

RECONSTITUTION OF *PYROCOCCUS FURIOSUS* RNA POLYMERASE
AND
IDENTIFICATION OF POTENTIAL TRANSCRIPTIONAL REGULATORS RESPONDING
TO IRON DEPRIVATION IN *PYROCOCCUS FURIOSUS*

by

MINNIE MURUGESAN

(Under the Direction of Robert A. Scott)

ABSTRACT

As part of a large-scale effort to discover regulatory mechanisms in the hyperthermophilic archaeon *Pyrococcus furiosus* (*Pf*), (1) *Pf* RNA polymerase (RNAP) subunits (10) were cloned, expressed, purified and reconstituted to form *Pf* RNAP complex; (2) DNA-affinity protein capture method was used to identify PF0851p as a potential regulatory transcription factor associated with PF1281, PF1282, PF1283 operon (PF1282, encoded as rubredoxin, down-regulated under iron limitation based on Microarray expression profiling) and PF1287, PF1285, PF1286 operon (PF1285, PF1286, putative *sufD*, *sufB* respectively, up-regulated under iron limitation based on Microarray expression profiling). The binding potential of PF0851p on PF1283 UOR and PF1287 UOR has been determined by electromobility shift assay and the putative binding site for both the UORs were determined by DNase I footprinting assay.

INDEX WORDS: Archaea, *Pyrococcus furiosus*, RNA polymerase, Regulatory transcription factor, Binding site, Electromobility shift assay, DNase I footprinting assay, In vitro transcription, Digoxigenin labeling.

RECONSTITUTION OF *PYROCOCCUS FURIOSUS* RNA POLYMERASE
AND
IDENTIFICATION OF POTENTIAL TRANSCRIPTIONAL REGULATORS RESPONDING
TO IRON DEPRIVATION IN *PYROCOCCUS FURIOSUS*

by

MINNIE MURUGESAN

M.S., St. Josephs College, Bangalore, India, 1994

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2007

© 2007

MINNIE MURUGESAN

All Rights Reserved

RECONSTITUTION OF *PYROCOCCUS FURIOSUS* RNA POLYMERASE
AND
IDENTIFICATION OF POTENTIAL TRANSCRIPTIONAL REGULATORS RESPONDING
TO IRON DEPRIVATION IN *PYROCOCCUS FURIOSUS*

by

MINNIE MURUGESAN

Major Professor:	Robert A. Scott
Committee:	Michael W. W. Adams Jonathan I. Amster

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2007

DEDICATION

To my mother and father, Rathna B. Nair and A. V. Bhargavan Nair whose love and support has helped me complete this program. My sister, Vandana Nair and my son Shrideesh Nair who have endured a lot of sacrifice through tough times being away from home for such a long period of time.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Professor Robert A. Scott for his guidance and support. I thank my committee, Dr. Michael W. W. Adams, Dr. Jonathan I. Amster and Dr. Marly Eidnsess for their guidance and help. I would like to acknowledge Dr. Marly Eidnsess for teaching me the basics of molecular biology and training me adequately specially during my initial stages of research.

I wish to thank Dr. Dennis Philips (Mass Spectrometer facility, Dept. of Chemistry), Mr. Jeff Wagner (IBL, UGA), Ms. Vicki Bauer (Dept. of Chemistry), Dr. Francis Jenney (Adam's Group), Dr. Angeli L. Menon (Adam's Group), Dr. M. Roy, M. Tern's group, Dr. R. Viswanathan (Kurtz's group), Dr. F. Sugar (BioXpress), Mr. Robert (CSS) and all associated with the administration of the Department of Chemistry for helping me several times.

I wish to thank all my friends.

I would like to thank my family. My mother and father, Rathna B. Nair and A. V. Bhargavan Nair who have supported and encouraged me throughout my life. I thank my sister Vandana Nair who took care of my family in my absence and has inspired me not to give up even when times are rough. I wish to thank my son for motivating me in his own little way.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	viii
LIST OF FIGURES	x
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW	1
1.1 Hyperthermophiles	1
1.2 Transcription	2
1.3 Metals – essential for cell growth and function	7
1.4 SUF system	8
1.5 Rubredoxin, Ruberythrin, Superoxide reductase	14
1.6 Metal stress response	16
2 RECONSTITUTION OF <i>PYROCOCCUS FURIOSUS</i> DNA-DEPENDENT RNA POLYMERASE	39
2.1 Objective	39
2.2 Materials.....	39
2.3 Methods.....	44
2.4 Results and Discussion	60
3 IDENTIFICATION OF POTENTIAL TRANSCRIPTIONAL REGULATORS RESPONDING TO IRON DEPRIVATION IN <i>PYROCOCCUS FURIOSUS</i>	99

3.1 Objective	99
3.2 Materials.....	100
3.3 Methods.....	104
3.4 Results and Discussion	118
REFERENCE	148

LIST OF TABLES

	Page
Table 2.1: LC-MS data obtained from a single-step chromatographic separation of <i>Pf</i> RNAP complex from <i>Pf</i> cells.....	65
Table 2.2: Gene details and nomenclature of the <i>Pf</i> RNAP subunits	66
Table 2.3: Details of the sequence of the primers used to amplify the PCR inserts of subunits of <i>Pf</i> RNAP	67
Table 2.4: Conditions for cloning of <i>Pf</i> RNAP genes into pET151/d-TOPO vector	68
Table 2.5: Fragments of <i>Pf</i> RNAP genes generated by restriction endonucleases.....	69
Table 2.6: Calculated protein content of all the purified <i>Pf</i> RNAP subunits	70
Table 3.1a: Microarray data of <i>Pf</i> ORFs up-regulated by more than 2.5-fold in iron-limited growth conditions.....	124
Table 3.1b: Microarray data of <i>Pf</i> ORFs down-regulated by more than 2.5-fold in iron-limited growth conditions.....	125
Table 3.2a: Occurrence of DNA-binding motif (fur box consensus) of Fur in <i>Pf</i> UOR identified using MEME and correlated to iron-regulated ORFs from Microarray profiling	126
Table 3.2b: Occurrence of DNA-binding motif (DtxR operator consensus) of DtxR in <i>Pf</i> UOR identified using MEME and correlated to iron-regulated ORFs from Microarray profiling	127
Table 3.3a: Details of the UOR DNA probe used for EMSA	128
Table 3.3b: EMSA conditions using PF1283 and PF1287 UOR DNA with PF0851p.....	128

Table 3.4a: Details of the UOR DNA used for DNase I footprinting.....	129
Table 3.4b: DNase I footprinting conditions using PF1283 and PF1287 UOR DNA with PF0851p.....	129

LIST OF FIGURES

	Page
Figure 1.1: Three domains of life	19
Figure 1.2: Phylogenetic tree of Archaea.....	20
Figure 1.3a: <i>Pyrococcus furiosus</i> , coccoid in shape with monopolar peritrichous flagella.....	21
Figure 1.3b: <i>Pyrococcus furiosus</i> cells	21
Figure 1.4: Transcription process	22
Figure 1.5: Eukaryotic transcription initiation complex	23
Figure 1.6: Yeast RNA polymerase II consisting of 12 subunits	24
Figure 1.7: Comparison map of sequence homology between yeast RNAP II subunits and <i>Pf</i> RNAP subunits obtained using the BLASTP program on NCBI database.....	25
Figure 1.8: Three-dimensional model of the core of yeast RNAP II and <i>E. coli</i> RNAP	26
Figure 1.9: Model comparing the subunits of yeast RNAP II and <i>Pf</i> RNAP.....	27
Figure 1.10: Model of mechanism of transcriptional repression.....	28
Figure 1.11: Number of transcriptional regulators in archaeal genomes.....	29
Figure 1.12: Model of iron assimilation and repair of [Fe–S] clusters under oxidative stress and iron limitation in <i>E. coli</i> . Under iron limitation and oxidative stress, the reductase converts (Fe ³⁺)-loaded siderophore (chrysobactine) to its apo form and repairs the damaged [4Fe–4S] cluster	30
Figure 1.13: <i>suf</i> homologs present in <i>E. coli</i> , archaea, and <i>T. maritima</i>	31
Figure 1.14: Map of <i>suf</i> operon present in archaea.....	32

Figure 1.15: The occurrence of sufS in archaea	33
Figure 1.16: Proposed mechanism of Fe-S cluster formation in SUF system in <i>E. coli</i>	34
Figure 1.17: Three dimensional model of Fur dimer-DNA complex constructed from X-ray structure of Fur (<i>P. aeruginosa</i>).....	35
Figure 1.18: Deduced consensus sequences of DtxR binding sites.....	36
Figure 1.19: DNA-binding domain and crystal structure of DtxR in <i>Corynebacterium diphtheriae</i>	37
Figure 1.20: Metal-binding sites of DtxR	38
Figure 2.1: SDS-PAGE gel of His ₆ -RpoD over-expressed with the addition of IPTG	71
Figure 2.2: Purified recombinant His ₆ -RpoD.....	72
Figure 2.3: Zinc affinity purification of <i>Pf</i> cell extract treated with polymine P combined with His ₆ -RpoD.....	73
Figure 2.4: Native-PAGE gel of fractions eluted from Zn chelating column	74
Figure 2.5: LC-MS chromatograms of putative native <i>Pf</i> RNAP complex extracted from <i>Pf</i> cells using recombinant His ₆ -RpoD.....	75
Figure 2.6: Native-PAGE gel and western blot of <i>Pf</i> RNAP cell extract	78
Figure 2.7: Schematic representation of the vector used to clone <i>Pf</i> RNAP subunits.....	79
Figure 2.8: PCR product and restriction endonuclease digest map of <i>Pf</i> RNAP subunits.....	80
Figure 2.9: Restriction endonuclease Nsi I digest map of plasmid of <i>Pf</i> RNAP subunits.....	83
Figure 2.10: Electropherogram of RpoA.....	86
Figure 2.11: SDS-PAGE gel of <i>Pf</i> RNAP subunits	88
Figure 2.12: MALDI mass spectra and sequence obtained using MS-Fit program	89
Figure 2.13: RNA transcripts generated using T7 RNAP and pGEM template.....	96

Figure 2.14a: RNA transcripts generated by in vitro transcription reaction using MBH template and reconstituted <i>Pf</i> RNAP.....	97
Figure 2.14b: 10% TBE-urea gel and Nylon membrane blot of RNA generated using T7 RNAP and reconstituted <i>Pf</i> RNAP.....	98
Figure 3.1: Microarray scatter plot of iron-sufficient and iron-limited <i>Pf</i> ORFs with 2-fold regulation boundaries	130
Figure 3.2: Sequence of the UOR DNA.....	131
Figure 3.3: Schematic representation of DNA-affinity protein capture by direct and indirect approach.....	132
Figure 3.4a: SDS-PAGE gel of proteins bound to PF1283 UOR DNA probe.....	133
Figure 3.4b: SDS-PAGE gel of proteins bound to PF1287 UOR DNA probe.....	134
Figure 3.5: Conserved domain database search results for PF0851p	135
Figure 3.6: Secondary structure prediction of PF0851p using PSIPRED view	136
Figure 3.7: A-Primary structure alignment of DtxR homologs regulated by Mn(II) or Fe(II); B- Compilation of 19 bp DtxR-like binding site; C-Compilation of 19 bp MntR-like binding site.....	137
Figure 3.8: Comparison map of sequence homology of PF0851p (query) with known metal- dependent repressors	138
Figure 3.9: Validation of sequence of PF0851 plasmid.....	139
Figure 3.10: EMSA of UOR DNA with PF0851p.....	140
Figure 3.11a: CE chromatograms with footprints of PF0851p on PF1283 UOR DNA	141
Figure 3.11b: CE chromatograms with footprints of PF0851p on PF1287 UOR DNA	143
Figure 3.12: Binding site of PF0851p on PF1283 and PF1287 UOR DNA	145

Figure 3.13: Sequence conservation of PF0851p with other homologs in archaea.....	146
Figure 3.14: 3D model of PF0851p built based on 16 known X-ray structure of DtxR-like proteins using Geno3D.....	147

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 **Hyperthermophiles:** Hyperthermophiles are microorganisms found in marine environment, deep-sea volcanic hydrothermal vents and in shallow geothermal-heated vents. They are also found in acidic environments, sulfur (S°) containing solfataric fields and in alkaline hot water springs [1, 2]. High concentrations of hyperthermophiles are also present in oil reservoirs [3, 4]. *Aquifex pyrophilus* and *Thermotoga maritima* are hyperthermophiles that belong to the bacterial domain; all the other hyperthermophiles are classified as archaea.

1.1.1 **Archaea:** Archaea represent the third domain of life. The phylogenetic tree of life based on 16S rRNA sequences proposed by Woese and coworkers constitutes three main domains: bacteria, eukaryotes and archaea [5, 6] (Figure 1.1). Archaea are single-celled prokaryotes like bacteria, but they possess cell walls, cell membranes, transcription and translation machinery that resemble that of the eukaryotes. The capacity of many archaea to thrive under extreme conditions (e.g. temperature) that usually cause protein denaturation and DNA breaks presents a unique opportunity for study. The archaeal domain is divided into three phyla: euryarchaeaota, crenarchaeaota and nanoarchaeaota (Figure 1.2). Euryarchaeaota, from the Greek word “wide” represents phenotypic and metabolic diversity. This phylum is composed of halophiles, methanogens, thermoacidophiles and hyperthermophiles [7]. Most hyperthermophilic archaea

are obligate anaerobes with the exception of a few species, which require oxygen. Many hyperthermophiles fall into the “sulfur-dependent” category since they are able to reduce elemental sulfur (S°) to H_2S . With the exception of *Pyrococcus* and *Thermococcus*, which have the ability to grow in the absence of S° , all the others utilize sulfur for their metabolic activity [8].

1.1.1.1 *Pyrococcus furiosus*: *Pyrococcus furiosus*, (*Pf*; literally "angry fireball") a member of the Euryarchaeota branch Thermococcaceae, is a hyperthermophilic marine archaeon. It was first isolated from geothermal heated marine sediments off the coast of Volcano, Italy by Stetter and coworkers [4, 9]. *Pf* represents a novel genus of marine heterotrophic archaea growing optimally around 100°C. The size of this hyperthermophilic archaeon ranges from 0.8 to 2.5 μm in diameter, is coccoid in shape and possesses monopolar peritrichous flagella (Figure 1.3a - 1.3b). *Pf* is an obligate anaerobe and inhabits environments with temperatures ranging from 70°C to 104°C, pH values ranging from 5.0 to 9.0 and salinity range of 0.5 to 5%. In the laboratory, *Pf* has an optimal growth temperature of 100°C. It has the ability to grow both proteolytically on peptides (in the absence or presence of S°) and saccharolytically utilizing starch, glycogen or cellobiose as its primary carbon source. Acetate, H_2 and CO_2 are produced as metabolic end products [10-12].

1.2 **Transcription:** The post-genomic research presents a challenging task to construct the complete transcriptional regulatory network encoded by a genome sequence [13, 14]. The transcriptional network of the gram-negative bacterium *Escherichia coli* is currently the best understood regulatory system of a single

cell. The key components in regulation of bacterial gene expression are DNA-binding transcription factors that are able to sense the changing environmental conditions and modulate the expression of relevant target genes. In order to understand the regulation of gene expression in a bacterial cell in its entirety, the key factor is to identify the regulatory proteins and all the operons that are under the transcriptional control of the identified regulators [15, 16]. However, only a small fraction of the transcriptional regulatory interactions are currently known even in a well-researched organism like *E. coli*. Transcription is a fundamental biological process that uses the DNA-dependent enzyme RNA polymerase (RNAP) to convert the genetic information from DNA to RNA, which then encodes proteins responsible for specific functions (Figure 1.4). In eukarya, the transcriptional process involves the RNAP along with other general transcription factors such as TFIIA, TFIIB, TFIID, TFIIIE, TFIIF, and TFIIF to synthesize messenger RNA (mRNA) [17]. Since proteins control the phenotypic expression of the organism, the selection of the amount and time at which an RNA molecule is transcribed from a particular gene is tightly regulated. In eukarya, transcription initiation involves a cascade of protein assemblies at several regulatory DNA sequences collectively referred to as the promoter. One of the most easily recognized eukaryal elements is the TATA or Goldberg-Hogness box. It is an AT-rich region of DNA with the consensus $\text{TATA}^{\text{A}}/\text{T}^{\text{A}}/\text{T}$ located approximately 25-30 bases upstream of the transcription start. In contrast to the bacterial homolog, the Pribnow box, mutation or deletion of the TATA box does not halt transcription. Archaeal transcription requires a relatively small number of

components compared to that in eukarya. The transcriptional machinery associated with basal transcription in archaea is composed of RNA polymerase (RNAP), TATA binding protein (TBP) and transcription factor B (TFB). The TATA binding protein (TBP) recognizes the promoter of a gene by making sequence-specific contacts with the TATA box [18]. Transcription factor B (TFB) binds to the TBP-promoter complex in a sequence-specific manner as well, recognizing a second conserved region, the BRE. The RNA polymerase is then recruited to the promoter forming the pre-initiation complex and facilitating transcription initiation [19-21] (Figure 1.5).

1.2.1 DNA-dependent RNA Polymerase: RNA polymerase is the key component of the transcription machinery and is recruited to the promoters by TATA-binding protein (TBP) and transcription factor IIB (TFIIB) in eukaryotes [22]. Similar to eukaryal polymerases, the archaeal versions are composed of 7-12 polypeptides [23-25]. DNA-dependent RNA polymerase of *Pf* also has 12 polypeptides [26, 27]. Jean Marx precisely states the significance of RNA polymerase, "If an enzyme does the cell's heavy lifting, it's RNA polymerase II. Its job: getting the synthesis of all the proteins in higher cells under way by copying their genes into RNAs, and doing it at just the right time and in just the right amounts. As such, pol II, as the enzyme is called, is the heart of the machinery and controls everything that cells do..." [28]. Archaeal RNAP share close resemblance to that of eukarya compared to the bacterial RNAP (four subunits $\alpha_2\beta\beta'$ and σ which forms the core). In yeast, DNA-dependent RNAP II is a multisubunit complex consisting of 12 polypeptides [29-31] (Figure 1.6). Sequence conservation

between archaeal RNA polymerase and the yeast RNA polymerase II has been demonstrated [26, 27] (Figure 1.7). Due to the presence of a core RNA polymerase in the yeast [31-33] and in *E. coli* [34] (Figure 1.8), it can be hypothesized that a similar core exists in *P. furiosus* RNAP, demonstrated by the sequence conservation between subunits [22, 31, 35, 36]. Archaeal RNAPs have their corresponding homologs in eukaryal RNAP II. The gene encoding the largest subunit in eukaryal RNAP, RpoA/RPB1 and α_2 in *E. coli* are homologous to two genes encoding subunits A and A' (or labeled as A' and A'') in Archaea. RpoB/RPB2 in eukaryal RNAP and the corresponding bacterial homolog represented by β and β' are denoted as two subunits B and B' (or labeled as B' and B'') in methanogens and extreme halophilic archaea [37]. RpoD/RPB3, RpoH/RPB5, RpoK/RPB6, RpoM/RPB9, RpoN/RPB10, RpoL/RPB11, RpoP/RPB12 are present in all the three eukaryotic RNAP I, RNAP II, RNAP III and in archaeal RNAP, however they do not have any bacterial homologs. RPB4 and RPB7 of eukaryal RNAP II (also represented as RpoF and RpoE respectively), have unique corresponding homologs in archaea. RpoE/ RPB7 is represented as RpoE/RpoE' in archaea and the second gene, RpoE'' which is highly conserved in archaea does not have a corresponding homolog in eukaryal RNAP. The subunit composition of RNAP in crenarchaeota and euryarchaeota is made up of BAA'DEFLHNKP subunits [38, 39] (Figure 1.9).

1.2.2 Regulated Transcription: Transcription in archaea is tightly regulated by various regulatory transcription factors, which include activators and repressors that respond to environmental challenges or other stress-related conditions.

Comparative analysis of several archaeal genomes to that in bacteria and eukarya reveal that archaea contain a significant proportion of bacterial transcriptional regulators in addition to a few others in eukarya [40]. Transcriptional regulation by prokaryotic transcriptional activators possibly occurs by either promoting the recruitment of RNAP and other basal transcription factors, by stabilization of the transcriptional pre-initiation complex or by interaction with the promoter region of the DNA in a sequence-specific manner (also called the operator). On the other hand, the repression mechanism in archaea could follow either of the two pathways: In one of the proposed mechanisms, the regulatory transcription factor which functions as a repressor binds to a region overlapping the TATA/BRE promoter element and blocking the TBP/TFB interaction with the promoter (e.g. TrmB) [41]. Negative regulation could result due to the repressor binding close to the transcription start site on the promoter region thereby preventing recruitment of RNAP (e.g. LrpA) [42, 43] (Figure 1.10). A comparative study of 15 archaeal genomes based on sequence homology indicates the presence of iron-dependent repressors (Pfam: PF01325) in archaea [40] (Figure 1.11).

1.2.2.1 *Transcription in Pyrococcus furiosus:* Transcription under *Pf* growth conditions (100°C) requires an unusual and perhaps subtle array of thermoadaptations. Regulation of expression of genes in response to cellular environment often occurs at the level of transcription. The significance of obtaining a reconstituted *Pf* RNAP is to validate the newly identified transcriptional regulatory factors responding to environmental stress by in vitro transcription reaction.

1.3 Metals – essential for cell growth and function: Transition metals such as iron are involved in electron-transport mechanisms, in redox and non-redox catalysis, in gene regulation and as sensors within all living organisms, hence iron is an essential nutrient for cell growth [44, 45]. Alteration of iron pools in an organism result in a network of events, acting at the transcriptional level to change the expression of proteins involved in transport, cellular uptake, utilization and storage of iron [45]. Iron homeostasis is the ability of a cell to regulate iron levels by utilizing iron uptake, transport and storage proteins to maintain a stable dynamic equilibrium of iron within the cell [46] (Figure 1.12). Thus, the tendency of a cell to maintain Fe homeostasis by regulating transcription of key transport proteins is a prominent mechanism to be investigated in *Pf* cells. Iron is the most difficult of all the metals to investigate experimentally because it is so ubiquitous as a trace contaminant. The goal is to achieve conditions under which there is little if any growth of *Pf* so that metal-limited (batch) and metal-shock (kinetic) experiments can be conducted. For iron, this involved adding a chelator like bathophenanthroline disulphonate (BPS). The goal is to obtain conditions where addition of the target metal dramatically stimulates growth. The primary objective is to define the metal ion concentration required to obtain a change in the growth rate.

1.3.1 *Iron-sulfur clusters:* Fe-S clusters form an integral part of diverse biological processes such as energy conversion, the citric acid cycle, nitrogen fixation, amino acid biosynthesis, heme and biotin biosynthesis, DNA synthesis, DNA repair, and regulation of gene expression [47-49]. However, the biogenesis of

iron-sulfur clusters is still not fully understood. Fe–S proteins are involved in various cellular processes ranging from electron transfer to transcription regulation [48, 50]. Fe–S clusters can assemble in vitro spontaneously under favorable conditions in the presence of free iron and sulfide whereas biosynthesis of Fe–S clusters in vivo is a complex process involving many specific proteins. In bacteria, NIF, ISC and SUF are three systems involved in Fe–S cluster assembly machinery [51]. The SUF system was first identified in *Escherichia coli* and is found in both prokaryotes including archaea and higher plant plastids [44, 52].

1.4 SUF system: The SUF system in *E. coli* comprises 6 genes in *suf*ABCDSE operon. It is evident that *sufB* and *sufC* genes are conserved in all the three domains of life. Components of the SUF system such as SufS, SufD, SufA, and SufE are present in bacteria; however, the archaeal system lacks some of these counterparts. In bacteria, the expression of the gene cluster *suf*ABCDSE is stimulated by cellular oxidative stress and iron starvation, implying that the *suf* genes are probably involved in repairing the damaged iron-sulfur clusters in cells [53, 54].

1.4.1 SufC: SufC is probably a key component in the SUF system, because SufC-deficient mutants in *Erwinia chrysanthemi* exhibit phenotypes sensitive to oxidative stress or iron limitation [55, 56]. SufC forms a complex with SufB and SufD, but the SufBCD complex has no transmembrane region, existing as a soluble stable complex. The functional roles of SufB and SufD remain unclear [56-58]. The SufBCD complex may also function, however, in another cluster-

building step using the energy from ATP hydrolysis as a scaffolding system for assembly of transient Fe-S clusters and transfer to target proteins since the complex contains four highly conserved and six partially conserved cysteines [57, 59]. The sequence of SufB and SufD are distinct, however, the C-terminal domain of SufB and SufD is conserved (< 30% identity). Some organisms like *Thermoplasma acidophilum* and *Bacillus subtilis* possess a sufB-sufB pair instead of a sufB-sufD pair (Figure 1.13, 1.14). It is therefore likely that sufD was the result of gene duplication of sufB and subsequent domain shuffling. The crystal structures of SufC from *Thermus thermophilus* HB8, *Methanocaldococcus jannaschii* have been determined [60-65]. The crystal structure of SufD from *E. coli* predicts that the α -helices of the conserved C-terminal are involved in the interaction with SufC [66]. From the structural data obtained, it has been proposed that SufB/SufD forms a heterodimer interacting with the Q-loop of SufC and the activity of ATP hydrolysis of SufC is increased in the presence of complex formation of SufBCD [58, 60, 67].

1.4.2 SufA: SufA has high sequence homology with IscA, including three conserved cysteine residues, and has been shown in vitro to act as a scaffold for both [2Fe-2S] and [4Fe-4S] clusters as previously demonstrated for IscA and IscA/NifA [68, 69]. However, SufA and IscA do not have corresponding homologs in *Pyrococcus furiosus*.

1.4.3 SufS: The SufS protein is the best biochemically characterized. The NifS/IscS group of proteins in *E. coli* [57, 70], *Erwinia chrysanthemi* [46], and *Synechocystis* [71] has been extensively studied. SufS functions as a cysteine

desulfurase producing sulfur required for Fe–S assembly by the conversion of L-cysteine to L-alanine [72, 73]. The SufS reaction is slower compared to that of NifS/IscS with specific activities in the range 8–25 mU-mg⁻¹ attributed to the ineffective cleavage of the C–S bond of the cysteine substrate [46, 57, 70]. The closest predicted homolog of SufS in *Pyrococcus furiosus* based on sequence conservation is PF0164 (Figure 1.15).

1.4.4 *SufE*: SufE interacts with SufS and enhances its desulfurase activity. The conserved cysteine residue of SufE has been shown to accept a sulfur atom from SufS [46, 57]. The corresponding homolog of SufE is absent in *Pyrococcus furiosus* based on sequence conservation.

1.4.5 *SUF system and transcription factors*: Most of the *suf* genes are found in many organisms, but the complete *suf* operon is seldom conserved [74]. The most conserved *suf* genes are *sufB* and *sufC*, which are found in eubacteria, archaea, plants and parasites. Many species of archaea contain the SufBCD homologs but no IscA/SufA or IscU/SufU scaffold proteins [44, 59]. For example in organisms like *Thermatoga maritima* (*Tm*), which is, classified as hyperthermophilic bacteria, the SUF system appears to be the only system for Fe-S cluster assembly. It does not have the ISC and the NIF system. Moreover the *suf* operon in *Tm* does not have a SufA scaffold protein. Using sequence homology, the corresponding *suf* genes that occur as an operon in *Pyrococcus furiosus* are annotated as PF1285 (putative *sufD*), PF1286 (putative *sufB*) and PF1287 (putative *sufC*). Mutations in *suf* genes in *E. coli* proved to be lethal. It has been shown by microarray profiling that *suf* transcription is up regulated as part of the *E. coli* response to

hydrogen peroxide, which emphasizes the importance of the Suf proteins during oxidative stress conditions [44, 75]. In addition, iron starvation also induces physiological challenge, which activates Suf production [54]. Transcriptional control of the *suf* operon in *E. coli* uses regulators of gene expression such as OxyR and IHF (integration host factor) during oxidative stress and the metalloregulatory protein FUR induced in the presence of excess iron [52-54] (Figure 1.16). Additionally, the Suf proteins play an important role in sulfur metabolism during stress in *E. coli* and other organisms like *Synechocystis*, *B. subtilis* [76] and *Mycobacterium smegmatis* [77].

1.4.5.1

Fur and DtxR: Excess iron can also be toxic to the cells. Iron uptake is therefore regulated to maintain the intracellular concentration of the metal within desirable limits. This regulation of iron was first demonstrated in *E. coli* and the regulator was named FUR (Ferric Uptake Regulator). The Fur protein in *E. coli* is a transcriptional repressor of iron-regulated promoters [78, 79]. Under iron-rich conditions, Fur binds the divalent ion, acquires a configuration to bind target DNA sequences (Fur boxes or iron boxes), and inhibits transcription of all the genes repressed by the metal. However, under iron-limiting conditions, the equilibrium is displaced to release Fe^{2+} , the RNA polymerase accesses the genes for the biosynthesis of siderophores and other iron-related functions are expressed [80-82]. The crystal structure and modeling of the FUR system in *Pseudomonas aeruginosa* reveals that the DNA-binding domain of Fur consists of a winged helix (WH) motif and the consensus sequence of fur operator is a tandem hexamer repeat of 5'-NAT(A/T)AT-3' [83, 84] (Figure 1.17). Like Fur, DtxR also controls

iron uptake functions and thereby plays a similar role to that of the ferric uptake regulator [82]. In many gram-positive species, e.g., *Corynebacterium diphtheriae*, *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and *Streptomyces coelicolor*, proteins of the DtxR family function as global iron regulators. The DtxR protein has a major role in controlling the expression of genes involved in iron metabolism and exerts dual regulatory function, as repressor of genes participating in iron uptake, utilization, and an activator of genes responsible for iron storage and DNA protection (e.g. *M. tuberculosis*) [85, 86]. When not complexed with iron, the DtxR proteins appear to be inactive [87]. DNase I footprinting analysis was used to identify the DtxR binding site on the basal promoter of the operon containing a 27-bp interrupted palindrome, while in vitro affinity selection has identified the minimal DtxR operator to be a 19-bp palindrome separated by a single base pair [88] (Figure 1.18). In addition, binding of DtxR to its operator site has been shown to occur in the presence of divalent metal ions (e.g., Mn^{2+} and Co^{2+}) other than Fe^{2+} . The crystal structure of dimeric DtxR indicates the presence of two metal binding sites and a helix-turn-helix DNA-binding domain at the N-terminal. The divalent cation controls the motion of the HTH motif and facilitates DNA recognition. Molecular modeling of DtxR protein with the tox operator suggests that the bound DNA adopts a bent conformation with $\alpha 3$ helix of DtxR interacting with the major grooves [89] (Figure 1.19). The metal ions in DtxR along with Cys102 (disulfide bridge) are involved in the formation of the dimer (Figure 1.20). SufR like DtxR, functions as a repressor of the *sufBCDS* operon, and it has been proposed that SufR senses

the levels of iron-sulfur clusters in the cell through its own unstable iron-sulfur cluster. In the presence of Fe-S cluster, SufR binds to its operator and represses the *suf* operon. Repression is relieved, if the Fe-S cluster is oxidatively damaged or absent [90].

1.4.5.2 *OxyR*: The known regulator of the *suf* operon in *E. coli* is the oxidative stress-dependent OxyR activator [52, 55, 91]. Mutational studies led to the discovery of OxyR regulatory protein [92]. Some other OxyR-activated genes include *katG* (encoding hydroperoxidase I), *ahpCF* (encoding alkyl hydroperoxide reductase), *oxyS* (encoding a small regulatory RNA), *dps* (encoding a nonspecific DNA binding protein), *gorA* (encoding glutathione reductase), *grxA* (encoding glutaredoxin 1), *trxC* (encoding thioredoxin 2), *fur* (encoding the Fur repressor of ferric ion uptake), and *dsbG* (encoding disulfide chaperone-isomerase) [93, 94]. OxyR also has been shown to be a repressor of its own expression as well as that of *fluF* (encoding a ferric ion reductase) and *flu* (encoding the antigen 43 outer membrane protein) under oxidative stress. Based on sequence conservation (<http://string.embl.de>), *Pf* does not contain the corresponding homolog of OxyR, however other OxyR activated genes such as *aphC* and *aphF* encoded, as PF0722 and PF1422 respectively are present.

1.4.5.3 *IHF*: Regulation of gene transcription in both prokaryotes and eukaryotes involves formation of specific protein–DNA complexes of higher order, by bringing distant regions of DNA together [95]. Integration host factor (IHF) is a transcriptional regulator associated with the SUF system involved in DNA bending, a contributing factor to the specificity of binding in *E. coli* [95]. IHF in

E. coli plays an essential role in several cellular processes that include site-specific recombination, transcription and DNA replication. *Pf* does not have a corresponding homolog of IHF.

1.5 Rubredoxin, Ruberythrin, Superoxide reductase: Study of hydrothermal vent system has shown that exposure of anaerobic hyperthermophilic archaea including *P. furiosus* to significant levels of oxygen or mixing of hot anaerobic vent fluids with cold, oxygen-saturated seawater can cause cellular damage. An oxygen detoxification system, constituting the superoxide reductase (SOR) reduces (rather than dismutates) superoxide to hydrogen peroxide in anaerobes like *Pf* [8, 96, 97]. Subsequent studies have shown that the SOR-related detoxification system is also present in *Archaeoglobus fulgidus*, *Desulfovibrio gigas*, *Desulfoarculus baarsii*, and *Treponema pallidum* [98-100]. For the superoxide reduction process, rubredoxin (Rd), a small (6 kDa) iron-containing protein, serves as the electron donor in *Pf* and is reduced in vivo by NADH-dependent peroxidase and rubredoxin oxidoreductase (NROR) [8, 101, 102]. NROR provides the reduced rubredoxin as an electron donor for superoxide reductase to reduce superoxide during oxidative stress in *Pf*. Hydrogen peroxide produced by the reduction of superoxide is removed through reduction by enzymes such as peroxiredoxin and NADH peroxidase [103]. Rubrerythrin (Rr), another nonheme iron-containing protein unique to anaerobes, has been shown to function as an NADH-dependent peroxidase in *P. furiosus* [104-106]. The discovery of superoxide reductase (SOR) activity generated by the non-heme iron proteins (neelaredoxin, desulfoferrodoxin), postulates the presence of a reactive

oxygen species (ROS) detoxification pathway in anaerobes instead of the superoxide dismutase (SOD)-based pathway present in aerobes [107]. Recent studies have shown that Rd serves as an electron donor for Rr, a peroxidase enzyme involved in the oxidative stress response of *Pf*. Consistent with a physiological relationship between SOR, Rd and Rr is the fact that the genes encoding these three proteins are clustered together in the *Pf* genome and are labeled as PF1281, PF1282, and PF1283 respectively.

1.5.1 Ruberythrin: Rubrerythrin (Rr) is a non-heme iron protein that was originally isolated from the cytoplasm of the anaerobic bacterium *Desulfovibrio vulgaris* and later in many other organisms [108]. Ruberythrin (Rr) is a homodimer (monomer molecular weight of 19.4 kDa) with two separate metal binding domains. Rr derives its name from these two words: RUBREdoxin and hemeRYTHRIN. The rubredoxin-like domain in Rr (*D. vulgaris*) contains two iron atoms and one zinc atom per monomer coordinated by four cysteine residues [109, 110]. The other domain is comprised of four helices and contains a binuclear center that is coordinated by one His, five Glu residues, and an oxo bridge. The metal content and redox state of the purified forms of Rr from *D. vulgaris* affect its catalytic activity. The oxidized form of Rr shows inorganic pyrophosphatase activity and the reduced form is inactive. Rr participates in the SOR-pathway combining with Rd in the reduction of the peroxide generated by SOR [111, 112]. The crystal structure of Rr in *P. furiosus* has been determined and contains two metal binding sites (both occupied by Fe) responsible for the peroxidase activity [113]. Oxygen tends to attack iron sulfur clusters [4Fe-4S],

oxidatively inactivating them resulting in the release of large amounts of iron into the cytosol. H_2O_2 released forms $\text{HO}\cdot$, a powerful oxidant that can react with most biomolecules (DNA, protein, membrane lipids). The regulation of this system in *E. coli* is carried out by SoxRS transcription factors (in response to O_2) and the OxyR transcription factors (in response to H_2O_2). OxyR and SoxR are typically not present in anaerobic organisms. Identification of similar transcriptional regulators in *Pf* under iron-limited conditions would provide information on its response to oxidative stress and ultimately help in constructing the regulatory network.

1.6 Metal stress response: Fe-S clusters are abundant in almost all organisms and Fe-S biosynthesis has been shown to be remarkably conserved in eukarya, bacteria and archaea [59, 114]. The effect of Fe limitation on batch cultures of *Pf* was investigated using DNA microarrays by A. L. Menon in the laboratory of M. W. W. Adams, UGA. In the ‘standard’ *Pf* medium, Fe is added to a final concentration of $7.4\ \mu\text{M}$. Fe-limited medium ($\text{Fe} < 0.8\ \mu\text{M}$) was obtained by omitting Fe (and using acid-washed glassware and high-purity reagents). Cells were grown at 95°C under otherwise standard conditions using maltose as the carbon source in the absence of elemental sulfur [115]. The following results obtained were from three Fe-sufficient and three Fe-limited batch cultures. A total of 34 ORFs were up-regulated > 2.5 -fold and 61 ORFs were down-regulated > 2.5 -fold by Fe limitation. The ORFs that were up-regulated include: PF0723 (7.5-fold), putative high-affinity iron transporter which has significant homology to the Ftr1 family of high-affinity iron permeases; PF0857 (4.7-fold), which is

homologous to FeoB, a high-affinity ferrous iron transporter protein that serves as a virulence factor in *Helicobacter pylori* [116, 117]; PF1285 (5.0-fold) and PF1286 (6.1-fold), the putative Fe-S cluster assembly/repair proteins, which are homologous to sufD and sufB, respectively [44, 118]; PF1657-PF1666 (up to 5.8-fold), two amino acid biosynthetic clusters which are potentially involved in histidine biosynthesis; PF0935-PF0942 (up to 4.5-fold), which are potentially involved in branched-chain amino acid biosynthesis; and PF0725-PF0728 (up to 3.5-fold, a four-gene cluster of conserved hypothetical and hypothetical ORFs. The ORFs that were down-regulated include: PF0153 (up to 5.0-fold, annotated as phosphoribosylaminoimidazole-succinate); PF0339-PF0340 (up to 4.9-fold, annotated as methyltransferase and putative HTH transcription regulator respectively); PF0422 (up to 4.6-fold, annotated as phosphoribosylamine-glycine ligase); PF0430-PF0432 (up to 7.2-fold, annotated as phosphoribosylglycinamide formyl transferase, phosphoribosylformylglycinamide cyclo-ligase, and putative sugar-catabolism phosphotransferase respectively); PF0971 (3.1-fold, annotated as 2-ketovalerate ferredoxin oxidoreductase); and PF1282 (5.5-fold, annotated as rubredoxin (Rd)). Among these regulated genes, the putative *suf* operon (PF1285, PF1286, PF1287) and the operon that included genes PF1281, PF1282, PF1283 (encoding SOR, Rd, Rr, respectively) were selected for the study. The present study focused on discovery of regulatory transcription factor(s) in the archaeon *Pyrococcus furiosus* (*Pf*) responding to iron limitation as a stress response. An approach developed by the Scott group called STRES (Survey of Transcriptional Response to Environmental Stress) uses techniques like microarray expression

profiling, bioinformatic analysis, DNA-affinity protein capture, electromobility shift assay, footprinting, SELEX other biophysical and analytical methods to identify and characterize these regulatory transcriptional factors [119].

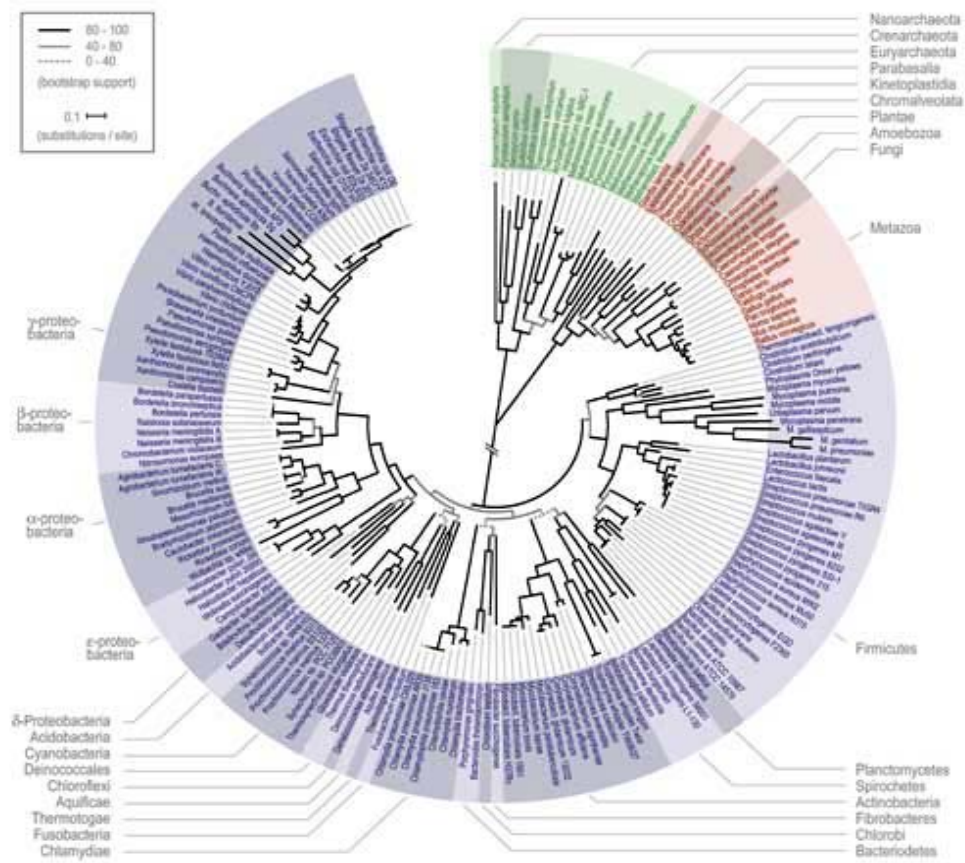


Figure 1.1: Three domains of life [120].

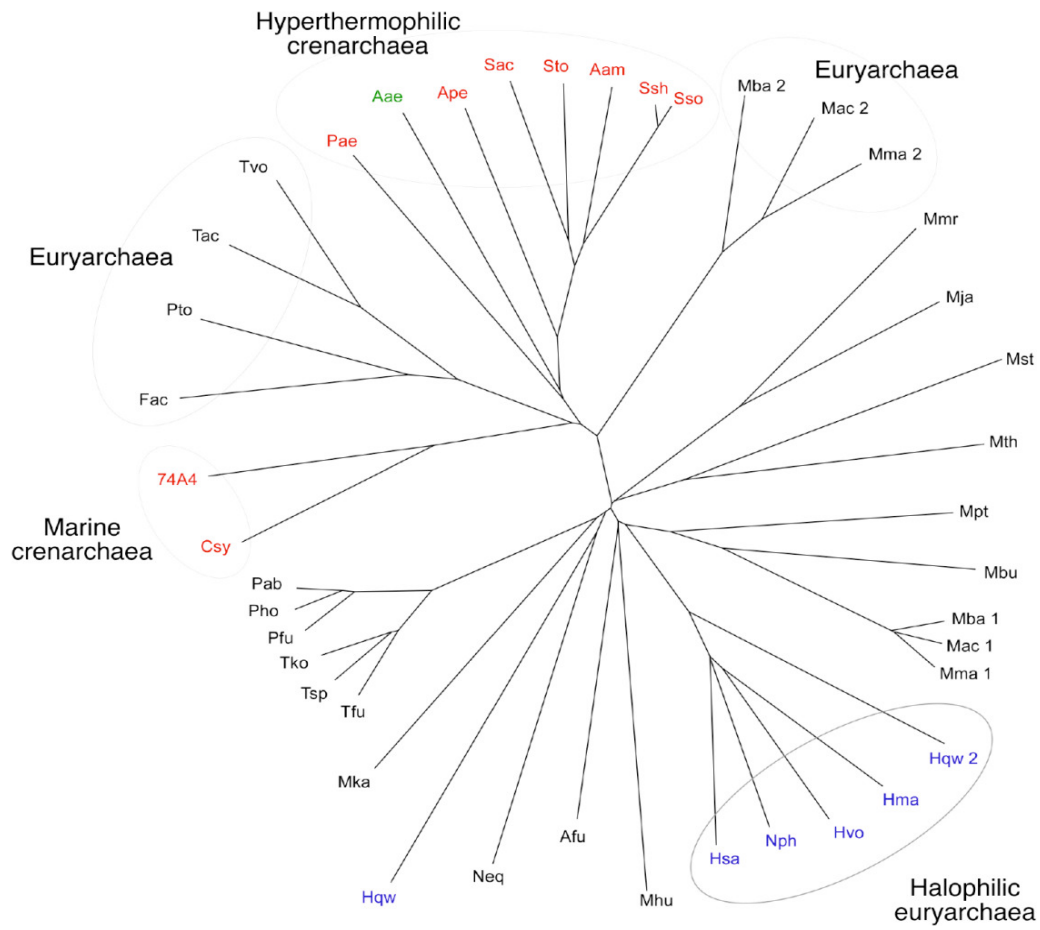


Figure 1.2: Phylogenetic tree of Archaea [121].

Labels: Pfu – *Pyrococcus furiosus*, Pab – *Pyrococcus abyssii*, Pho – *Pyrococcus horikoshii* shinkaj OT3.

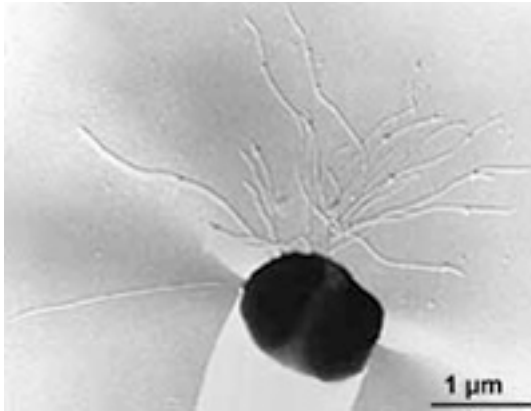


Figure 1.3a: *Pyrococcus furiosus*, coccoid in shape with monopolar peritrichous flagella [122].

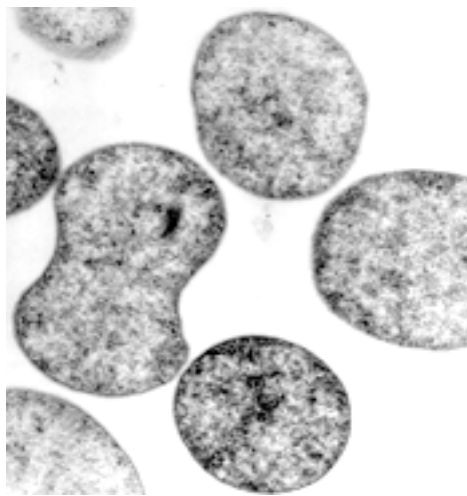


Figure 1.3b: *Pyrococcus furiosus* cells [123].

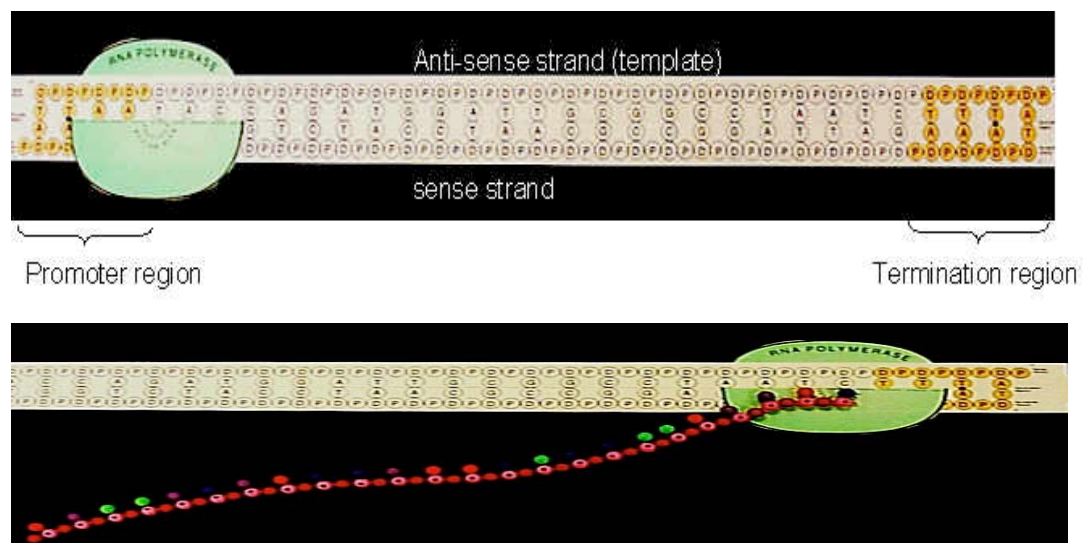
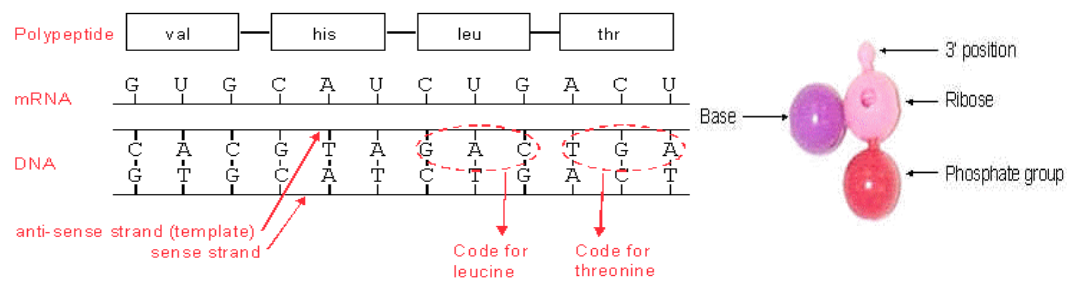


Figure 1.4: Transcription process [124].

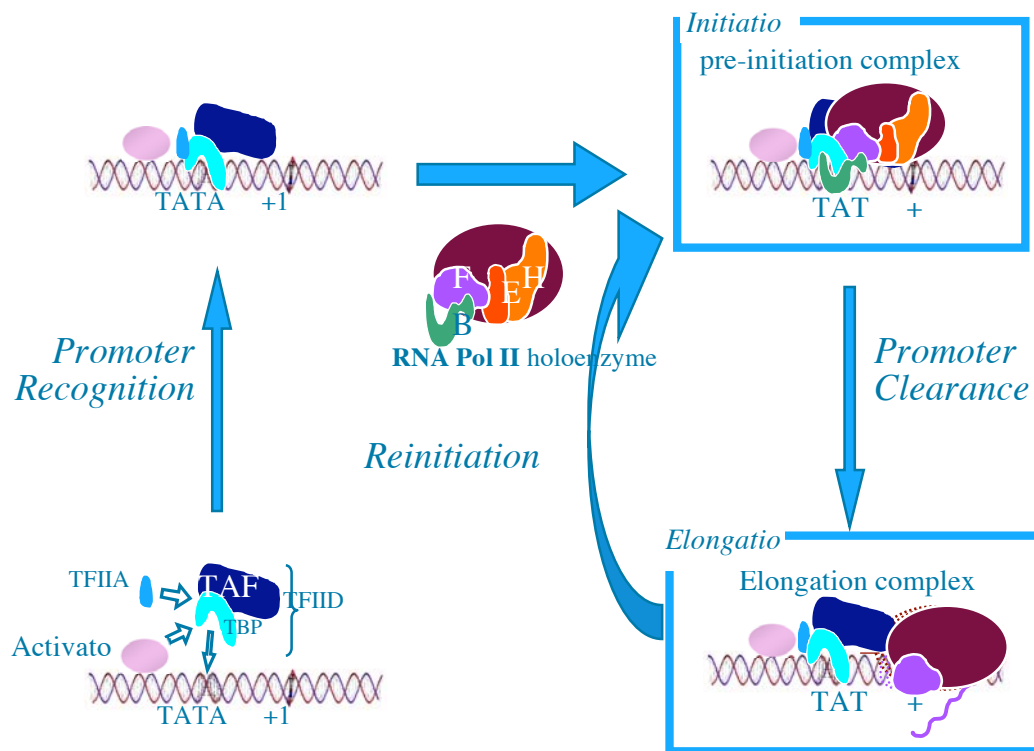


Figure 1.5: Eukaryotic transcription initiation complex.

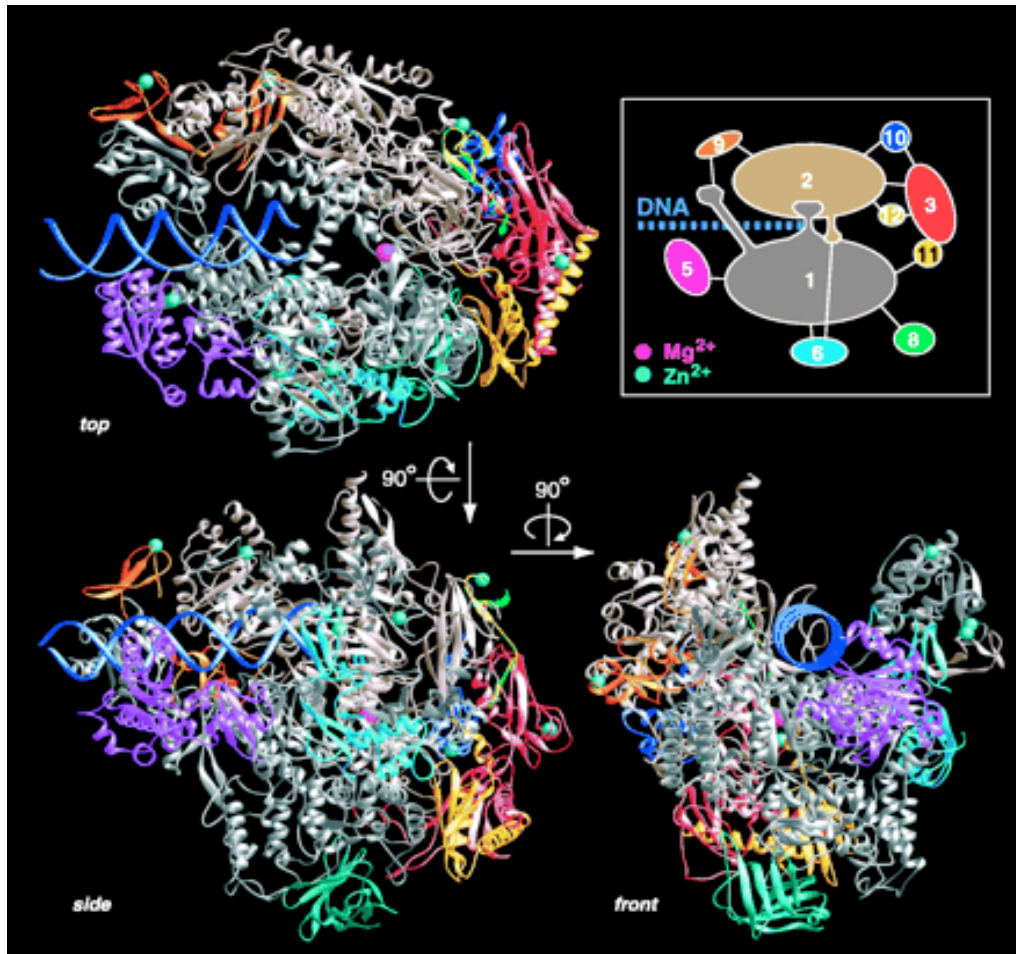


Figure 1.6: Yeast RNA polymerase II consisting of 12 subunits [125].

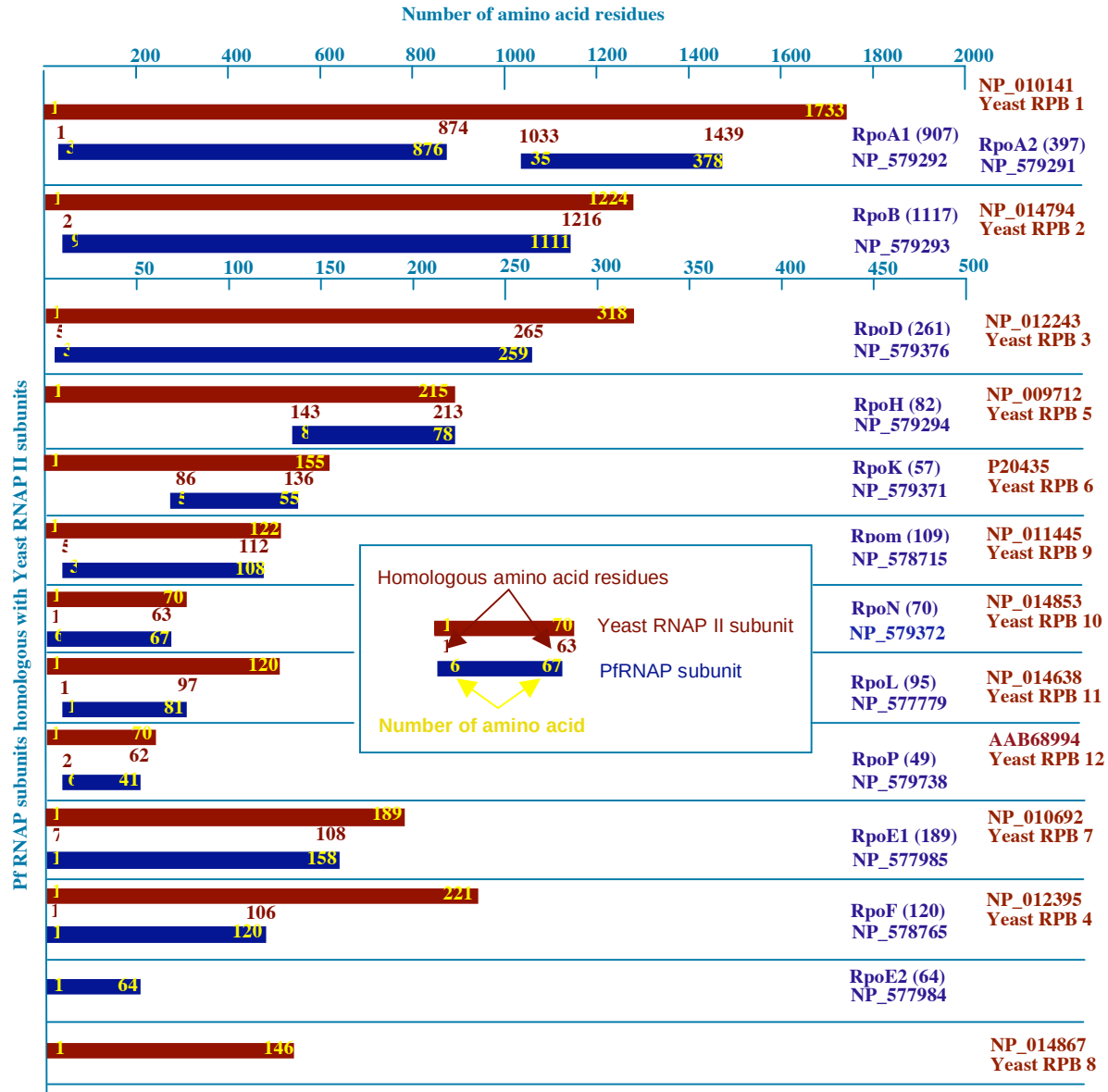


Figure 1.7: Comparison map of sequence homology between yeast RNAP II subunits and *Pf* RNAP subunits obtained using the BLASTP program on NCBI database (www.ncbi.nlm.nih.gov).

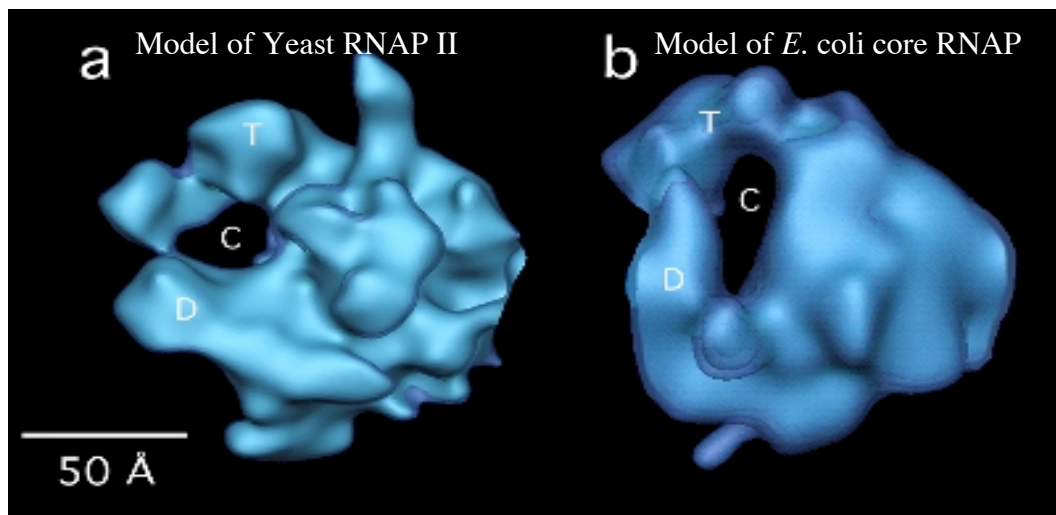


Figure 1.8: Three-dimensional model of the core of yeast RNAP II [32] and *E. coli* RNAP [34].

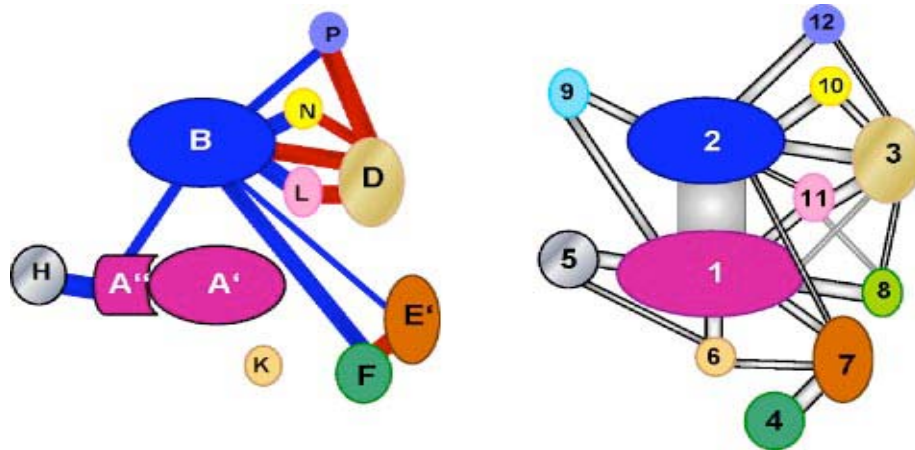
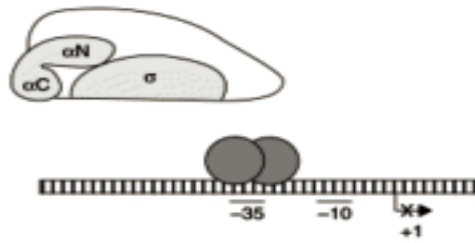


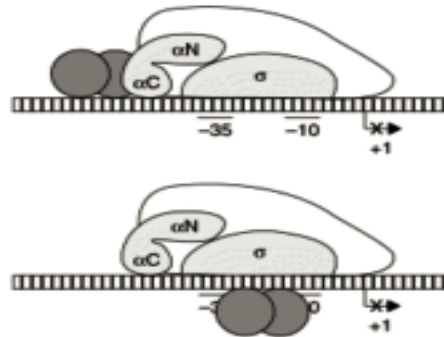
Figure 1.9: Model comparing the subunits of yeast RNAP II and *Pf* RNAP [39].

Labels: The corresponding homologous subunits are coded with the same color. The thick blue lines represent the strength of interaction between the subunits.

Repression by steric hindrance



Inhibition of transition from closed to open complex



Inhibition of promoter clearance

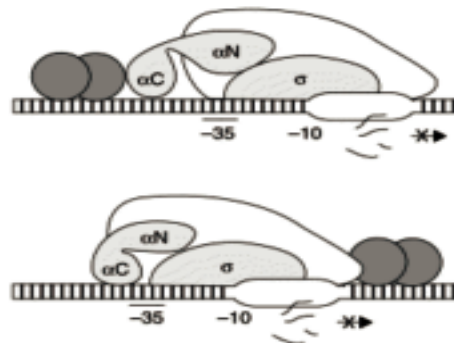


Figure 1.10: Model of mechanism of transcriptional repression [43].



Figure 1.11: Number of transcriptional regulators in archaeal genomes [40].

Labels: Euryarchaeota genomes: Ta, Tv, Mk, Pa, Mj, Mt, Ph, Af, H, Mm, Ma (green).
e.g. Transcriptional regulator, Iron-dependent repressor: Pfam accession no. PF01325; 21
homologous sequences (No. on the left); maximum and minimum no. of genes are 12.5 and
4.6 (No. on the right in green); and mean number of match $\pm$ standard deviation is 8.0 ± 2.9 .

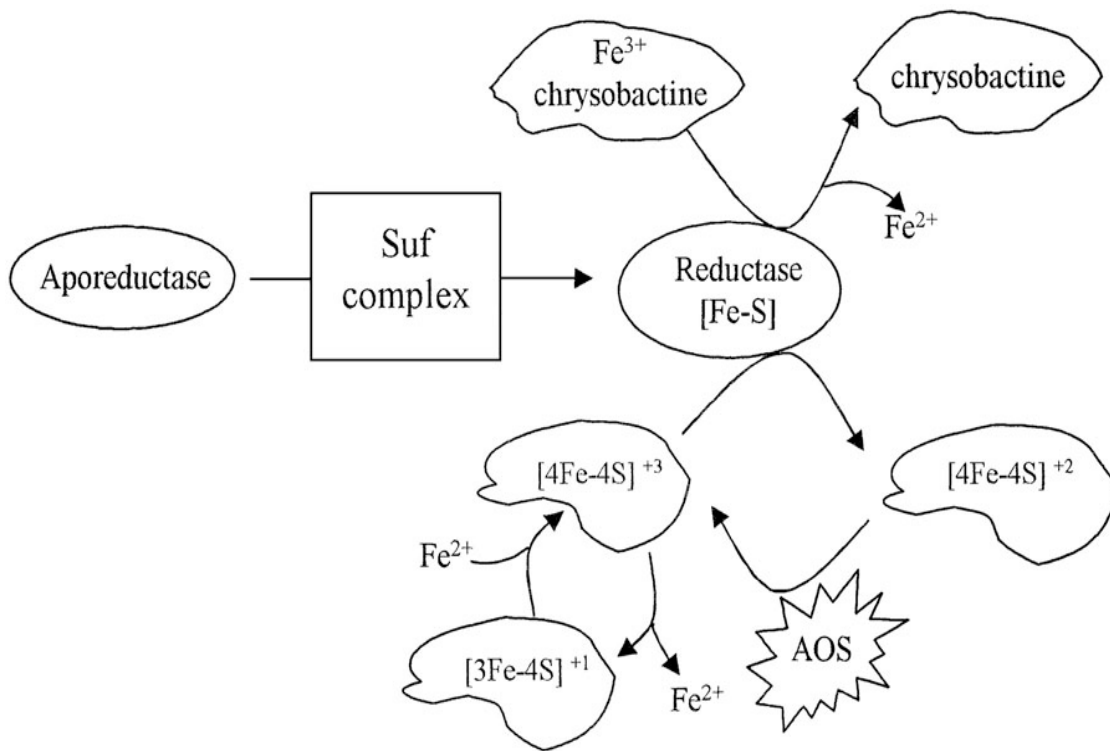


Figure 1.12: Model of iron assimilation and repair of [Fe-S] clusters under oxidative stress and iron limitation in *E. coli*. Under iron limitation and oxidative stress, the reductase converts (Fe³⁺)-loaded siderophore (chrysobactine) to its apo form and repairs the damaged [4Fe-4S] cluster [56].

Labels: AOS-activated oxygen species.

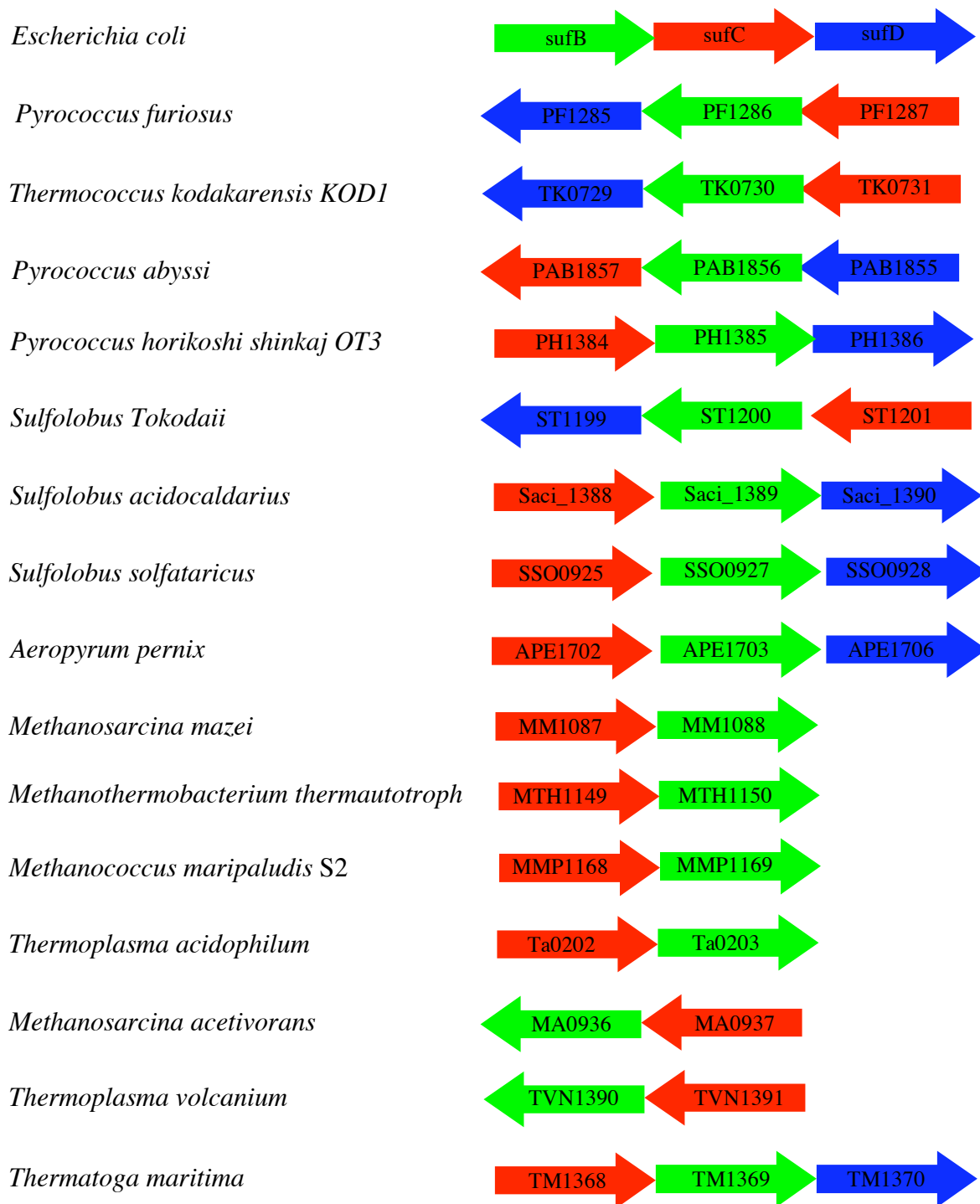


Figure 1.13: *suf* homologs present in *E. coli*, archaea, and *T. maritima* (<http://string.embl.de>).

Legends: sufC (red), sufB (green) and sufD (blue).

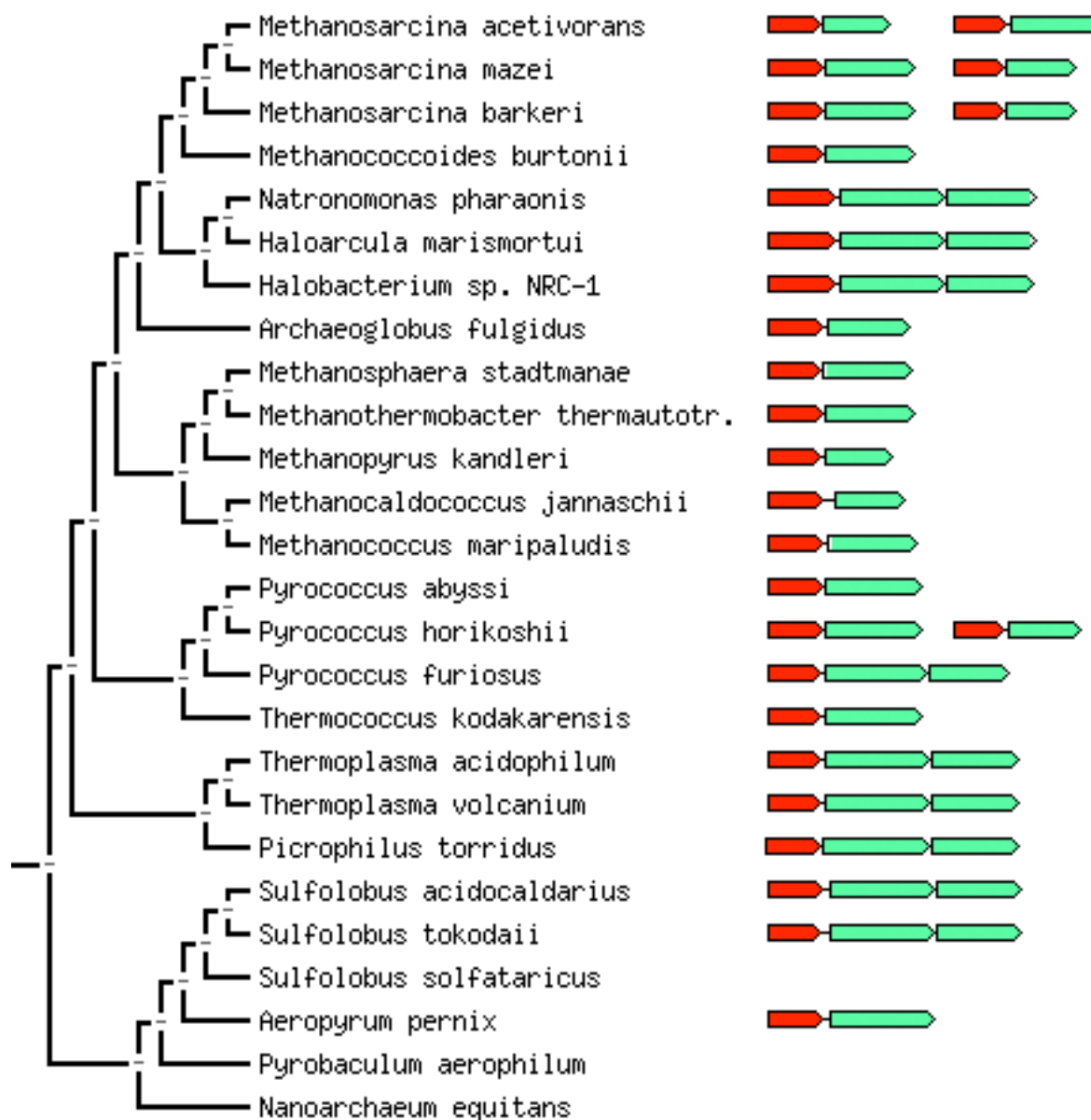


Figure 1.14: Map of *suf* operon present in archaea (<http://string.embl.de>).

Legends: *sufC* homologs (red), *sufB* and *sufD* homologs (green) with either gene duplication (e.g. *sufB*-*sufB* pair) or one of the *sufBCD* genes absent in some of the genomes.

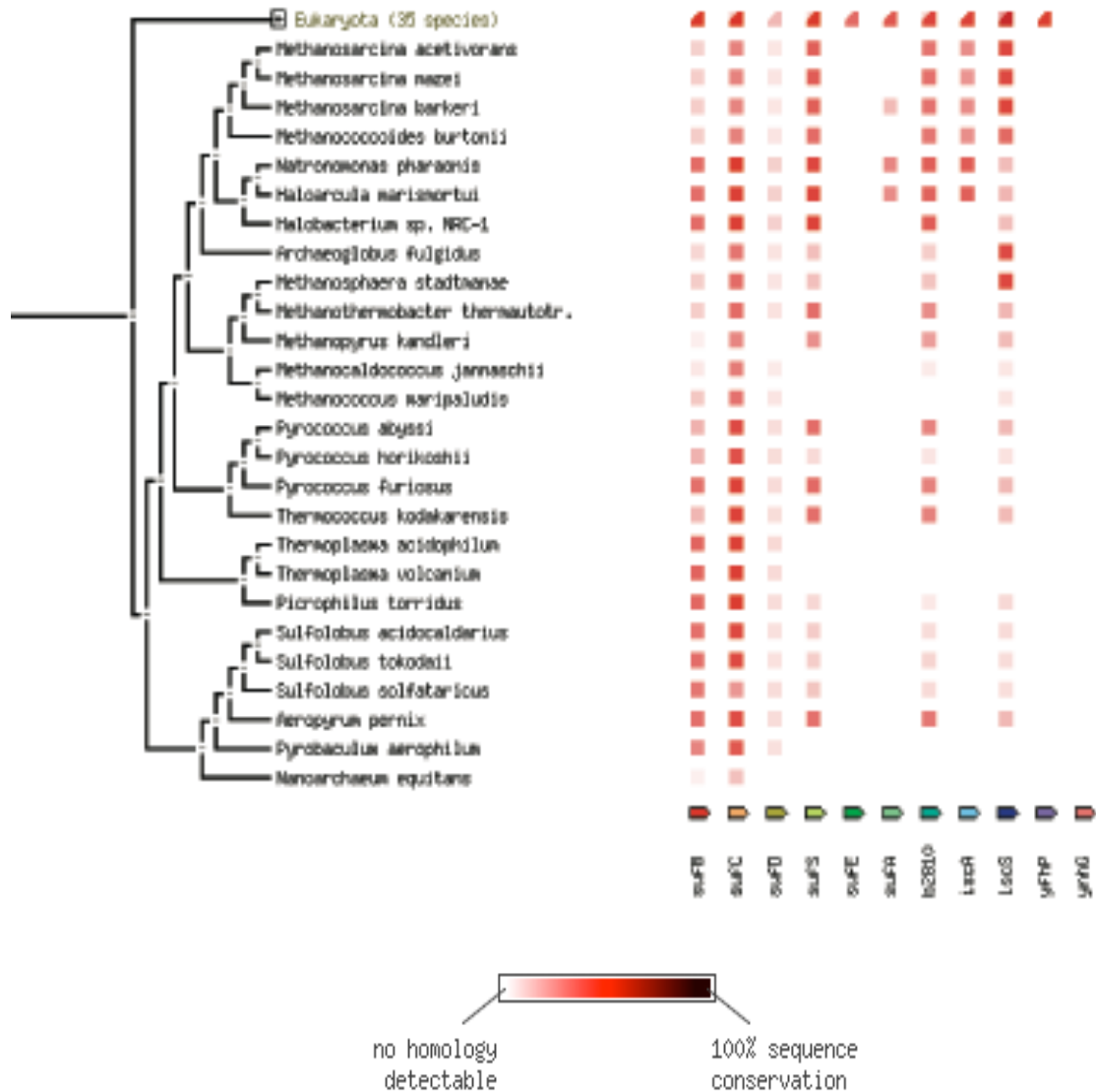


Figure 1.15: The occurrence of *sufS* in archaea (<http://string.embl.de>).

Labels: White to Black color and intensity is an indicator of % sequence homology. Genes represented from left to right – Archaeal homologs of *sufB*, *sufC*, *sufD*, *sufS*, *sufE*, *sufA*, *b2810* (cysteine sulfinatase desulfinate), *iscA*, *iscS*, *yfnhP* (hypothetical protein), *ynhG* (hypothetical protein).

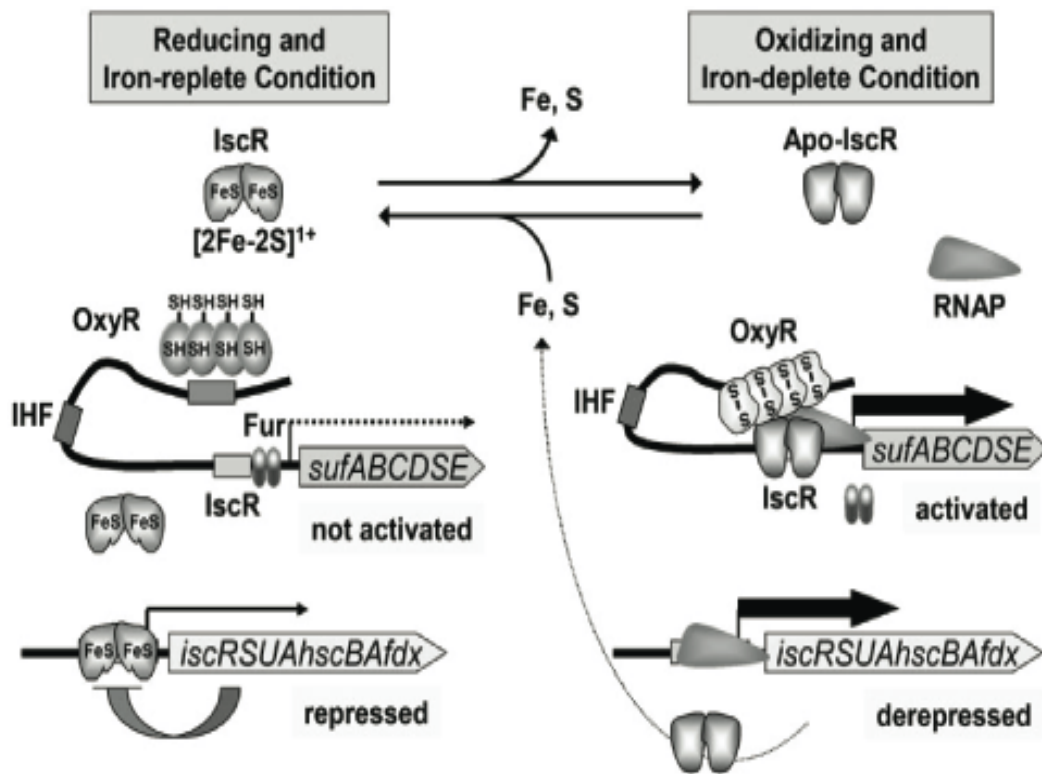
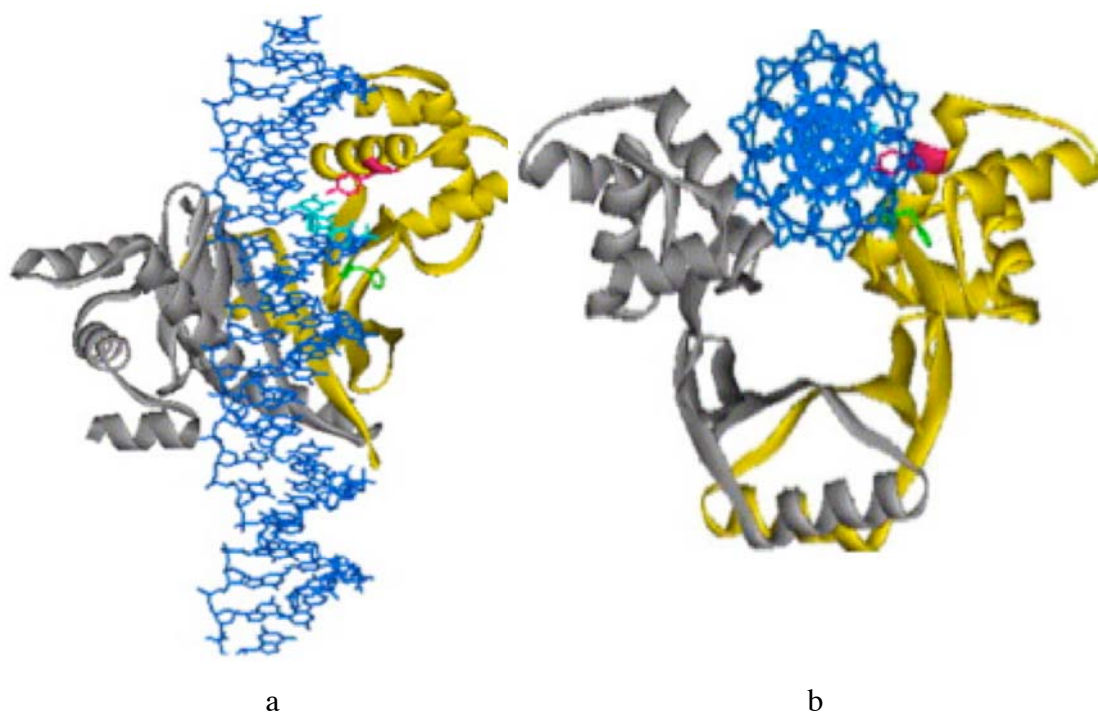


Figure 1.16: Proposed mechanism of Fe-S cluster formation in SUF system in *E. coli* [126].



Fur box	<i>E. coli</i>	tGATAATGATAATCATTATCa
Fur box	<i>P. aeruginosa</i>	tGATAATGATAATCATTATCa
Fur box	<i>B. subtilis</i>	TGATAATNATTATCA
PerR box	<i>B. subtilis</i>	TTATAATNATTATAA
Zur box	<i>B. subtilis</i>	aaaTCGTAATNATTACGAtt
Zur box	<i>E. coli</i>	AAGTGTGATATTATAACATTTCatg

c

Figure 1.17: Three dimensional model of Fur dimer-DNA complex constructed from X-ray structure of Fur (*P. aeruginosa*) [83].

Labels: a-Lengthwise view of the Fur-DNA complex; b-axial view of the complex; c-DNA consensus sequence boxes of Fur and Zur from *E. coli*, PerR, Fur and Zur from *B. subtilis*, and Fur from *P. aeruginosa*. Cyan-Thymines (DNA), Red-Tyrosine (Fur), Green-Lysine (Fur).

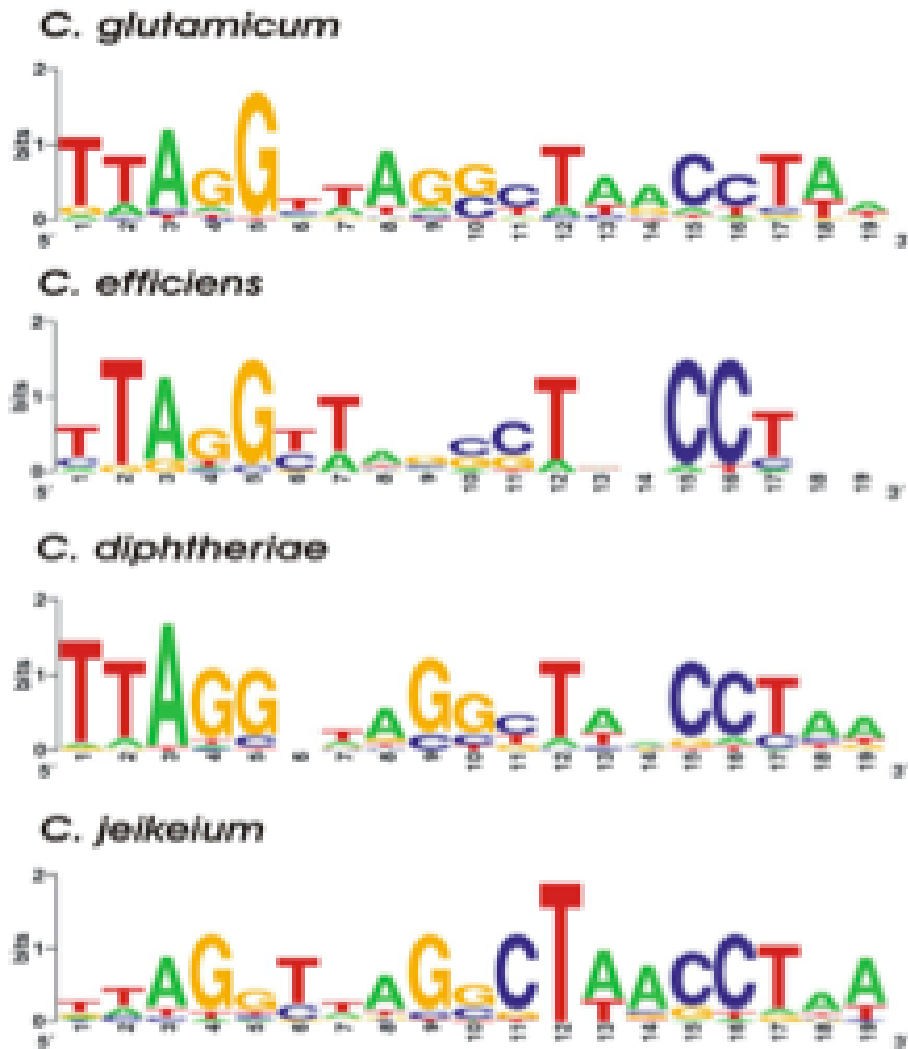
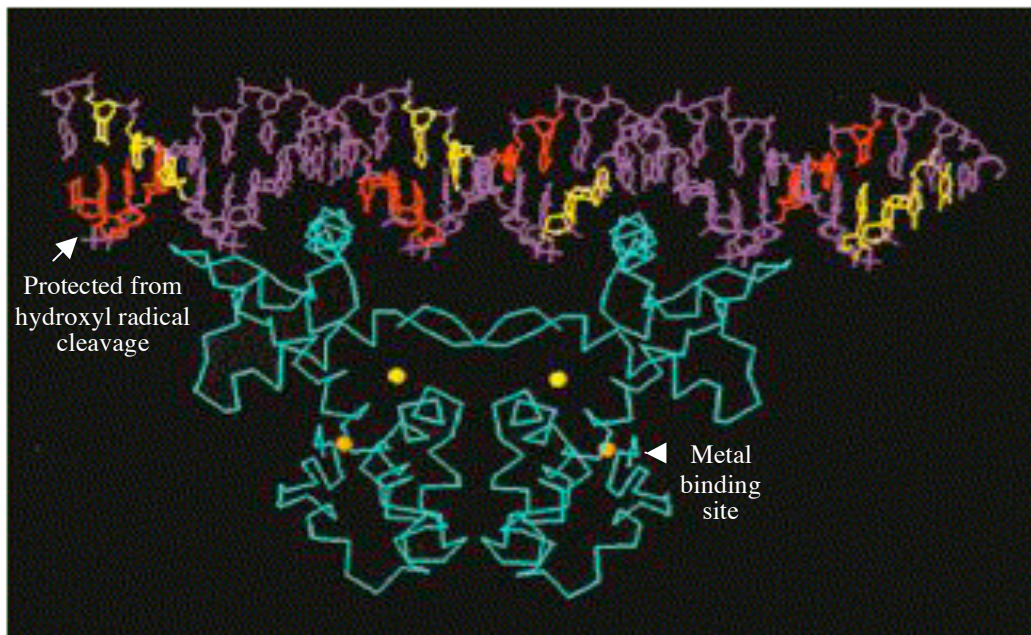
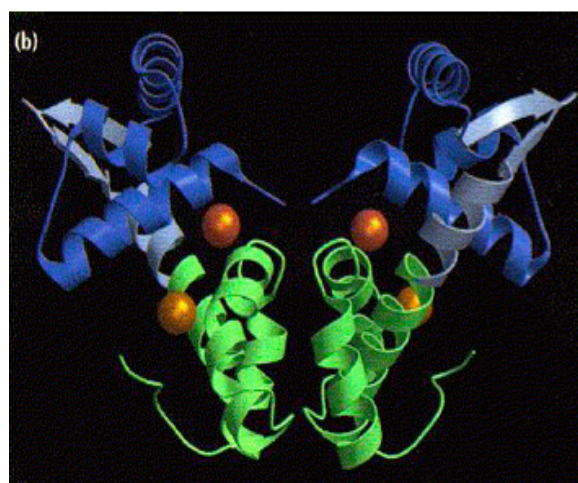


Figure 1.18: Deduced consensus sequences of DtxR binding sites [86].

Labels: Genomes represented are *C. glutamicum* (Cg), *C. efficiens* (Ce), *C. diphtheriae* (Cd), *C. jeikeium* (Cj).

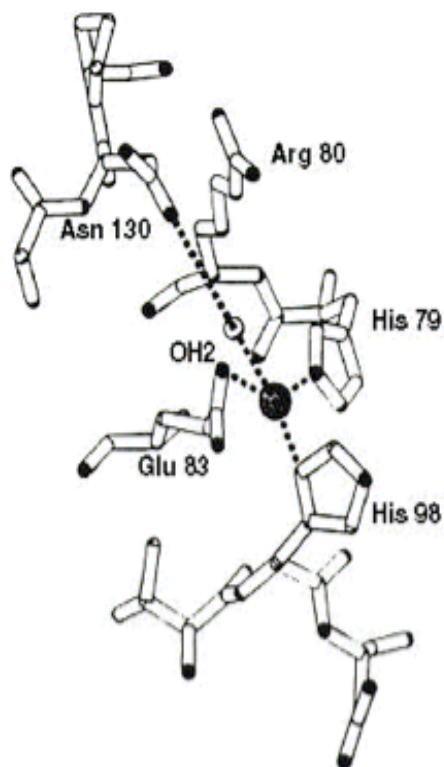


Model of DNA-binding domain of DtxR based on crystal structure and footprinting studies.

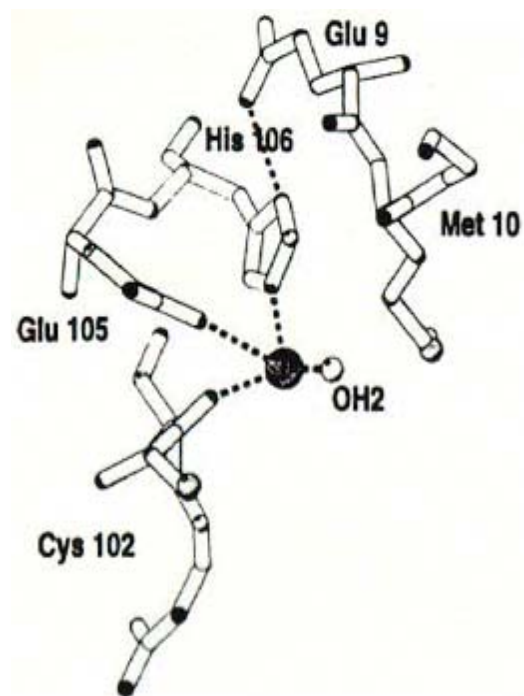


Crystal structure of DtxR (homodimer) in *Corynebacterium diphtheriae*

Figure 1.19: DNA-binding domain and crystal structure of DtxR in *Corynebacterium diphtheriae* [89].



Metal binding site 1
Might affect the dimerization of DtxR



Metal binding site 2
Near the major grooves of HTH domain ($\alpha 3$)

Figure 1.20: Metal-binding sites of DtxR [89].

CHAPTER 2

RECONSTITUTION OF *PYROCOCCUS FURIOSUS* DNA-DEPENDENT RNA POLYMERASE

2.1 Objective: DNA-dependent RNA polymerase is the main component of the transcriptional machinery, which synthesizes the mRNA transcript thereby facilitating selective expression of the genetic information. Regulation of expression of genes in response to cellular environment often occurs at the level of transcription. The significance of obtaining a reconstituted *Pf* RNAP is to validate the newly identified transcriptional regulatory factors responding to environmental stress by *in vitro* transcription reaction. Initially, the possibility of extracting the native RNA polymerase from *Pf* cells in a single-step chromatographic separation was conceived. Technical difficulties to detect the complex using polyclonal antibodies and literature evidence of reconstitution of recombinant RNAP subunits in other organisms like *Methanocaldococcus jannaschii* substantiated the proposal of adopting a similar approach to obtain a reconstituted RNAP in *Pf*. The work described in this chapter was directed at cloning, expression, and purification of all major *Pf* RNAP subunits and reconstitution of an active RNAP for *in vitro* transcription assays.

2.2 Materials

2.2.1 Cells, bio-reagents and chemicals

2.2.1.1 BL21-AI one shot chemically competent *E. coli* (Cat No. C6070-03, Invitrogen)

- 2.2.1.2** BL21* (DE3) one shot chemically competent *E. coli* (Cat. No. C6010-03, Invitrogen)
- 2.2.1.3** pET151/d-TOPO vector (Cat. No. K151-01, Invitrogen)
- 2.2.1.4** Top10 chemically competent *E. coli* (Cat. No. C4040-10, Invitrogen)
- 2.2.1.5** α -cyano-4-hydroxycinnamic acid (Matrix for MALDI), Cat. No. C2020-10G, Sigma-Aldrich, saturated solution in 50% acetonitrile, 0.1% TFA
- 2.2.1.6** 1x TAE buffer: 20mM Tris (pH=7.8), 10mM sodium acetate, 0.5mM EDTA
- 2.2.1.7** 1x TBE buffer: 50mM Tris (pH=8.3), 50mM sodium borate, 0.5mM EDTA
- 2.2.1.8** 1x Tris/Glycine/SDS buffer: 25mM Tris, 192mM glycine, 0.1% (w/v) SDS (sodium dodecyl sulfate)
- 2.2.1.9** 1x Tris/Glycine buffer: 25mM Tris, 192mM glycine (pH=8.3)
- 2.2.1.10** AcTEV (Enhanced form of Tobacco Etch Virus) protease, Cat. No. 12575-015, Invitrogen
- 2.2.1.11** Acetonitrile (100% v/v)
- 2.2.1.12** Agarose (1% w/v)
- 2.2.1.13** Ammonium bicarbonate (NH_4HCO_3), J. T. Baker, (25 mM)
- 2.2.1.14** Antibiotics: Ampicillin, Cat. No. E477, Amresco, (100 $\mu\text{g}/\text{mL}$); Carbenicillin disodium salt, Cat. No. 10177-012, Invitrogen, (100 $\mu\text{g}/\text{mL}$); Kanamycin sulfate Cat. No. K400-5G, Sigma-Aldrich, (50 $\mu\text{g}/\text{mL}$)
- 2.2.1.15** Binding buffer: 20 mM HEPES ((N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]) potassium salt Cat. No. H-0527, Sigma-Aldrich), 150 mM NaCl (sodium chloride, J. T. Baker)

- 2.2.1.16** Elution buffer: 20 mM HEPES ((N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]) potassium salt Cat. No. H-0527, Sigma-Aldrich), 150 mM NaCl (sodium chloride, J. T. Baker), 500 mM Imidazole
- 2.2.1.17** Ethanol (70%, 100% v/v in water)
- 2.2.1.18** Gel Stain: (i) Protein - Coomassie blue stain (0.15% w/v coomassie brilliant blue G250, 17% w/v ammonium sulfate, 34% v/v methanol, 3% v/v phosphoric acid; (ii) DNA – SYBR green I (1:10,000 dilution in 1x TBE buffer, Cat. No. BP1755-10, Invitrogen), ethidium bromide: 0.5 $\mu\text{g/mL}$ in water; (iii) RNA – SYBR green II (1:10000 dilution in 1x TBE buffer, Cat. No. BP1755-10, Invitrogen)
- 2.2.1.19** Isopropyl-beta-D-thiogalactopyranoside (IPTG) Cat. No. BP1755-10, Fisher, (0.4 mM)
- 2.2.1.20** LB agar plates: 2.5% LB broth, 1.5% Bacto agar, autoclave for 30 min, add antibiotic, and pour into 10 cm plates, store at 4°C until required
- 2.2.1.21** Media (Culture growth): *E. coli* - LB (Luria-Bertani) Broth, Miller, Cat. No. B1426-2, Fisher Chemicals, (2.5% (w/v)); ZYP 5052 (Autoinduction): 1 mM MgSO_4 , 1x metals mix (50 μM Fe, 20 μM Ca, 10 μM Mn, 10 μM Zn, 2 μM Co, 2 μM Ni, 2 μM Mo, 2 μM Se), 1x NPS (5 mM PO_4^{2-} , 1.25 mM SO_4^{2-} , 2.5 mM NH_4^+ , 5 mM Na^+ , 2.5 mM K^+), 1x 5052 (0.01% glycerol, 0.001% glucose, 0.004% α -lactose) in 1000 mL; *Pyrococcus furiosus* - (Standard defined medium) 75 g maltose, 7.5 g yeast, 1x salts (420 g NaCl, 52.50 g MgSO_4 , 40.50 g MgCl_2 , 4.95 g KCl, 3.75 g NH_4Cl , 2.10 g CaCl_2), 1x trace minerals (0.03 g FeCl_3 , 0.0008 g H_3BO_3 , 0.0008 g ZnCl_2 , 0.0005 g CuCl_2 , 0.0008 g MnCl_2 , 0.0008 g $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24}$, 0.0008 g $\text{AlK}(\text{SO}_4)_2$, 0.0008 g CoCl_2 , 0.0008 g NiCl_2), 0.0495 g

Na₂WO₄, 7.5 g Cysteine HCl, 7.5 g Na₂S, 15 mL of 1 M potassium phosphate buffer (pH 6.8).

- 2.2.1.22** Molecular weight markers: Protein - Precision Blue SDS standard; HMW calibration kit for native electrophoresis, Cat. No. 17-0445-01, Amersham Biosciences-GE; DNA – 1 kb, 100 bp, 25 bp step ladder (Promega); RNA – 0.1 to 2 kb RNA ladder, Cat. No. 15623-100, Invitrogen
- 2.2.1.23** Northern Blot: Washing buffer – 0.1 M Maleic acid, 0.1 M NaCl (pH 7.5), 0.3% (v/v) Tween 20; 1x Blocking buffer – 1:10 of 10x Blocking buffer (Roche) in 0.1 M Maleic acid, 0.1 M NaCl (pH 7.5); Anti-DIG antibody solution: 1:10,000 dilution in 1x Blocking buffer; Detection buffer – 0.1 M Tris HCl, 0.1 M NaCl (pH 9.5); Chemiluminescent reagent – CDP-star, ready-to-use, Cat. No. 2041677
- 2.2.1.24** NTP Mix: dNTP - 2.5 mM dATP, 2.5 mM dGTP, 2.5 mM dCTP, 2.5 mM dTTP; rNTP - 1.0 mM ATP, 1.0 mM GTP, 1.0 mM CTP, 1.0mM UTP, Invitrogen
- 2.2.1.25** Optizyme ribonuclease inhibitor, Cat. No. BP3222-1, Fisher
- 2.2.1.26** Opti-4CN detection kit, Goat-anti-rabbit, Cat. No. 170-8236, Biorad
- 2.2.1.27** Penta-His HRP conjugate, Cat No. 34460, Qiagen
- 2.2.1.28** Protease inhibitor cocktail, Cat. No. P8849, Sigma-Aldrich
- 2.2.1.29** *Pyrococcus furiosus* genomic DNA (S. Brehm, M. W. W. Adam's Group, UGA)
- 2.2.1.30** Qiagen plasmid maxi kit, Cat. No. 12163, Qiagen
- 2.2.1.31** QIAquick PCR purification kit, Cat. No. 28106, Qiagen
- 2.2.1.32** Sample loading buffer: Protein – 4x Laemmli buffer (160 mM Tris-HCl (pH 6.8), 50% (v/v) glycerol, 8% (w/v) SDS, 0.1% (w/v) bromophenol blue; DNA – 6x loading dye (Promega); RNA – 2x TBE-urea buffer (Biorad)

- 2.2.1.33** S. N. A. P. gel purification kit, Cat. No. 45-0078, Invitrogen
- 2.2.1.34** S. O. C medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose)
- 2.2.1.35** Sodium acetate, J. T. Baker, 3M (NaOAc)
- 2.2.1.36** Trifluoroacetic acid (TFA), 0.1%, 0.5% v/v
- 2.2.1.37** Tris (Tris hydroxy methyl aminomethane), Fisher Biotech
- 2.2.1.38** Trypsin, 20 µg in 200 µL of trypsin dilution buffer, Promega
- 2.2.1.39** Turbo *Pf* polymerase (2.5 U/µL)
- 2.2.2** *Equipment and glassware*
- 2.2.2.1** AKTA prime plus FPLC (Fast Protein Liquid Chromatography) system, Amersham Biosciences-GE
- 2.2.2.2** Centrifuge with JA-5.3 and JA-25.5 rotors, Avanti J-E, Beckman Coulter
- 2.2.2.3** Centrifugal filters: YM-3, YM-10, YM-30, YM-100 membrane filter (MWCO = 3,000, 10,000, 30,000, 100,000 respectively), Millipore
- 2.2.2.4** DC protein assay kit, Biorad
- 2.2.2.5** Electrophoresis system (DNA, Protein, Mini Trans-blot (for western transfer)), Biorad
- 2.2.2.6** Membrane filter: Syringe filter (0.2 µ), Nitrocellulose (0.45 µ), Millipore
- 2.2.2.7** Fluorochem 8000 Digital Imager (Alpha Innotech Corp.)
- 2.2.2.8** Gels: 4-20% Tris gel, 10-20% Tris gel (Criterion, Biorad); 10% TBE-urea (Ready gels, Biorad)
- 2.2.2.9** Hoefer DyNA Quant 200 fluorometer (Amersham Pharmacia Biotech)

- 2.2.2.10** Mass spectrometry: LC-ESIMS (Liquid chromatography-electron spray ionization mass spectrometry) (Bruker), MALDI ToF (Matrix Assisted Laser Desorption Ionization Time of Flight)
- 2.2.2.11** Membranes: RNA - Nylon membranes, positively charged Cat. No. 1209 299, Roche; Western blot transfer - Polyvinylidene difluoride (PVDF) membrane
- 2.2.2.12** Nickel affinity columns: HP chelating column (1mL, 5mL), HistrapFF (1 mL), Amersham Biosciences-GE
- 2.2.2.13** NuTip C-18 tips (Glygen Inc.)
- 2.2.2.14** PCR Thermocycler
- 2.2.2.15** Sterile test tubes (15 mL), Erlenmeyer flask (250 mL), Fernbach flask (2.8 L)
- 2.2.2.16** Shaker-Incubator
- 2.2.2.17** Sonic dismembranator (Fisher)
- 2.2.2.18** Waterbath (temperature set at 42°C)
- 2.2.2.19** Vacuum evaporator
- 2.2.2.20** X-ray film (BioMax MR imaging film) 13 x 18 cm., Cat. No. 8941114, Kodak

2.3 Methods

2.3.1 *Extraction of native Pf RNAP complex using His₆-RpoD*

2.3.1.1 *Plasmid purification of RpoD (PF1647):* The gene encoding RpoD (PF1647) was cloned into pET21b vector with a His₆-tag and the cells containing the plasmid were obtained (M. Lewis and M. K. Eidsness). His₆-RpoD plasmid and a positive control using pUC18 were transformed into XL1 blue cells separately following standard protocol [127]. 100 μ L and 200 μ L of the culture was transferred onto two LB-agar plates and the plates were incubated at 37°C for 16 h. Colonies (8-

10) were selected and step cultures of 3 mL and 100 mL were grown for 16 h at 37°C with shaking at 225 rpm in LB media containing 100 µg/mL ampicillin and then transferred to 1 L LB media in a 2.8 L fernbach flask with 100 µg/mL of ampicillin, and incubated at 37°C for 16 h. The cell extract was collected and centrifuged at 3,000 xg for 20 min using a JA-5.3 rotor. The supernatant was collected and the plasmid was purified using Qiagen plasmid purification kit following the manufacturer's protocol. The plasmid was collected in 5 mL eluting buffer. Plasmid was purified by ethanol precipitation by adding 1/10th volume of 3 M sodium acetate (500 µL), mixed well, followed by addition of 2.5x of cold 100% ethanol and mixed well. The resulting solution is centrifuged at 20,000 xg for 15 min using a JA-25.5 rotor. The supernatant is carefully decanted without disturbing the precipitated plasmid DNA at the bottom of the tube followed by the addition of 1 mL of cold 70% ethanol and centrifuged at 20,000 xg for 20 min using JA-25.5 rotor. The supernatant was again carefully decanted without disturbing the precipitate (invisible, slightly gel-like) at the bottom of the tube. The tube was allowed to air dry for 20 to 30 min until all the residual ethanol evaporated. The purified plasmid was re-suspended in 10 mM Tris buffer (pH 8.0). The sequence of the gene was verified (Sequetech Corp., CA).

2.3.1.2 *Protein expression, purification and characterization of RpoD (PF1647)*

2.3.1.2.1 *Protein expression:* The purified plasmid was transformed into the expression strain of *E. coli* BL21* (DE3) gold competent cells [128]. Step cultures of 3 mL and 100 mL of 2.5% LB media containing 100 µg/mL ampicillin were grown for 16 h at 37°C with shaking at 225 rpm. The culture was transferred to 1 L LB

media containing 100 $\mu\text{g/mL}$ of ampicillin and incubated at 37°C for 3.5 h with shaking at 225 rpm until $\text{OD}_{600} = 0.6 - 0.8$. The culture was induced for over-expression of the recombinant protein by the addition of 0.4 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) and incubated for 8 h at 37°C (Figure 2.1). The cells were harvested by centrifugation at 3,000 xg for 20 min using a JA-5.3 rotor, the cell pellet was collected and stored at -20°C until further purification.

2.3.1.2.2 *Preparation of E. coli cell extract:* Frozen cell pellet (5 g, wet weight) was thawed with the addition of 140 μL of protease inhibitor cocktail and re-suspended in 25 mL of 20 mM HEPES, 150 mM NaCl (pH 7.55). The cells were then disrupted by means of a sonicator in the continuous mode, the amplitude gradually increased from 0 to 9, and continuously sonicated with the power set to 20% for 1 min. The amplitude knob is turned back to zero for 30 sec and the cycle was repeated for 10 min. The cell debris was removed using a centrifuge at 35,000 xg for 30 min. The clear supernatant was collected and filtered through a 0.2-micron syringe filter and loaded into a nickel affinity column.

2.3.1.2.3 *Protein purification:* The HP chelating column was loaded with 10 mL of 0.1 M nickel sulfate and washed with 5 column volumes of degassed nanopure water until all the residual nickel is removed from the column. The column was connected to the AKTA prime plus FPLC (Fast Protein Liquid Chromatography) system. The UV detector was set to 0.1 AUFS (Absorbance Unit Full Scale) and the wavelength was 280 nm. The column was equilibrated with the binding buffer: 20 mM HEPES, 150 mM NaCl. The flow rate was set to 2.0 mL/min. The clarified supernatant (28 mL) containing the recombinant protein was loaded

onto a 5 mL chelating nickel column equilibrated with binding buffer and the His₆-RpoD protein was eluted with a linear gradient of buffer containing 100-500 mM imidazole (Elution buffer: 20 mM HEPES, 150 mM NaCl, 500 mM imidazole). Fractions (15 mL) were collected and concentrated using YM-10 membrane centrifugal filter (MWCO: 10,000) to 100 μ L. The amount of protein was estimated using Biorad DC assay.

2.3.1.2.4 *SDS-PAGE analysis of purified recombinant protein:* The protein was separated using SDS-PAGE on a 4-20% Tris gel in 1x Tris/glycine/SDS buffer. The power supply for electrophoresis was set to 200 V for 60 min. Samples were heated at 95°C for 5 min in 4x Laemmli buffer prior to loading. The gels were stained using coomassie blue stain and the image is recorded using a Fluorchem 8000 imaging system (Figure 2.2).

2.3.1.2.5 *Protein characterization using MALDI ToF:* The gel band was excised and transferred to a microcentrifuge tube, destained with 6% acetic acid. The gel slice was rinsed quickly thrice with nanopure water to remove the residual coomassie blue stain. The gel slices were then dehydrated using a vacuum evaporator. Trypsin digestion was carried out with the addition of 15 μ L of 10 ng/ μ L trypsin and incubated for 16 h at 37°C. Trypsin cleaves the protein after every lysine (K) and arginine (R). Tryptic peptides from the gel slice were extracted in four 10 min steps (1x with 30 μ L of 25 mM NH₄HCO₃, and 2x with 30 μ L of 50% acetonitrile, 25 mM NH₄HCO₃ and 1x with 75% acetonitrile, 0.1% TFA). The supernatant from each extraction was combined and collected. The extracted peptides in solution were concentrated using a vacuum evaporator to 4 μ L [129].

To the resulting solution, 1 μ L of 0.5% TFA was added. The tryptic peptides were desalted and concentrated using NuTip C-18 tip and spotted directly on a MALDI plate (3 μ L, containing the matrix α -cyano-4-hydroxycinnamic acid). Data analysis is performed using Protein Prospector v. 3.2.1 using MS-Fit program [130]. The sequence of the recombinant protein is verified using BLASTP on NCBI database.

2.3.1.3 *Attempt to capture native Pf RNA polymerase using purified His₆-RpoD (PF1647)*

2.3.1.3.1 *Growth of Pf cells: Pyrococcus furiosus* cells were grown by A. L. Menon according to previously reported methods [131]. Briefly, the cells were grown at 95°C with maltose as the carbon source, a standard defined medium in the absence of elemental sulfur in a 600 L fermenter. The pH of the culture was not adjusted during growth. The cells were harvested when the cell density reached 2×10^8 cells/mL to yield about 2 kg of cell paste. The cells were stored in small aliquots at –80°C until required.

2.3.1.3.2 *Preparation of Pf cell extract:* Frozen cell pellet (5 g, wet weight) was thawed with the addition of 15 mL of 20 mM HEPES, 150 mM NaCl for 15 min at room temperature (RT). The cell suspension was sonicated for 10 min, and centrifuged at 50,000 xg for 60 min at 4°C. The supernatant was collected. Amount of protein in the cell extract estimated by Biorad DC assay was 60 mg/mL.

2.3.1.3.3 *Extraction of RNAP complex from Pf cell extract:* An attempt to purify native RNA polymerase from *Pf* cell extract in a single-step was carried out [132]. *Pf* cell extract (5 mL) was treated with 0.4% polymine P and 70% saturated ammonium sulfate, centrifuged and the supernatant was collected [133]. Purified

recombinant His₆-RpoD (0.28 mg) was added to the supernatant and loaded on a 5 mL chelating zinc column equilibrated with binding buffer (20 mM HEPES and 150 mM NaCl, pH 7.55) washed with 5 column volumes of binding buffer. The hypothesis was that the native RNA polymerase (with native RpoD) would exchange and/or bind with the His₆-RpoD. The protein complex along with His₆-RpoD was eluted with a linear gradient of elution buffer containing imidazole (Figure 2.3). Three fractions were collected and concentrated to 500 μ L using YM-100 centrifugal membrane filter (MWCO: 100,000 Da). The concentrated fractions containing the protein complex was separated on a 4-20% Tris gel under native conditions (without SDS) (Figure 2.4).

2.3.1.3.4 *Characterization of proteins captured using His₆-RpoD:* Reverse phase C8 (octyl) column connected to a LC-ESIMS system was used to identify the proteins eluted with His₆-RpoD in fraction 3. The chromatograms obtained (Figure 2.5a - 2.5j) and results are tabulated (Table 2.1). The chromatograms obtained had relatively low resolution, so the m/z values obtained did not match the exact molecular weight of the individual subunits. Hence, to determine if His₆-RpoD formed a complex with the native *Pf* RNAP, detection of His₆-RpoD in the complex using polyclonal antibodies of His₆-RpoD was employed. The antibody could be used to immunoprecipitate the complex, had the test indicated positive results.

2.3.1.3.5 *Detection of His₆-RpoD in the protein complex using polyclonal antibodies of His₆-RpoD:* Polyclonal antibodies were generated using His₆-RpoD protein at IBL, UGA. Briefly, His₆-RpoD was used as the antigen along with Freund's complete adjuvant and 100 μ g of the protein was injected into rabbits. The titre

value is a measure of the quality of the antibodies in the serum; analyzed by indirect ELISA. The titre value of the polyclonal antibody generated was 1:10,000 (IBL, UGA). The concentration of the protein complex (separated on the Zn chelating column, section 2.2.1.3.3) was too low to be detected by polyclonal antibodies on a western blot. An attempt to test the binding of His₆-RpoD to native *Pf* RNAP complex in the whole *Pf* cell extract was carried out. *Pf* cell extract was treated with 0.4% polymine P, the supernatant combined with 10-30 μ g of His₆-RpoD, incubated at 70°C for 5 min and loaded on a 4-20% Tris gel in 1x Tris/glycine and the gel was run under native conditions (without SDS) at a constant voltage of 400 V for 40 min (Figure 2.6a). Western blot transfer of the proteins from the gel onto a polyvinylidene difluoride (PVDF) membrane was carried out. Briefly, the gel was placed on a PVDF membrane and a sandwich was prepared with thick filter paper on either side. The gel sandwich was loaded on a cassette in a mini trans blot and the unit was placed in a refrigerator (4°C). Western transfer of proteins from the gel to the membrane was carried out in 1x Tris/glycine buffer at a constant voltage of 100 V for 1 h. The blot was treated with blocking buffer, followed by incubation with polyclonal antibody (primary antibody), generated using recombinant His₆RpoD for 30 min. Addition of the horseradish peroxidase (secondary antibody), generated color that was detected chromogenically (Figure 2.6b). The polyclonal antibody binds to pure recombinant His₆-RpoD, however in *Pf* cell extract, only one band corresponding to His₆-RpoD was detected. Since this was not a conclusive result, detection using anti-his antibody was carried out.

2.3.1.3.6 *Detection of His₆-RpoD in the protein complex using Pentahis antibody:* An alternative approach using anti-his antibody (Pentahis antibody HRP conjugate). Opti4CN was used for chromogenic detection of the western blot.

Both the western blots using polyclonal antibody and pentahis antibody did not detect the complex. The possible reasons could be:

1. The native gel and western blot transfer under native conditions with the absence of detergents might have reduced the possibility of exposure of the antibody to the His₆-RpoD.
2. Perhaps the native *Pf* RNAP complex was not bound to His₆-RpoD due to the presence of the His₆-tag.
3. The proposed exchange/binding amount of His₆-RpoD in the native *Pf* RNAP complex was below the detection limit by the antibody method.

2.3.2 *Reconstitution of Pyrococcus furiosus RNA polymerase and in vitro transcription*

2.3.2.1 *Preparation and purification of plasmid for the Pf RNAP subunits*

2.3.2.1.1 *Design Primers for eleven subunits of Pf RNAP:* Primers were designed for RpoB (PF1564), RpoA' (PF1563), RpoA (PF1562), RpoE' (PF0256), RpoM (PF0986), RpoL (PF0050), RpoH (PF1565), RpoN (PF1643), RpoE'' (PF0255), RpoK (PF1642). The sequence of the primers and the gene details are described (Table 2.2, 2.3). The vector, pET151/d-TOPO enables directional cloning in a single step method without the need for ligation. The vector has an engineered His₆-tag followed by a TEV/AcTEV cleavage site (Glu, Asn, Leu, Tyr, Phe, Gln/Gly) (Figure 2.7a, 2.7b). Addition of TEV protease after purification of the protein

facilitates removal of His₆-tag. The forward primers for all the mentioned genes were designed such that it contains 'cacc' sequence at the 5' end and the reverse primers had no modifications. The 'cacc' sequence serves as the recognition site that enables the directional cloning of the insert in the vector.

2.3.2.1.2 *PCR amplification of Pf RNAP genes (inserts):* The annealing temperature for all the inserts were optimized by gradient PCR. Blunt end PCR products were prepared using 2.5 U (1 μ L) of *Pf* DNA polymerase, 200 ng (2 μ L) of *Pf* genomic DNA, 10 mM (2 μ L) of dNTPs, 1x *Pf* polymerase buffer, and 10 μ M (5 μ L) of the corresponding forward and reverse primers for each insert per 100 μ L of the reaction mixture [134]. The PCR conditions were: 94°C for 3 min followed by 30 cycles of 1 min at 94°C (denaturation), 1 min at optimized annealing temperature, and 1 min at 72°C (extension) per cycle, with a final extension of 20 min at 72°C.

2.3.2.1.3 *Purification of the PCR product:* The PCR product was purified using the QIAquick PCR purification kit. The purified DNA was loaded on a 1% agarose gel and the gel was run under native running conditions using 1x TAE buffer, at a constant voltage of 80 V for 45 min. The sample was mixed with 6x loading dye in a 6:1 proportion prior to loading on the gel. The gel was stained with 0.5 μ g/mL ethidium bromide for 10 min and the image was recorded. The gel band was excised and purified further using the S. N. A. P. gel purification kit. The amount of gel-purified DNA was estimated using a Hoefer DyNA Quant 200 fluorometer, and the DNA amount was expressed in terms of ng/ μ L. Following the same protocol, the control PCR product (primers were obtained from the

manufacturer) was also purified and the image recorded. Restriction digest for the PCR products was also performed (Figure 2.8a - 2.8h).

2.3.2.1.4 *TOPO Cloning Reaction:* The reaction mixture included 1:1 molar ratio of gel purified PCR product: pET151/d-TOPO vector. Salt solution and sterile water was also added to a total volume of 10 μ L (Table 2.4). The solution was mixed gently and incubated at room temperature for 10 min. For large PCR inserts, like RpoA' and RpoB, overnight incubation at room temperature was recommended by the manufacturer (personal communication, Invitrogen). The control PCR insert (750 bp) was used as a positive control for the cloning reaction.

2.3.2.1.5 *Transformation into Top10 cells:* The vector with the insert was transformed into one shot Top10 chemically competent *E. coli* cells [127] (www.invitrogen.com). The cells were subjected to heat-shock at 42°C for 30 sec without shaking, the tubes were immediately transferred to ice and 250 μ L of S. O. C. medium was added. The tubes were incubated with shaking at 200 rpm at 37°C for 1 h. The transformation was spread on LB-agar plates with 100 μ g/mL of carbenicillin and incubated at 37°C for 16 h. Colonies (8-10) were selected and step culture of 3 mL and 100 mL were grown for 16 h at 37°C with shaking at 225 rpm in LB medium containing 100 μ g/mL carbenicillin and then transferred to 1 L LB medium with carbenicillin, and incubated at 37°C for 16 h. The cell extract was collected and centrifuged at 3,000 xg for 20 min using a JA-5.3 rotor. The supernatant was collected; the plasmid was purified from the supernatant using Qiagen plasmid purification kit. The plasmid was collected in 5 mL eluting buffer (QF buffer, Qiagen).

2.3.2.1.6 *Purification of plasmid DNA of Pf RNAP subunits:* The plasmid was purified by ethanol precipitation by adding 1/10th volume of 3 M sodium acetate (500 μ L), mixed well, followed by the addition of 2.5x cold 100% ethanol and mixed well. The resulting solution is centrifuged at 20,000 xg for 15 min using a JA-25.5 rotor. The supernatant was carefully decanted and 1 mL of cold 70% ethanol was added. The solution was centrifuged at 20,000 xg for 20 min. The supernatant was removed and the tube was allowed to air dry for 30 min until all the residual ethanol evaporated. The purified plasmid was resuspended in 10 mM Tris buffer (pH 8.0). The fragment pattern of the plasmid digested using restriction endonuclease substantiates successful directional cloning of the inserts into the vector (Table 2.5) (Figure 2.9a - 2.9f). These results were further corroborated with electropherograms obtained (Figure 2.10a, 2.10b). The sequence of the gene obtained (IBL, UGA) was compared to the corresponding sequence from the NCBI database and verified (www.ncbi.nlm.nih.gov).

2.3.2.2 *Recombinant Protein expression, purification and characterization of Pf RNAP subunits*

2.3.2.2.1 *Protein expression:* About 10 ng of purified plasmid DNA was transformed into BL21* [DE3] cells. Briefly, the cells were heat-shocked at 42°C for 30 sec and immediately transferred to an ice bath for 1 min. S. O. C. medium (200 μ L) was added and the culture was incubated with shaking at 37°C for 1 h. An aliquot of 50 μ L and 100 μ L cultures were streaked on LB-agar plates (www.invitrogen.com). The plates were incubated at 37°C for 16 h. Colonies (8-10) were picked and the viable ones were selected by inducing expression in pilot

scale (3 mL) with IPTG. Step cultures of 3 mL and 100 mL were grown in LB media with 100 $\mu\text{g/mL}$ carbenicillin, followed by growth in 1 L ZYP-5052 media with 100 $\mu\text{g/mL}$ carbenicillin. ZYP-5052 is a rich medium for growth with little or no induction during log phase and auto-induction of expression as the culture approaches saturation [135]. The conventional method using IPTG for induction produced very low yield of recombinant proteins (approximately 1-10 μg). Autoinduction using ZYP-5052 media provided about 100 μg - 1 mg protein. The culture was incubated for 8 h at 37°C, after which, the cells were harvested by centrifugation at 3,000 $\times g$ for 20 min, the cell pellet was collected and stored at – 20°C until further purification. Protein expression for RpoA' and RpoB resulted in very low yield (< 0.1 mg). Transformation into BL21 AI cells and protein expression was performed. BL21 AI cells tightly regulate expression, the inducer being 0.2% arabinose (personal communication, Invitrogen technical support). This method produced about 0.25 mg of protein. Probably these genes were toxic, or perhaps the low expression could be related to the large size of the insert (gene).

2.3.2.2.2 *Preparation of E. coli cell extract:* Frozen cell pellet (2 g - 5 g, wet weight) was thawed with the addition of 140 μL of protease inhibitor cocktail and re-suspended in 25 mL of 20 mM HEPES, 150 mM NaCl (pH=7.55). The cells were disrupted by means of a sonicator in the continuous mode, the amplitude gradually increased from 0 to 9, and continuously sonicated with the power set to 20% for 1 min. The amplitude knob was turned back to zero for 30 sec, and the cycle was repeated for 10 min. The cell suspension was clarified using a

centrifuge at 35,000 xg for 30 min. The clear supernatant was collected and filtered through a 0.2-micron syringe filter and the extract was ready for injection onto a nickel affinity column.

2.3.2.2.3 *Nickel affinity protein purification:* HP chelating column was connected to AKTA prime plus FPLC system. The UV detector was set to 0.1 AUFS and the wavelength was 280 nm. The column was loaded with 5 mL of 0.1 M nickel sulfate and washed with 5 column volumes of degassed nanopure water until all the residual nickel was removed from the column. The column was equilibrated with binding buffer and the flow rate was set to 2.0 mL/min. The clarified supernatant (28 mL) containing the recombinant protein was loaded onto a 5 mL chelating nickel column equilibrated with 20 mM HEPES and 150 mM NaCl (pH 7.55) and the His₆-tag protein was eluted with a linear gradient of elution buffer containing 100 mM - 500 mM imidazole. The eluted His₆-tag protein (15 mL) was concentrated to 100 μ L using YM-3/ YM-10/ YM-30/ YM-100 membrane centrifugal filter depending on the calculated molecular weight of the protein. The amount of protein was estimated using Biorad DC assay.

2.3.2.2.4 *SDS-PAGE analysis of purified protein:* The proteins were separated using SDS-PAGE on a 4-20% Tris gel in 1x Tris/glycine/SDS. The gel was run at 200 V for 60 min. Samples were heated at 95°C for 5 min in 4x Laemmli buffer prior to loading. The gels were stained using coomassie blue stain and the image recorded (Figure 2.11).

2.3.2.2.5 *Validation of the protein sequence by tryptic mass fingerprinting:* The gel band was excised and transferred to a microcentrifuge tube, destained with 6% acetic

acid. The gel slice was rinsed quickly three times with nanopure water to remove the residual coomassie blue stain. The gel slices were then dehydrated using a vacuum evaporator. Trypsin digestion was carried out with the addition of 15 μL of 10 ng/ μL trypsin and incubated for 16 h at 37°C. Tryptic peptides from the gel slice were extracted in four 10 min steps (1x with 30 μL of 25 mM NH_4HCO_3 , and 2x with 30 μL of 50% acetonitrile, 25 mM NH_4HCO_3 and 1x with 75% acetonitrile, 0.1% TFA). The supernatant from each extraction were combined and collected. The extracted peptides in solution were concentrated using a vacuum evaporator to 4 μL and 1 μL of 0.5% TFA was added. The tryptic peptides for each *Pf* RNAP subunit were purified separately using NuTip C-18 tips and spotted directly on a MALDI plate (3 μL , containing the matrix α -cyano-4-hydroxycinnamic acid). Data analysis is performed using Protein Prospector v 3.2.1 using MS-Fit program [130] (Figure 2.12a- 2.12t). The sequence obtained from the program was compared using BLASTP on NCBI database.

2.3.2.2.6 *Cleavage of His₆-tag:* The His₆-tag was cleaved for each purified subunit. TEV protease (1 μL , 10 U) was added to 20 μg of recombinant protein, 7.5 μL of 20x TEV buffer, 1.5 μL of 0.1 mM DTT, in a total volume of 150 μL . The mixture was incubated at 30°C for 6 h (manufacturer's protocol, Invitrogen). The solution was then loaded on a nickel affinity column equilibrated with the binding buffer. The protein without the His₆-tag is eluted in the flow-through fraction and the His₆-tag detached by the TEV protease is eluted using a linear gradient of elution buffer containing imidazole (100 mM - 500 mM). The fractions were collected

and concentrated to 250 μ L using the corresponding membrane filter as previously described.

2.3.2.2.7 *Reconstitution of Pf RNAP:* Total reconstitution of *Pf* RNA polymerase from isolated subunits was attempted by mixing stoichiometric molar ratios of each subunit. Subunits RpoA', RpoB, RpoA, RpoD were mixed with the other subunits in the ratio 1:10 (Table 2.6). Urea dialysis was performed. Briefly, the *Pf* RNAP recombinant subunits were mixed in 20 mM HEPES, 150 mM NaCl, and 6 M urea and loaded in a dialysis membrane (slide A lyzer, MWCO: 10,000 Da). The dialysis membrane was suspended in a reservoir containing 6 M urea, the buffer was changed every 3 h. Step dialysis was achieved by decreasing the concentration of urea in six steps [25]. The final buffer was 20 mM HEPES, 150 mM NaCl. The resulting solution was injected into a heparin column and eluted with 20 mM HEPES, 150 mM NaCl, 0.5 M $(\text{NH}_4)_2\text{SO}_4$. The fraction was collected and concentrated using YM-100 membrane filter. The functionality of the reconstituted RNA polymerase was tested by non-specific in vitro transcription assay.

2.3.2.2.8 *In-vitro Transcription Reaction:* Non-specific transcription assay conditions using *Pf* RNAP have been optimized previously using radioactive labeled nucleotides [136, 137]. The optimal pH of the transcription buffer was 9.0 and temperature was 70°C. The template used for in vitro transcription reaction with reconstituted *Pf* RNAP was a fragment of the *Pf* membrane-bound hydrogenase gene (PF1423), 213 bp upstream of the translation start site and 150 bp downstream of the transcription initiation site and PF1287 (sufC), 380 bp upstream of the translation

start site and 50 bp downstream of the operon reading frame (ORF). The template was amplified by PCR. A positive control for DIG labeling and detection was carried out using T7 RNAP and pGEM as the template (Figure 2.13). Two negative controls incorporated were the reaction mixture with all the components except the template and the other without the RNAP.

2.3.2.2.9 *DIG labeling of the RNA:* A non-radioactive method was used to test the transcriptional activity of the reconstituted *Pf* RNAP. Digoxigenin (DIG) labeled UTP (DIG-11-UTP) was used to label the generated mRNA transcripts. Transcription assay conditions and DIG labeling of RNA transcript was carried out following the manufacturer's protocol for controls involving T7 polymerase and pGEM template (www.roche-applied-science.com). The transcription assay conditions with reconstituted *Pf* RNAP was slightly modified, the reaction mixture (25 μ L) contained 250 ng of purified PF1423 UOR DNA as the template, 40 mM HEPES (pH 7.9), 1x transcription buffer, 1 mM each of ATP, CTP and GTP, 0.65 mM UTP, 0.35 mM DIG-11-UTP, 2 μ L of RNase inhibitor, 1 μ M each of *Pf* TFB and *Pf* TBP and 10 μ L of the heparin fraction of *Pf* RNAP. The mixture was incubated at 70°C for 30 min, cooled for 1 min on ice and the reaction was stopped by the addition of 2 μ L of 0.2 M EDTA (pH 8.0). The positive control reaction mixture included 1 μ L (1 μ g) of linearized pGEM template, 2 μ L of 10x transcription buffer (400 mM Tris-HCl (pH 8.0), 60 mM MgCl₂, 100 mM DTT), 2 μ L 10x DIG RNA labeling mix (1 mM each of ATP, CTP and GTP, 0.65 mM UTP, 0.35 mM DIG-11-UTP), 2 μ L (20 U/ μ L) T7 RNAP, sterile DEPC water to make a total reaction volume of 20 μ L. The

negative controls were also treated in a similar way. The reaction mixture was incubated at 42°C for 2 h and the reaction was stopped by the addition of 2 μ L of 0.2 M EDTA (pH 8.0) (Figure 2.14).

2.3.2.2.10 *Analysis of RNA transcript:* The RNA samples generated were mixed 1:1 with 2x TBE loading buffer, heated at 95°C for 5 min and cooled on ice for 5 min. 10 μ L of denatured sample was loaded on a 10% TBE-urea gel (pre-washed and pre-run) in 1x TBE buffer run at 400 V for 30-45 min. 5 μ L of RNA marker and 5 μ L of DIG-labeled Actin RNA was also loaded on the same gel. The gel was removed and placed on a positively charged nylon membrane and loaded in a blot transfer cassette and the transfer was carried out at 100 V for 30 min.

2.3.2.2.11 *Detection:* The RNA transcript on the membrane was fixed by exposing the membrane to UV for 30 sec. The membrane was transferred to a tray and incubated with 10 mL of the washing buffer for 5 min followed by incubation with 20 mL of 1x blocking buffer for 30 min with intermittent shaking. About 10 mL of anti-DIG antibody solution was added to the membrane and incubated for 30 min followed by 2x incubation with 20 mL of washing buffer for 20 min. The membrane is then equilibrated with 10 mL of detection buffer for 10 min. The membrane is transferred to a hybridization bag with the RNA side facing up and about 4-6 drops of CDP-star solution was added, air bubbles were removed, incubated for 5-30 min and the membrane was not allowed to dry completely. Visualization of RNA bands was carried out using an X-ray film.

2.4 Results and Discussion

2.4.1

Extraction of native *Pf* RNAP using His₆-RpoD: Methods to purify native RNAP from cells have been developed in many organisms including *Pf* [133, 138]. These methods employ conventional column chromatographic methods and although an active RNAP could be purified, it is time-consuming, involves various separation steps and contains extra polypeptides bound to RNAP. An attempt to purify native *Pf* RNAP from *Pf* cells in a single-step chromatographic separation method using a bait like His₆-RpoD was carried out [132]. His₆-RpoD was expressed and purified from cells containing the plasmid obtained (L. M. Lewis and M. K. Eidsness). Recombinant His₆-RpoD was bound to a zinc affinity column, the *Pf* cell extract was loaded on the same column and the His₆-RpoD (and proteins bound to His₆-RpoD) was eluted and analyzed by LC-ESIMS. The presence of His₆-RpoD in any high-molecular weight complex so obtained was then analyzed using polyclonal antibodies of His₆-RpoD and with pentahis (His-tag) antibodies. Approximately 0.1 mg (1.22 nmol) of protein could be extracted from 5 g (wet weight) of *Pf* cells. Based on empirical calculations assuming this was all RNAP complex, the estimated amount of native RpoD present in *Pf* RNAP complex would be 0.03 mg. The amount of pure recombinant His₆-RpoD added was 0.80 mg (25 nmol), detected by western blot using polyclonal antibody. Neither the western blots using polyclonal antibody or those using pentahis antibody were able to detect the complex. The possible reasons could be:

1. The native gel and western blot transfer under native conditions with the absence of detergents might have reduced the possibility of exposure of the antibody to the His₆-RpoD.
2. The native *Pf* RNAP complex may not have bound to His₆-RpoD due to the presence of the His₆-tag.
3. The amount of exchanged His₆-RpoD in the native *Pf* RNAP complex was below the detection limit by the antibody method.

Thus, the exchange or binding of His₆-RpoD to native RNA polymerase was not detected by this method. So I attempted cloning and reconstitution of the subunits of *Pf* RNAP so that RNA polymerase could be prepared in large scale for *in vitro* transcription experiments for the group.

2.4.2 ***Reconstitution of *Pf* RNAP from recombinant subunits:*** *Pf* RNAP subunits exhibit significant homology to other archaeal and yeast homologs (Figure 1.7, 1.9). Based on the gene details of all the subunits of *Pf* RNAP (Table 2.2) obtained from NCBI database, primers were designed for individual subunits (A, A', B, E', E'', H, K, L, m, N) (Table 2.3). Blunt-end PCR products for individual subunits were prepared with CACC at the 5' end which enables directional cloning of the fragment into pET151/D-TOPO vector, the overhang in the cloning vector (GTGG) invades the 5' end of the PCR product, anneals to the added bases, and stabilizes the product in the correct orientation (Figure 2.7) (www.invitrogen.com). PCR fragments were analyzed using restriction endonucleases like Hind III and Nsi I to confirm the size of the gene amplified (Table 2.5, Figure 2.8) before performing the cloning reaction. The conditions for

cloning of the larger subunits were slightly altered (Table 2.4). The pET TOPO construct with the insert was transformed into One shot Top10 chemically competent *E. coli* cells and the cells were grown on a plate. About 5-10 colonies were picked, plasmid DNA was isolated and the plasmid was analyzed using restriction endonucleases (Nsi I) to confirm the presence and correct orientation of the insert (Table 2.5, Figure 2.9). Further, the sequence was also verified by sequence analysis (IBL, UGA) (Figure 2.10a, 2.10b). The purified plasmid was then transformed into BL21 DE3 cells. His-tag and a TEV protease cleavage site on the pET151/D-TOPO vector enables metal-affinity purification of the expressed recombinant protein followed by cleavage of the His-tag using TEV protease. Individual subunits were expressed and purified separately (Table 2.6, Figure 2.11). I encountered difficulties in cloning and purifying the larger subunits (RpoA', 103 kDa; RpoB, 127 kDa) using the pET 151/D-TOPO vector. Efforts to improve the yield of the overexpressed recombinant protein using different expression strains of BL21 DE3 (BL21 DE3 pLysS and BL21 AI) competent cells, varying growth conditions (with different inducers like IPTG, arabinose; growth temperature), and growth media (2.5 % LB, 1% glucose, ZYP-5052 (Studier growth media)) were carried out. The probable cause of low expression yield of these proteins could be the toxicity of the gene (personal communication, Invitrogen). Tryptic peptides of all the subunits generated separately were analyzed by MALDI ToF to verify the sequence of the protein (Figure 2.12).

Reconstitution of *Pf* RNA polymerase from isolated subunits was attempted by mixing stoichiometric molar ratios of larger subunits (RpoA', RpoB, RpoA, RpoD) with 10-fold higher molar amounts of smaller subunits (RpoE', RpoE'', RpoH, RpoN, RpoK, RpoL, RpoM) (Table 2.6). To enable optimal folding of the *Pf* RNAP complex, step-dialysis using urea was performed [25]. The *Pf* RNAP complex was purified using a heparin column and concentrated using YM-100 membrane filter. The functionality of the reconstituted RNA polymerase was analyzed by non-specific *in vitro* transcription assay using a non-radioactive method to label the mRNA transcripts with digoxigenin (DIG)-labeled UTP (DIG-11-UTP). The mRNA transcripts generated were transferred onto a positively charged nylon membrane and analyzed using anti-DIG AP. A non-specific *in vitro* transcription assay using reconstituted *Pf* RNAP, TFB, and TBP was performed using *mbh* and PF1287 UOR (Upstream of ORF Region) DNA as the template (Figure 2.14a, 2.14b). T7 polymerase and pGEM template containing the T7 promoter element was used as a positive control (Figure 2.13) of the protocol. The results were inconclusive since the mRNA transcripts were not completely denatured; no distinct bands are seen.

Table 2.1: LC-MS data obtained from a single-step chromatographic separation of *Pf* RNAP complex from *Pf* cells.

Yeast RNAP Subunits	<i>Pf</i> RNAP Subunits	Expected Mol. wt. (kDa)	LC-MS Mol. wt. (kDa)
RPB12	RpoP	5760	
RPB6	RpoK	6240	
	RpoE"	6950	7150
RPB10	RpoN	7790	7274
RPB5	RpoH	9250	
RPB11	RpoL	11120	10254
RPB9	Rpom	12860	
RPB7	RpoE'	21700	21605
RPB3	RpoD	29780	28615
	RpoD6xHis	31760	31803
	RpoM (hypothetical)	41230	41614
RPB1	RpoA	44400	49585
	RpoA'	103000	90475
RPB2	RpoB	127000	

Table 2.2: Gene details and nomenclature of the *Pf* RNAP subunits (NCBI database).

No	<i>Pf</i> subunit	GI	Region		Common Name	Protein Mol. wt. (Da)	Yeast Equivalent	<i>Pf</i> Accession Number (NCBI)
			From	To				
1	Subunit l	18976422	57460	57747	RpoL	11114	RPB11	NP_577779
2	Subunit e" or e	18976627	263306	263491	RpoE2	6920		NP_577984
3	Subunit e'	18976628	263488	264057	RpoE1	21696	RPB7	NP_577985
3*	Hypothetical M	18977357	940306	941363		41206		
4	Subunit m	18977358	940318	940650	Rpom	12863	RPB9	NP_578715
5	Subunit a or a"	18977934	1455699	1456892	RpoA2	44404	RPB1	NP_579291
6	Subunit a'	18977935	1456903	1459626	RpoA1	103114	RPB1	NP_579292
7	Subunit b	18977936	1459639	1462992	RpoB	127003	RPB2	NP_579293
8	Subunit h	18977937	1463013	1463261	RpoH	9212	RPB5	NP_579294
9	Subunit k	18978014	1532854	1533027	RpoK	6236	RPB6	NP_579371
10	Subunit n	18978015	1533155	1533367	RpoN	8198	RPB10	NP_579372
11	Subunit d	18976372	1534790	1535575	RpoD	29781	RPB3	NP_579376
12	Subunit P	28380128	1857791	1857940	RpoP	5756	RPB12	NP_579738
13	Subunit F	18977408	991090	991452	RpoF	14095	RPB4	NP_578765

Table 2.3: Details of the sequence of the primers used to amplify the PCR inserts of subunits of *Pf* RNAP.

<i>Pf</i> RNAP Subunits	Primary Locus	Forward Primer	Reverse Primer	Gene-length (bp)
RpoA	NP_579291	5'-CACCATGGTCTCTCTTCTACTATTAAATC-3'	5'-TCACACCTCCTCCTTTTCTTTCTGTGGTCT-3'	1194
RpoA'	NP_579292	5'-CACCATGAAAAAAGTTATTGGAAGTATTGAG-3'	5'-TCACACCTTCGCCTTGT TATTTCTCTCATCTTTAG-3'	2724
RpoB	NP_579293	5'-CACCATGAGAGGTCCGACTGTTGTAGATGTTACT-3'	5'-TCACACCTCTCTGAGAGGTTTAACTTAGGTCTA-3'	3354
RpoH	NP_579294	5'-CACCGTGGCGGGGAAAAAGGAATT-3'	5'-TTAGTCTTCAACAACAACCCTATAGTAG-3'	249
RpoE''	NP_577984	5'-CACCGTGAGTGAAAAAGCCTGCAG-3'	5'-TCAGCGCACCCCTTATGGCAT-3'	186
RpoK	NP_579371	5'-CACCATGTTCAAGTATACGAGGTTTGA-3'	5'-TCAGCTCGGTCTGATTACTGTTATTGG-3'	174
RpoL	NP_579372	5'-CACCATGAAGATAGAAGTGATAAAGAAGG-3'	5'-TCAGCTCTTCACCGCCTTCTCCCATGCC-3'	288
Rpom	NP_578715	5'-CACCATGGTGAAATTCTGCCCCAAATG-3'	5'-TTATTCGTAACCTCTCCATACGTATCC-3'	333
RpoN	NP_579372	5'-CACCATGGGGGCAAGTCCCGTGATTGT-3'	5'-TCAATACACTCTGTAATGCATTATGTC-3'	213
RpoE'	NP_577985	5'-CACCATGTACAAGATAGTCACCGTAAAGG-3'	5'-TCACTTCTTACCCTCTCCTTTTCTTTC-3'	570

Table 2.4: Conditions for cloning of *Pf* RNAP genes into pET151/d-TOPO vector (www.invitrogen.com).

	Small size inserts (< 2 kb)	Large size inserts (< 4 kb)
Gel purified PCR insert	0.5-4.0 mL (e.g.: 1152 bp, 2 ng)	0.5-4.0 mL (e.g.: 3456 bp, 6 ng)
Salt solution	1 mL	3.3 mL
Water	Final volume = 6 mL	Final volume = 20 mL
pET151/d-TOPO vector	1 mL	1 mL
Incubation at room temp.	5 min	Overnight

Table 2.5: Fragments of *Pf* RNAP genes generated by restriction endonucleases.

<i>Pf</i> RNAP subunits	PCR insert size / Gene length (bp)	PCR digest map using HindIII (bp)	Plasmid size (bp)	Plasmid digest map Nsi I cuts vector (bp)
RpoA	1194	86, 1108	6954	6954
RpoA'	2724	-	8484	2194, 2476, 3814
RpoB	3354	396, 431, 772, 1755	9114	9114
RpoH	249	96, 153	6009	6009
RpoE"	186	-	5946	5946
RpoK	174	62, 112 (Nco I)	5934	5934
RpoL	288	82, 206 (Bsa I)	6048	6048
Rpom	333	49, 284 (Ssp I)	6093	6093
RpoN	213	197, 16 (Nsi I)	5973	197, 5776

Table 2.6: Calculated protein content of all the purified *Pf* RNAP subunits.

Subunit	Mol. wt. (kDa) w/o (with His ₆ -tag)	Conc. in mg/mL	Total amount (mg)	Amount (nmoles)
RpoA'	103 (106)	0.82	624	6.1
RpoB	127 (130)	0.78	630	5.01
RpoA	44.4 (47.4)	1.17	580	13.1
RpoD	30 (34)	0.51	1440	48.3
RpoH	9 (12)	1.98	240	25.9
RpoE''	7 (10)	1.10	254	36.5
RpoN	8 (11)	0.92	385	46.1
RpoK	6 (9)	0.99	197	31.6
RpoL	11 (14)	0.65	130	11.7
Rpom	13 (16)	1.54	600	46.6
RpoE'	22 (25)	0.50	150	6.9
Complex	382 (412)			

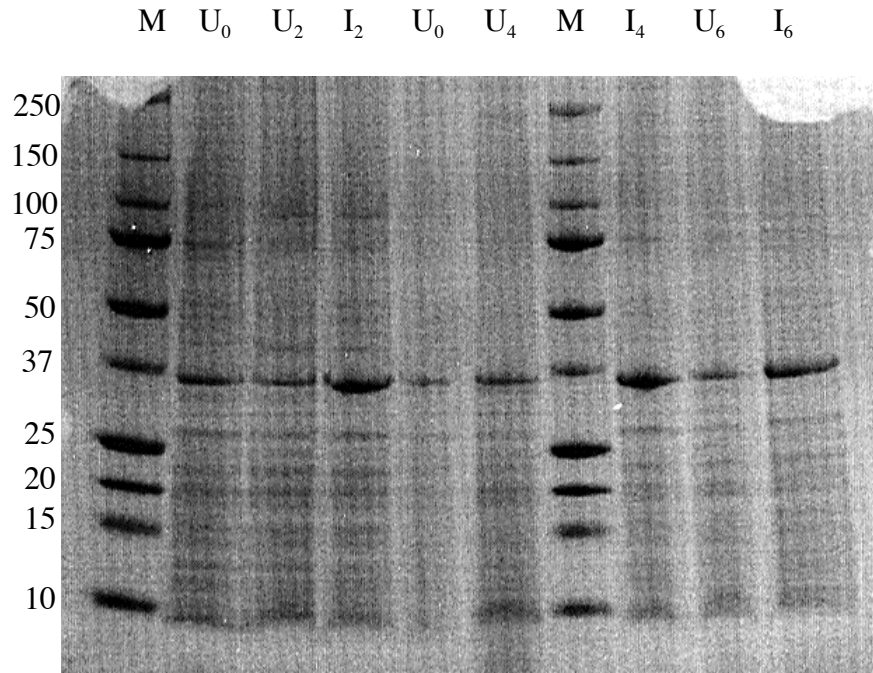


Figure 2.1: SDS-PAGE gel of His₆-RpoD over-expressed with the addition of IPTG.

Labels: U₀ – Uninduced 0 h, U₂ – Uninduced 2 h, I₂ – Induced 2 h, U₄ – Uninduced 4 h, I₄ – Induced 4 h, U₆ – Uninduced 6 h, I₆ – Induced 6 h, M – Precision plus markers.

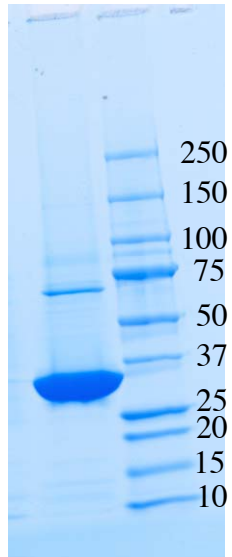


Figure 2.2: Purified recombinant His₆-RpoD.

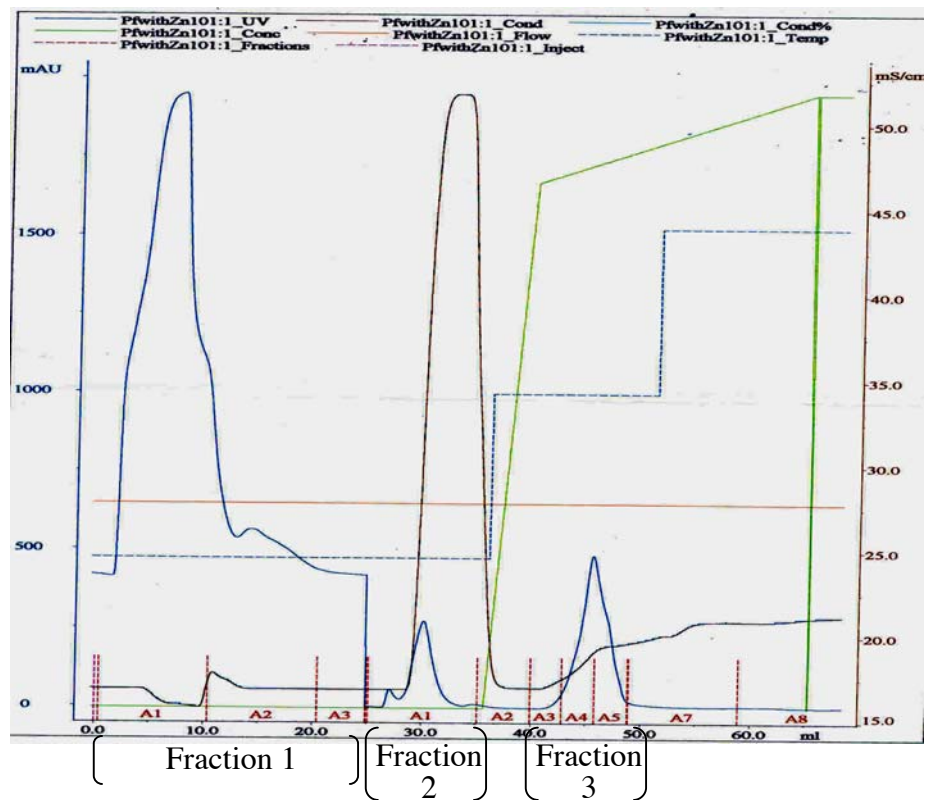


Figure 2.3: Zinc affinity purification of *Pf* cell extract treated with polymine P combined with His₆-RpoD.

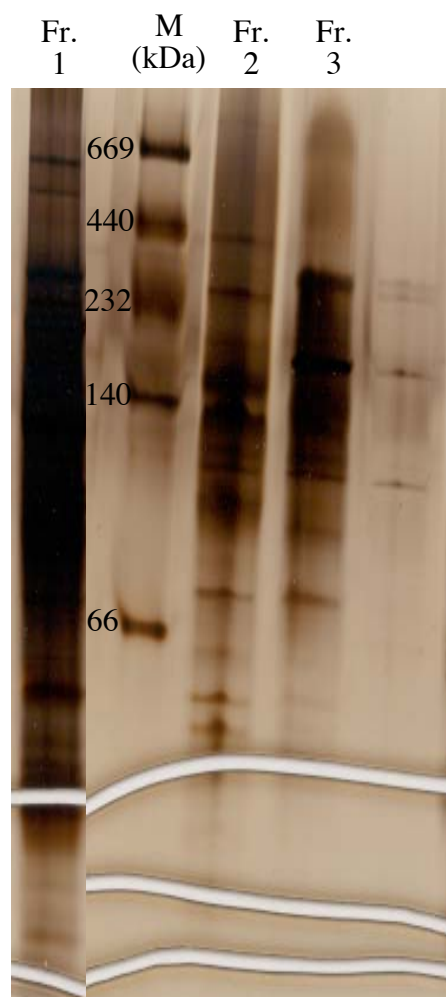
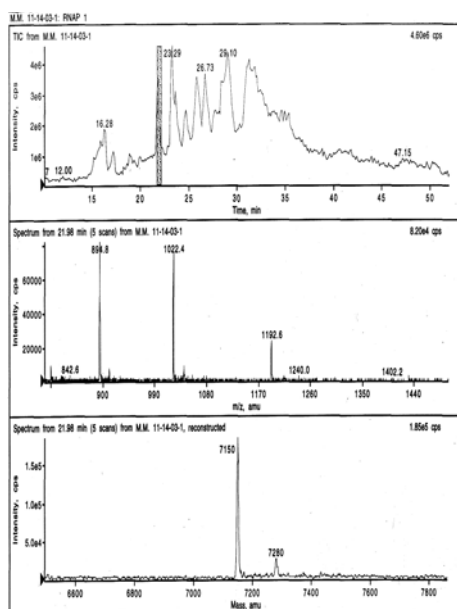
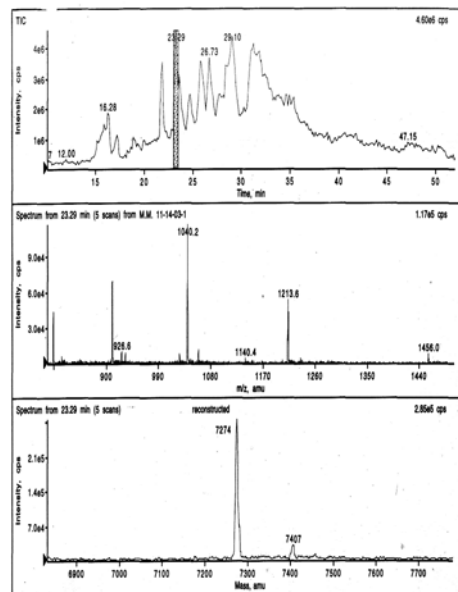


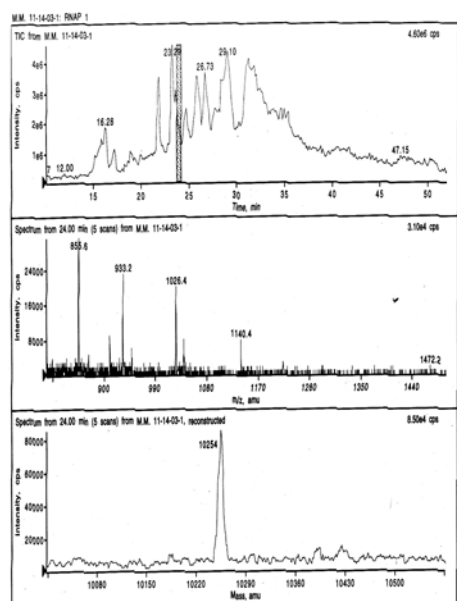
Figure 2.4: Native-PAGE gel of fractions eluted from Zn chelating column.



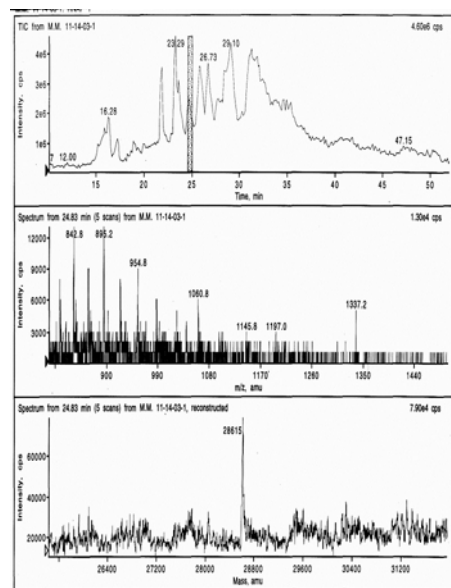
a



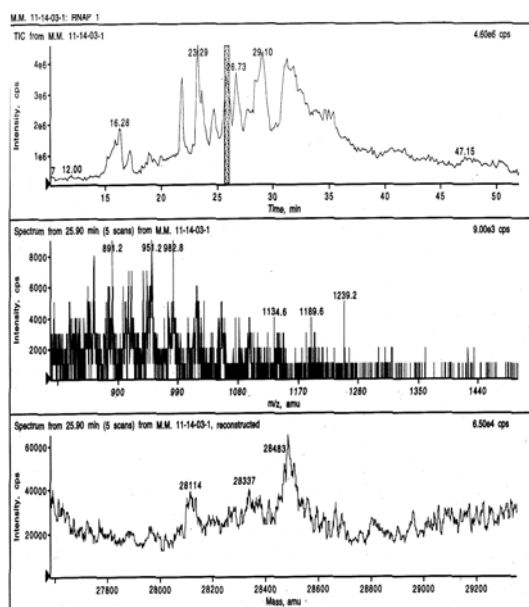
b



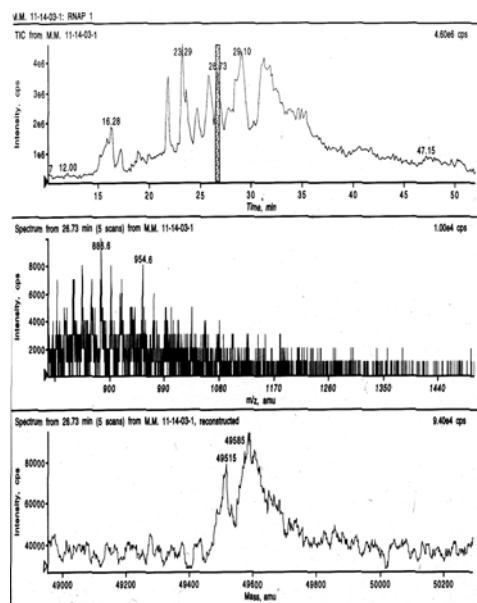
c



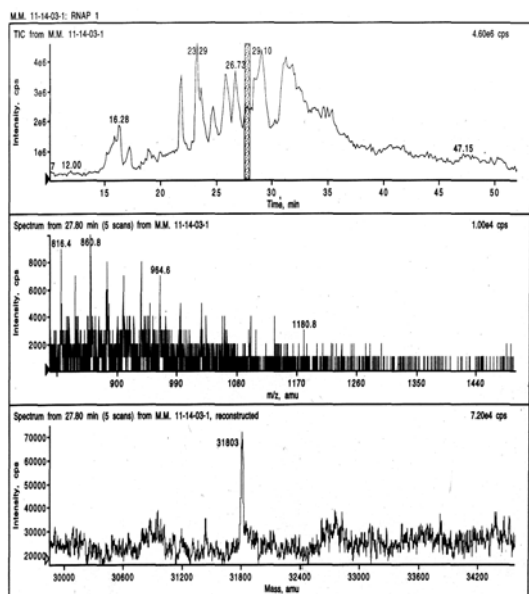
d



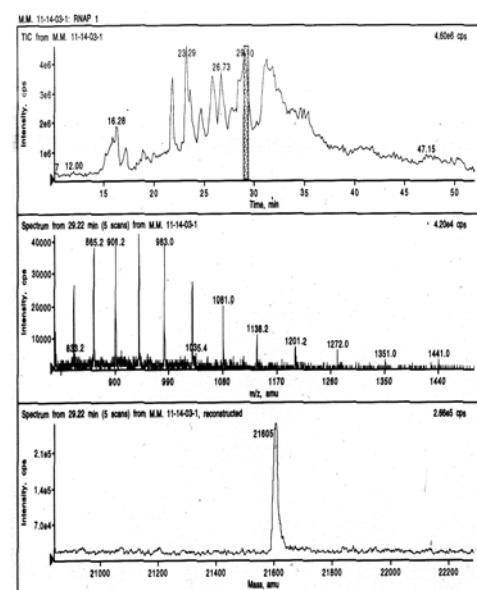
e



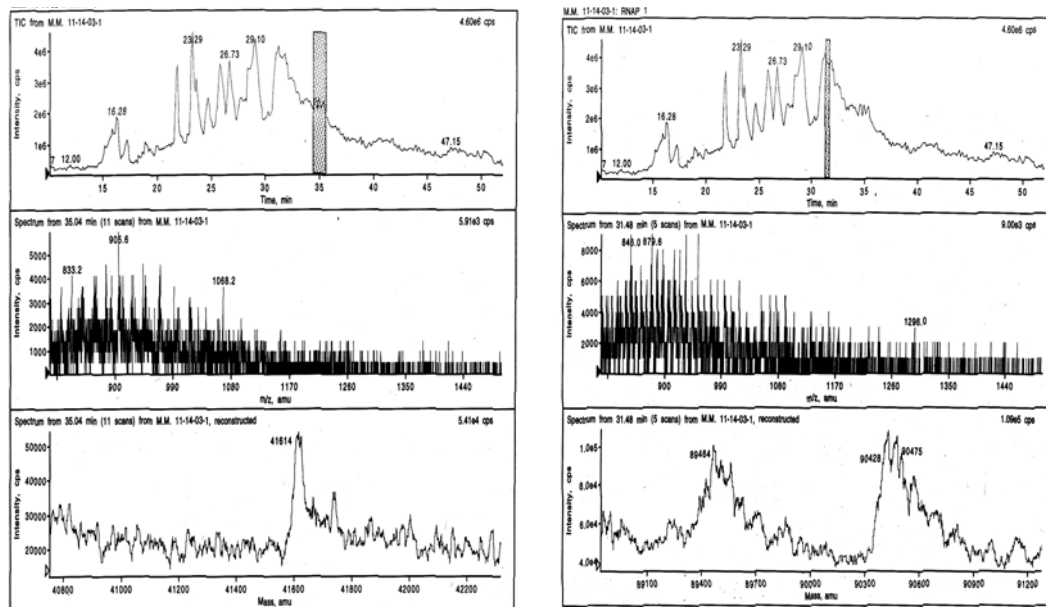
f



g



h

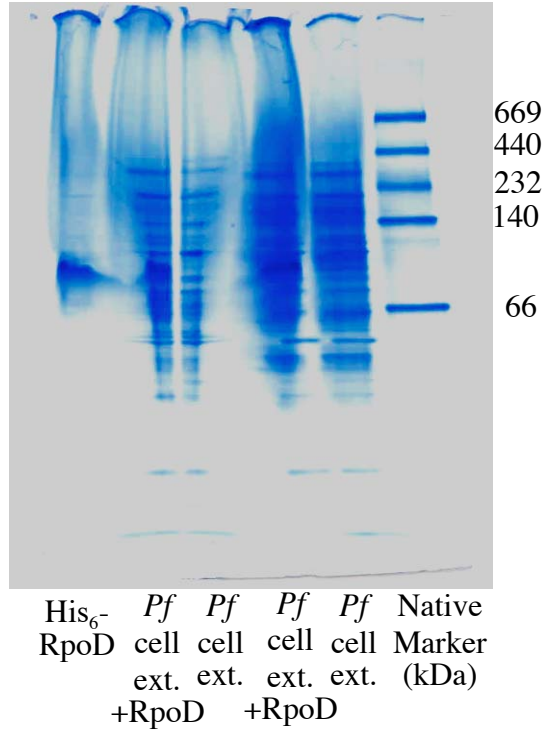


I

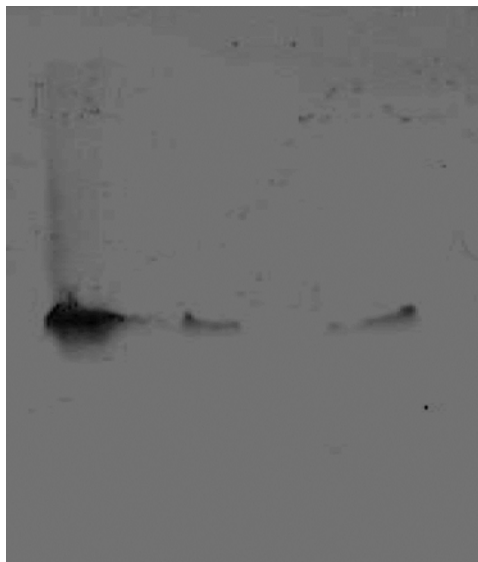
j

Figure 2.5: LC-MS chromatograms of putative native *Pf* RNAP complex extracted from *Pf* cells using recombinant His₆-RpoD.

Labels: a-RpoE'', b-RpoN, c-RpoL, d-RpoD, e-unknown, f-RpoA, g-His₆-RpoD, h-RpoE', i-RpoM, j-RpoA'.

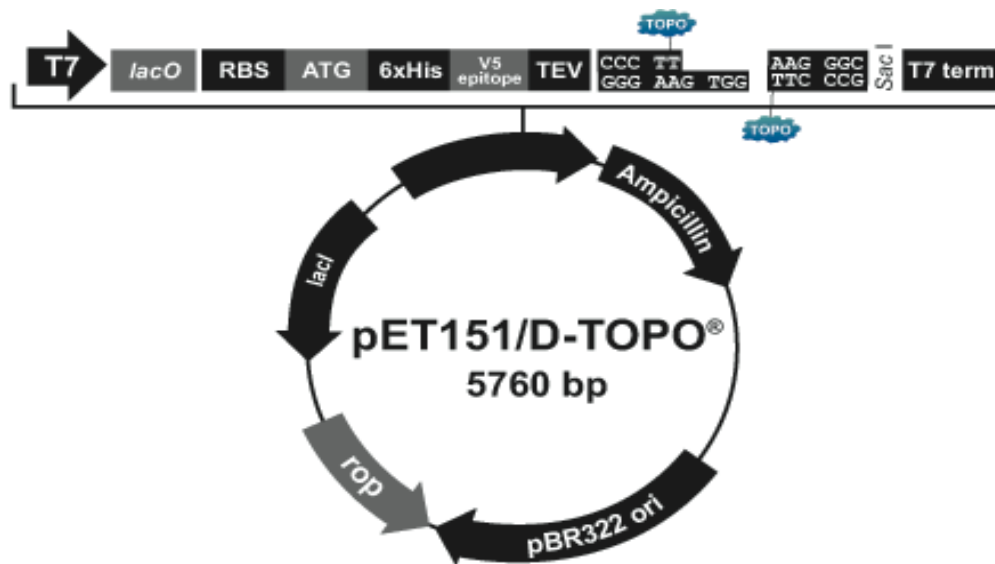


Native-PAGE gel of *Pf* RNAP cell extract combined with His₆-RpoD.



Western blot treated with polyclonal antibody of His₆-RpoD.

Figure 2.6: Native-PAGE gel and western blot of *Pf* RNAP cell extract.



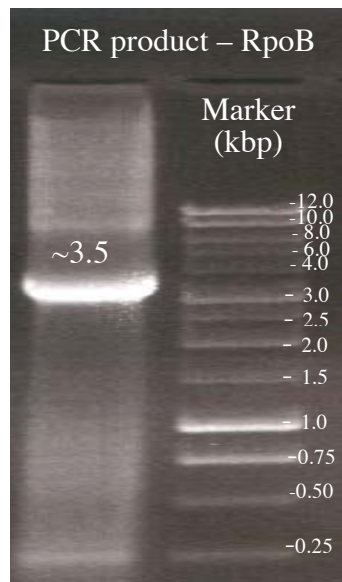
pET 151/D-TOPO vector

```

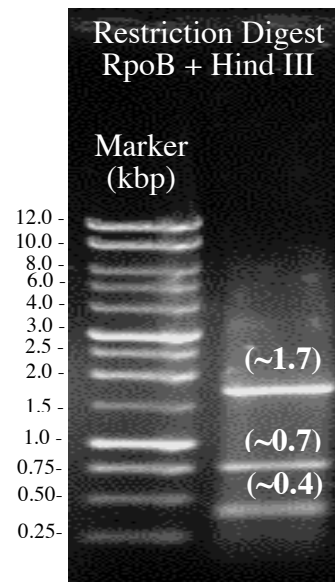
121 ATAGGGGCGCA GCAACCGCAC CTGTGGCGGC GGTGATGCCG GCCACGATGC GTCGGCGTA GAGGATCGAG ATCTCGATCC
      T7 promoter/priming site
      T7 promoter          lac operator
201 CCGGAAATTA ATACGACTCA CTATAGGGGA ATTGTGAGCG GATAACAATT CCCCTCTAGA AATAATTTTG TTAACTTTA
      RBS          Polyhistidine region          V5 epitope
281 AGAAGGAGAT ATACAT ATG CAT CAT CAC CAT CAC CAT GGT AAG CCT ATC CCT AAC CCT CTC CTC GGT CTC
      Met His His His His His His Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu
      TEV recognition site
351 GAT TCT ACG GAA AAC CTG TAT TTT CAG GGA ATT GAT CCC TT AAG AAGGG CGAGCTCAGA
      Asp Ser Thr Glu Asn Leu Tyr Phe Gln Gly Ile Asp Pro Phe Thr
      TEV cleavage site          T7 reverse priming site
411 TCCGGCTGCT AACAAAGCCC GAAAGGAAGC TGAGTTGGCT GCTGCCACCG CTGAGCAATA ACTAGCATAA
  
```

Sequence of His₆-tag and TEV protease cleavage site in pET151/D-TOPO vector

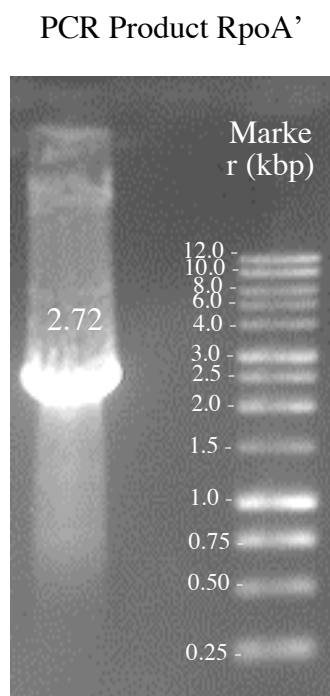
Figure 2.7: Schematic representation of the vector used to clone *Pf* RNAP subunits (www.invitrogen.com).



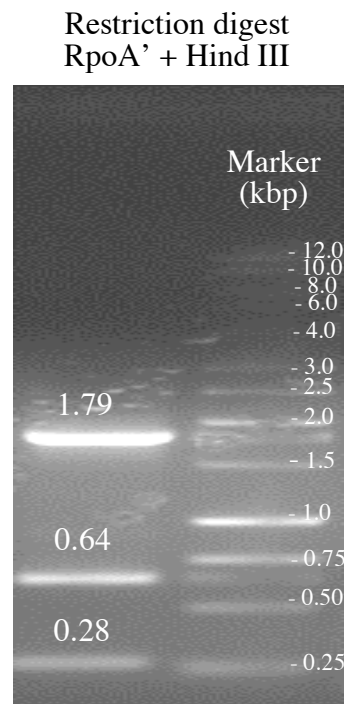
a



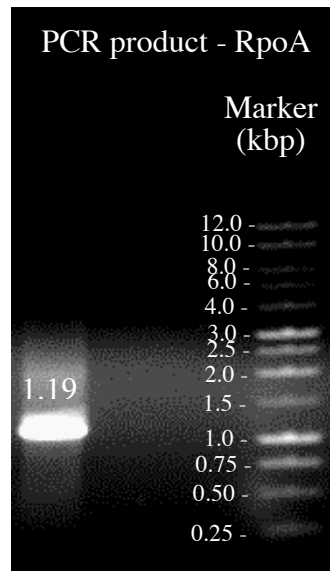
b



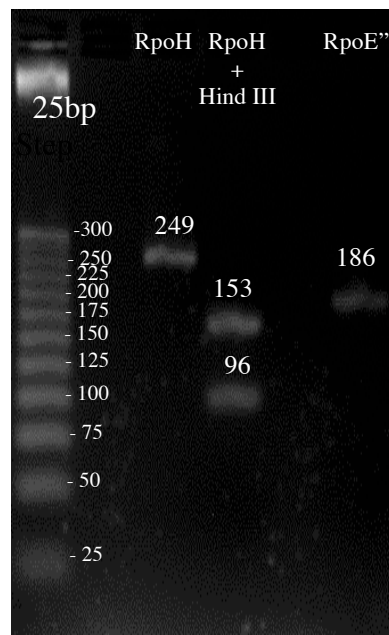
c



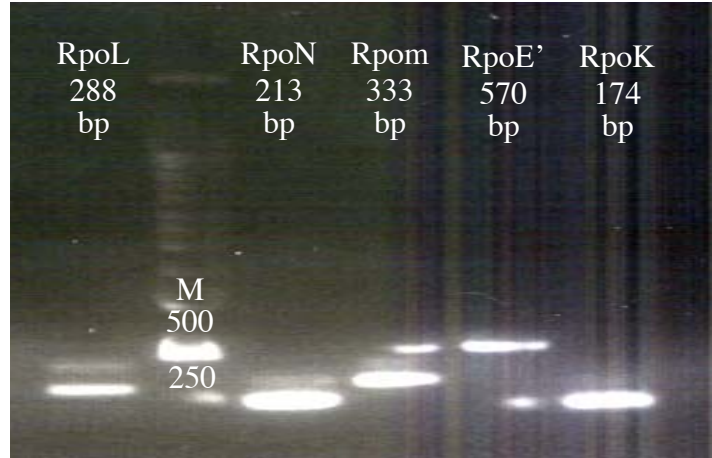
d



e

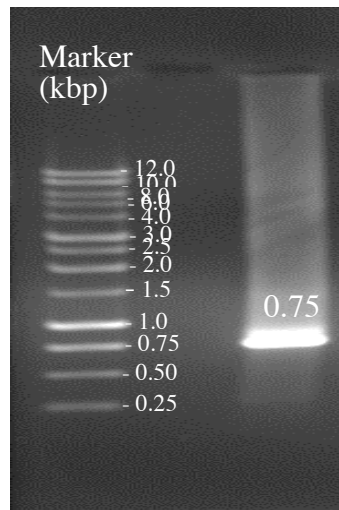


f



g

Control PCR insert

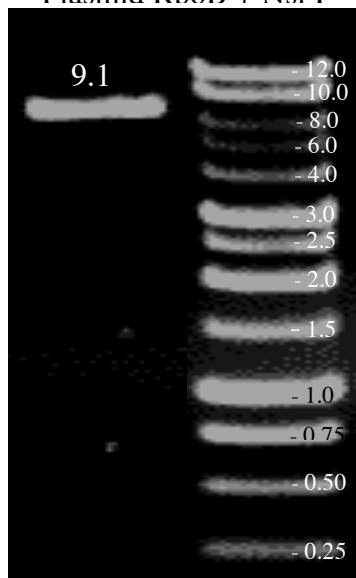


h

Figure 2.8: PCR product and restriction endonuclease digest map of *Pf* RNAP subunits.

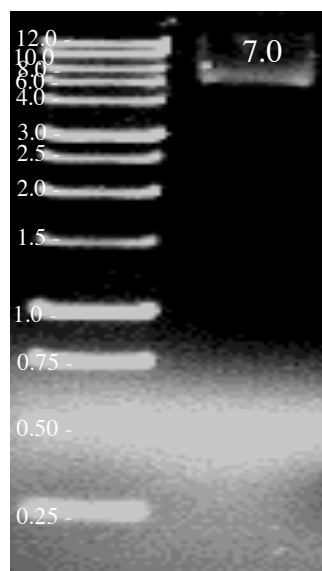
Labels: a-RpoB, c-RpoA', e-RpoA, h-control insert, g-RpoL, RpoN, RpoM, RpoE', RpoK PCR products; b-RpoB, d-RpoA', PCR product digested using the restriction endonuclease Hind III; f-RpoH PCR (lane 2), RpoH gene digested using Hind III (lane 3), RpoE'' PCR (lane 4).

Restriction digest
Plasmid RboB + Nsi I

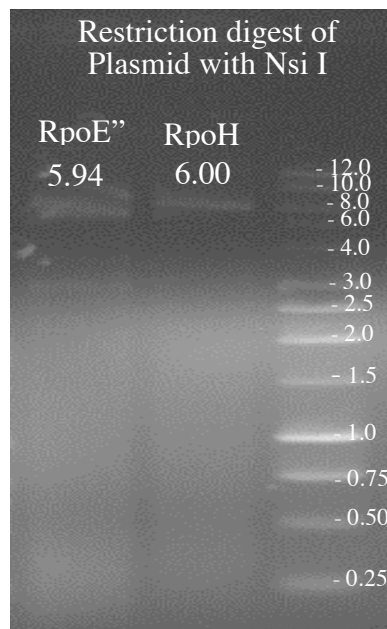


a

Restriction digest
Plasmid RboA + Nsi I

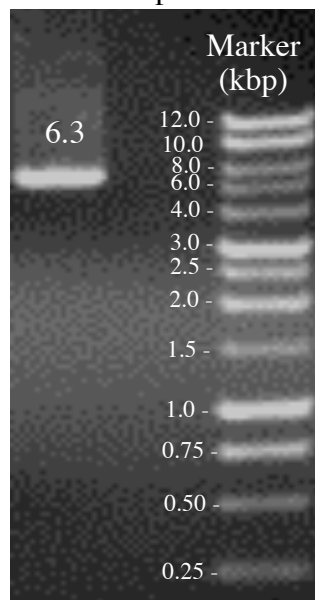


b

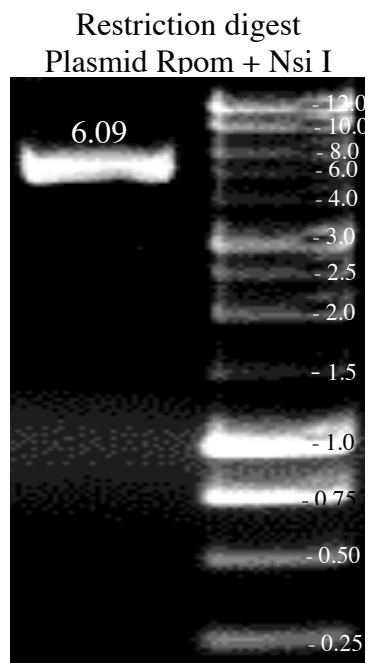


c

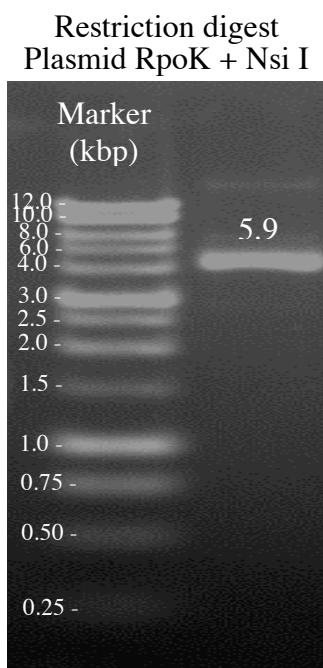
Restriction digest
Plasmid RpoE' + Nsi I



d



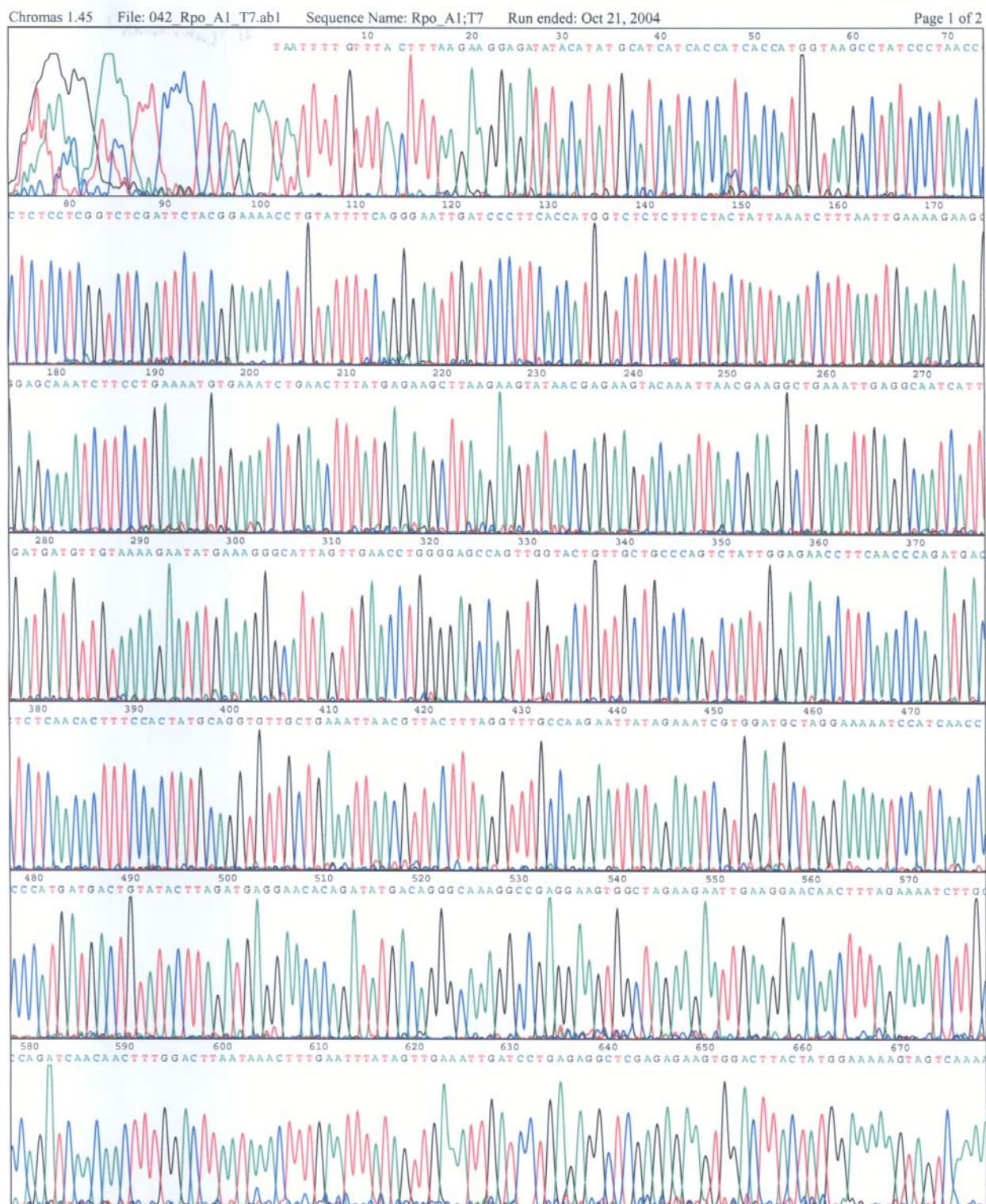
e



f

Figure 2.9: Restriction endonuclease Nsi I digest map of plasmid of *Pf* RNAP subunits.

Labels: a-RpoB, b-RpoA, c-RpoE'' & RpoH, d-RpoE', e-Rpom, f-RpoK.



a: Sequencing analysis carried out using T7 primer

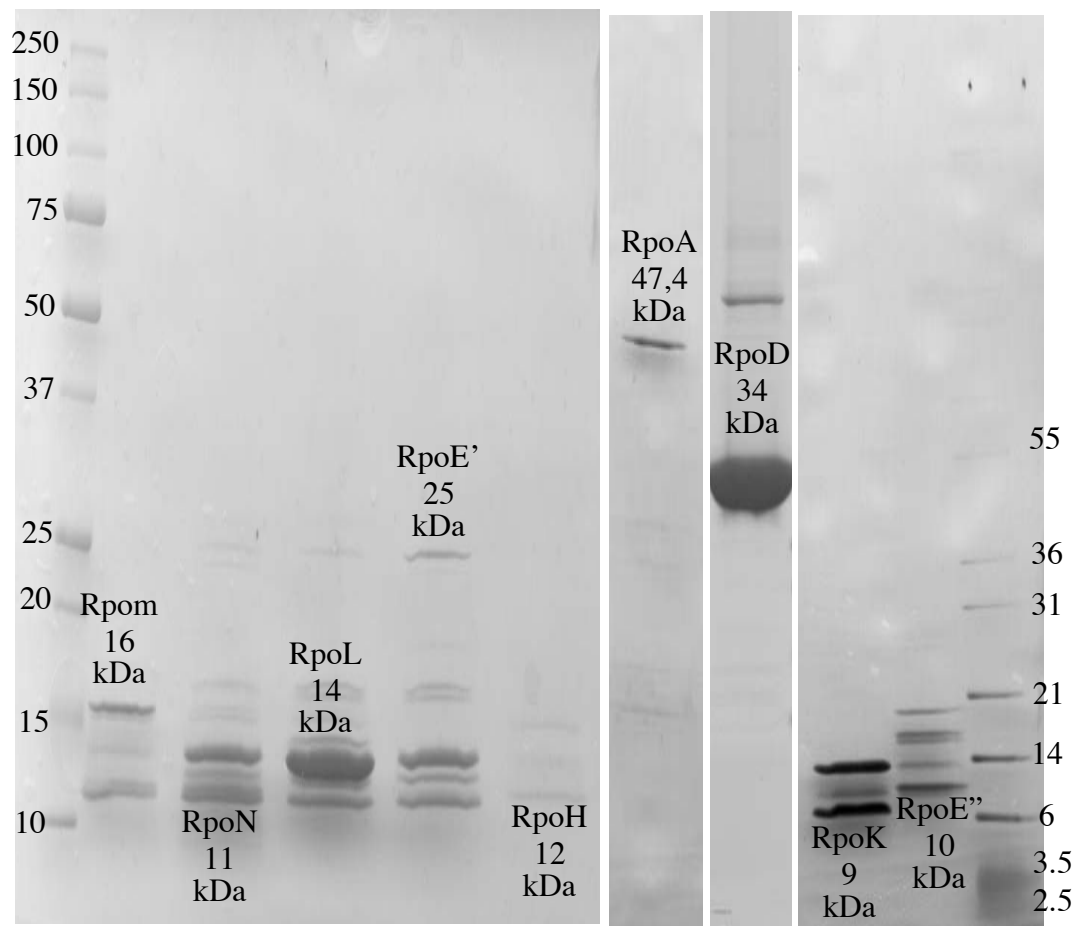
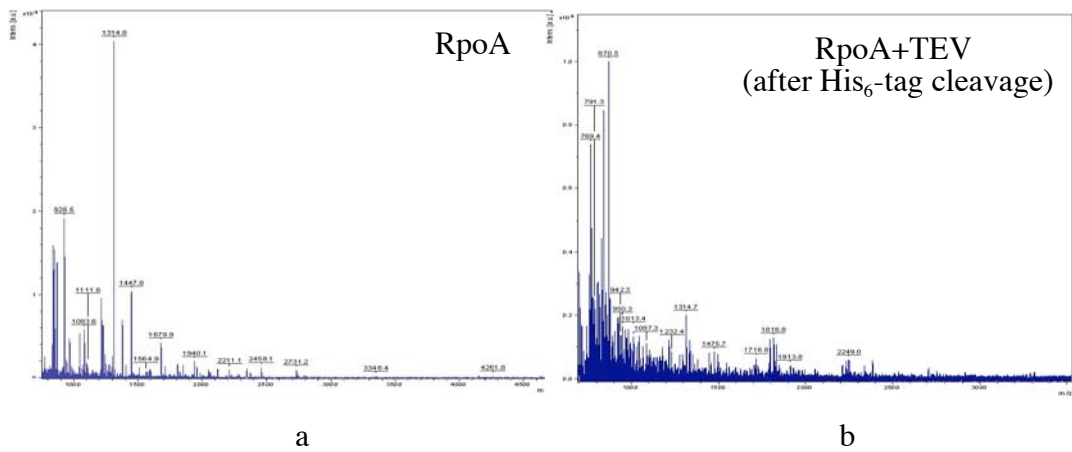
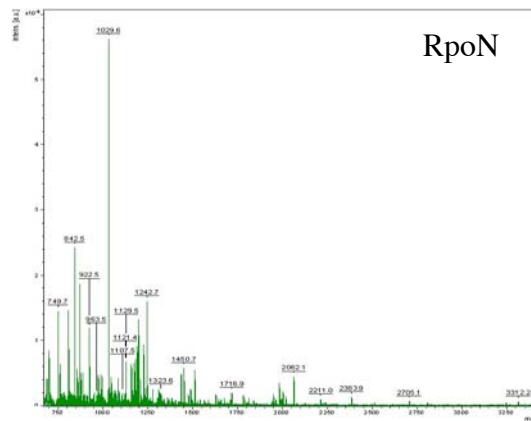


Figure 2.11: SDS-PAGE gel of *Pf* RNAP subunits.



MVSLSTIKSL IEK**KGANLPE** NVKSELYEKL KKYNEKYKLT **KAEIEAII**DD **VVKEYERALV**
 EPGEVGTVA AQSIGEPSTQ MTLNTFHYAG VAEINVTLGL **PRIIEIVDAR** **KNPSTPMMTV**
YLDEEHRYDR AKAEEVARRI **EGTTLENLAR** STTLDLINFE FIVEIDPERL ERSGLTMEKV
 VKKLESSFKS **AEFEVDGYTL** **IVRPK**KADKI **SDLRRFAEKI** **KKHRLKGLSG** **VGKTIVRKEG**
 DEYVIYTEGS NFKQVLKVPV VDPTRTRTNN **IHEIAEVLGI** **EAARNAIIDE** **IVSTMQEQGL**
EVDIRHIMLV ADMMTLDGIV RPIGRHGVVG EKSSVLARAA **FEITVQHLFE** **AAEKGEVDNL**
 NGVIENVLIG **QVPVPGTGMV** **KLTMKLPLRP** **QKEKEEV**

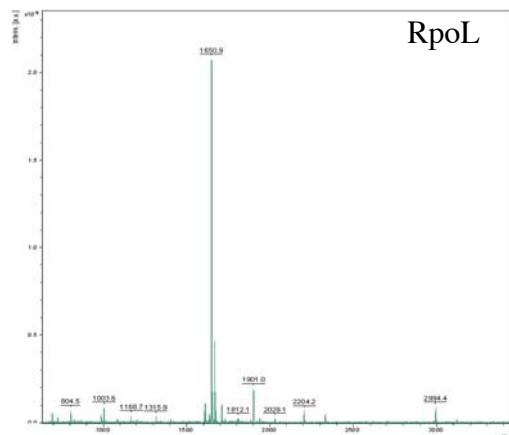
c



d

M**GASPVIVPV** **RCFTCGK**VIG DKYYEFKRR**V** **EAGEDPEKVL** **DDLGLERYCC**
RRMLLSHVEL **IDDIMHYRVY**

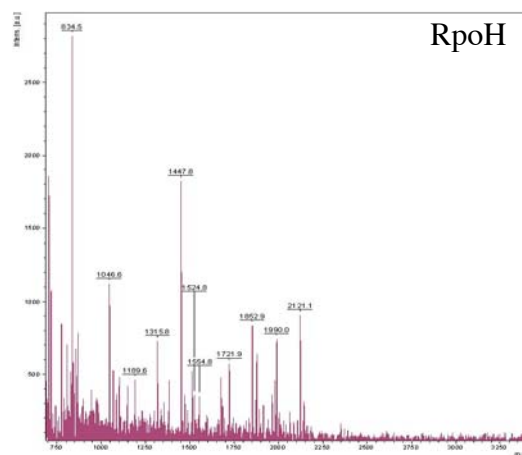
e



f

MKIEVIKKEE NLEFYLEGE DHTFANLLVE TLRENPHVKF TAYTIEHPIT
 MARKPRFRVV TDGEITPEEA LEEAAKKIFE RAKEEVLEAWE KAVKS

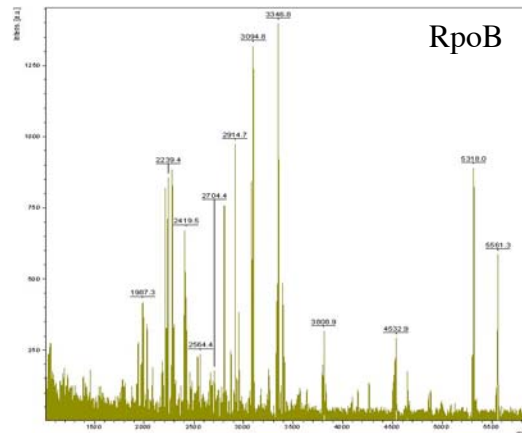
g



h

MAGKKEFSIF DHVLVPEHRV LSEEEKKALL EKYKITLAQL PQIKASDPAV
 KALGAKPGDV IEIKRKSPTA GVYYYYRVVV ED

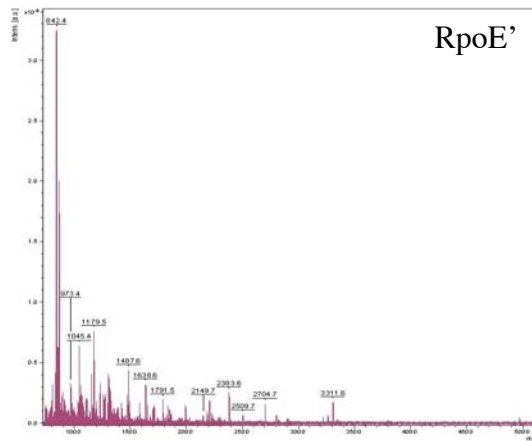
i



j

MRGPTVVDVT	PDDLWIVMES	YWQEKGLVRQ	HLDSYNAFIE	RGLQEVVNEF
GGIKPDIPDF	EVKFGKIRVG	EPEFQEPHGQ	RKPLYPMDAR	IRNLTYAAPL
YLEMIPVIKG	VEQDPVEVRI	GELPIMLKSK	VCRLYGLSDE	ELIKLGEDPK
DPGGYFIING	SERVIVSIED	IAPNRTLVEK	DERQEKYIAK	VFSYRHGYRA
LVAVERRKDG	ILYVDIPNVP	KPVKFVYVMR	ALGLERDKDI	VEAVGNDPEV
QQIMFDNLED	ASDITTQQA	LEFIGKLVAP	GQAREYRLKR	AEYVIDNNLL
PHMGVSPEDR	IKKAYYLGM	ALKVIELSLG	RRPEDDKDHY	ANKRLRLAGD
LLKDLFRVAF	SQLVKDIQYQ	MTKTYQRKGD	KYTFGNIHRF	IRNSIRPDVL
TERIEHALAT	GAWPGGRTGV	SQLLDRTNYM	STLSHLR RVT	SPLD REQPHF
EARDLHGTHW	GRICPTETPE	GPNCGLVKNL	ALMAQITTAV	PEEEIREYLM
SLGIVPIEVR	RPDPTLWRVY	LNGVLIGTIE	DGKALVQRIR	QDRRSGKISD
VINVAYYEDV	KEYVNSDDG	RVRRPLIIVE	NGVPKLTKEH	VQAVKEGRLK
WSDLIK MGII	EYLDAAAAA	AYVAMWPWEV	TEEHTHLELM	PAAILGLPAS
LVPYPEHNAA	PRNTYGAGMA	KQSLGLGWAN	FRIRVDTRGH	LLHYPQIPLV
NSRIMKAVGY	EDRPAGQNFV	VAVLAYQGYN	MEDAIINKA	SIERGLARST
FFRTYEAEEK	KYLGQTDKF	EIPDPTVRGY	RGEKYRNL	EDGLIFPESK
VEGKDVLVGR	TSPPR FLEER	GLGTVSLQER	RETSVAVRPS	EKGVDKVII
TETGDGTKLV	KVTVRDLR IP	ELGDKFASRH	GQKGVIGLIV	PQEDMPWTES
GIVPDLIVNP	HGIPSRMTVG	QLIEAIGGKV	ASLKGRVDG	TAFIGEPEEK
LRK ELEELGF	KHTGREIMYD	GITGKRLEAD	IFIGVIYYQR	LHMHVADK IH
ARSRGPVQVL	TKQPTTEGRAR	EGGLRFGEME	RDVLVGHGAA	MLLIERLLEE
SDKTEVWVCE	NCGHIALEDK	RRRRVYCPVC	GEEERISKVE	MSYAFKLLLD
ELKAMVIRPK	LNLSE			

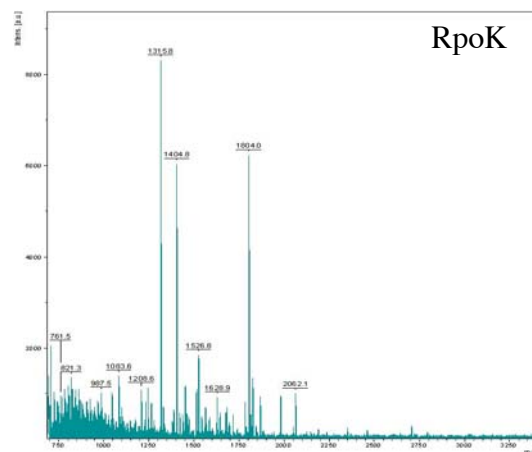
k



l

MYKIIVTVKDV VR**IPPTMFTM** **DPKE**AAKIIL RETYEGTYDK DEGVILSILE
 VKDIKDGIII PGDGATYHEV VFDVLVWEPK IHEVVEGYVA DVMPFGAFIR
 IGPIDGLVHI SQLMDDYVVY DERNK**QFVGK** **EKKYLLKIGD** LVRAR**IINIS**
AKSKVIRENR IGLTMRQPGL GKFEWIEKEK KKEKEEGKK

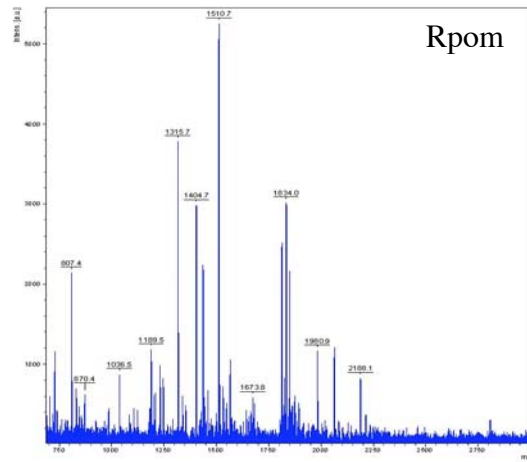
m



n

MFKYTRFEKA RIIGARALQI SMGAPVLIDV PPGITPLEAA ILEFEKGVIP
 ITVIRPS

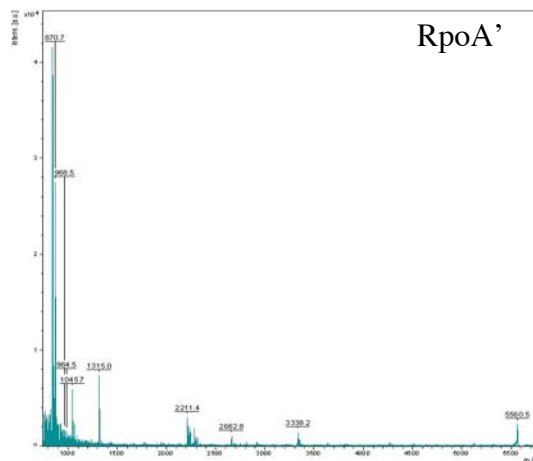
o



p

MVKFCPKCGS IMIPDRRRGV FVCRKCGYEE PINPEDTKAY RRTEEVKHRP
 DEGVVVIEQE VSTLPTAKVT CPKCGHNEAW WWELQTRAGD EPSTIFYKCK
 KCGYVWRSYE

q



I

MKKVIGSIEF GILSPQEIRK MSAVEVTVPD TYDDDGYPPIE GGVMDKRMGV
 IDPGLRCETC GGRAGECPGH FGHIELARPV IHVGFAKTIY **RILESTCREC**
 GRIKLTDEEI EEYMKKIELA RNRRSEFNEI IKEIHKKAKE RMVCPHCGAP
 QYPIKF EKPT IYWEIRK**DEQ GNEYR**HRLMP TEVRDWLEKI PDKDLPLLGL
 DPEKSRPEWM VLTVLPVPPV TARPSITLET GIRAEDDLTH **KLVDIIRINN**
 RLKQNI EAGA PQLIIEDLWD LLQYHVTTYI NNEAPGVPPA **KHKSGRPLKT**
 LAQRLKGKEG RFRGNLSGKR VNFSARTVIS PDPMISINEV GVPVEVAMEL
 TVPEKVTEFN IERLR**KMVLN GPDKYPGANY VIDPEGR**RRR IMDSNKETLA
 NQLDIGWTVE RHLMDGDIVL FNRQPSLHRM **SIMAHVRVM PYRTFR**LNLA
 VCPPYNADFD GDEMNLHVPQ TEEAQAEARI LMEVQNHIIS PRYGGPIIGG
 IQDHISGGYL LTREGAYFTR DEVEQMLMFA GVDITELPEP DKYDENG NPL
 WSGKTIF SLL LPEDLTVWYR NKLCDEPERC EALEKLIEEK LMPDPEEVRK
 LAYDGFVYIQ NGKLLSGAID KKAYGREDGI ILDLIVREYG VERARQFLDQ
 VTKLTIWVIT HKGFTTGIDD EDLPEEARDR IREIIREAEE RVNK**LIEAYK**
RGELEPLPGK SLEDTLESLI MAVLAEARDN AGAVAEKYL G MDNHTVIMAK
 TGARGKILNI TQMAALLGQQ SIRGKRLYRG FRGRVLSHF PGDLGARAKG
 FVVNSYK SGL TPQEYFFHAM GGREGLVDTA VRTAQSGYMQ RRLINALQDL
 KVEYDGT VRS **PEGIIVQFKY GEDGVDP**MKS WRGKTVDVDR IIVRT**LTKMR**
 GNNKAKV

S

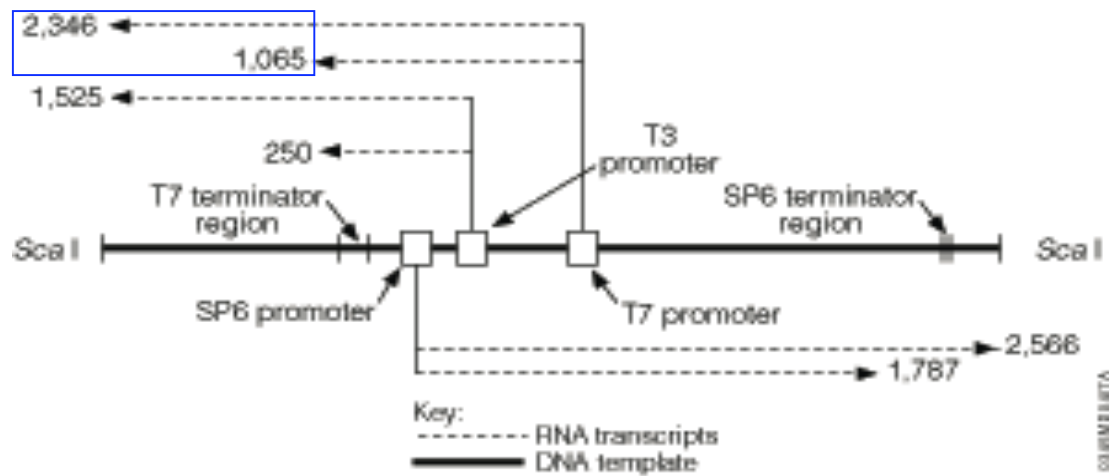
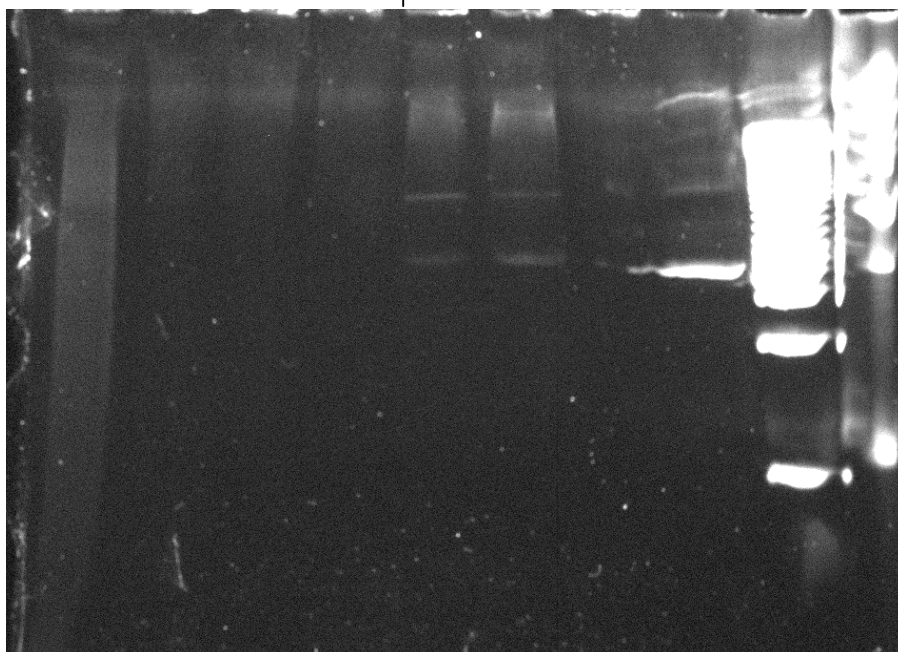


Figure 2.13: RNA transcripts generated using T7 RNAP and pGEM template (www.promega.com).

pGEM Template	+	-	-	+	+	+	-	+	MBH template
T7 RNAP	+	+	-	-	+	+	-	-	<i>Pf</i> RNAP
DIG labeling mix	+	+	+	+	+	+	-	-	DIG labeling mix
					70	55			Temperature (°C)



DNA 100bp ladder
RNA 0.1-2kB ladder

Figure 2.14a: RNA transcripts generated by in vitro transcription reaction using MBH template and reconstituted *Pf* RNAP.

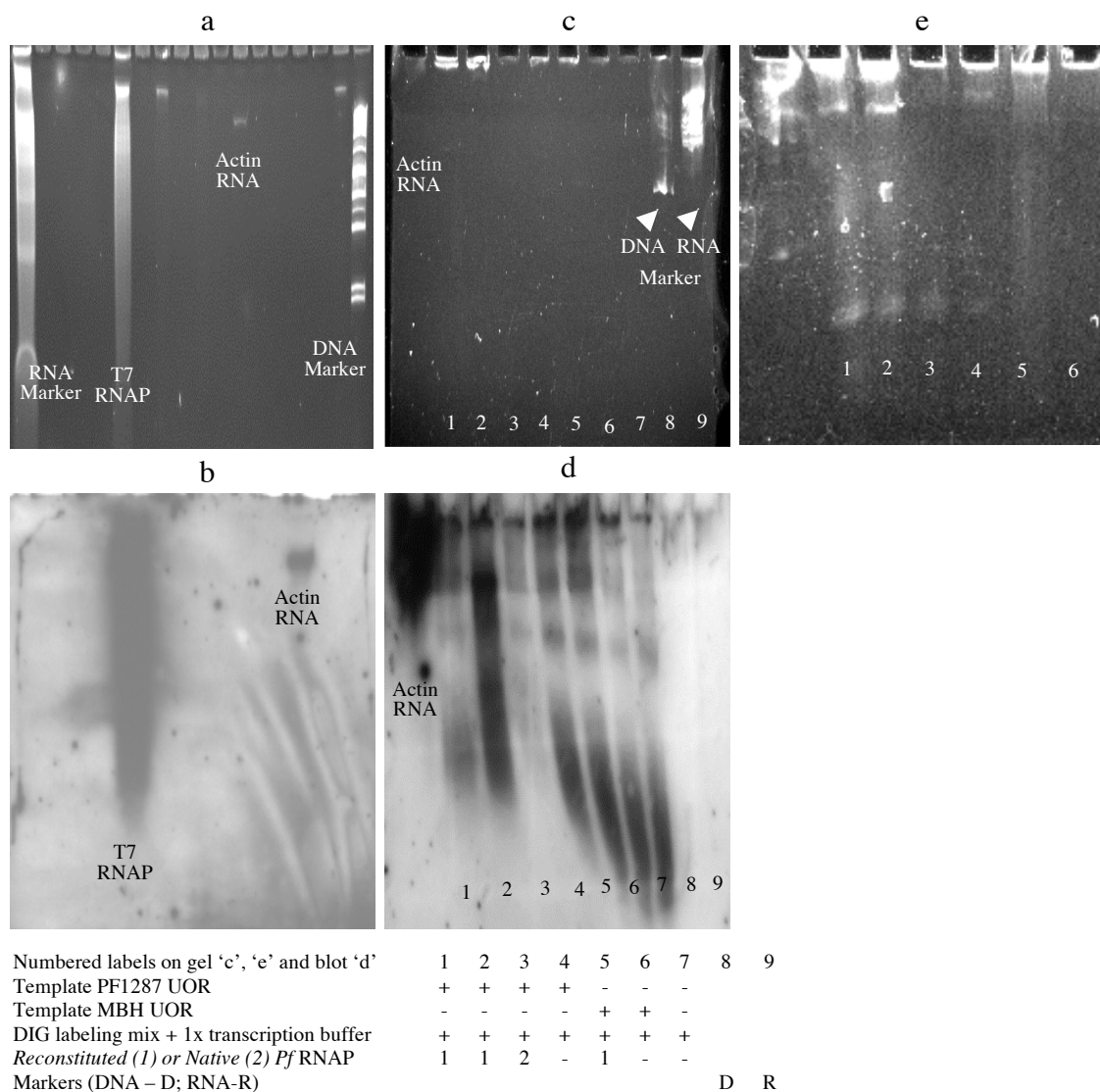


Figure 2.14b: 10% TBE-urea gel and Nylon membrane blot of RNA generated using T7 RNAP and reconstituted *Pf* RNAP.

Labels: a- Gel of DIG labeled RNA generated using T7 RNAP; b-Blot of gel a treated with anti-DIG AP; c-gel of DIG labeled RNA generated using reconstituted *Pf* RNAP after transfer onto the blot; d-blot of gel c treated with anti-DIG AP; e- gel of DIG labeled RNA generated using reconstituted *Pf* RNAP.

CHAPTER 3

IDENTIFICATION OF POTENTIAL TRANSCRIPTIONAL REGULATORS RESPONDING TO IRON DEPRIVATION IN *PYROCOCCUS FURIOSUS*

3.1 Objective: The presence of metals in hydrothermal vent fluids influences the physiology and ecology of many microbes including that of *Pyrococcus furiosus*. Microorganisms respond to excess or depleted levels of metal ions through several intrinsic mechanisms to regulate intracellular concentrations [139]. Of these, [Fe-S] clusters are involved in electron transport, and as substrate binding site for a wide range of redox and nonredox enzymes and act as sensors for various regulatory processes [44]. The tendency of a cell to maintain Fe homeostasis by regulating transcription of key transport proteins is a prominent mechanism to be investigated in *Pf* cells. Iron is the most difficult of all the metals to investigate experimentally because it is so ubiquitous as a trace contaminant. The goal is to conduct metal-limited (batch) and metal-shock (kinetic) experiments, such that the regulatory transcription factors could be identified and the mechanism of iron regulation in *Pf* cells can be elucidated. For iron, this involved adding a chelator like bathophenanthroline disulphonate (BPS). The effect of Fe limitation on batch cultures of *Pf* was investigated using DNA microarray expression profiles by A. L. Menon in the laboratory of M. W. W. Adams, UGA. The regulation of genes differentially expressed in standard *Pf* growth medium (Fe, 7.4 μ M), Fe-limited medium (Fe < 0.8 μ M), and Fe-depleted

medium (Fe ~0.5 μ M) were analyzed by microarray. Among 34 ORFs that were up-regulated > 2.5-fold and 61 ORFs that were down-regulated > 2.5-fold by Fe limitation, PF1285 (up-regulated, 5.0-fold) and PF1286 (up-regulated, 6.1-fold), the putative Fe-S cluster assembly/repair proteins homologous to sufD and sufB, respectively [44, 118], and PF1282 (down-regulated, 5.5-fold) encoded as rubredoxin (Rd) were selected for the study to identify regulatory transcription factor(s) responding to iron limitation as a stress response. The SUF system in *E. coli* is relatively well characterized and the existence of sequence homology of *suf* genes (sufD, sufB, sufC) in *E. coli*, other bacterial organisms and in *Pf* (PF1285, PF1286, PF1287), indicate that the occurrence of a similar regulatory mechanism could exist in *Pf*. Mutations in *suf* genes in *E. coli* proved to be lethal and microarray profiling reveals that *suf* transcription is up-regulated as part of the *E. coli* response to iron starvation, which emphasizes the importance of the Suf proteins [54]. An approach developed in the Scott group called STRES (Survey of Transcriptional Response to Environmental Stress), which uses techniques like microarray expression profiling, bioinformatic analysis, DNA-affinity protein capture, electromobility shift assays, footprinting, SELEX, and other biophysical and analytical methods was applied to identify and characterize these regulatory transcriptional factors [119].

3.2 Materials

3.2.1 *Cells, bio-reagents and chemicals*

3.2.1.1 BL21* (DE3) chemically competent *E. coli* cells (Cat. No. C6010-03, Invitrogen)

- 3.2.1.2** α -cyano-4-hydroxycinnamic acid (Matrix for MALDI), Cat. No. C2020-10G, Sigma-Aldrich, saturated solution in 50% acetonitrile, 0.1% TFA
- 3.2.1.3** 1x TAE buffer: 20 mM Tris (pH=7.8), 10 mM sodium acetate, 0.5 mM EDTA
- 3.2.1.4** 1x TBE buffer: 50 mM Tris (pH=8.3), 50 mM sodium borate, 0.5 mM EDTA
- 3.2.1.5** 1x Tris/Glycine/SDS buffer: 25 mM Tris, 192 mM glycine, 0.1% (w/v) SDS (sodium dodecyl sulfate)
- 3.2.1.6** 5x EMSA buffer: 100 μ L of 1 M HEPES, 100 μ L of 5 M NaCl, 250 μ L of glycerol, 5 mM DTT were added to a final volume of 1 mL, aliquots of 100 μ L in separate tubes were stored at -80°C
- 3.2.1.7** Buffers used for immobilization and protein capture experiment: (i) 2x Binding and Washing (B&W) buffer: 10 mM Tris, 2.0 M NaCl, (pH=7.0), 1 mM EDTA; (ii) Heparin buffer: 0.1 M sodium heparin, 20 mM HEPES, 150 mM NaCl (pH 7.55), 5% glycerol, 1 mM DTT, 0.1% Triton X-100; (iii) 1x incubation buffer: 20 mM HEPES, 150 mM NaCl, (pH 7.55), 5% glycerol, 1 mM DTT, 0.1% Triton X-100
- 3.2.1.8** Acetonitrile (100% v/v)
- 3.2.1.9** Agarose (1% w/v) in water
- 3.2.1.10** Ammonium bicarbonate (NH_4HCO_3), J. T. Baker, (25 mM)
- 3.2.1.11** Antibiotic: Kanamycin sulfate, Cat. No. K400-5G, Sigma-Aldrich, (50 $\mu\text{g/mL}$)
- 3.2.1.12** Binding buffer: 20 mM HEPES ((N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]) potassium salt Cat. No. H-0527, Sigma-Aldrich), 150 mM NaCl, J. T. Baker

- 3.2.1.13** Elution buffer: 20 mM HEPES ((N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulfonic acid]) potassium salt Cat. No. H-0527, Sigma-Aldrich), 150 mM NaCl (J. T. Baker), 500 mM Imidazole
- 3.2.1.14** Ethanol (70%, 100% v/v in water)
- 3.2.1.15** Gel Stain: (i) Protein - Coomassie blue stain (0.15% w/v coomassie brilliant blue G250, 17% w/v ammonium sulfate, 34% v/v methanol, 3% v/v phosphoric acid; Neutral silver stain (90 min incubation with 30% ethanol, 10% acetic acid, 10 min incubation with 20% ethanol, 10 min incubation with water, 1 min incubation with 0.2% sodium thiosulfate, 30 min incubation with 0.25% silver nitrate, 3x wash with nanopure water, 5-10 min incubation and gentle rotation with 6% sodium carbonate, 0.004% sodium thiosulfate, 37% formaldehyde (10 μ L per 100 mL of the mix); (ii) DNA – SYBR green I (1:10,000 dilution in 1x TBE buffer, Cat. No. BP1755-10, Invitrogen); (iii) Ethidium bromide: 0.5 μ g/mL in water.
- 3.2.1.16** Isopropyl-beta-D-thiogalactopyranoside (IPTG) Cat. No. BP1755-10, Fisher, (0.4 mM)
- 3.2.1.17** LB agar plates: 2.5% LB broth, 1.5% Bacto agar, autoclave for 30 min, add antibiotic, and pour into 10 cm plates, store at 4°C until required
- 3.2.1.18** Lysis buffer: 20 mM HEPES, 150 mM NaCl, 0.1 mM PMSF (polymethylsulfonyl fluoride
- 3.2.1.19** Media (Culture growth): *E. coli* - LB (Luria-Bertani) Broth, Miller, Cat. No. B1426-2, Fisher Chemicals, (2.5% (w/v)); *Pyrococcus furiosus* - (Standard defined medium) 75 g maltose, 7.5 g yeast, 1x salts (420 g NaCl, 52.50 g MgSO₄, 40.50 g MgCl₂, 4.95 g KCl, 3.75 g NH₄Cl, 2.10 g CaCl₂), 1x trace minerals (0.03

g FeCl₃, 0.0008 g H₃BO₃, 0.0008 g ZnCl₂, 0.0005 g CuCl₂, 0.0008 g MnCl₂, 0.0008 g (NH₄)₆Mo₇O₂₄, 0.0008 g AlK(SO₄)₂, 0.0008 g CoCl₂, 0.0008 g NiCl₂, 0.0495 g Na₂WO₄, 7.5 g Cysteine HCl, 7.5 g Na₂S, 15 mL of 1 M potassium phosphate buffer (pH 6.8).

- 3.2.1.20** Molecular weight markers: Protein - Precision plus protein standard, Cat. No. 161-0373, Biorad; HMW calibration kit for native electrophoresis, Cat. No. 17-0445-01, Amersham Biosciences-GE; DNA – 1 kb, 100 bp, 25 bp step ladder (Promega)
- 3.2.1.21** NTP Mix: dNTP - 2.5 mM dATP, 2.5 mM dGTP, 2.5 mM dCTP, 2.5 mM dTTP
- 3.2.1.22** *Pyrococcus furiosus* genomic DNA (S. Brehm, M. W. W. Adam's Group)
- 3.2.1.23** Qiagen plasmid maxi kit, Cat. No. 12163, Qiagen
- 3.2.1.24** QIAquick PCR purification kit, Cat. No. 28106, Qiagen
- 3.2.1.25** Sample loading buffer: Protein – 4x Laemmli buffer (160 mM Tris-HCl (pH 6.8), 50% (v/v) glycerol, 8% (w/v) SDS, 0.1% (w/v) bromophenol blue); DNA – 6x loading dye (xylene cyanol, bromophenol blue, xylene orange), Promega
- 3.2.1.26** S. N. A. P. gel purification kit, Cat. No. 45-0078, Invitrogen
- 3.2.1.27** S. O. C. medium (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 20 mM glucose)
- 3.2.1.28** Sodium acetate, J. T. Baker, 3 M (NaOAc)
- 3.2.1.29** Trifluoroacetic acid (TFA), 0.1%, 0.5% v/v
- 3.2.1.30** Tris (Tris hydroxy methyl aminomethane), Fisher Biotech
- 3.2.1.31** Trypsin: 20 µg in 200 µL of trypsin dilution buffer, Promega
- 3.2.1.32** Turbo *Pf* polymerase (2.5 U/µL)

3.2.2 *Equipment and glassware*

- 3.2.2.1** AKTA prime plus FPLC (Fast Protein Liquid Chromatography) system
(Amersham Biosciences-GE)
- 3.2.2.2** Centrifuge (Avanti J-E, Beckman Coulter) with JA-5.3 and JA-25.5 rotors
- 3.2.2.3** Amicon membrane filter (MWCO = 10,000, 30,000 Da), Millipore
- 3.2.2.4** DC protein assay kit (Biorad)
- 3.2.2.5** Electrophoresis system (DNA, Protein), Biorad
- 3.2.2.6** Membrane filter: Syringe filter (0.2 μ), Nitrocellulose (0.45 μ), Millipore
- 3.2.2.7** Fluorochem 8000 Digital Imager (Alpha Innotech Corp.)
- 3.2.2.8** Gels: 4-20% Tris gel, 10-20% Tris gel (Criterion, Biorad); 5% TBE (Ready gels, Biorad)
- 3.2.2.9** Hoefer DyNA Quant 200 fluorometer (Amersham Pharmacia Biotech)
- 3.2.2.10** MALDI ToF (Matrix Assisted Laser Desorption Ionization Time of Flight) mass spectrometer, Autoflex, Bruker
- 3.2.2.11** Nickel affinity column: HistrapFF (1 mL), Amersham Biosciences-GE
- 3.2.2.12** NuTip C-18 tips (Glygen Corp.)
- 3.2.2.13** PCR Thermocycler
- 3.2.2.14** Sterile test tubes (15 mL), Erlenmeyer flask (250 mL), Fernbach flask (2.8 L)
- 3.2.2.15** Shaker-Incubator
- 3.2.2.16** Sonic dismembranator (Fisher)
- 3.2.2.17** Waterbath
- 3.2.2.18** Vacuum evaporator

3.3 **Methods**

3.3.1

Growth of *Pyrococcus furiosus* cells: The effect of Fe limitation on batch cultures of *Pf* cells was investigated using DNA microarray (unpublished, A. L. Menon, M. W. W. Adam's group, UGA). *Pyrococcus furiosus* DSM 3638 was grown in the presence and absence of iron with maltose as the primary carbon source. In the standard defined growth medium, as previously reported [131], Fe is added to a final concentration of 7.4 μM as $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$. Fe-limited medium ($\text{Fe} < 0.8 \mu\text{M}$) was obtained by omitting Fe in the growth medium and care was taken to eliminate iron as a contaminant by using acid-washed glassware and nanopure water for preparation of all the buffers and the growth medium. Cells were grown at 95°C in the absence of S° . Additionally, growth experiments were conducted to limit iron in the growth medium by the incorporation of bathophenanthroline bisulphonate (BPS), a chelator of residual iron. In a separate experiment, *Pf* cultures were grown in standard medium with the elimination of iron in the media, and incorporation of BPS until they reached mid-log phase (0.8×10^8 cells/mL) and Fe was added to a final concentration of 7.4 μM . This was done to monitor the number of iron-regulated genes, which are adaptive to the presence of iron in the medium. All growth experiments were carried out in a 20 L custom fermenter. At various time points, the cell growth was monitored and harvested when the cell count reached $\sim 10^8$ cells/mL. The cell culture was pumped from the 20 L fermenter through a rapidly cooled glass coil and the cell pellet was collected by centrifugation (10,000 $\times g$, 15 min) [140]. The cells were frozen at -80°C until further analysis. Samples were removed at various time points during the growth for RNA extraction and analysed by DNA microarray

for the following four conditions (A. L. Menon, M. W. W. Adam's group) (Table 3.1a, 3.1b; Figure 3.1).

1. Standard defined medium (Fe = 7.4 μ M)
2. Fe-limited medium (Fe < 0.8 μ M)
3. Fe-limited medium + BPS (chelates the residual Fe)
4. Fe-limited medium + BPS + Fe added back (Fe = 7.4 μ M)

The results obtained suggest that about 34 ORFs were up-regulated > 2.5-fold with limited iron. This includes ORFs encoding the putative Fe-S cluster assembly/repair proteins PF1285 (5.1-fold) and PF1286 (6-fold), which are homologous to sufD and sufB, respectively. Conversely, about 61 ORFs were down-regulated by more than 2.5-fold by Fe limitation. This includes an ORF encoding abundant iron-containing protein rubredoxin, PF1282 (5.5-fold).

3.3.1.1 *Preparation of Pyrococcus furiosus cell extract:* Frozen cell pellet (2 g, wet weight) was thawed with 6 mL of 20 mM HEPES, 150 mM NaCl, 1 mM DTT for 15 min at room temperature. The cell extract is then sonicated for 10 min and centrifuged at 50,000 xg for 60 min at 4°C using JA-25.5 rotor. The supernatant was collected and aliquots of 0.5 mL were transferred into separate tubes pre-flushed with argon under anaerobic conditions and secured with stoppers. The tubes were stored at –80°C until required. The protein amount in the cell extract was estimated to be 6 mg/mL by Biorad DC assay.

3.3.2 *Selection of UOR DNA probe for DNA-affinity protein capture:* Promoter identified for rubredoxin (PF1282, down-regulated under iron-limiting growth conditions), is upstream of PF1283 ORF. PF1281 (superoxide reductase),

PF1282 (rubredoxin) and PF1283 (ruberythrin) are predicted to occur as an operon (www.tigr.org). Similarly, the promoter identified for the hypothetical protein encoded by PF1285 (sufD homolog, up-regulated under iron limiting growth conditions of *Pf*), hypothetical protein encoded by PF1286 (sufB homolog, up-regulated under iron limiting growth conditions of *Pf*) is upstream of the PF1287 ORF (Figure 3.2). PF1285 (sufD-homolog), PF1286 (sufB-homolog), PF1287 (sufC-homolog) occur as a single operon. In the STRES approach, we assume that regulatory transcription factors (rTFs) have binding sites that lie in the upstream region of the ORF, referred to as UOR (Upstream of ORF (Open ReadFraMe) Region) assumed to contain the promoter. For the DNA-protein capture experiment, the promoter region selected was a few hundred base pairs upstream of the translation start site of the gene (PF1287 and PF1283).

3.3.2.1 *Primer design and PCR amplification of PF1283 UOR and PF1287 UOR DNA probes:* Primers were designed with biotin-TEG tag (IDT) at the 5' end of the UOR DNA probes. The forward primer sequence for PF1283 UOR was 5'-biotinTEG-AAGACTCTCTATGCAAGGCATG-3' and the reverse primer was 5'-biotinTEG-GCTTTCGCCTGCAAAGGCTTC-3'. The forward primer sequence for PF1287 UOR DNA was 5' biotinTEG-GACAAGGAAAGATGACTGAATT 3' and the reverse primer, 5' biotinTEG-CTTAATGCCCTTGACTATTTCTTTATC 3'. The TEG (Triethylene glycol, non-reactive spacer arm) attached to the biotin probably provides an efficient mechanism to capture the proteins that adhere to the promoter compared to that

with a biotin tag alone because it alleviates the problem of steric hindrance during the binding process (personal communication, Integrated DNA Technology).

3.3.2.1.1 Gradient PCR was carried out to optimize the annealing temperature for the pair of primers such that the yield and the purity of PCR product were optimal. The annealing temperature range tested for PF1283 UOR DNA was 40°C - 60°C. The annealing temperature range tested for PF1287 UOR DNA was 35°C - 55°C. The biotin-TEG tagged UOR DNA probes were amplified by PCR. The PCR reaction mixture (100 μ L) included 1 μ M of forward primer, 1 μ M of reverse primer, 100 ng *Pf* genomic DNA, 2.5 mM NTP, 10 μ L 10x cloned *Pf* buffer, and 2.5 U of turbo *Pf* polymerase [134]. The PCR conditions were set up as follows: 94°C for 3 min, followed by 30 cycles of 1 min at 94°C (denaturation), 1 min at 42°C (annealing temperature for PF1283 UOR) or 1 min at 55°C (annealing temperature for PF1287 UOR), and 1 min at 72°C (extension) per cycle, with a final extension of 20 min at 72°C [134]. The UOR DNA probes were purified using the QIAquick PCR purification kit. The purified DNA was loaded on a 1% agarose gel and the gel was run under native running conditions using 1x TAE buffer, at a constant voltage of 80 V for 45 min. The sample was mixed with 6x loading dye in 5:1 proportion prior to loading on the gel. The gel was stained with 0.5 μ g/mL of ethidium bromide for 10 min and the image was recorded using a Fluorchem 8000 imaging system. The amount of purified UOR DNA probe was estimated using a Hoefer DyNA Quant 200 fluorometer, expressed in terms of ng/ μ L.

3.3.2.2

Immobilization of the UOR DNA probe using Dynabeads: Magnetic capture and separation of proteins from the *Pf* cell extract bound to the UOR DNA probe is carried out using the M-280 streptavidin-coated Dynabeads. These beads are uniform superparamagnetic, polystyrene beads with a monolayer of streptavidin (MW 66 kDa, four subunits of 16 kDa each) covalently attached to the hydrophobic bead surface. The beads have a diameter of 2.8 μm , specific surface area of 4-8 m^2/g , and density of 1.4 g/cm^3 (www.invitrogen.com). Capture of proteins on the UOR DNA probe was carried out by the indirect approach (Figure 3.3). In this approach, proteins bound to the biotin-TEG tagged DNA probe were incubated with the Dynabeads at room temperature, thereby preventing the dissociation of the biotin-streptavidin bond at higher temperatures [119]. Moreover, this approach is useful when the concentration of the target protein (rTF) is low (www.invitrogen.com).

In a sterile microcentrifuge tube, 100 μL suspension (1 mg containing $\sim 6 \times 10^7$ Dynabeads) of homogenous suspension of the streptavidin-coated Dynabeads was added. The suspension was washed three times with gentle vortex for 10 min (two times with 100 μL of 2x B&W buffer, once with 100 μL of 1x B&W buffer) to remove the preservative NaN_3 in which the beads were suspended. The tubes were placed on the magnetic stand for 1 min. The supernatant was removed using a pipette, special care was taken not to touch the inside wall of the tube where the beads were attracted to the magnet. The ratio of optimal binding of UOR DNA to the streptavidin-coated Dynabeads is dependent on the length and the amount of biotin-TEG tagged UOR DNA (400 pmol of 300

bp biotinylated oligonucleotide typically binds to 1 mg of Dynabeads (www.invitrogen.com). Accurately calculated amount of purified biotin-TEG tagged UOR DNA was used for immobilization on the Dynabead.

For the DNA-affinity protein capture experiment, in a separate microcentrifuge tube, 100 μ L (5.0 mg/mL) of *Pf* cell extract (control: maltose medium) and 350 ng (5.92 nmol) biotinTEG-UOR DNA were added. The mixture was incubated at 55°C for 30 min using a tabletop heater, cooled to room temperature and the condensate was collected.

The next step was to immobilize the biotinTEG-UOR DNA (with the proteins bound) on the pre-washed Dynabeads. The UOR DNA with protein mixture was added to the pre-washed Dynabeads, and incubated with gentle vortex for 10 min at room temperature (RT). The tubes were placed on a magnetic stand and the supernatant (containing the unbound proteins) was removed using a pipette. The beads were treated with 100 μ L of 0.1 M heparin solution, a non-specific DNA competitor. The tubes were placed on the magnetic stand and the supernatant was removed. The contents in the tubes were washed twice with 100 μ L of 1x incubation buffer to remove the proteins with non-specific interactions [119].

Proteins (including probable rTFs) bound to the UOR DNA immobilized on the Dynabeads, were extracted with the addition of 40 μ L of 4x Laemmli buffer, vortexed for 10 min at RT, and incubated at 55°C for 10 min. The tubes were placed on the magnetic stand and the supernatant was collected. To ensure that all the bound proteins were dissociated from the magnetic Dynabeads, a

second extraction was carried out with the addition of 40 μL of 4x Laemmli buffer, and incubated at 95°C for 10 min. The supernatant was collected.

DNA-protein capture experiments for both the bioTEG-PF1283 UOR DNA and bioTEG-PF1287 UOR DNA were performed separately. These experiments were simultaneously conducted and compared for three different growth conditions of *Pf* cells (standard defined medium (Fe = 7.4 μM), Fe-limited medium (Fe < 0.8 μM), and Fe-limited medium + BPS (bathophenanthroline bisulphonate).

3.3.3 *Identification and characterization of prospective transcriptional regulator:* The proteins in the supernatant were separated using SDS-PAGE on a 4-20% Tris gel in 1x Tris/glycine/SDS buffer. The gels were run at a constant voltage of 200 V for 60 min, stained using silver stain and the image recorded (Figure 3.4a, 3.4b).

The identified gel bands (corresponding to putative regulatory transcriptional factor (rTF)) were excised. The gel slices were transferred to separate labeled microcentrifuge tubes, destained with 30 μL 1:1 mix of 30 mM potassium ferricyanide and 100 mM sodium thiosulfate. The gel slices were rinsed quickly three times with nanopure water to remove the residual potassium ferricyanide and sodium thiosulfate and dehydrated using a vacuum evaporator [141].

Trypsin digestion was carried out according to standard protocol [142]. The gel slices were soaked in 12 μL of 10 ng/ μL of trypsin and incubated for 16 h at 37°C. Trypsin cleaves the protein after lysine (K) and arginine (R). Tryptic peptides from the gel slices were extracted in four 10-min steps (1x with 15 μL of

25 mM NH_4HCO_3 , and 2x with 15 μL of 50% acetonitrile, 25 mM NH_4HCO_3 and 1x with 15 μL of 75% acetonitrile, 0.1% trifluoroacetic acid (TFA)). The supernatant from each extraction was collected and combined. The solution was concentrated using a vacuum evaporator to a final volume of 4 μL . Finally, 1 μL of 0.5% TFA in water was added [129]. The extracted tryptic peptides were desalted and concentrated further using NuTip C-18 tips (Glygen Inc.). The peptide mixture was eluted from the NuTip with 3 μL of freshly prepared saturated matrix solution (α -cyano-4-hydroxycinnamic acid in 50% acetonitrile, 0.1% TFA), and spotted directly onto a MALDI plate. The samples were analyzed by MALDI-ToF mass spectrometry.

The mass spectrum was acquired over a range of m/z 800 to 5000 and calibrated using trypsin autolysis products as internal standards. Data analysis was performed using Protein Prospector v. 3.2.1 using MS-Fit program [130]. Among the proteins bound to PF1283 UOR DNA and PF1287 UOR DNA (section 3.3), one potential rTF bound to PF1283 UOR DNA and PF1287 UOR DNA was identified as PF0851p.

3.3.3.1 *Bioinformatics analysis of identified UOR DNA bound protein:* The protein sequence of PF0851 (iron-dependent repressor) was aligned with other homologous protein sequences obtained from NCBI non-redundant database using the web-based NCBI BLASTP and the conserved domain database (CDD) (www.ncbi.nlm.nih.gov). The results indicate that PF0851p, a homolog of DtxR (Fe-dependent repressor in *Corynebacterium diphtheriae*), and TroR (Mn-dependent repressor in *Bacteria subtilis*) has a N-terminal Fe-dependent

repressor-binding domain (Figure 3.5). The secondary structure prediction using PSIPRED view indicates the presence of helix-turn-helix DNA binding domain and the sequence is conserved in many metal-dependent repressors [143] (Figure 3.6, 3.7, 3.8). MEME was used to determine the occurrence of identified motif (binding sites) in other *Pf* UOR based on the conserved Fur box and tox operator sequence well characterized in other organisms. The list of UOR with the motif was compiled based on the iron regulation obtained from microarray (program link developed by D. Cowart, Scott's group, UGA) (Table 3.2a, 3.2b).

3.3.3.2 *Plasmid purification of the putative transcription factor:* The gene encoding PF0851 was cloned into pET24dBam vector with a His₆-tag and the cells containing the plasmid were obtained from (Francis Jenney in M. W. W. Adam's group). Step cultures of 3 mL and 100 mL were grown for 16 h at 37°C with shaking at 225 rpm in LB media containing 50 µg/mL kanamycin and then transferred to a 2.8 L Fernbach flasks (1L LB media containing 50 µg/mL of kanamycin), and incubated at 37°C for 16 h. The cells were harvested using a centrifuge with JA-5.3 rotor at 3,000 xg for 20 min. The cell pellet was collected, lysed and the plasmid was purified using Qiagen plasmid purification kit. The plasmid was collected in 5 mL eluting buffer and was further purified by ethanol precipitation with the addition of 1/10th volume of 3 M sodium acetate (500 µL), mixed well, followed by addition of 2.5x cold 100% ethanol. The resulting solution was centrifuged at 20,000 xg for 15 min using a JA-25.5 rotor. The supernatant was carefully decanted without disturbing the precipitated plasmid DNA at the bottom of the tube. Then, 1 mL of cold 70% ethanol was added,

mixed well and centrifuged at 20,000 xg for 20 min. The supernatant was again carefully decanted or removed using a pipette without disturbing the precipitate at the bottom of the tube. The tube was allowed to air dry for 20 to 30 min until all the residual ethanol evaporates. The purified plasmid was resuspended in 100 μ L of 10 mM Tris buffer (pH 8.0). The sequence of the gene was verified (IBL, UGA) (Figure 3.9).

3.3.3.3 *Protein expression, purification, and characterization:* The purified plasmid was transformed into the expression strain of *E. coli* BL21(DE3). Step cultures of 3 mL and 100 mL of 2.5% LB media containing 50 μ g/mL kanamycin were grown for 16 h at 37°C with shaking at 225 rpm. The culture was transferred to a 2.8 L fernbach flask (1 L, 2.5% LB media, 50 μ g/mL of kanamycin) and incubated at 37°C for 2-3.5 h with shaking at 225 rpm until $OD_{600} = 0.75$, before induction with 0.4 mM isopropyl-beta-D-thiogalactopyranoside (IPTG). The culture was incubated, post-induction for 16 h at 17°C, and the cells were harvested by centrifugation at 3,000 xg for 20 min using a JA-5.3 rotor, the cell pellet was collected and stored at -20°C until further purification. Frozen cell pellet (2-5 g, wet weight) was thawed and resuspended in 25 mL of lysis buffer. The cells were disrupted using a sonicator in the continuous mode. The amplitude was gradually increased from 0 to 9, and continuously sonicated with the power set to 20% for 1 min. The amplitude knob was turned back to zero for 30 sec, and the cycle was repeated for 10 min. The cell suspension was centrifuged at 35,000 xg for 30 min. The clear supernatant was collected and filtered through a 0.2-micron syringe filter.

- 3.3.3.4** *Nickel-affinity protein purification using HisTrapFF column:* The hisTrapFF column (1 mL, pre-loaded with nickel) was connected to the AKTA prime FPLC, washed with 5 mL of filtered and degassed 20% ethanol, 5 mL of filtered and degassed nanopure water, and then equilibrated with the binding buffer (20 mM HEPES, 150 mM NaCl). The flow rate was set to 2.0 mL/min, and the UV detector was set to 0.1 AUFS (absorbance unit full scale), 280 nm. The clarified supernatant (~25 mL) containing the recombinant protein was loaded onto a 1 mL column, washed with the 5 mL of binding buffer and the His₆-tag protein was eluted with a linear gradient of buffer containing 100-500 mM imidazole. The fraction (15 mL) was collected and concentrated to 250 μ L using Amicon (MWCO: 10,000) centrifugal membrane filter unit. Protein estimation was performed using Biorad DC assay.
- 3.3.3.5** *Desalting of the purified protein:* The protein concentrate from the previous section was injected into a 5 mL G-25 sepharose column to remove the imidazole. The column was equilibrated with binding buffer (20 mM HEPES, 150 mM NaCl) and the flow-through collected contained the protein without the imidazole. The protein amount was estimated using Biorad DC assay.
- 3.3.3.6** *SDS-PAGE analysis of purified protein:* The protein was separated on a 4-20% Tris gel in 1x Tris/glycine/SDS buffer at 200 V for 60 min. Samples were heated at 95°C for 5 min in 4x Laemmli buffer prior to loading. The gels were stained using coomassie blue stain and the image was recorded.
- 3.3.4** *Determination of binding potential of PF0851p by EMSA (Electromobility Shift Assay):* This method was used to validate sequence-specific DNA binding

proteins. The protein-DNA complex migrates through a nondenaturing polyacrylamide gel more slowly than free DNA fragments. With increasing protein/DNA ratio, the DNA shifts and at an optimal protein concentration, no free DNA remains [144]. The binding potential of PF0851p using PF1283 and PF1287 UOR probe DNA was determined using EMSA (also known as gel shift).

For the gel shift assay, the reaction mixture contained ~30 ng of PF1283 UOR DNA, 2 μ L of 5x EMSA buffer, and the appropriate amount of His₆-tag PF0851p was added to achieve protein/DNA ratio from 0 to 30 in separate tubes (Table 3.3a, 3.3b). Water was added to a final volume of 10 μ L. The tubes were incubated at 55°C for 20 min. The resulting mixture was loaded directly (without a loading dye) on a 5% TBE gel in 1x TBE buffer and was run at 200 V for 25 min. The gel was stained using SYBR green I stain and the image was recorded. A negative control without protein was also subjected to a similar treatment and loaded on the same gel. Similarly, the gel shift assay with PF1287 UOR DNA was performed (Figure 3.10).

3.3.5 *Determination of binding site by nonradioactive DNase I footprinting assay:*

Footprinting enables identification of particular base sequence within a given DNA sequence to which the protein is bound (binding site). In this protection assay, the protein bound to the DNA sequence shields the DNA from enzymatic (DNase I) or chemical modifying agents (OH). This facilitates the identification of protein interaction with specific bases in a DNA sequence [145, 146]. DNase I footprinting was performed using capillary electrophoresis (CE) to determine the specific binding site of PF0851p on PF1283 and PF1287 UOR DNA.

3.3.5.1 *Primer design:* Primers for PF1283 and PF1287 UOR probe DNA were designed with different fluorescent tags at the 5' and 3' ends. The 5' end of the forward primer was labeled with a 6FAM (6-carboxyfluorescein) fluorescent tag and the 5' end of the reverse primer was labeled with a HEX (hexachloro fluorescein) fluorescent tag. The forward and reverse primers for PF1283 UOR DNA were 5' 6FAM-CAAGGATATGAGAGTAAGCCCGGAATTC 3', 5' HEX-TATGGGCCATGCTTTTCGCCTGCAAAGGC 3' and that for PF1287 UOR DNA were 5' 6FAM-GGAATAGTTACGGGTGACTCCCTGGG 3', 5' HEX-GTTAATGCCCTTGAGTATTTCTTTATCTTCG 3' respectively. PCR amplification of the tagged PF1283 and PF1287 UOR DNA was performed following similar protocol to that described in section 3.2.2.1.

3.3.5.2 *Footprinting experiment:* The footprinting reaction mix consisted of calculated amount of PF1283 UOR DNA probe, 5X EMSA buffer and PF0851p (Table 3.4a-3.4b). The mixture was incubated at 55°C for 20 min, after which, 50 μ L of Ca^{2+} , Mg^{2+} solution (5 mM CaCl_2 , 10 mM MgCl_2) was added and incubated for 1 min at room temperature. DNA cleavage was performed with the addition of 3 μ L of DNase I (0.05 U/ μ L of RQ1 RNase-free DNase) followed by incubation for 5 min at room temperature. The reaction was stopped with the addition of 200 μ L of phenol: chloroform: isoamyl alcohol (25:24:1). The mixture was immediately vortexed. The clear supernatant (~100 μ L of aqueous layer) was extracted and transferred to a sterile microcentrifuge tube. DNA was precipitated with the addition of 1/10th volume of 3 M sodium acetate, followed by addition of 2.5x cold 100% ethanol and thorough mixing. The mixture was centrifuged at 17,000

xg for 10 min using a tabletop microcentrifuge. The supernatant was carefully decanted without disturbing the precipitated DNA pellet at the bottom of the tube and 1 mL of cold 70% ethanol was added, and centrifuged at 17,000 xg for 5 min. The supernatant was again carefully decanted. The tube was allowed to air dry for 30 min until all the residual ethanol evaporated. A master mix of high purity deionized formamide solution, and GeneScan-500-ROX size standard (Applied Biosystems) in a ratio 9.8:0.2 was freshly prepared. The precipitated DNA was resuspended in 10 μ L of the mixture prepared. The sample (10 μ L) was transferred onto a 96-well optical bar-coded plate (Applied Biosystems), and covered with a lint-free sheet. The samples were analyzed using an auto sequencer (IBL, UGA). A footprint of PF1283 and PF1287 UOR DNA obtained indicates that the binding site is close to the translation start site of both the ORFs (Figure 3.11a, 3.11b).

3.4 Results and Discussion

3.4.1 *DNA-affinity protein capture experiment:* The effect of iron limitation on batch cultures of *Pf* cells was investigated using DNA microarray expression profiling in M. W. W. Adam's group (Figure 3.1). From these experiments, about 34 ORFs were up-regulated > 2.5-fold and 61 ORFs were down-regulated > 2.5-fold by Fe limitation (Table 3.1a, 3.1b). Among these, an ORF that was down-regulated (5.0-fold), identified as PF1282, belongs to PF1281, PF1282, PF1283 operon and another set of up-regulated ORFs were PF1285 (5.0-fold, sufD), and PF1286 (6.1-fold, sufB) grouped as a single operon consisting of PF1285, PF1286, and PF1287 (Figure 1.13). Primers for DNA upstream of genes PF1283 and PF1287 (referred

to as Upstream of QRF Region, UOR and presumably containing the operon promoter and any operator sequence) were designed with a biotin-TEG tag on both 5' and 3' ends (Figure 3.2) for DNA-affinity protein capture experiments (Figure 3.3). The proteins bound to the UOR DNA were separated using SDS-PAGE. Several gel bands were excised and tryptic digests of the proteins were analyzed by MALDI mass spectrometry to identify the proteins. Using the MS-FIT database, the proteins bound to PF1283 UOR DNA were identified as PF0732 (annotated as daunorubicin resistance ATP-binding protein), PF1772 (annotated as 2-ketoacid: ferredoxin oxidoreductase subunit beta), PF1690 (annotated as 2-dehydro-3-deoxyphosphoheptonate aldolase), PF0905 (annotated as putative ABC transporter), PF1199 (annotated as ferritin homolog), and PF0851 (annotated as iron-dependent repressor) (Figure 3.4a). Similarly, the proteins bound to PF1287 UOR DNA were identified as PF1952 (annotated as iron-sulfur protein), PF0205 (annotated as glutamate synthase), PF0851 (the same iron-dependent repressor), and PF1794 (annotated as inosine-5'-monophosphate dehydrogenase related protein). Among these, the potential transcriptional regulator identified for both promoters was PF0851p (Figure 3.4b). Cells containing the plasmid of PF0851 were obtained (from Francis Jenney, M. W. W. Adam's group) (Figure 3.9), and PF0851 was expressed and purified by Ni-affinity chromatography.

3.4.2 DNA-protein interaction: To establish the binding potential of PF0851p on the selected UOR DNA (PF1283 and PF1287 promoters), electromobility shift assay (EMSA) experiments were performed. Formation of the DNA-protein complex

changed the mobility of DNA. With 0.343 μM (4.2 mole ratio over PF1283 UOR DNA concentration) of PF0851p added, no free PF1283 UOR DNA remained. Similarly, the PF0851p concentration required for complete complex formation with PF1287 UOR DNA was 0.685 μM (12 mole ratio over PF1287 UOR DNA concentration) (Table 3.3a, 3.3b; Figure 3.10). This tight binding affinity strongly suggests that PF0851p binds sequence-specifically to some sequence within the PF1283 and PF1287 UORs. DNase I footprinting using non-radioactive fluorescence labeling and CE sequencing was used to investigate the binding region within these UORs. In the footprinting experiment, the protein-DNA complex formed prevents DNase I digestion of part of the UOR DNA to smaller fragments. A forward primer with 6FAM fluorophore and a reverse primer with HEX fluorophore attached at 5' ends were designed and PF1283 and PF1287 UOR DNA were amplified by PCR (Table 3.4a, section 3.3.5.1). The protein/DNA ratio (Table 3.4b) was varied for the footprinting experiment. The footprint obtained with a protein/DNA ratio of 2.0 (Figure 3.11a, 3.11b) indicates that the binding site of PF0851p on PF1283 UOR DNA extends approximately from -30 to -5 bp and the binding site on PF1287 UOR DNA extends approximately from -28 to -5 bp with respect to the translation start site of the ORFs (Figure 3.12). Decreased fluorescence intensity was observed for the reverse strand (PF1283 and PF1287 UOR DNA) in the presence of the protein compared to control (without protein). The binding affinity of PF0851p on both PF1283 and PF1287 UOR DNA were strong as indicated by the EMSA experiment. The DNase I protected region (footprint) obtained for both the UORs

is relatively short approximately 25 bp long binding site. The probable binding site on PF1283 and PF1287 UOR indicated by the footprints also contain the DtxR box and fur box consensus sequences respectively. PF0851p is similar to the Fe-dependent repressor Fur (Chapter 1) of well characterized *suf* operon in bacterial organisms and Fur binds to tandem hexamer of 5'-NAT(A/T)AT-3' on the promoter [84] suggesting that the binding of PF0851p to a specific sequence and/or nucleotides could be validated by hydroxyl radical footprinting.

3.4.3 *Bioinformatic analysis:*

3.4.3.1 *PF0851p*: PF0851p is composed of 133 amino acid residues and the conserved domain database search indicates the presence of Fe-dependent repressor domain at the N-terminal end of PF0851p. The PF0851 gene is homologous to TroR, a Mn-dependent transcriptional regulator; the region of sequence homology extends from 11-133 amino acid residues. The MarR domain (homolog of IscR, a regulator of *suf* operon in *E. coli*) extends from 5-75 amino acid residues. PF0851p has a helix-turn-helix DtxR-like domain; the binding domain of DtxR (another iron-dependent repressor) and its interaction with the *tox* operator is well characterized in many bacterial genomes (Chapter 1). The secondary structure prediction using PSIPRED indicates the presence of helix-turn-helix DNA binding domain (Figure 3.6) and the sequence is conserved in many metal-dependent repressors (Figure 3.7). Based on literature evidence of DtxR and other iron-dependent repressors (Chapter 1) and the conserved domains identified in PF0851p, the sequence of PF0851p was compared to other iron-dependent repressors including a few archaeal homologs (AF0245, DesR (*A. fulgidus*);

AF1984, TroR (*A. fulgidus*); MA3468 (*M. acetivorans* C2A); MA4338, Idr (*M. acetivorans* C2A); MTH936, Irep (*M. thermautotrophicus* str *Delta H*). Conserved amino acid residues C89 (cysteine), E92 (glutamic acid), and H93 (histidine), of PF0851p are predicted to be the dimerization and metal-binding domains (Figure 3.8). Other archaeal homologs of PF0851p with more than 75% sequence homology are PAB0714 (*P. abyssi*), PH1163 (*P. horikoshii* OT3), and TK0107 (*T. kodakarensis*) (Figure 3.13). Considering sequence homology between PF0851p and other DtxR-like proteins, a 3D model was constructed using Geno3D program [147] (Figure 3.14).

3.4.3.2 *Binding site (putative operator)*: Based on the binding site tentatively identified from footprint and evidence of the known *dtxR* operator (TTAGG, Figure 3.7 (B)), and the fur box consensus sequence GATATAT [84, 148], the UOR database developed by Scott group (D. Cowart) was used to find similar motifs in other *Pf* UORs. Results obtained from a MEME motif search of all UORs was then correlated to iron-regulated genes from microarray data. A few of the *Pf* UORs containing the motif GATATAT (fur box consensus sequence) selected also show > 2.5-fold regulation and the occurrence of the motif is close to the translational start site (Table 3.2a). Similarly, *Pf* UORs containing the TTAGG (DtxR consensus) motif that show > 2.5-fold regulation is present close to the translational start site (Table 3.2b). Compiling results from the motif search, EMSA, and footprints with supporting literature evidence, I tentatively predict that PF0851p functions as an iron-dependent repressor binding to PF1287 UOR (putative *suf* promoter) and PF1283 UOR. Further work should be directed at

identifying the specific binding site using SELEX and/or genomic SELEX. This would enable the identification of similar binding sites for PF0851p in promoter regions of other ORFs in *Pf*.

Table 3.1a: Microarray data of *Pf* ORFs up-regulated by more than 2.5-fold in iron-limited growth conditions (unpublished data, A. L. Menon, M. W. W. Adam's group).

Operon	<i>Pf</i> ORF Label	Fold	Annotation
PF0529 PF0531	PF0531	3.2	Cobalamin biosynthesis protein m
	PF0693	2.5	Conserved hypothetical protein
	PF0723	7.6	Iron permease (FTR1)
PF0725 PF0729	PF0725	2.6	CoA-binding
	PF0726	3.3	Conserved hypothetical protein
	PF0727	3.5	Heavy metal transport/detoxification protein
	PF0728	2.7	Protein Kinase (ATP)
	PF0755	2.5	NAD-dependent aldehyde dehydrogenase (DHA)
PF0857 - PF0859	PF0857	4.7	Ferrous iron transport protein b (FeoB)
PF0911 - PF0917	PF0913	4.5	Formyl methanofuran dehydrogenase, subunit E
	PF0935	4.0	Acetolactate synthase
PF0936 - PF0942	PF0936	3.8	Acetohydroxy acid isomeroreductase (IlvC)
PF0936 - PF0942	PF0937	4.4	Pyruvate carboxyltransferase (HMGL-like)
PF0936 - PF0942	PF0938	4.5	3-isopropylmalate dehydratase large subunit
PF0936 - PF0942	PF0940	3.1	Isocitrate/isopropylmalate dehydrogenase (Isodh)
PF0936 - PF0942	PF0941	3.4	2-isopropylmalate synthase/homocitrate synthase
PF0936 - PF0942	PF0942	2.9	Dihydroxy acid dehydratase
PF1144 PF1145	PF1145	2.5	Nucleotide-binding protein (mrp/nbp35 family)
PF1285 – PF1287	PF1285	5.1	Conserved hypothetical protein (sufD homolog)
PF1285 – PF1287	PF1286	6.2	Conserved hypothetical protein (sufB homolog)
PF1528 – PF1530	PF1528	2.5	SNO glutamine aminotransferase (SNO)
PF1657 – PF1665	PF1657	2.5	Histidyl-tRNA synthetase
PF1657 – PF1665	PF1658	4.2	ATP phosphoribosyltransferase
PF1657 – PF1665	PF1659	5.0	Histidinol dehydrogenase
PF1657 – PF1665	PF1660	5.8	Imidazoleglycerol-phosphate dehydratase
PF1657 – PF1665	PF1661	4.1	Glutamine amidotransferase hisH
PF1657 – PF1665	PF1662	4.5	Histidine biosynthesis
PF1657 – PF1665	PF1663	3.6	Imidazoleglycerol-phosphate synthase
PF1657 – PF1665	PF1664	4.2	Phosphoribosyl-AMP cyclohydrolase
PF1657 – PF1665	PF1665	4.0	Probable histidinol-phosphate aminotransferase
	PF1666	3.0	Conserved hypothetical protein
	PF1678	4.8	3-isopropylmalate synthase (AIPN)
	PF1679	2.8	3-isopropylmalate dehydratase large subunit

Table 3.1b: Microarray data of *Pf* ORFs down-regulated by more than 2.5-fold in iron-limited growth conditions (unpublished data, A. L. Menon, M. W. W. Adam's group).

Operon	<i>Pf</i> ORF Label	Fold	Annotation
PF0529 PF0531	PF0001	3.2	Nitroreductase
	PF0153	5.0	SAICAR synthetase
	PF0154	2.9	Amidophosphoribosyltransferase
PF0191 - PF0199	PF0197	3.8	Phosphoribosylformylglycinamide synthetase
	PF0198	2.8	Phosphoribosylformylglycinamide synthase I
PF0214 - PF0216	PF0215	2.6	Enolase (2-phosphoglycerate dehydratase)
	PF0339	4.4	Methyltransferase
	PF0340	4.9	Putative HTH transcriptional regulator
PF0344 - PF0346	PF0345	2.7	Molybdopterin converting factor, subunit 1
PF0344 - PF0346	PF0346	2.5	Aldehyde ferredoxin oxidoreductase (AFOR_C)
	PF0422	4.6	Phosphoribosylamine-glycine ligase (GARS)
	PF0430	5.7	Phosphoribosylglycinamide formyltransferase
PF0431 PF0432	PF0431	7.2	AIR synthase related protein (AIRS)
PF0431 PF0432	PF0432	3.1	Haloacid dehalogenase-like hydrolase
PF0524 PF0525	PF0524	4.2	Conserved hypothetical protein
	PF0793	2.8	Hypothetical protein
	PF0970	3.1	4Fe-4S ferredoxin, iron-sulfur binding protein
	PF1013	2.7	Hypothetical protein
PF1281 – PF1283	PF1282	5.5	Rubredoxin-type Fe(Cys) ₄ protein
	PF1497	2.7	Alanine aminotransferase
	PF1500	3.4	Hypothetical protein
	PF1515	3.0	Glutamine aminotransferase
PF1516 PF1517	PF1516	2.2	GMP synthase, C-terminal (GMP_synt_C)
PF1516 PF1517	PF1517	5.4	ATP-grasp fold (ATP_GRASP)
	PF1532	2.6	NADH oxidase
PF1545 – PF1547	PF1547	2.8	Endoglucanase
	PF1687	3.6	Conserved hypothetical protein
	PF1701	3.2	Chorismate mutase
PF1786 PF1787	PF1787	2.8	Acetyl-CoA synthetase
PF1888 PF1890	PF1890	4.3	Conserved hypothetical protein
	PF1893	3.0	Transcriptional regulator, asnC family

Table 3.2a: Occurrence of DNA-binding motif (fur box consensus) of Fur in *Pf* UOR identified using MEME and correlated to iron-regulated ORFs from Microarray profiling.

UOR	GATATAT Motif coordinates		Microarray Data			Annotation
	From	To	Up-regulated (+)	Down-regulated (-)	Fold (up to)	
PF0154	-82	-76		-	2.9	Amidophosphoribosyltransferase
PF0401	-38	-32		-	4.5	Methyltransferase
PF0754	-60	-54	+		2.0	2-keto acid:ferredoxin oxidoreductase
PF0911	-96	-90	+		4.5	Iron (III) ABC transporter, ATP-binding
PF0948	-28	-22		-	1.9	Conserved hypothetical protein
PF1080	-78	-72		-	2.4	Hypothetical protein
PF1190	-39	-33		-	1.1	Conserved hypothetical protein (Ruberythrin)
PF1283	-20	-14		-	5.5	Ruberythrin
PF1396	-37	-31		-	1.2	2-dehydropantoate 2-reductase
PF1397	-9	-3	+		1.1	Hypothetical protein
PF1654	-14	-8		-	1.9	Hypothetical ABC transporter
PF1685	-15	-9		-	2.1	Acetylornithine/acetyl-lysine aminotransferase
PF1827	-40	-34		-	2.1	Conserved hypothetical protein
PF1835	-57	-51		-	1.7	Thiamine biosynthesis protein ThiI
PF1890	-62	-56		-	4.3	Conserved hypothetical protein
PF1938	-60	-54		-	2.2	Putative malE-like sugar binding protein

Table 3.2b: Occurrence of DNA-binding motif (DtxR operator consensus) of DtxR in *Pf*UOR identified using MEME and correlated to iron-regulated ORFs from Microarray profiling.

UOR	TTAGG Motif coordinates		Microarray Data		Annotation
	From	To	Up-regulated (+)	Down-regulated (-)	
PF0062	-13	-9	-	1.6	Riboflavin-specific deaminase
PF0067	-48	-44	-	1.8	Cobalt transport ABC transporter
PF0094	-80	-76	-	2.0	Glutaredoxin-like protein
PF0137	-20	-16	-	1.5	Putative methyltransferase
PF0166	-13	-9	-	1.5	Oxidoreductase
PF0299	-7	-3	-	1.1	Cobalamin synthase
PF0355	-39	-35	-	2.2	Hypothetical protein
PF0376	-28	-24	-	2.1	LSU ribosomal protein LXA
PF0432	-6	-2	-	3.1	Putative sugar-catabolism phosphotransferase
PF0526	-13	-9	-	2.3	soj-homolog (CbiA)
PF0529	-33	-29	+	2.4	Hypothetical protein (CbiQ)
PF0531	-17	-13	+	3.2	Cobalamin biosynthesis protein m
PF0553	-13	-9	-	1.7	Arsenate reductase
PF0722	-53	-49	+	1.5	Alkyl hydroperoxide reductase subunit c
PF0722	-34	-30	+	1.5	Alkyl hydroperoxide reductase subunit c
PF0723	-15	-11	+	7.6	Hypothetical protein (Iron permease, Ftr1)
PF0796	-17	-13	-	2.2	Galactoside o-acetyltransferase
PF0858	-26	-22	-	1.1	Conserved hypothetical protein (FeoA)
PF0889	-15	-11	-	2.0	Hypothetical protein
PF0937	-63	-59	+	4.4	2-isopropylmalate synthase
PF0938	-15	-11	+	4.5	3-isopropylmalate dehydratase
PF1033	-16	-12	-	1.6	Alkylhydroperoxide reductase
PF1042	-47	-43	-	2.3	Hypothetical protein (Ruberythrin)
PF1194	-10	-6	-	1.2	Conserved hypothetical protein (Furr family)
PF1199	-21	-17	-	1.9	Conserved hypothetical protein (Ferritin-like)
PF1222	-13	-9	-	2.7	Hypothetical protein
PF1230	-26	-22	-	2.3	Putative atpase
PF1246	-29	-25	-	1.6	Sarcosine oxidase, beta subunit (DAO)
PF1287	-18	-14	+	1.7	Putative ABC transporter
PF1552	-15	-11	-	2.1	Hypothetical protein
PF1588	-38	-34	-	2.8	Hypothetical protein
PF1602	-20	-16	-	2.1	Glutamate dehydrogenase
PF1659	-15	-11	+	5.0	Histidinol dehydrogenase
PF2026	-17	-13	+	1.6	Putative polyferredoxin

Table 3.3a: Details of the UOR DNA probe used for EMSA.

Locus	Gene length (bp)	Promoter length (bp)	Mol. wt. g/mole	Stock conc. of promoter
PF1283	516	191	60038.6	70 ng/ μ L, 1.17 μ M
PF1287	732	225	70789.8	100 ng/ μ L, 1.41 μ M

Table 3.3b: EMSA conditions using PF1283 and PF1287 UOR DNA with PF0851p.

Expt. No.	1	2	3	4	5	6	7
PF1283 DNA amount in μ M	0.082	0.082	0.082	0.082	0.082	0.082	0.082
PF0851 protein amount in μ M	0	0.171	0.343	0.685	1.028	1.371	1.713
Protein/ DNA ratio	0	2.10	4.19	8.40	12.60	16.79	20.99
PF1287 DNA amount in μ M	0.059	0.059	0.059	0.059	0.059	0.059	0.059
PF0851 protein amount in μ M	0	0.171	0.343	0.685	1.028	1.371	1.713
Protein/ DNA ratio	0	2.89	5.77	11.55	17.32	23.10	28.88

Table 3.4a: Details of the UOR DNA used for DNase I footprinting.

Locus	Gene length (bp)	Promoter length (bp)	Mol. wt. g/mole	Stock conc. of promoter
PF1283	516	301	189342.1	143 ng/ μ L, 0.74 μ M
PF1287	732	431	269635.7	118 ng/ μ L, 0.44 μ M

Table 3.4b: DNase I footprinting conditions using PF1283 and PF1287 UOR DNA with PF0851p.

Expt. No.	1	2	3	4	5
PF1283 DNA amount in μ M	0.040	0.040	0.040	0.040	0.040
PF0851 protein amount in μ M	0	0.076	0.152	0.305	0.495
Protein/ DNA ratio	0	1.89	3.78	7.56	12.29
PF1287 DNA amount in μ M	0.029	0.029	0.029	0.029	0.029
PF0851 protein amount in μ M	0	0.076	0.152	0.305	0.495
Protein/ DNA ratio	0	2.61	5.22	10.44	16.96

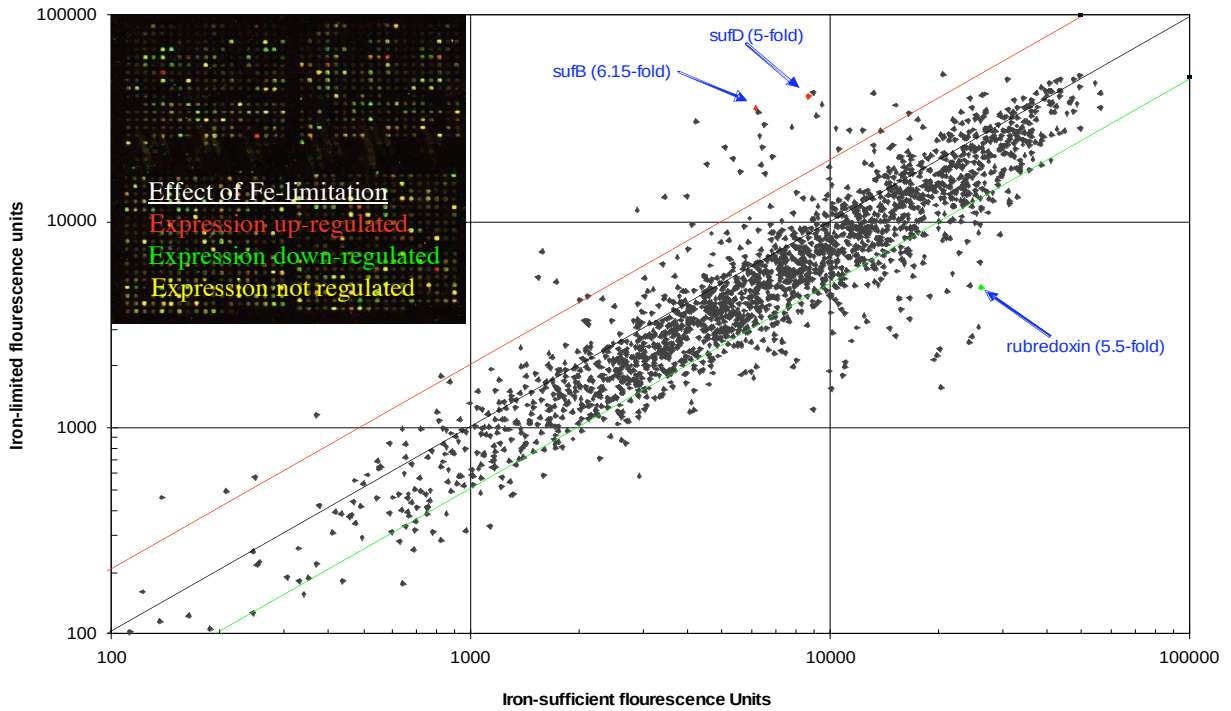


Figure 3.1: Microarray scatter plot of iron-sufficient and iron-limited *Pf* ORFs with 2-fold regulation boundaries (unpublished data, A. L. Menon, M. W. W. Adam's Group).

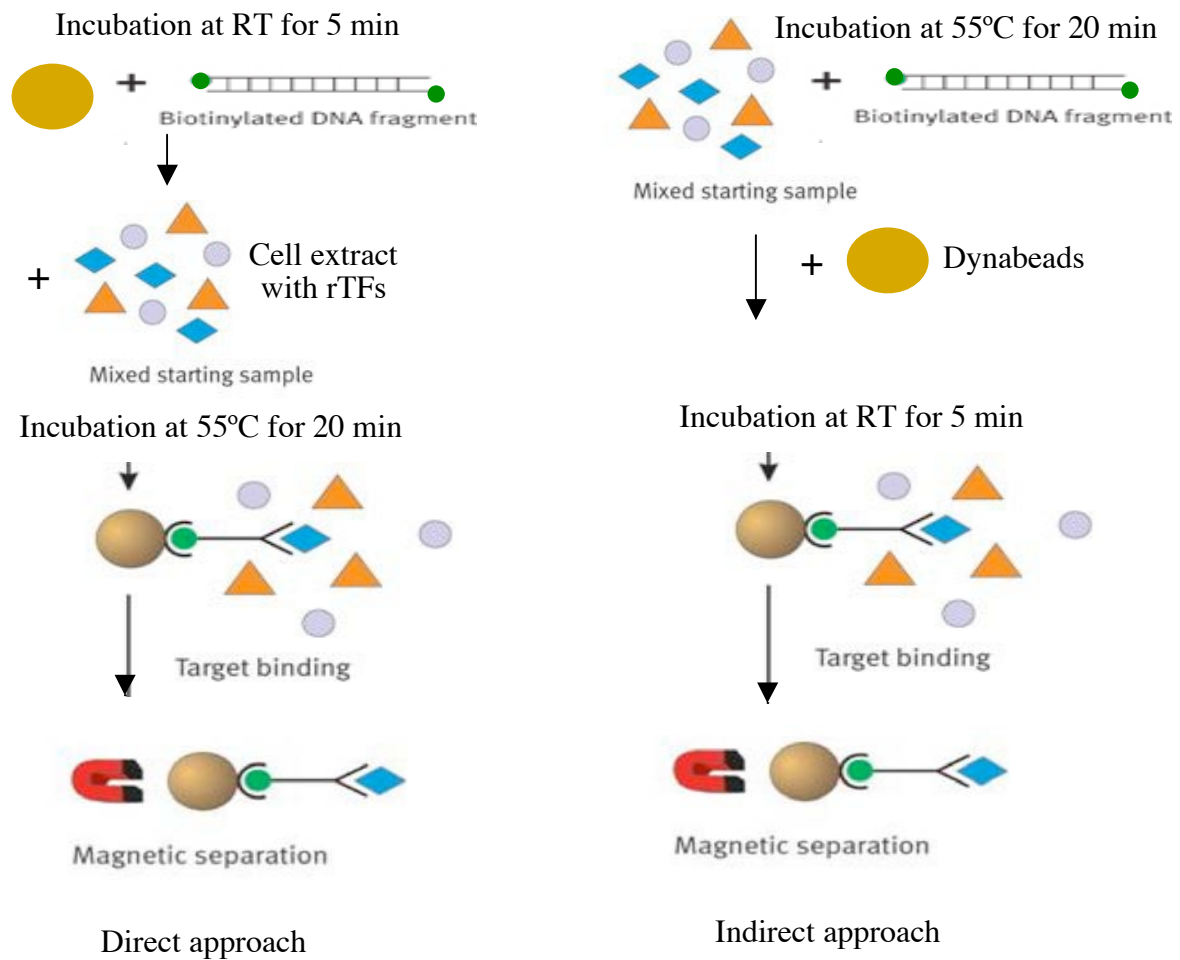


Figure 3.3: Schematic representation of DNA-affinity protein capture by direct and indirect approach (www.invitrogen.com).

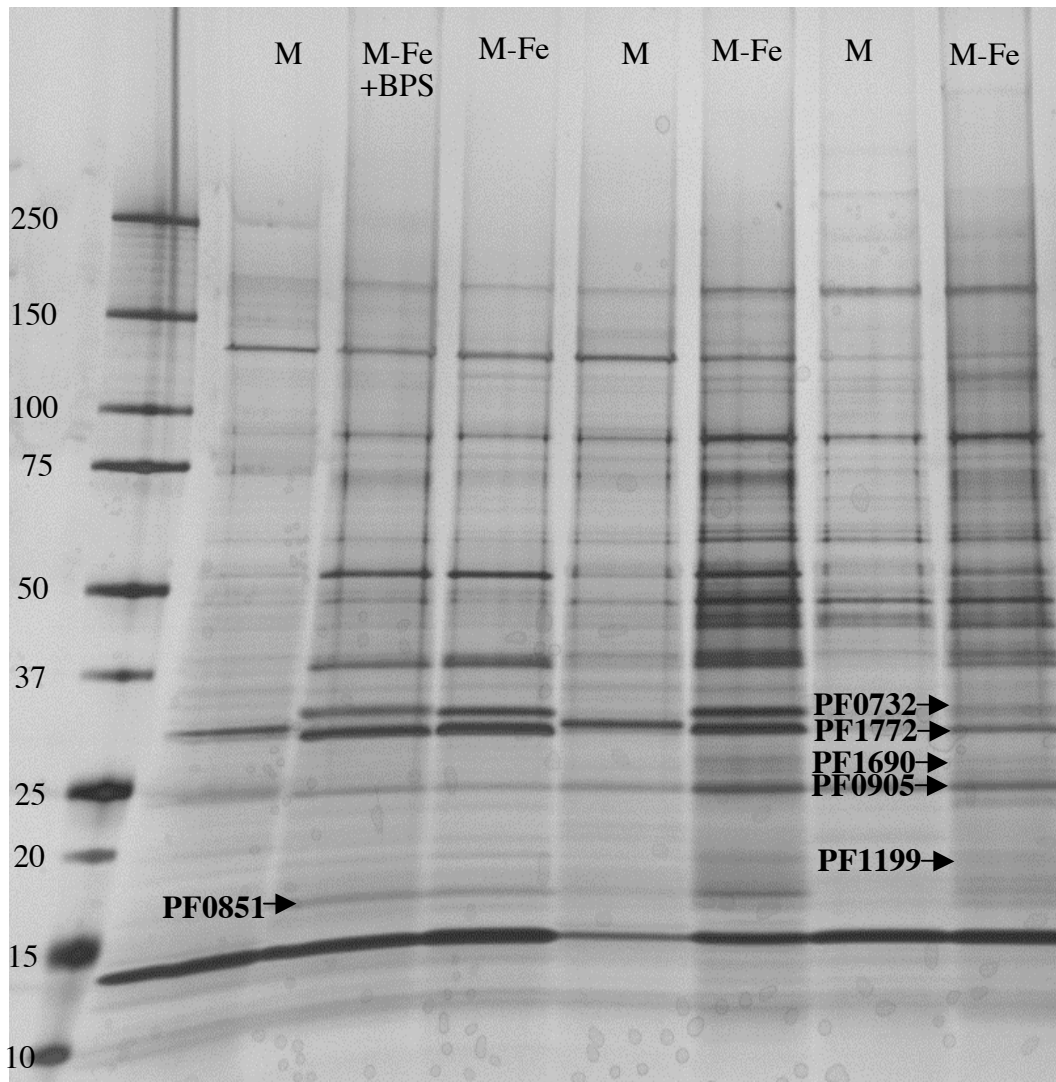


Figure 3.4a: SDS-PAGE gel of proteins bound to PF1283 UOR DNA probe.

Legends: M-Maltose control *Pf* cell extract (Fe 7.4 μM); M-Fe+BPS (Fe < 0.8 μM) *Pf* cell extract; M-Fe (Fe $\sim$ 0.8 μM) *Pf* cell extract. PF0732, Daunorubicin resistance ATP-binding protein; PF1772, 2-keto acid:ferredoxin oxidoreductase subunit beta; PF1690, 2-dehydro-3-deoxyphosphoheptonate aldolase; PF0905, putative ABC transporter; PF1199, Ferritin homolog; PF0851, iron-dependent repressor.

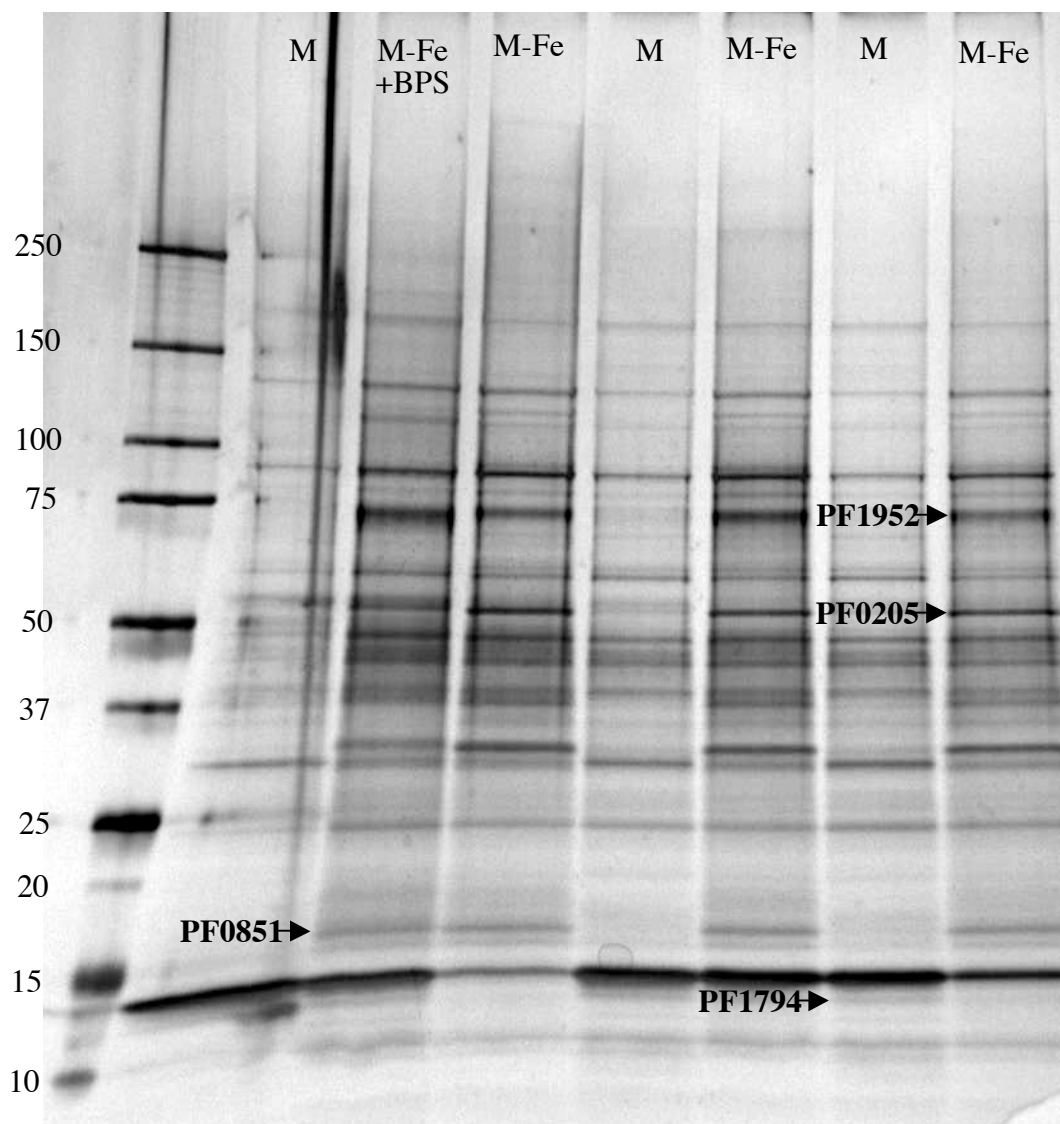


Figure 3.4b: SDS-PAGE gel of proteins bound to PF1287 UOR DNA probe.

Legends: M-Maltose control *Pf* cell extract (Fe 7.4 μM); M-Fe+BPS (Fe < 0.8 μM) *Pf* cell extract; M-Fe (Fe $\sim$ 0.8 μM) *Pf* cell extract. PF1952, iron-sulfur protein; PF0205, glutamate synthase; PF0851, iron-dependent repressor; PF1794, inosine-5'-monophosphate dehydrogenase related protein.



Figure 3.5: Conserved domain database search results for PF0851p (www.ncbi.nlm.nih.gov).

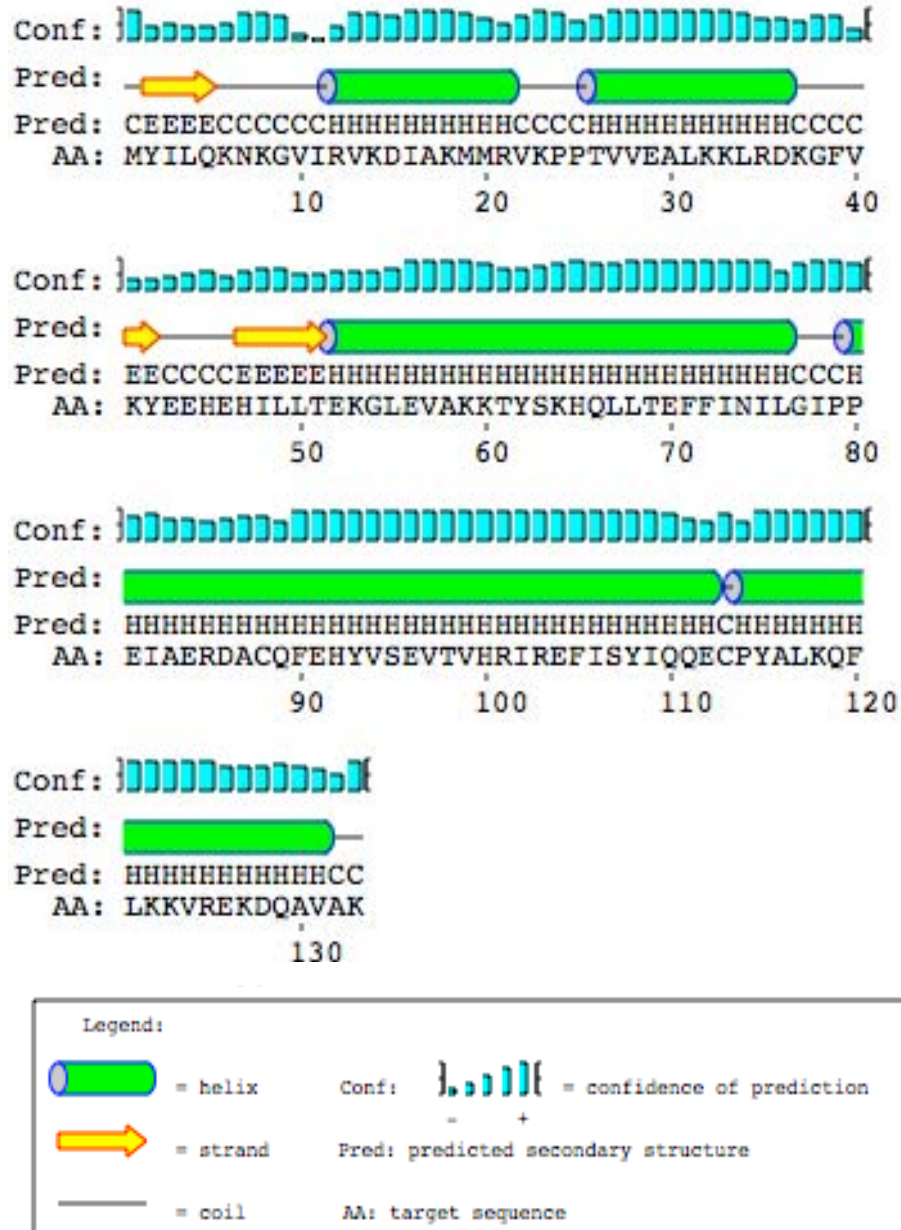


Figure 3.6: Secondary structure prediction of PF0851p using PSIPRED view [143].

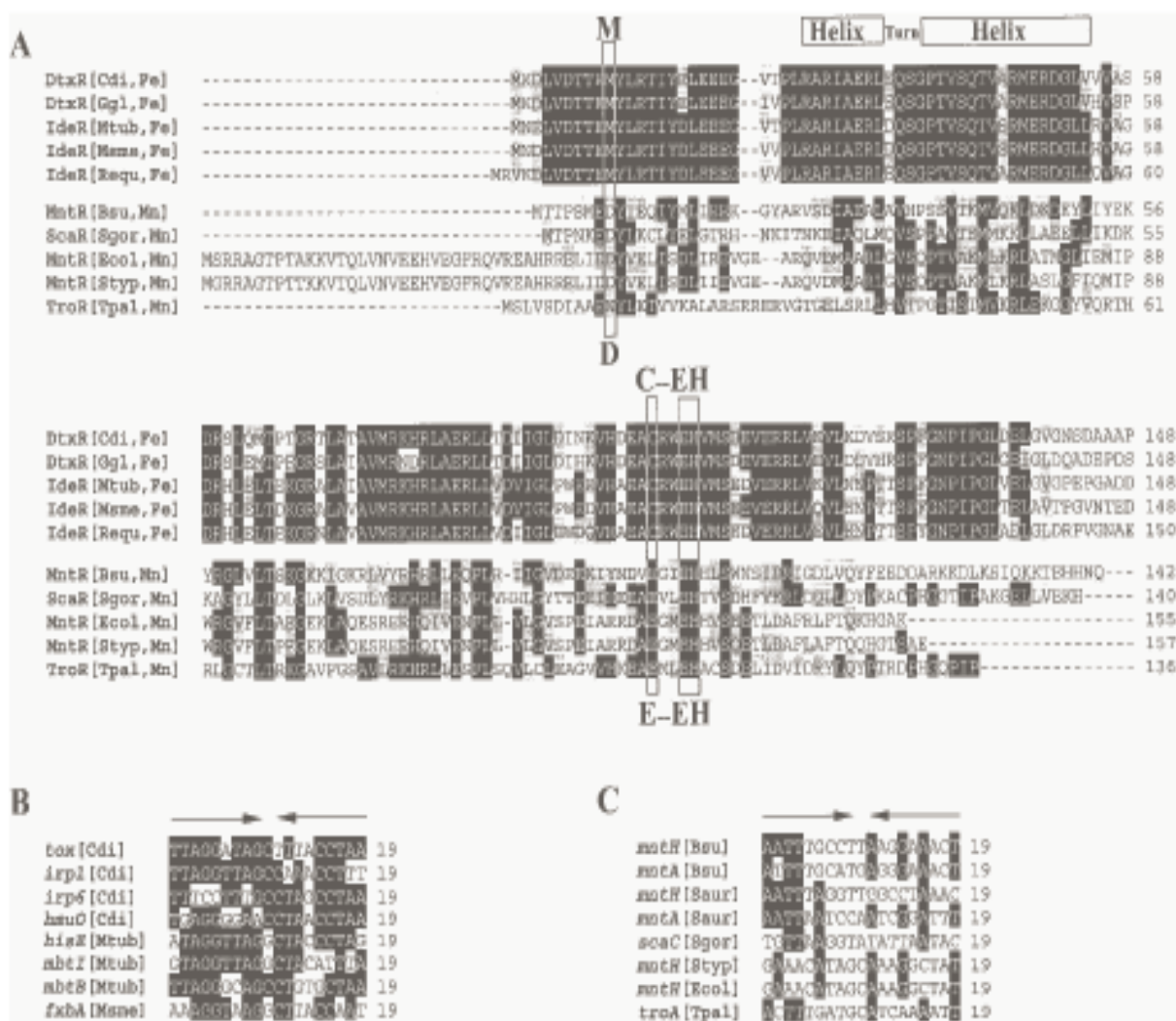


Figure 3.7: A-Primary structure alignment of DtxR homologs regulated by Mn(II) or Fe(II); B-Compilation of 19 bp DtxR-like binding site; C-Compilation of 19 bp MntR-like binding site [82].

Labels: The putative residues involved in the coordination of the regulatory metal are boxed. Sequences in A from: *C. diphtheriae* (Cdi), *C. glutamicum* (Cgl), *M. tuberculosis* (Mtub), *M. smegmatis* (Msme), and *Rhodococcus equi* (Requ), MntR homologs from *B. subtilis* (Bsu), *S. gordonii* (Sgor), *E. coli* K12 (*E. coli*), *Salmonella typhimurium* LT2, and TroR from *T. pallidum* (Tpal). B, C: Arrows show the inverted repeat.

```

1FX7_A      1 -[25]-LRARIAERLDQSGPTV.[1]-QTVSRMERDGLLRV.[1]-GDRHLELTEKGRALAIAMRKHRLAERLLV 87
query       1 -[11]-RVKDIAMMRVKPPTV.[1]-EALKKLDRKGFVKY.[1]-EMEHILLTEKGLEVAKKYTKHQLLEFFI 73
gi 11497861 1 -[1]-MAEEEGKRLKVDVSVET EALDELKRECLTV DCCIIKLTDRGRKKAEXIVRLHRLAERLLS 60
gi 11499566 5 -[26]-GPTQVAEELGVSKSTA.[1]-KMLNELSKACYGVY.[1]-PKKGLVLNERCKMEARAAMHRLIECLLA 92
gi 1731043   1 -[23]-RVSDIAEALAVHPSSV.[1]-KMWOKLDRKDEYLIY.[1]-KYRGLVLTSGKKIKRLVYRHELDDQFLR 85
gi 15807523 2 -[24]-STQALAAALGVAPASV.[1]-GMLRKLTEQGLVSH.[1]-PYQGARLTAEGERVAVLEVRHRLLELFLH 87
gi 20088921 19 -[27]-KTTEISSCLQVDPSTT.[1]-KTLNELAAACYLNH.[1]-PYRGVDLTENGKEYAEFLIRRHRLSLFT 107
gi 20092281 1 -[23]-KVKDVANELELGPSTV.[1]-GMFKKLDRECYINY.[1]-KYCGVLTTEKGMETARKTREKYSMLKDFLM 85
gi 20093126 2 -[26]-KTKQIAEELNVSSPSV.[1]-EMIKKLHSMGFVEY.[1]-PYQGVLTTEKGAEAVRIKRNHQLVETFLA 89
gi 15678956 9 -[24]-KISDLSSRMNVRPASV.[1]-QMLKRLHENGVIY.[1]-PYLGAELTDEGREIAENVLRHRLHILECFPH 94

1FX7_A      88 -[3]-GLPWEEVH.[2]-ACRNEHVMSEDVRRRLVKVLNPTTSP.[45]- 172
query       74 -[3]-GIPPEIAE.[2]-ACQPEHVVSEVTVHRIREFISYIQEC.[20]- 133
gi 11497861 61 -[3]-GPKDVEEH ACKTEHLDIDDEAEAICTLLGHPOVCP.[37]- 135
gi 11499566 93 -[2]-DLGVEDFC.[2]-AEKISMWAGSSSLVDALEARYGDREFCP.[10]- 139
gi 1731043   86 -[2]-GVDEEKIY.[2]-VEGIEHLSMNSIDRIGDLVQYFEEDD.[18]- 142
gi 15807523 88 -[3]-GVPLDEVH.[2]-AEALEHLSERLEARIAAWLGDPTHDP.[33]- 160
gi 20088921 100 -[2]-GLSTEEAC.[2]-VSRTEAFVSRDAVNKICNSMGHPNVGV.[22]- 168
gi 20092281 86 -[2]-GIDEKTAE.[2]-ACKIEHILNPKTAQALKKFIEPTDKED.[36]- 160
gi 20093126 90 -[3]-DFDRKEAH.[2]-ACELEHVAVSDSVLERLYDFLCNPEYCP.[42]- 171

```

Figure 3.8: Comparison map of sequence homology of PF0851p (query) with known metal-dependent repressors (COG1321: TroR) (www.ncbi.nlm.nih.gov).

Labels: 1FX7_A-IdeR (*M. tuberculosis*); query-PF0851 (*P. furiosus*); gi 11497861-AF0245, desR (*A. fulgidus*); gi 11499566-AF1984, troR (*A. fulgidus*); gi 1731043-BSU24520, mntR (*B. subtilis*); gi 15807523-DR_2539, idr (*D. radiodurans*); gi 20088921-MA3468 (*M. acetivorans* C2A); gi 20093126-MA4338, idr (*M. acetivorans* C2A); gi 15678956-MTH936, irep (*M. thermautotrophicus str Delta H*). Predicted dimerization and metal binding domain are boxed.

T7 promoter

Score = 715 bits (372), Identities = 396/400 (99%), Gaps = 1/400 (0%), Strand=Plus/Plus

```

Query 78  TATATCCTTCAAAAGAACAAGGGTGTAAATAAGGGTAAAGGACATAGCCAAGATGATGAGG 137
          |||
Sbjct 4    TATATCCTTCAAAAGAACAAGGGTGTAAATAAGGGTAAAGGACATAGCCAAGATGATGAGG 63

Query 138  GTAAAACCCCGACCGTTGTTGNGGCATTAAAAAAGCTTANGGATAAGGGGTTTGTAAAA 197
          |||
Sbjct 64    GTAAAACCCCGACCGTTGTTGAGGCATTAAAAAAGCTTAGGGATAAGGGGTTTGTAAAA 123

Query 198  TATGAGGAGCACGAGCACATTCTTCTCACTGAAAAAGGTCTGGAAGTTGCAAAGAAAAACA 257
          |||
Sbjct 124   TATGAGGAGCACGAGCACATTCTTCTCACTGAAAAAGGTCTGGAAGTTGCAAAGAAAAACA 183

Query 258  TACTCGAAGCATCAGCTTTTGACGGAGTTTTTCATAAACATTTTAGGCATTCCACCCGAA 317
          |||
Sbjct 184   TACTCGAAGCATCAGCTTTTGACGGAGTTTTTCATAAACATTTTAGGCATTCCACCCGAA 243

Query 318  ATAGCTGAGAGAGATGCGTGCCAATTTGAGCACTACGTTAGTGAGGTTACAGTACATANG 377
          |||
Sbjct 244   ATAGCTGAGAGAGATGCGTGCCAATTTGAGCACTACGTTAGTGAGGTTACAGTACATA-G 302

Query 378  GATTAGAGAATTCATAAGCTATATTTCAGCANGAATGTCCCTATGCGTTAAAGCAGTTCTT 437
          |||
Sbjct 303   GATTAGAGAATTCATAAGCTATATTTCAGCAAGAATGTCCCTATGCGTTAAAGCAGTTCTT 362

          Query 438  AAAAAAAGTTAGAGAAAAGGATCAAGCTGTGGCTAAGTAG 477
                      |||
          Sbjct 363  AAAAAAAGTTAGAGAAAAGGATCAAGCTGTGGCTAAGTAG 402

```

T7 Terminator

Score = 681 bits (354), Identities = 388/400 (97%), Gaps = 1/400 (0%), Strand=Plus/Minus

```

Query 38  CTACTNACCCACAGCTTGATCCTTTTCTCTAACTTTTTTTAAGAACTGCTTTAACGCATA 97
          |||
Sbjct 402  CTACTTAGCCACAGCTTGATCCTTTTCTCTAACTTTTTTTAAGAACTGCTTTAACGCATA 343

Query 98  GGGACATTCTTGCTGAATATAGACTTATGAATTCTCTAATCCTATGTACTGNAACCTCAC 157
          |||
Sbjct 342  GGGACATTCTTGCTGAATATAG-CTTATGAATTCTCTAATCCTATGTACTGTAACCTCAC 284

Query 158  TAACGTAGTGCTCAAATTGGCAGCATCTCTCTCAGCTATTTTCGGGNGGAATGCCTAAAA 217
          |||
Sbjct 283  TAACGTAGTGCTCAAATTGGCAGCATCTCTCTCAGCTATTTTCGGGTGGAATGCCTAAAA 224

Query 218  TGNTTATGAAAAACTCCGTCAAAAGCTGATGCTTCGAGTATGTTTTCTTTGCAACTTCGA 277
          |||
Sbjct 223  TGTTTATGAAAAACTCCGTCAAAAGCTGATGCTTCGAGTATGTTTTCTTTGCAACTTCCA 164

Query 278  GACCTTTTTCAGTGAGAAGAATGTGCTCGNGCTCCTCATATTTTACAAACCCCTTATCCC 337
          |||
Sbjct 163  GACCTTTTTCAGTGAGAAGAATGTGCTCGTCTCCTCATATTTTACAAACCCCTTATCCC 104

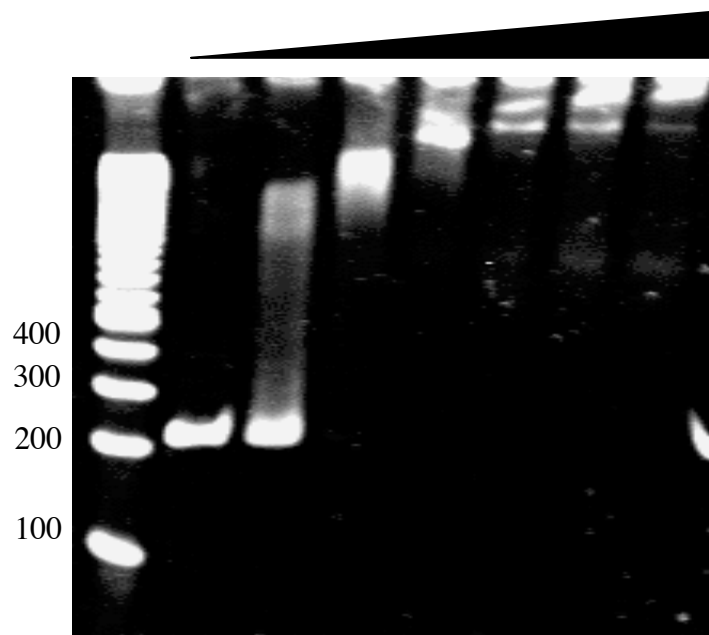
Query 338  TAAGCTTNTTTAATGCCTCAACAACGGTCGGGGGTNTTACCNCATCATCTTGGCTATGT 397
          |||
Sbjct 103  TAAGCTTTTNTAATGCCTCAACAACGGTCGGGGGTNTTACCCTCATCATCTTGGCTATGT 44

          Query 398  CCTTTACCCTTATTACACCCTTGNTCTTTTGAAGGATATA 437
                      |||
          Sbjct 43  CCTTTACCCTTATTACACCCTTGTTCTTTTGAAGGATATA 4

```

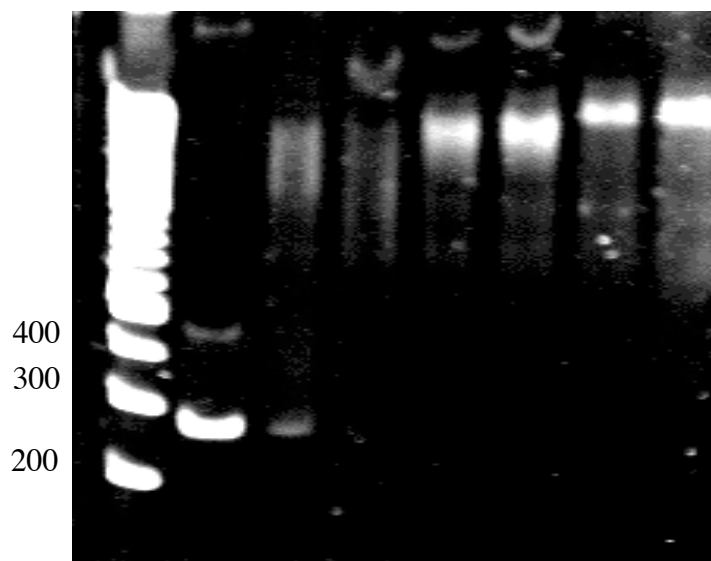
Figure 3.9: Validation of sequence of PF0851 plasmid (www.ncbi.nlm.nih.gov).

Protein/DNA ratio 0 2.1 4.2 8.4 12.6 16.8 21



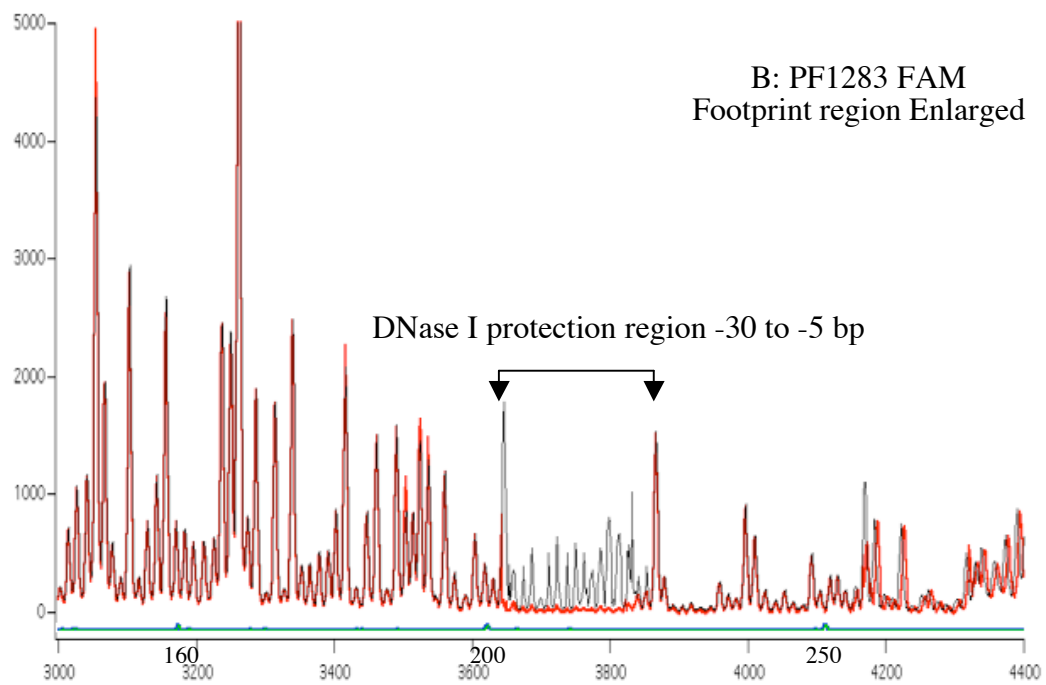
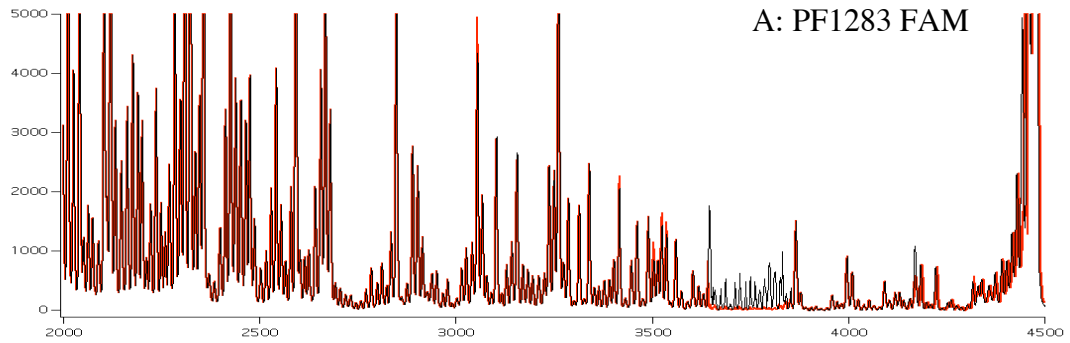
PF1283 UOR DNA (0.082 μ M)

Protein/DNA ratio 0 2.9 5.8 11.6 17.3 23.1 28.9



PF1287 UOR DNA (0.059 μ M)

Figure 3.10: EMSA of UOR DNA with PF0851p.



PF1283 UOR DNA (-231 to +70 bp)

5' 6FAM



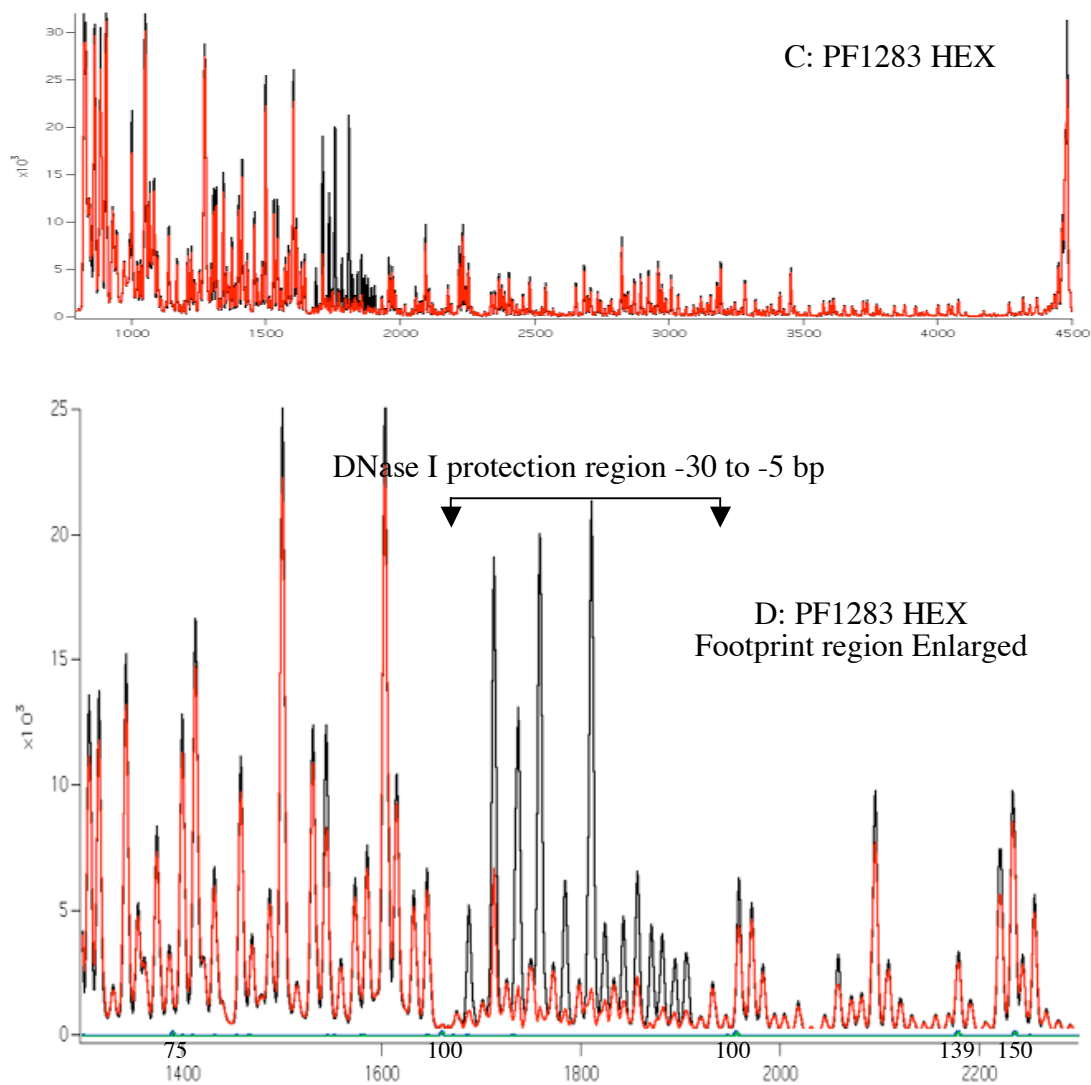
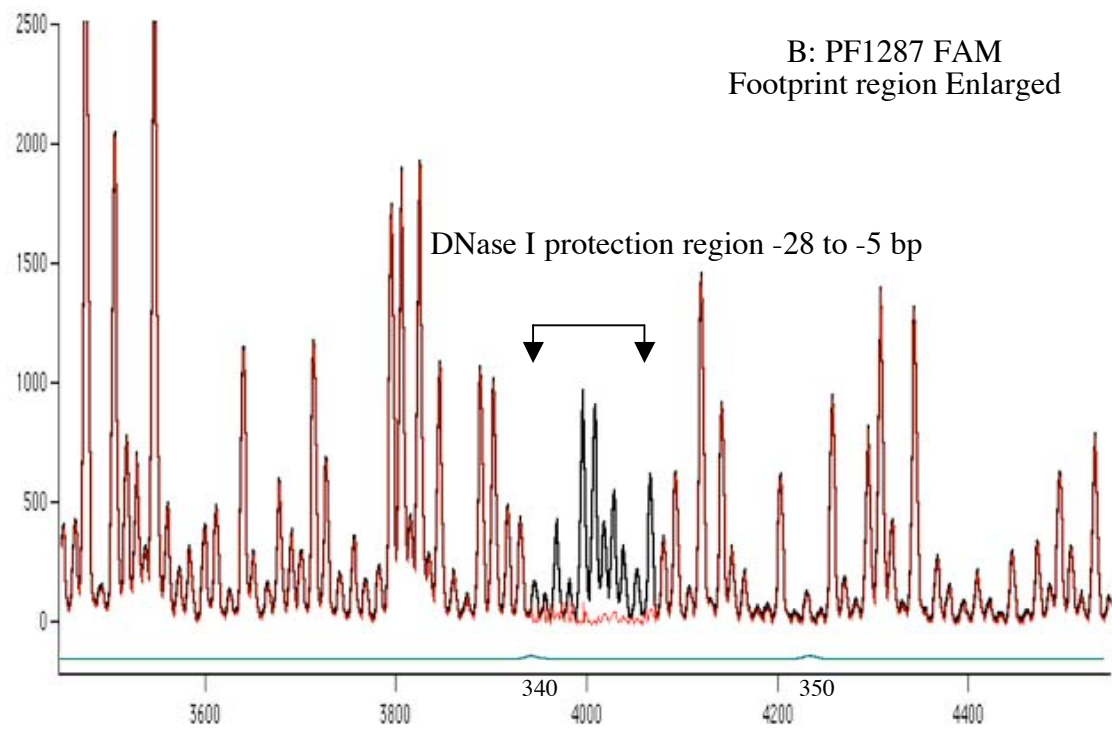
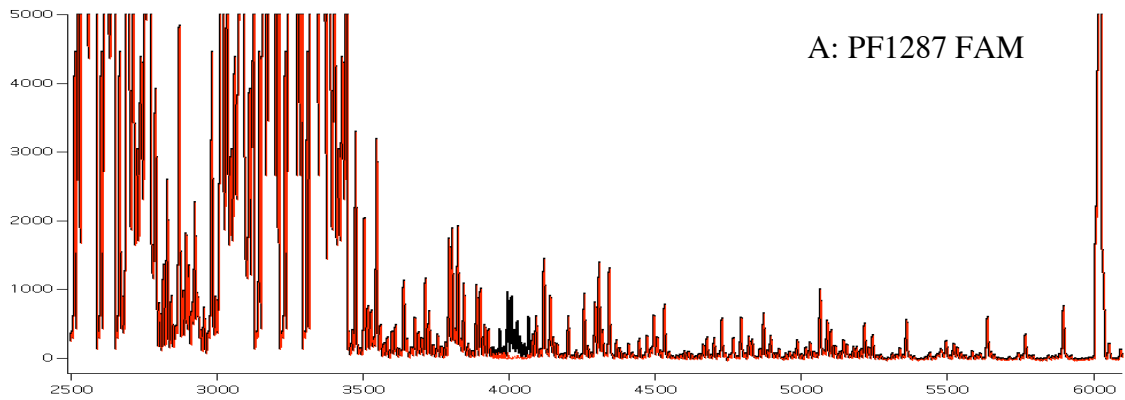


Figure 3.11a: CE chromatograms with footprints of PF0851p on PF1283 UOR DNA.

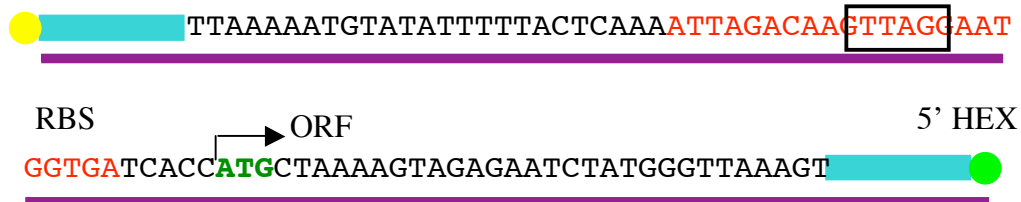
Labels: Each chromatogram compares control (black plot, without protein) and sample (red plot, with PF0851p). DNase I protected region on PF1283 UOR is indicated. Chromatograms (A, B) is the footprint of the forward strand (6FAM labeled), and chromatograms (C, D) is the footprint of the reverse strand of PF1283 UOR DNA and the putative binding site of PF0851p is -30 to -5 bp (relative to translation start site).



PF1287 UOR DNA (-368 to +63 bp)

5' 6FAM

Predicted DtxR binding site



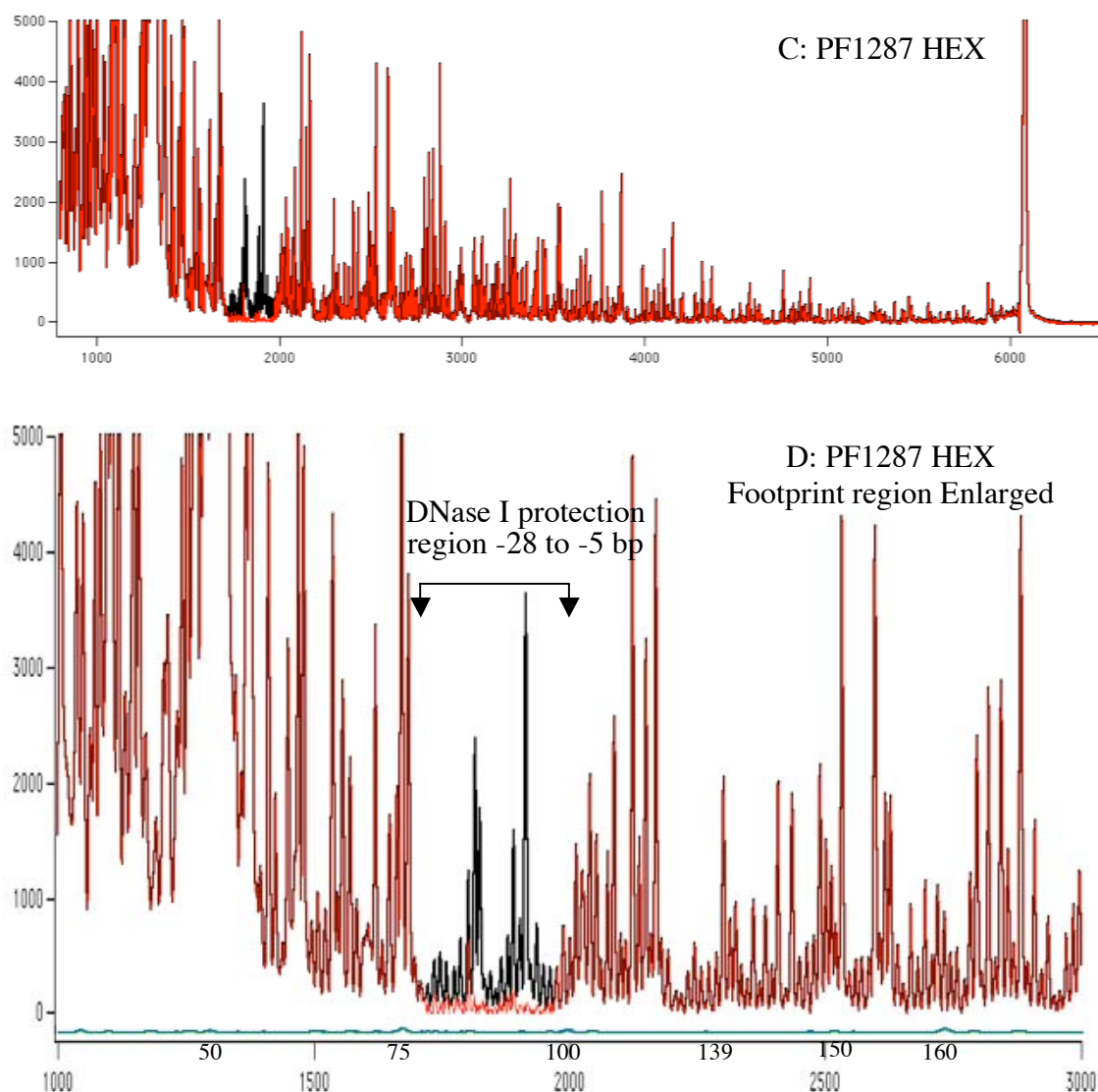


Figure 3.11b: CE chromatograms with footprints of PF0851p on PF1287 UOR DNA.

Labels: Each chromatogram compares control (black plot, without protein) and sample (red plot, with PF0851p). DNase I protected region on PF1287 UOR is indicated. Chromatograms (A, B) is the footprint of the forward strand (6FAM labeled), and chromatograms (C, D) is the footprint of the reverse strand of PF1287 UOR DNA and the putative binding site of PF0851p is -28 to -5 bp (relative to translation start site).

PF1283 UOR DNA sequence (1211260 - 1211560 Reverse Strand)

-231
 ▼
 CAAGGATATGAGAGTAAGCCCGGAATTCTATGAAGCCCTCGAAGCTGAAGTCA
 AGGCCCTCATCGAGAAAGCCGTTAAGAGGGCCCAAGCTGAAGGCAGAAAGACT
 CTCTATGCAAGGCATGTTTAACTTCCTTCTACCTATTTTTTTCAGTTTTATTCA
 CGGCATATTTTTTGAAGTGAATTTAAGAAAAAGTTAAATATTGAAGGATATAG
 ▲₋₃₀
 ATATATTCTGGTGTCCACATGTCGTGAAAAGAACAATGACTAAAAAGTTCT
 ▲₊₁
 TGGAAGAAGCCTTTGCAGGCGAAAGCATGGCCCATA
 ▲₊₇₀

PF1287 UOR DNA sequence (1214920 - 1215350 Reverse Strand)

-368
 ▼
 GGAATAGTTACGGGTGACTCCCTGGGTCAAGTTGCTTCTCAAACGCTTAGCAA
 CTTATTTCATAGAAACCATCAGTGTGGATTATCCAATATACAGGCCATTAATCG
 GTTTTGATAAAGAAGAAATTGTTTCAATAGCAAAAAGAATAGGTACCTACGAT
 GCGTTCTTGGAGTACCCCTATTGTGAATGTCCCTTTAGACCCAACAGAGTTAT
 TACACAAGGAAAGATCACTGAATTTTATTTTATAAGAGATGAGCTTAGAAAAG
 AGGGTCTTCTGTAACCCTCTATATTTTACCCAAATTTGGGTAAGAAAGACTTA
 AAAATGTATATTTTTACTCAAAATTAGACAAGTTAGGAATGGTGATCACCATG
 ▲₋₂₈ ▲₊₁
 CTAAAAGTAGAGAATCTATGGGTAAAGTCGAAGATAAAGAAATACTCAGGGC
 ATTAAC
 ▲₊₆₃

Figure 3.12: Binding site of PF0851p on PF1283 and PF1287 UOR DNA (red).

Legends: **GGTG** – Ribosome binding site, **ATG** – Translational start site.

Binding site (footprint) on PF1283 UOR DNA extends from -30 to -5 bp (red bar).

Binding site (foorprint) on PF1287 UOR DNA extends from -28 to -5 bp (red bar).

GENE ID: 1496428 PAB0714 | iron-dependent repressor [Pyrococcus abyssi GE5]
(10 or fewer PubMed links)

Score = 177 bits (449), Expect = 2e-43, Method: Composition-based stats.
Identities = 95/129 (73%), Positives = 117/129 (90%), Gaps = 0/129 (0%)

Query	1	MYILQKNKG VIRVKDI AKMMRVK PPTVVEALKKLRD KGFVKYEEHEHILLTEKGLEVAKK	60
Sbjct	13	MY+LQKNKG VIRVKDIA+++VKPP+VV+ALKKL KG ++YE+++ I LTE+G +A++ MYLLQKNKG VIRVKDIARVLKVKPPSVVDALKKLSKKGLIEYEKYDRITL TEEGKRIAE	72
Query	61	TYSKHQLL TEFFINILGIPPEIAERDACQFEHYVSEVTVHRIREFISYIQQEC PYALKQF	120
Sbjct	73	TYSKH LLTEFF+N+LGIPPEIAE DACQFEHYVSEVTV RIREFISYIQ+ CPY LKQF TYSKHVLLTEFFVNLGIPPEIAEDACQFEHYVSEVTVQRIRREFISYIQEHCPYVLKQF	132
Query	121	LKKVREKDQ	129
Sbjct	133	LKKV+E+D+ LKKVQEEDK	141

GENE ID: 1443483 PH1163 | hypothetical protein [Pyrococcus horikoshii OT3]
(10 or fewer PubMed links)

Score = 176 bits (447), Expect = 3e-43, Method: Composition-based stats.
Identities = 97/129 (75%), Positives = 117/129 (90%), Gaps = 0/129 (0%)

Query	1	MYILQKNKG VIRVKDI AKMMRVK PPTVVEALKKLRD KGFVKYEEHEHILLTEKGLEVAKK	60
Sbjct	13	MY+LQKNKG VIRVKDIAK+++VKPP+VV+ALKKL KG V+YE+H+ I LTE+G ++A++ MYLLQKNKG VIRVKDIAKILKVKPPSVVDALKKLSKKGLVEYEKHDRITL TEEGRKIAE	72
Query	61	TYSKHQLL TEFFINILGIPPEIAERDACQFEHYVSEVTVHRIREFISYIQQEC PYALKQF	120
Sbjct	73	TYSKH LLTEFF N+LGIPPEIAE DACQFEHYVSEVTV RIREFI+YIQ+ CPY LKQF TYSKHILLTEFFTNVLGIPPEIAEDACQFEHYVSEVTVKRIRREFITYIQENC PYVLKQF	132
Query	121	LKKVREKDQ	129
Sbjct	133	LKKV+E+D+ LKKVQEEDK	141

GENE ID: 3235244 TK0107 | manganese-dependent transcription regulator
[Thermococcus kodakarensis KOD1] (10 or fewer PubMed links)

Score = 166 bits (421), Expect = 3e-40, Method: Composition-based stats.
Identities = 87/122 (71%), Positives = 107/122 (87%), Gaps = 0/122 (0%)

Query	1	MYILQKNKG VIRVKDI AKMMRVK PPTVVEALKKLRD KGFVKYEEHEHILLTEKGLEVAKK	60
Sbjct	14	MYIL KNKG VIRVKDIAK + V+PP+VV+ALKKL +KG ++YE+++ ILLTEKG E+A+ MYILHKNKG VIRVKDIAKALNVRPPSVVDALKKLA EKGLIEYEKYDRILLTEKGREIAEN	73
Query	61	TYSKHQLL TEFFINILGIPPEIAERDACQFEHYVSEVTVHRIREFISYIQQEC PYALKQF	120
Sbjct	74	TYSKH LT+PF +ILGIPPEIAE+DACQFEHYVSE+TV R+REF +IQ++CPY LKQF TYSKHVFLTKPFTDILGIPPEIAEQDACQFEHYVSEITVKRMREFAFIQEQCPYVLKQF	133
Query	121	LK	122
Sbjct	134	LK	135

Figure 3.13: Sequence conservation of PF0851p with other homologs in archaea (www.ncbi.nlm.nih.gov).

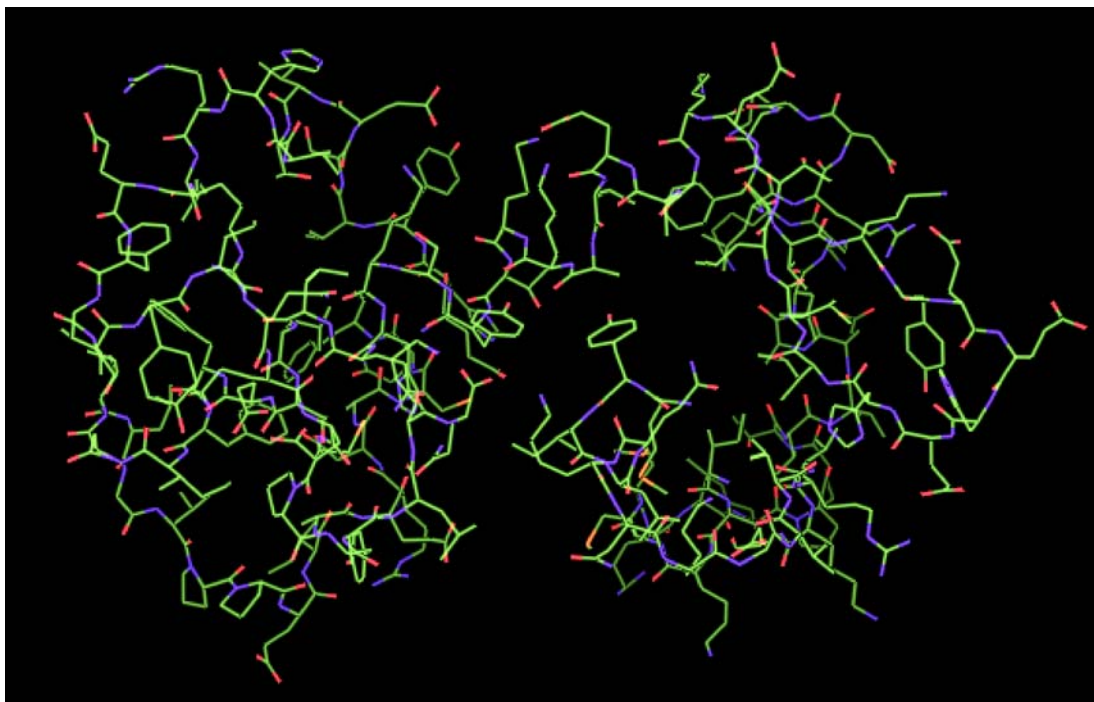


Figure 3.14: 3D model of PF0851p built based on 16 known X-ray structure of DtxR-like proteins using Geno3D [147].

REFERENCE

1. Segerer, A.H., et al., *Original Life Evolution Biosphere*. 1993. **23**: p. 77-90.
2. H. Kristjansson, G.O.H., *World Journal of Microbiology and Biotechnology*, 1995. **11**(17-25).
3. Huber, R., Huber, H., Stetter, K. O., *Towards the ecology of hyperthermophiles: biotopes, new isolation strategies and novel metabolic proerties*. FEMS Microbiology Reviews, 2000. **24**(5): p. 615-623.
4. K. O. Stetter, R.H., E. Blochi, M. Kurr, R. D. Eden, M. Fielder, H. Cash, I. Vance, *Hyperthermophilic archaea are thriving in deep North sea and Alaskan oil reservoirs*. *Nature*, 1993. **365**: p. 743-745.
5. S. Winker, C.R.W., et al., *System Applied Microbiology*, 1991. **14**: p. 305-310.
6. C. R. Woese, G.E.F., *Phylogenetic structure of the prokaryotic domain: the primary kingdom*. *Proceedings of National Academic Sciences*, 1977. **74**: p. 5088-5090.
7. P. Forterre, C.B., H. Philippe *Evolution of the Archaea*. *Theoretical Population Biology*, 2002. **61**: p. 409-422.
8. K. Ma, M.W.A., *A hyperactive NAD(P)H: rubredoxin oxidoreductase from the hyperthermophilic archaeon Pyrococcus furiosus*. *Journal of Bacteriology*, 1999. **181**: p. 5530-5533.

9. Fiala, G., Stetter, K. O., *Pyrococcus furiosus* sp-nov represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100C. Arch Microbiology, 1986. **145**(): p. 56-61.
10. W. M. de Vos, S.W.K., W. G. Voorhorst, J. van der Oost, Extremophiles, 1998. **2**: p. 201-205.
11. Kengen, S.W., Stams, A. J. M., de Vos, W. M., *Sugar metabolism of hyperthermophiles* FEMS Microbiology Review, 1996. **18**: p. 119-137.
12. Sapra, R., *Respiratory pathways and mechanisms of energy conservation in the hyperthermophilic archaeon pyrococcus furiosus*, in *Department of Biochemistry and Molecular Biology*. 2003, University of Georgia: Athens. p. 249.
13. M. J. Herrgard, M.W.C., B. O. Palsson, *Reconstruction of microbial transcriptional regulatory networks*. Current Opinions in Biotechnology, 2004. **15**(1): p. 70-77.
14. K. Tan, G.M.-H., J. Collado-Vides, G. D. Stormo, *A comparative genomics approach to prediction of new members of regulons*. Genome Res., 2001. **11**(4): p. 566-584.
15. G. D. Stormo, K.T., *Mining genome databases to identify and understand new gene regulatory systems*. Current Opinions in Microbiology, 2002. **5**(2): p. 149-153.
16. E. Perez-Rueda, J.C.-V., *The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12*. Nucleic Acids Research, 2000. **28**(8): p. 1838-1847.

17. S. D. Bell, a.S.P.J., *Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features*. Trends in Microbiology, 1998. **6**(6): p. 222-228.
18. Rowland T., B., P., and Jackson, S. P., *The TATA-binding protein: a general transcription factor in eukaryotes and archaeobacteria*. Science, 1994. **264**(5163): p. 1326-1329.
19. S. D. Bell, C.J., M. Nadel, P. F. Kosa, S. P. Jackson, *Temperature, template, topology and factor requirements of archaeal transcription*. Proceedings of National Academic Sciences, USA, 1998. **95**: p. 15218-15222.
20. S. D. Bell, C.P.M., S. P. Jackson, *Basal and regulated transcription in Archaea*. Biochemical Society Transactions, 2001. **29**(4): p. 392-395.
21. S. A. Qureshi, B.K., P. Baumann, S. P. Jackson, *Molecular cloning of the transcription factor TFIIB homolog from Sulfolobus shibatae*. Proceedings of National Academic Sciences, USA, 1995b. **92**(13): p. 6077-6081.
22. Cramer, P., *Multisubunit RNA Polymerases*. Current Opinions in Structural Biology, 2002. **12**: p. 89-97.
23. Hampsey, M., *Molecular genetics of the RNA polymerase II general transcriptional machinery*. Microbiology and Molecular Biology Reviews, 1998. **62**(2): p. 465-503.
24. Karim-Jean Armache, H.K., Cramer, P., *Architecture of initiation competent 12-subunit RNA polymerase II*. Proceedings of National Academic Sciences, USA, 2003. **100**(12): p. 6964-6968.
25. Finn Werner, R.O.J.W., *A recombinant RNA polymerase II-like enzyme capable of promoter-specific transcription*. Molecular Cell, 2002. **10**: p. 143-166.

26. E. Bright, R.H., *RNA polymerase: Structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II*. Journal of Molecular Biology, 2000. **304**: p. 687-698.
27. E. Peter Geiduschek, a.M.O., *Archaeal transcription and its regulators*. Molecular Microbiology, 2005. **56**(6): p. 1397.
28. Marx, J., *X-ray Crystallography: Transcription enzyme structure solved*, in *Science*. 2001. p. 411-414.
29. F. Todone, R.O.J.W., P. Brick, S. Onesti, *Crystal structure of RPB5, a universal eukaryotic RNA polymerase subunit and transcription factor interaction target*. Proceedings of National Academic Sciences, USA, 2000. **97**(12): p. 6306-6310.
30. I. Maillet, J.M.B., A. Sentenac, J. Labarre, *Rpb4p is necessary for RNA polymerase II activity at high temperature*. Journal of Biological Chemistry, 1999. **274**(32): p. 22586-22590.
31. M. Kimura, A.I., A. Ishihama, *RNA polymerase II subunits 2, 3, and 11 form a core subassembly with DNA binding activity*. Journal of Biological Chemistry, 1997. **272**(41): p. 25851-25855.
32. Darst, S.A., Edwards, A. M., Kubalek, E. W., Kornberg, R. D., *Three-dimensional structure of yeast RNA polymerase II at 16A resolution*. Cell, 1991. **66**(1): p. 121-128.
33. Akira Ishiguro, M.K., Kiyoshi Yasui, Akira iwata, Susumu Ueda, and Akira Ishihama, *Two large subunits of the fission yeast RNA polymerase II provide platforms for the assembly of small subunits*. Journal of Molecular Biology, 1998. **279**: p. 701-712.

34. Polyakov, A., Severinova, E., Darst, S. A., *Three-dimensional structure of E.coli core RNA polymerase: Promoter binding and elongation conformations of the enzyme*. Cell, 1995. **83**: p. 365-373.
35. A. L. Gnatt, P.C., J. Fu, D. A. Bushnell, R. D. Kornberg, *Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution*. Science, 2001. **292**: p. 1876-1882.
36. Meinhart, A., Cramer, P., *Recognition of RNA polymerase II carboxy terminal domain by 3'-RNA processing factors*. Nature, 2004. **430**: p. 223-226.
37. Schnabel, R., Thomm, M., Gerardy-Schahn, R., Zillig, W., Stetter, K. O., Huet, J., EMBO Journal, 1983. **2**: p. 751-755.
38. Lanzendo Orfer, M., Langer, D., Hain, J., Klenk, H. P., Holz, I., Arnold, I., Zillig, W., System Applied Microbiology, 1994. **16**: p. 654-664.
39. Goede, B., Naji, S., von Kampen, O., Ilg, K., Thomm, M., *Protein-Protein interactions in the archaeal transcriptional machinery*. The Journal of Biological Chemistry, 2006. **281**(41): p. 30581-30592.
40. R. M. R. Coulson, N.T., C. A. Ouzounis, *Lineage-specific partitions in archaeal transcription*. Archaea, 2006. **2**: p. 117-125.
41. Lee, S., J., Engelmann, A., Horlacher, R., Qu, Q., Vierke, G., Hebbeln, C., Thomm, M., Boos, W., *TrmB, a sugar-specific transcriptional regulator of the trehalose/Maltose ABC transporter from the hyperthermophilic archaeon Thermococcus litoralis* Journal of Biological Chemistry, 2003. **278**: p. 983-990.

42. Dahlke, I., Thomm, M., *A Pyrococcus homolog of the leucine-responsive regulatory protein, LrpA, inhibits transcription by abrogating RNA polymerase recruitment.* Nucleic Acids Research, 2002. **30**: p. 701-710.
43. Rojo, F., *Mechanisms of transcriptional repression.* Current Opinion in Microbiology, 2001. **4**(2): p. 145-151.
44. Takahashi, Y., Tokumoto, U., *A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids.* Journal of Biological Chemistry, 2002. **277**(32): p. 28380-28383.
45. Lind, M.I., *Characterization of some iron proteins in Drosophilla melanogaster and Pacifastacus leniusculus,* in *Faculty of Science and Technology.* 2000, Uppsala University: Uppsala.
46. L. Loiseau, S.O.-d.-C., L. Nachin, M. Fontecave, F. Barras, *Biogenesis of Fe-S cluster by the bacterial Suf system: SufS and SufE from a new type of cysteine desulfurase.* Journal of Biological Chemistry, 2003. **278**(40): p. 38352-38359.
47. H. Beinert, R.H.H., E. Munck, *Iron-sulfur clusters: nature's modular multipurpose structures.* Science, 1997. **277**: p. 653-659.
48. J. Frazzon, D.R.D., *Formation of iron-sulfur clusters in bacteria: an emerging field in bioinorganic chemistry.* Current Opinions in Chem. Biol., 2003. **7**: p. 166-173.
49. B. Ding, E.S.S., H. Ding, *Mobilization of the iron center in IscA for the iron-sulfur cluster assembly in IscU.* Biochemical Journal, 2005.
50. H. Beinert, P.J.K., *Fe-S proteins in sensing and regulatory functions.* Current Opinions in Chem. Biol., 1999. **3**: p. 152-157.

51. L. Zheng, V.L.C., D. H. Flint, D. R. Dean, *Assembly of iron-sulfur clusters. Identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii.* Journal of Biological Chemistry, 1998. **273**: p. 13264-13272.
52. S. I. Patzer, K.H., *SufS is a NifS-like protein, and sufD is necessary for the stability of the [2Fe-2S] FhuF protein in Escherichia coli.* Journal of Bacteriology, 1999. **181**: p. 3307-3309.
53. J. H. Lee, W.S.Y., J. H. Roe, *Induction of the sufA operon encoding Fe-S assembly proteins by superoxide generators and hydrogen peroxide: involvement of OxyR, IHF and an unidentified oxidant-responsive factor.* Molecular Microbiology, 2004. **51**: p. 1745-1755.
54. Wayne F. Outten, O.D., and Gisela Storz., *A suf operon requirement for the Fe-S cluster assembly during iron starvation in Escherichia coli.* Molecular Microbiology, 2004. **52**(3): p. 861.
55. L. Nachin, M.E.h., L. Loiseau, D. Expert, F. Barras, *SoxR-dependent response to oxidative stress and virulence of Erwinia chrysanthemi: the key role of SufC, an orphan ABC ATPase.* Molecular Microbiology, 2001. **39**: p. 960-972.
56. L. Nachin, L.L., D. Expert, F. Barras, *SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress.* EMBO J., 2003. **22**: p. 427-437.
57. Wayne F. Outten, M.J.W., F. M. Munoz, G. Storz, *The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in Escherichia coli.* Journal of Biological Chemistry, 2003. **278**: p. 45713-45719.

58. K. Rangachari, C.T.D., J. F. Eccleston, E. M. Hirst, J. W. Saldanha, M. Strath, R. J. Wilson, *SufC hydrolyzes ATP and interacts with SufB from Thermotoga maritima*. FEBS Letters, 2002. **514**: p. 225-228.
59. D. C. Johnson, D.R.D., A. D. Smith, M. K. Johnson, *Structure, function and formation of biological iron-sulfur clusters*. Annual Rev. Biochem., 2005. **74**: p. 247-281.
60. S. Watanabe, A.K., Kunio Miki, *Crystal structure of atypical cytoplasmic ABC-ATPase SufC from Thermus thermophilus HB8*. Journal of Molecular Biology, 2005. **353**: p. 1043-1054.
61. Y. R. Yuan, S.B., Martsinkevich, O., L. Millen, P. J. Thomas, J. F. Hunt, *The crystal structure of the MJ0796 ATP-binding cassette. Implications for the structural consequences of ATP hydrolysis in the active site of an ABC transporter*. Journal of Biological Chemistry, 2001. **276**: p. 32313-32321.
62. Locher, K.P., *Structure and mechanism of ABC transporters*. Current Opinions in Structural Biology, 2004. **14**: p. 426-431.
63. L. W. Hung, I.X.W., K. Nikaido, P.Q. Liu, G. F. Ames, S. H. Kim, *Crystal structure of the ATP-binding subunit of an ABC transporter*. Nature, 1998. **396**: p. 703-707.
64. K. Diederichs, J.D., G. Greller, C. Muller, J. Breed, C. Schnell et al., *Crystal structure of MalK, the ATPase sbunit of the trehalose/maltose ABC transporter of the archaeon Thermococcus Litoralis*. EMBO J., 2000. **19**: p. 5951-5961.
65. G. Verdon, S.V.A., B. W. Dijkstra, A. J. Driessen, A. M. Thunnissen, *Crystal structures of the ATPase subunit of the glucose ABC transporter from Sulfolobus*

- solfataricus: nucleotide-free and nucleotide-bound conformations*. Journal of Molecular Biology, 2003. **330**: p. 342-358.
66. J. Badger, J.M.S., J. M. Adams, S. Antonysamy, K. Blain, Bergseid, M. G. et al, *Structural analysis of a set of proteins resulting from a bacterial genomics project*. Proteins: Structural Functional Genetics, 2005. **60**: p. 787-796.
67. S. Kitaoka, K.W., Y. Hasegawa, Y. Minami, K. Fukuyama, Y. Takahashi, *Crystal structure of Escherichia coli SufC, an ABC-type ATPase component of the SUF iron-sulfur cluster assembly machinery*. FEBS Letters, 2006. **580**: p. 137-143.
68. S. Ollagnier-de-Choudens, L.N., Y. Sanakis, L. Loiseau, F. Barras, M. Fontecave, *SufA from Erwinia chrysanthemi. Characterization of a scaffold protein required for iron-sulfur cluster assembly*. Journal of Biological Chemistry, 2003. **278**: p. 17993-18001.
69. S. Ollagnier-de-Choudens, Y.S., M. Fontecave, *SufA/IscA: reactivity studies of a class of scaffold proteins involved in [Fe-S] cluster assembly*. Journal of Biological Inorganic Chemistry, 2004. **9**: p. 828-838.
70. S. Ollagnier-de-Choudens, D.L., L. Loiseau, F. barras, E. Forest, M. Fontecave, *Mechanistic studies of the SufS-SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE*. FEBS Letters, 2003. **555**: p. 263-267.
71. Kessler, D., *Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes*. FEMS Microbiology Review, 2006. **30**(825-840).
72. H. Mihara, T.K., T. Yoshimura, N. Esaki, *Kinetic and mutational studies of three NifS homologs from Escherichia coli: mechanistic difference between L-cysteine*

- desulfarase and L-selenocysteine lyase reactions*. Journal of Biochemistry (Tokyo), 2000. **127**: p. 559-567.
73. T. Kurihara, H.M., S. Kato, T. Yoshimura, M. Esaki, *Assembly of iron-sulfur clusters mediated by cysteine desulfurases. IscS, CsdB and CSD from Escherichia coli*. Biochim. Biophys. Acta, 2003. **1647**: p. 303-309.
 74. T. A. Ramelot, J.R.C., S. Goldsmith-Fischman, G. J. Kornhaber, R. Xiao, R. Shastri et al., *Solution NMR structure of the iron-sulfur cluster assembly protein U (IscU) with zinc bound at the active site*. Journal of Molecular Biology, 2004. **344**: p. 567-583.
 75. M. Zheng, X.W., L. J. Templeton, D. R. Smulski, R. A Larossa, G. Storz, *DNA microarray-mediated transcriptional profiling of the Escherichia coli response to hydrogen peroxide*. Journal of Bacteriology, 2001. **183**(15): p. 4562-4570.
 76. Tiirupati, B., Vey, J. L., Drennan, C. L. et al., *Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfarase from Synechocystis sp PCC 6803*. Biochemistry-US, 2004. **43**(38): p. 12210-12219.
 77. Huet, G., Daffe, M., Saves, I., *Identification of the Mycobacterium tuberculosis SUF machinery as the exclusive mycobacterial system of [Fe-S] cluster assembly; Evidence for its implications in the pathogen's survival*. Journal of Bacteriology, 2005. **187**(17): p. 6137-6146.
 78. Hantke, K., *Regulation of ferric iron transport in Escherichia coli K12: isolation of a constitutive mutant*. Mol. Gen. Genet, 1981. **182**: p. 288-292.

79. L. Escolar, J.P.-M., Victor de Lorenzo, *Opening the iron box: Transcriptional metalloregulation by the Fur protein*. Journal of Bacteriology, 1999. **181**(20): p. 6223-6229.
80. D. W. Griggs, B.B.T., J. Konisky, *Cloning and promoter identification of the iron regulated cir gene of the Escherichia coli*. Journal of Bacteriology, 1987. **169**: p. 5343-5352.
81. P. E. Klebba, M.A.M., J. B. Neilands, *Kinetics of biosynthesis of iron-regulated membrane proteins in Escherichia coli*. Journal of Bacteriology, 1982. **149**: p. 880-888.
82. E. Guedon, J.D.H., *Origins of metal ion selectivity in the DtxR/MntR family of metalloregulators*. Molecular Microbiology, 2003. **48**(2): p. 495-506.
83. Tiss, A., Barre, O., Michaud-Soret, I., Forest, E., *Characterization of the DNA-binding site in the ferric uptake regulator protein from Escherichia coli by UV crosslinking and mass spectrometry*. FEBS Letters, 2005. **579**(25): p. 5454-5460.
84. Lavrrar, J.L., Christoffersen, C. A., McIntosh, M. A., *Fur-DNA interactions at the bidirectional fepDGC-entS promoter region in Escherichia coli*. Journal of Molecular Biology, 2002. **322**(5): p. 983-995.
85. B. Gold, G.M.R., S. A. Marras, M. Pentecost, I. Smith, *The Mycobacterium tuberculosis IdeR is a dual functional regulator that controls transcription of genes involved in iron acquisition, iron storage and survival in macrophages*. Molecular Microbiology, 2001. **42**: p. 851-865.
86. I. Brune, H.W., A. T. Huser, J. Kalinowski, A. Puhler, A. Tauch, *The DtxR protein acts as dual transcriptional regulator directs a global regulatory network*

- involved in iron metabolism of Corynebacterium glutamicum*. BMC Genomics, 2006. **7**(21).
87. J. Wennerhold, M.B., *The DtxR regulon of Corynebacterium glutamicum*. Journal of Bacteriology, 2006. **188**(8): p. 2907-2918.
 88. Fourel, G., Phalipon, A., Kaczorek, M., *Evidence for direct regulation of diphtheria toxin gene transcription by an Fe-dependent DNA binding repressor, DtxR, in Corynebacterium diphtheriae*. Infections and Immunization, 1989. **57**: p. 3221-3225.
 89. Qiu, X., Verlinde, C. L. M. J., Zhang, S., Schnitt, M. P., Holmes, R. K., Hol, W. G. J., *Three-dimensional structure of the diphtheria toxin repressor in complex with divalent cation co-repressors*. Structure, 1995. **3**(1): p. 87-100.
 90. T. Wang, G.S., R. Balasubramanian, L. McIntosh, D. A. Bryant, J. H. Golbeck, *The sufR gene (sll0088 in Synechocystis sp. strain PCC 6803) functions as a repressor of the sufBCDS operon in iron-sulfur cluster biogenesis in cyanobacteria*. Journal of Bacteriology, 2004. **186**: p. 956-967.
 91. C. J. Schwartz, J.L.G., T. Patschowski, C. Luther, F. J. Ruzicka, H. Beinert, P. J. Kiley, Proc. Natl. Acad. Scie. USA, 2001. **98**(14895-14900).
 92. M. F. Christman, R.W.M., F. S. Jacobson, B. N. Ames, *Positive control of a regulon for defenses against oxidative stress and some heat-shock proteins in Salmonella typhimurium*. Cell, 1985. **41**: p. 753-762.
 93. M. Zheng, X.W., B. Doan, K. A. Lewis, T. D. Schneider, G. Storz, *Computation-directed identification of OxyR DNA binding sites in Escherichia coli*. Journal of Bacteriology, 2001. **183**: p. 4571-4579.

94. G. Storz, J.A.I., *Oxidative stress*. Current Opinions in Microbiology, 1999. **2**: p. 188-194.
95. S. V. Kuznetsov, S.S., P. Vivas, D. M. Crothers, A. Ansari, *Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF*. Proc. Natl. Acad. Scie. USA, 2006. **103**(49): p. 18515-18520.
96. E. D. Coulter, N.V.S., Z. M. Beharry, J. J. Smith, B. C. Prickril, D. M. Kurtz, *Ruberythrin-catalyzed substrate oxidation by dioxygen and hydrogen peroxide*. InorganicChim. Acta., 2000. **297**: p. 231-241.
97. V. Nivierre, M.F., *Discovery of superoxide reductase: An historical perspective*. Journal of Biological Inorganic Chemistry, 2004.
98. M. Li. M. Y. Lin, J.L., L. L. Gui, J. Liao, T. Jiang, J. Zhang, D. Liang and W. Chang, *Crystal structure studies on ruberythrin: Enzymatic activity in relation to the zinc movement*. Journal Biological Inorganic Chemistry, 2003. **8**: p. 149-155.
99. M. W. W. Adams, F.E.J., M. D. Clay, M. K. Johnson, *Superoxide reductase: fact or fiction?* Journal of Biological Inorganic Chemistry, 2002. **7**(647-652).
100. F. E. Jenney, M.F.J.M.V., X. Y. Cui, M. W. W. Adams, *Anaerobic microbes: Oxygen detoxification without superoxide dismutase*. Science, 1999. **286**: p. 306-309.
101. F. E. Jenney, M.W.W.A., *Rubredoxin from Pyrococcus furiosus*. Methods in Enzymology, 2001. **334**: p. 45-55.
102. E. D. Coulter, D.M.K., *A role for rubredoxin in oxidative stress protection in Desulfovibrio vulgaris: Catalytic electron transfer to ruberythrin and two iron superoxide reductases*. Arch. Biochem. Biophys., 2001. **394**: p. 76-86.

103. M. Kobayashi, T.i., M. Katayama, M. Kanehisa, M. Bhattacharyya-Pakrasi, H. B. Pakrasi, M. Ikeuchi, *Response to oxidative stress involves a novel peroxiredoxin gene in the unicellular cyanobacterium Synechoscystis sp. PCC 6803*. Plant Cell Physiology, 2004. **45**(3): p. 290-299.
104. M. Sztukowska, M.B., J. Potempa, J. Travis, D. M. Kurtz, *Role of ruberythrin in the oxidative stress response of Porphyromonas gingivalis*. Molecular Microbiology, 2002. **44**: p. 479-488.
105. M. V. Weinberg, F.E.J.J., X. Y. Cui, M. W. W. Adams, *Ruberythrin from the hyperthermophilic archaeon Pyrococcus furiosus is a rubredoxin-dependent, iron-containing peroxidase*. Journal of Bacteriology, 2004.
106. T. A. Geissmann, M.T., L. Meile, *Transcriptional analysis of the ruberythrin and superoxide dismutase genes of Clostridium perfringens*. Journal of Bacteriology, 1999. **181**: p. 7136-7139.
107. F. E. Jenney, M.F.J.M.V., M. W. W. Adams, *Neelaredoxin from Pyrococcus furiosus is a novel type of superoxide dismutase*. Journal of Inorganic Biochemistry, 1999. **74**: p. 181-181.
108. L. C. Sieker, M.H., I. Le trong, S. Turley, M. Y. Liu, R. E. Stenkamp, *The 1.9 angstrom crystal structure of the "as isolated" ruberythrin from Desulfovibrio vulgaris: some surprising results*. Journal of Biological Inorganic Chemistry, 2000. **5**: p. 505-513.
109. L. C. Sieker, M.H., I. Le trong, S. Turley, B. D. Santarsiero, M. Y. Liu, J. LeGall, R. E. Stenkamp, *Alternative metal-binding sites in ruberythrin*. Nature Structural Biology, 1999. **6**: p. 308-309.

110. J. Legall, B.C.P., I. Moura, A. V. Xavier, J. J. G. Moura, B. H. Huynh, *Isolation and characterization of ruberythrin, a non-heme iron protein from Desulfovibrio vulgaris that contains rubredoxin centers and a hemerythrin-like binuclear iron cluster*. Biochemistry, 1988. **27**: p. 1636-1642.
111. F. deMare, D.M.K., P. Nordlund, *The structure of Desulfovibrio vulgaris ruberythrin reveals a unique combination of rubredoxin-like FeS₄ and ferritin-like diiron domains*. Nature Structural Biology, 1996. **3**: p. 539-546.
112. F. Bonomi, D.M.K., X. Y. Cui, *Ferroxidase activity of recombinant Desulfovibrio vulgaris ruberythrin*. Journal of Biological Inorganic Chemistry, 1996. **1**: p. 67-72.
113. W. Tempel, Z.J.L., F. D. Schubot, M. V. Weinberg, F. E. Jenney Jr., M. W. W. Adams, J. P. Rose, B. C. Wang, *Structural genomics of Pyrococcus furiosus: X-ray crystallography reveals 3D domain swapping in ruberythrin*. Proteins Structure and Functional Genetics, 2004.
114. Lill, R., Muhlenhoff, U., Trends Biochemical Sciences, 2005. **30**: p. 133-141.
115. Weinberg, M.V., *Stress responses in the hyperthermophilic archaeon Pyrococcus furiosus*, in *Department of Biochemistry and Molecular Biology*. 2004, University of Georgia: Athens. p. 191.
116. Velayudhan, J., Hughes, N. J., McColm, A. A., Bagshaw, J., Clayton, C. L., Andrews, S. C., Kelly, D. J., *iron acquisition and virulence in Helicobacter pylori : a major role for FeoB, a high-affinity iron transporter*. Molecular Microbiology, 2000. **37**: p. 274-286.

117. Hantke, K., *Is the bacterial ferrous iron transporter FeoB a living fossil?* . Trends in Microbiology, 2003. **11**: p. 192-195.
118. Mansy, S.S., Cowan, J. A., *Iron-sulfur cluster biosynthesis: Toward an understanding of cellular machinery and molecular mechanism.* Acc. Chem. Res., 2004. **37**: p. 719-725.
119. Wang, M., *Discovery of new transcription factors in Pyrococcus furiosus with DNA-affinity protein capture and mass spectrometry*, in *Department of Chemistry*. 2006, University of Georgia: Athens. p. 155.
120. Peer Bork's Research Group, E.M.B.L. *Three domains of life*. 2006 [cited; Available from: http://www.scienceinschool.org/repository/images/issue2tree2_large.jpg.
121. Poidevin, e.a., BMC Molecular Biology, 2006. **7**: p. 44.
122. Stetter, K.O., Rachel, R. [cited; Available from: <http://www.microbeworld.org/microbes/archaea/>.
123. Aldrich, H. *Genome News Network*. 2002 [cited; Available from: http://genomenewsnetwork.org/resources/sequenced_genomes/genome_guide_p3.shtml.
124. Gregory, M. *Transcription and Translation Laboratory Exercise*. [cited; Available from: <http://faculty.clintoncc.suny.edu/faculty/Michael.Gregory/files/Bio%20101/Bio%20101%20Laboratory/Transcription,%20Translation/Transcription,%20Translation.htm>.

125. Patrick Cramer, D.A.B., and Roger D. Kornberg., *Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution*. Science, 2001. **292**(5523): p. 1863-1876.
126. W. Yeo. J. Lee, K.L., J. Roe, *IscR acts as an activator in response to oxidative stress for the suf operon encoding Fe-S assembly proteins*. Molecular Microbiology, 2006. **61**(1): p. 206-218.
127. Seidman, C.E., Struhl, K., Sheen, J., Jessen, T., *Introduction of plasmid DNA into cells*. 1997, Current Protocols in Molecular Biology. p. 1-10.
128. Schendel, P.F., *Expression of proteins in Escherichia coli*. 1998, Current Protocols in Molecular Biology. p. 1-13.
129. Labiomed. *Applied Biosystems: In-Gel Digestion Protocol for Coomassie & Silver Stained 1D & 2D Gels*. [cited; Available from: <http://labiomed.org/users/equip/MALDI-TOF/In-Gel-Digest.htm>.
130. UCSF Mass Spectrometry Facility, U.o.C., San Francisco. *MS-Fit*. [cited; Available from: <http://169.230.19.26:8080/prospector/4.27.1/cgi-bin/msform.cgi?form=msfitstandard>.
131. Bryant, F.O., and Adams, M. W. W., *Characterization of hydrogenase from the hyperthermophilic archaebacterium, Pyrococcus furiosus*. The Journal of Biological Chemistry, 1989. **264**: p. 5070-5079.
132. Jun Hasegawa, M.E., Takshi Narita, Tomoko Yamada, Yuki Yamaguchi, Tadashi Wada and Hiroshi Handa, *A rapid purification method for human RNA polymerase II by two-step affinity chromatography*. Journal of Biochemistry, 2003. **133**(1): p. 133-138.

133. Carina Hethke, G., A. C. M., Winfred Hausner, Willem M. de Vos, and Michael Thomm, *A cell-free transcription system for the hyperthermophilic archaeon pyrococcus furiosus*. Nucleic Acids Research, 1996. **24**(12): p. 2369-2376.
134. Janice Cline, B., J. C., Hogrefe, H. H., *PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases*. Nucleic Acids Research. **24**(18): p. 3546-3551.
135. Studier, F.W., Rosenberg, A. H., Dunn, J. J., Dubendorff, J. W., Kuo, M. H., Brownell, J. E., Sobel, R. E. Ranalli, T. A., and Cook, R. G., *Growth media for T7 E. coli strains*. Methods in Enzymology, 1990: p. 60-89.
136. Lewis, L.M., *Structural and functional characterization of the transcription preinitiation complex from Pyrococcus furiosus*, in *Department of Chemistry*. 2000, University of Georgia: Athens. p. 130.
137. M. B. Renfrow, N.N., M. Lewis, H. Chen, R. H. Ebright, R. A. Scott, *Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex*. 2003.
138. T. J. Darcy, W.H., D. E. Awery, A. M. Edwards, M. Thomm, J. N. Reeve, *Methanobacterium thermoautotrophicum RNA polymerase and transcription in vitro*. Journal of Bacteriology, 1999. **181**(14): p. 4424-4429.
139. J. F. Holden, M.W.W.A., *Microbe-metal interactions in marine hydrothermal environments*. Current Opinions in Chem. Biol., 2003. **7**: p. 160-165.
140. G. J. Schut, S.L.B., M. W. W. Adams, *Insights into the metabolism of elemental sulfur by the hyperthermophilic archaeon Pyrococcus furiosus: Characterization*

- of a Coenzyme A-dependent NAD(P)H sulfur oxidoreductase. Journal of Bacteriology*, 2007.
141. Gharahdaghi, F.e.a., *Mass spectrometric identification of proteins from silver-stained polyacrylamide gel; a method for the removal of silver ions to enhance sensitivity. Electrophoresis*, 1999. **20**: p. 601-605.
 142. Shevchenko, A.e.a., *Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Analytical Chemistry*, 1996. **68**: p. 850-858.
 143. McGuffin, L.J., Bryson, K., Jones, D. T., *The PSIPRED protein structure prediction server. Bioinformatics*, 2000. **16**: p. 404-405.
 144. Promega, *Gel shift assay system. Promega Technical Bulletin. TB110*: p. 1-17.
 145. Douglas O. Wilson, J., P., and Bruce R. McCord, *Nonradiochemical DNase I footprinting by capillary electrophoresis. Electrophoresis*, 2001. **22**: p. 1979-1986.
 146. Jantz, D., Berg, J. M., *Binding assays get into the groove. Nature Biotechnology*, 2002. **20**: p. 126-127.
 147. Combet, C., Jambon, M., Deleage, G., Geourjon, C. , *Pole Bioinformatique Lyonnais Geno3D. Bioinformatics*, 2002. **18**: p. 213-214.
 148. J. L. Giel, D.R., M. Liu, F. R. Blattner, P. J. Kiley, *IscR-dependent gene expression links iron-sulphur cluster assembly to the control of O₂-regulated genes in Escherichia coli. . Molecular Microbiology*, 2006. **60**(4): p. 1058-1075.