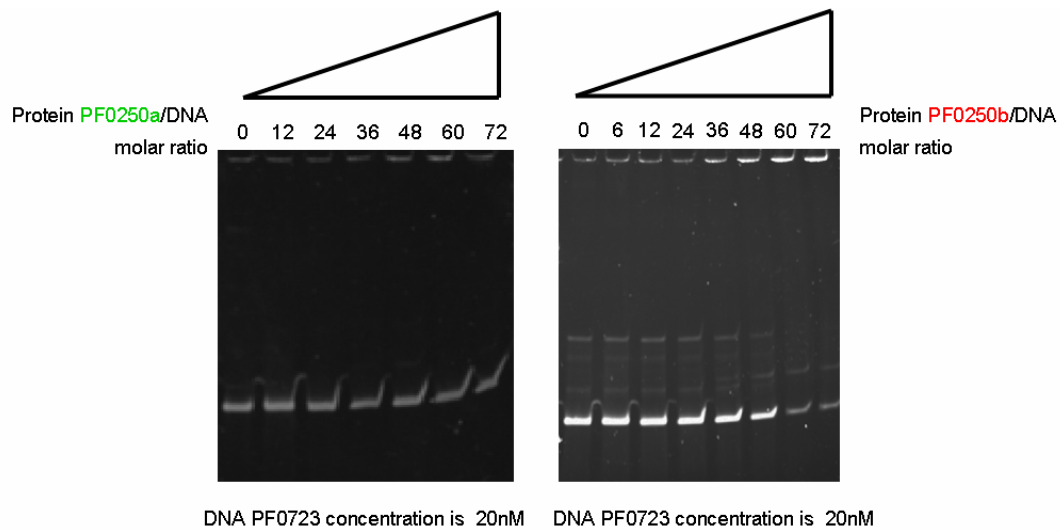
Figure 3.10. EMSA results for promoter PF0723 with proteins PF0250$_a$ (left) and PF0250$_b$ (right). Although there are no obvious shifted DNA bands in the lanes with PF0250p, more DNA moves from the free DNA band to the wells when PF0250$_b$ is used (right). This suggests that the insoluble PF0250$_b$ binds to the PF0723 promoter, but PF0250$_a$ does not.
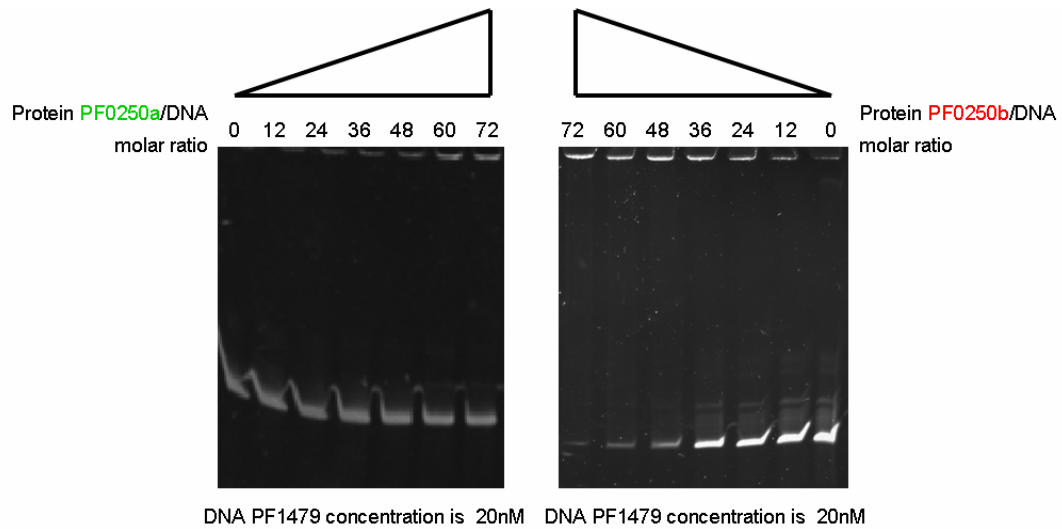
Figure 3.11. EMSA results for promoter PF1479 with proteins PF0250$_a$ (left) and PF0250$_b$ (right). The increasing DNA trapped in the wells as PF0250$_b$ is added (right) suggests again that PF0250$_b$ binds to the PF1479 promoter, but PF0250$_a$ does not.
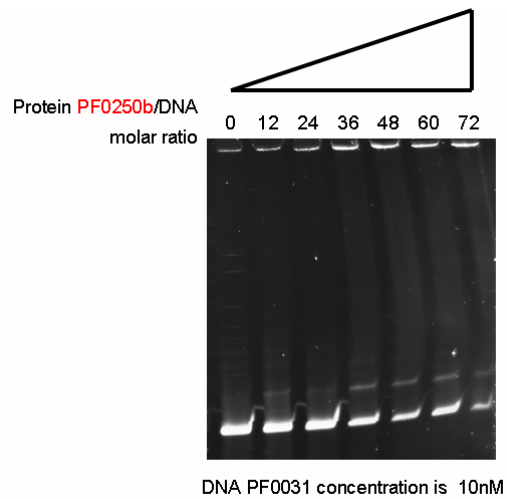
Figure 3.12. EMSA results for promoter PF0031 with PF0250$_b$. These results are similar to those in Figures 3.10 and 3.11, suggesting that the PF0031 promoter also binds PF0250$_b$.
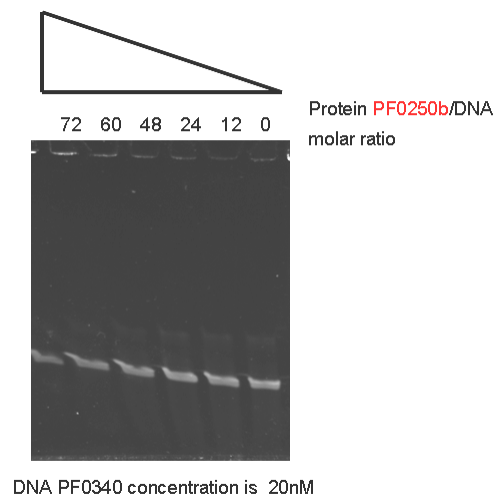
Protein PF0250b/DNA molar ratio

72 60 48 24 12 0

DNA PF0340 concentration is 20nM

Figure 3.13. EMSA results for promoter PF0340 with $PF0250_b$. There is no evidence of binding of $PF0250_b$ to the PF0340 promoter.

Protein PF0250c/DNA molar ratio

0   2   4   8   16
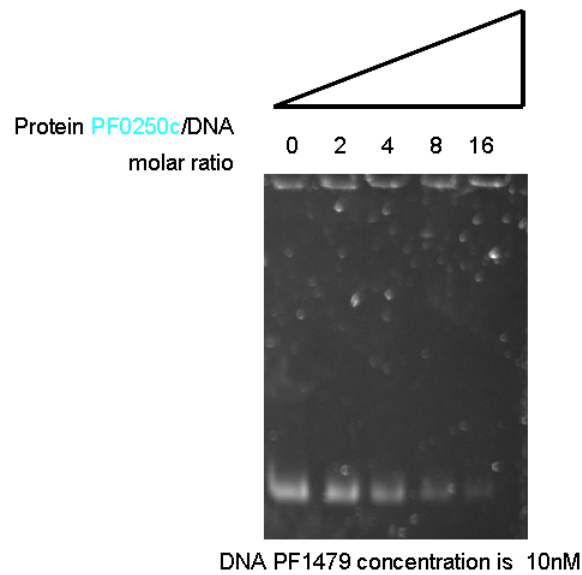
DNA PF1479 concentration is 10nM

Figure 3.14. EMSA results for the PF1479 promoter with $PF0250_c$. The protein/DNA molar ratio is from 0 to 16. Again, from the decrease in free DNA at higher protein/DNA ratios, it appears that $PF0250_c$ binds to the PF1479 promoter.

3.4.2. Method and Materials

Fluorescence DNase I footprinting was performed using the protocol from the Pratt lab [53] to identify the specific binding site of PF0250p on the PF1479 promoter. The PF1479 promoter sequence is displayed in Figure 3.15. The promoter sequence was amplified by PCR using a 5' primer with a 6-FAM fluorescent tag. The promoter spans -204 to + 95 related to the translation start site of PF1479. The dsDNA probes were PCR amplified, followed by purification using QIAEX II Gel Extraction Kit (QIAGEN). 200 ng of DNA probe was applied in 50 µL footprinting binding buffer (Table 3.3) and incubated with PF0250p at a gradient of protein-DNA ratios 0, 1, 4, 16. After incubation for 30 min at 55 °C (the same conditions as in the protein capture assay), the solution was mixed with 5 µL of DNase solution (0.01 U/µL, RNase-free) and incubated exactly 1 min at room temperature for every experiment. The cleavage reaction was stopped by applying 145 µL stop solution (Table 3.3) followed by adding 200 µL of phenol:chloroform:IAA (indoleacetic acid) (25:24:1) to extract the DNA fragments. Then, the extracted DNA was precipitated by adding 3 M sodium acetate and 2.5 volume of ethanol. The precipitated DNA was washed with 75% ethanol and finally re-dissolved in 10 µL formamide solution containing GeneScan-500-ROX size standard. The results of the fragments were analyzed on an ABI 3730x1 DNA sequencer.

3.4.3. Results and Conclusions

The footprinting analysis of PF0250p protection of DNA cleavage showed the specific binding sequence of DNA on the PF1479 promoter. Figure 3.16 shows the detailed information of the region protected by PF0250p. This region is probably located -38 to -4 relative to the PF1479 translation start site. This result suggests that our interpretation of the unusual EMSA results is probably correct and also identifies the binding sequence in the PF1479 promoter for PF0250p.

3.5. Summary

From the EMSA results, PF0250p binds to the PF0031, PF1479, and PF0723 promoters. On the other hand, no binding was observed of PF0250p to the PF0340 promoter, which is in agreement with the protein capture experiment. The footprinting showed that PF0250p binds specifically to the PF1479 promoter at a binding site -38 to -4 bp relative to the translation start site of the PF1479 gene.

To verify whether PF0250p also has a binding site in the PF0723 and PF0031 promoters, footprinting must be done with these promoters. To further determine the regulatory function of PF0250p in the transcription of PF0031, PF0723 and PF1479, *in vitro* transcription assays are needed. SELEX (selected evolution of ligands by exponential enrichment) is a possibility to further specify the consensus cognate DNA sequence for PF0250p.

```
5'-6fam-ATGAGAGTCCCAGTAGACATTTGTATAAAGGTATGGATATCAAGCATGT

CAAGAGATAGGGAGAATTAAAAAGGCAGAATTCTTGAAAAAGTTGAGGTTTAAAAAA

TATGTTAAAAACAACACTTTTTATGTTTAGAATTATATTTGAAGGTTTCAATAATCT

TTATATGAACTAACTCCCAACTATATTCCTGGTGATATAATATGAGCGAAGAAGTTC

AAGAAAGGATATGGATTCTAATAACTCCAGACAAATGTAGTGGATGTAGGCTATGTG

AAGTCACCTGTTCTCTGGAACAT
```

Figure 3.15. The *Pf* genome sequence in the region -204 to +95 relative to the translation start site of PF1479. This footprinting probe was labeled with the 6-FAM fluorescent tag on the 5' end, so we are examining the binding site(s) on the top strand. Also indicated are the TATA box (blue), ribosome binding site (underlined), and ATG as the translation start site (red).

Table 3.3. The reagents used in the fluorescence DNase I Footprinting

Binding Buffer
20 mM HEPES, pH 8.0
100 mM KCl
15 mM MgCl$_2$
5 mM CaCl$_2$
1 mM EDTA
1 mM DTT
5% glycerol

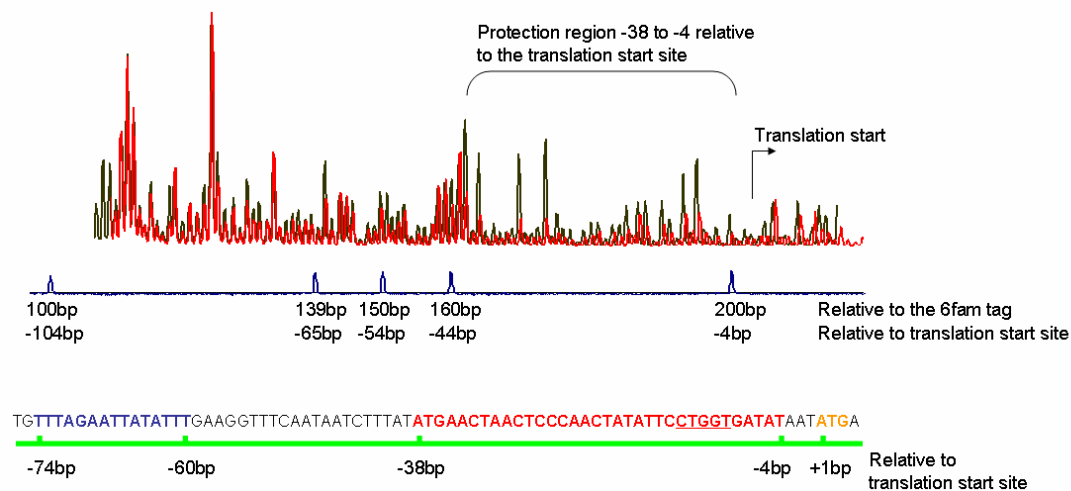| Stop Solutions |
|---|
| 130 mM NaCl |
| 20 mM EDTA |
| 0.6% SDS |

Figure 3.16. The capillary electrophoretogram of fluorescence DNase footprinting results of PF0250p interacting with the PF1479 promoter. The comparison is of control (black line, without protein) and sample (red line, with protein). The GeneScan-500-ROX size standard electrophoretogram is included (blue) to show the peak positions relative to the 6-FAM tag start site. The specific region of the PF1479 promoter protected by PF0250p is indicated. The DNA sequence highlights the TATA box (blue), the protected region (red), the ribosome binding site (underlined), and the translation start site (orange).

REFERENCES:

1. Hickey, A.J., E.C. de Macario, and A.J.L. Macario, *Transcription in the Archaea: Basal factors, regulation, and stress-gene expression.* Critical Reviews in Biochemistry and Molecular Biology, 2002. **37**(6): p. 537-599.
2. Woese, C.R., *Interpreting the universal phylogenetic tree.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(15): p. 8392-8396.
3. Ogunseitan, O., *Microbial Diversity.* 2004.
4. Bell, S.D. and S.P. Jackson, *Mechanism and regulation of transcription in archaea.* Current Opinion in Microbiology, 2001. **4**(2): p. 208-213.
5. Bell, S.D., C.P. Magill, and S.P. Jackson, *Basal and regulated transcription in Archaea.* Biochemical Society Transactions, 2001. **29**: p. 392-395.
6. Langer, D., et al., *Transcription in Archaea - Similarity to That in Eucarya.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(13): p. 5768-5772.
7. Zillig, W., K.O. Stetter, and D. Janekovic, *DNA-dependent RNA polymerase from the archaebacterium Sulfolobus acidocaldarius.* Eur J Biochem, 1979. **96**(3): p. 597-604.
8. Darcy, T.J., et al., *Methanobacterium thermoautotrophicum RNA polymerase and transcription in vitro.* J Bacteriol, 1999. **181**(14): p. 4424-9.
9. Hausner, W., et al., *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, 1996. **271**(47): p. 30144-8.
10. Hethke, C., et al., *A cell-free transcription system for the hyperthermophilic archaeon Pyrococcus furiosus.* Nucleic Acids Res, 1996. **24**(12): p. 2369-76.
11. Bell, S.D., et al., *Orientation of the transcription preinitiation complex in Archaea.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(24): p. 13662-13667.
12. Ouhammouch, M., et al., *Promoter architecture and response to a positive regulator of archaeal transcription.* Molecular Microbiology, 2005. **56**(3): p. 625-637.
13. Bell, S.D., *Archaeal transcriptional regulation - variation on a bacterial theme?* Trends in Microbiology, 2005. **13**(6): p. 262-265.
14. Bell, S.D., et al., *Transcriptional regulation of an archaeal operon in vivo and in vitro.* Molecular Cell, 1999. **4**(6): p. 971-982.
15. Bell, S.D. and S.P. Jackson, *Transcription in Archaea.* Cold Spring Harbor Symposia on Quantitative Biology, 1998. **63**: p. 41-51.
16. Ouhammouch, M., *Transcriptional regulation in Archaea.* Current Opinion in Genetics & Development, 2004. **14**(2): p. 133-138.
17. Lee, S.J., et al., *TrmB, a sugar-specific transcriptional regulator of the trehalose/maltose ABC transporter from the hyperthermophilic archaeon*

Thermococcus litoralis. Journal of Biological Chemistry, 2003. **278**(2): p. 983-990.

18. Napoli, A., et al., *An Lrp-like protein of the hyperthermophilic archaeon Sulfolobus solfataricus which binds to its own promoter.* J Bacteriol, 1999. **181**(5): p. 1474-80.

19. Dahlke, I. and M. Thomm, *A Pyrococcus homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment.* Nucleic Acids Research, 2002. **30**(3): p. 701-710.

20. Vierke, G., et al., *A novel archaeal transcriptional regulator of heat shock response.* Journal of Biological Chemistry, 2003. **278**(1): p. 18-26.

21. Brinkman, A.B., et al., *An Lrp-like transcriptional regulator from the archaeon Pyrococcus furiosus is negatively autoregulated.* J Biol Chem, 2000. **275**(49): p. 38160-9.

22. Bell, S.D., et al., *Transcriptional regulation of an archaeal operon in vivo and in vitro.* Mol Cell, 1999. **4**(6): p. 971-82.

23. Schut, G.J., J. Zhou, and M.W. Adams, *DNA microarray analysis of the hyperthermophilic archaeon Pyrococcus furiosus: evidence for anNew type of sulfur-reducing enzyme complex.* J Bacteriol, 2001. **183**(24): p. 7027-36.

24. Weinberg, M.V., et al., *Cold shock of a hyperthermophilic archaeon: Pyrococcus furiosus exhibits multiple responses to a suboptimal growth temperature with a key role for membrane-bound glycoproteins.* J Bacteriol, 2005. **187**(1): p. 336-48.

25. Fiala, G. and K.O. Stetter, *Pyrococcus-Furiosus Sp-Nov Represents a Novel Genus of Marine Heterotrophic Archaebacteria Growing Optimally at 100-Degrees C.* Archives of Microbiology, 1986. **145**(1): p. 56-61.

26. Poole, F.L., et al., *Defining genes in the genome of the hyerthermophilic archaeon Pyrococcus furiosus: Implications for all microbial genomes.* Journal of Bacteriology, 2005. **187**(21): p. 7325-7332.

27. Robb, F.T., et al., *Genomic sequence of hyperthermophile, Pyrococcus furiosus: Implications for physiology and enzymology.* Hyperthermophilic Enzymes, Pt A, 2001. **330**: p. 134-157.

28. Scott, R., *Grant Application:Metalloregulation of an Archaeal Genome.* 2005.

29. Hausner, W., G. Frey, and M. Thomm, *Control Regions of an Archaeal Gene - a Tata Box and an Initiator Element Promote Cell-Free Transcription of the Transfer Rnaval-Gene of Methanococcus-Vannielii.* Journal of Molecular Biology, 1991. **222**(3): p. 495-508.

30. Rees, D.C., *Great metalloclusters in enzymology.* Annu Rev Biochem, 2002. **71**: p. 221-46.

31. Outten, C.E. and T.V. O'Halloran, *Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis.* Science, 2001. **292**(5526): p. 2488-92.

32. Touati, D., *Iron and oxidative stress in bacteria.* Arch Biochem Biophys, 2000. **373**(1): p. 1-6.

33. Silver, S., *Bacterial resistances to toxic metal ions--a review.* Gene, 1996. **179**(1): p. 9-19.

34. Shi, W., J. Wu, and B.P. Rosen, *Identification of a putative metal binding site in a new family of metalloregulatory proteins.* J Biol Chem, 1994. **269**(31): p. 19826-9.

35.   Ochsner, U.A., et al., *GeneChip expression analysis of the iron starvation response in Pseudomonas aeruginosa: identification of novel pyoverdine biosynthesis genes.* Mol Microbiol, 2002. **45**(5): p. 1277-87.

36.   Pohl, E., et al., *Architecture of a protein central to iron homeostasis: crystal structure and spectroscopic analysis of the ferric uptake regulator.* Molecular Microbiology, 2003. **47**(4): p. 903-915.

37.   Al-Fageeh, M.B. and C.M. Smales, *Control and regulation of the cellular responses to cold shock: the responses in yeast and mammalian systems.* Biochemical Journal, 2006. **397**: p. 247-259.

38.   Jones, P.G., R.A. Vanbogelen, and F.C. Neidhardt, *Induction of Proteins in Response to Low-Temperature in Escherichia-Coli.* Journal of Bacteriology, 1987. **169**(5): p. 2092-2095.

39.   Jiang, W.N., Y. Hou, and M. Inouye, *CspA, the major cold-shock protein of Escherichia coli, is an RNA chaperone.* Journal of Biological Chemistry, 1997. **272**(1): p. 196-202.

40.   Caldas, T., S. Laalami, and G. Richarme, *Chaperone properties of bacterial elongation factor EF-G and initiation factor IF2.* Journal of Biological Chemistry, 2000. **275**(2): p. 855-860.

41.   Jones, P.G. and M. Inouye, *RbfA, a 30S ribosomal binding factor, is a cold-shock protein whose absence triggers the cold-shock response.* Molecular Microbiology, 1996. **21**(6): p. 1207-1218.

42.   Beran, R.K. and R.W. Simons, *Cold-temperature induction of Escherichia coli polynucleotide phosphorylase occurs by reversal of its autoregulation.* Molecular Microbiology, 2001. **39**(1): p. 112-125.

43.   Ermolenko, D.N. and G.I. Makhatadze, *Bacterial cold-shock proteins.* Cellular and Molecular Life Sciences, 2002. **59**(11): p. 1902-1913.

44.   Adams, M.W., et al., *Key role for sulfur in peptide metabolism and in regulation of three hydrogenases in the hyperthermophilic archaeon Pyrococcus furiosus.* J Bacteriol, 2001. **183**(2): p. 716-24.

45.   Lipscomb, G. and R.A. Scott, *Discovery of a new potential regulatory transcription factor in archaea.* Faseb Journal, 2005. **19**(4): p. A874-a875.

46.   Menon, A., *Metal stress Microarray Data.* 2005.

47.   Chen, Y., *The fishing Expedition: Searching for Unidentified Archaeal Transcription Factors.*, in *Chemistry Department.* 2003, University of Georgia: athens, GA.

48.   Wang, M., *Phd Dissertation.* 2006.

49.   Lipscomb, G., *Protein Expression and Purification Protocols.*

50.   Lee, S.J., et al., *TrmB, a sugar sensing regulator of ABC transporter genes in Pyrococcus furiosus exhibits dual promoter specificity and is controlled by different inducers.* Molecular Microbiology, 2005. **57**(6): p. 1797-1807.

51.   Lipscomb, G., *EMSA Protocols.*

52.   Brenowitz, M., et al., *Quantitative Dnase Footprint Titration - a Method for Studying Protein-DNA Interactions.* Methods in Enzymology, 1986. **130**: p. 132-181.

53.   Pratt, *Pratt's lab Fluorescence DNase Footprinting protocol.*