5-2015

Management of a Primordial Problem: Redox-sensitive Transcriptional Regulation in Methanosarcina acetivorans

Ryan Christopher Sheehan
University of Arkansas, Fayetteville

Management of a Primordial Problem:

Redox-sensitive Transcriptional Regulation in *Methanosarcina acetivorans*

Management of a Primordial Problem:

Redox-sensitive Transcriptional Regulation in *Methanosarcina acetivorans*

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Space & Planetary Sciences

by

Ryan Sheehan
Central Connecticut State University
Bachelor of Science in Chemistry & Biology, 2012

May 2015
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

_____

Dr. Daniel J. Lessner
Thesis Director

_____   _____

Dr. Ralph L. Henry   Dr. Timothy A. Kral
Committee Member   Committee Member

_____

Dr. John C. Dixon
Committee Member

Abstract

The primordial Earth which hosted the first forms of life was an environment free of oxygen.

Early organisms utilized metabolisms dependent upon anaerobic conditions and incorporated systems

to which oxygen is deleterious. As the content of oxygen in Earth's atmosphere increased, anaerobic

organisms had to acquire methods to sense and combat oxygen and reactive oxygen species. Several

mechanisms were advantageous to such anaerobic organisms which correlated transcriptional processes

with the redox state of the cell so that energy may be conserved and oxygen stress recovery genes

activated during periods of oxidative stress. Iron sulfur (Fe-S) cluster cofactors incorporated within RNA

polymerase (RNAP) may sense oxygen to globally regulate transcription. *Methanosarcina acetivorans*, a

methanogenic archaeon, offers an opportunity to study an RNAP with two Fe-S clusters within an

organism of a phylogenetically and metabolically diverse group. An *in vitro* transcription system for *M.

acetivorans* could be used to investigate the effects Fe-S cluster integrity on RNAP activity, which would

require the components involved in promoter-specific transcription: RNAP, TATA-binding protein (TBP),

and transcription factor B (TFB). This work describes the purification of *M. acetivorans* TBP and TFB for

the development of such a system. *M. acetivorans* also possesses the methanogen-specific redox-

sensitive transcriptional regulator MsvR. This work provides evidence of a physiologically-relevant

reducing partner for MsvR. As of yet, *M. acetivorans* MsvR has only been observed to bind to its own

gene. This work investigates the other potential gene targets for MsvR.

Acknowledgements

Special thanks Addison, Skippy, Tom, Stephen, Camice, and Jose- my lab family- for making my time inside and outside of research enjoyable.

Special thanks to Faith for being a lab mama and friend in good times and bad- someone to whom I could always turn for merriment or reassurance.

Special thanks to Dan for being a constant source of inspiration, for pushing me when necessary, for understanding my limits, for imparting criticism with a soft touch, for thinking big, and for his magnanimous support and dedication, without which I'd have been lost.

Special thanks to Ma and Pa- I wouldn't be where I am without my caregivers, biggest supporters, and best friends.

Table of Contents

List of Tables and Figures

Chapter One: Introduction

The purpose of this work is to investigate the mechanisms which sense oxygen and reactive oxygen species in a strictly anaerobic methane-producing archaeon, *Methanosarcina acetivorans*. The first objective is to investigate the role of oxygen-labile iron-sulfur clusters in the RNA polymerase of *M. acetivorans* by the development of an *in vitro* promoter-dependent transcription system with components from the organism. TATA-binding protein and transcription factor B were purified for this purpose. The second objective is it investigate thioredoxin as a physiologically-relevant reducing partner to the methanogen-specific regulatory protein MsvR and to identify potential gene targets for *M. acetivorans* MsvR.

Dismantling the conception of life as divided strictly into two monophyletic categories of prokaryotes (then thought to be solely composed of bacteria) and eukaryotes, *Archaea* was recognized as a domain of life in 1977 [1]. Today, life on Earth is organized into three domains- *Archaea*, *Bacteria*, and *Eukarya*. The phylogenetic and taxonomic decoupling of *Archaea* from *Bacteria* proved to be a major step in understanding the arrangement and evolution of life on Earth as more insight was gained on this third domain.

The discovery of the third domain required an interesting paradigm shift of then-current biology. There had hitherto been two types of cells: prokaryotes and eukaryotes. Classically, this separation was delineated as a distinction between 'simple' prokaryotic cells and more 'complex' eukaryotic cells. Furthermore, an evolutionary lineage was implicit in this model of life by which the prokaryotic cells had given rise to the eukaryotic cells. Before the advent of molecular sequencing, phylogenetic relations were determined by morphological comparisons, which yielded dubious results- particularly with regards to prokaryotes. The fine-tuning of sequencing analyses eventually produced

what is regarded as the most reliable method of phylogenetic characterization- comparative sequencing of ribosomal RNA (rRNA) components [2]. The rRNA molecule is universal (all cells engage in translation), ancient, and obdurate (it has changed slowly with time [3, 4]), making it a prime candidate for determining evolutionary lineages. Comparisons of the molecular sequences of ribosomal RNA genes (corroborated by comparisons of many other genes) have elucidated the paraphyletic nature of the prokaryotic urkingdom [5] and enhanced our overall understanding of the tree of life.

From these data, it has been determined that members of the domain *Archaea* share a more recent common ancestor with members of the domain *Eukarya* than with members of the domain *Bacteria* (**Fig. 1.1**). In such a manner, the conception of prokaryotes as a monophyletic group was refuted. Indeed, as morphologically similar archaea and bacteria may appear, members of *Archaea* exhibit aspects similar to *Bacteria* as well as *Eukarya* (**Table 1.1**), with notable unique features.

The similarities between bacteria and archaea are morphological and, largely, superficial. The demarcation of prokaryotic life was originally proposed by Edouard Chatton in the early 19th century. Although the eukaryote-prokaryote dichotomy isn't necessarily reflective of the evolutionary relations among life [5], important commonalities can be found between archaea and bacteria. Structurally, prokaryotes are defined by their organizational simplicity- a lack of membrane-bound organelles and, thus, the presence of all water-soluble cellular components within the one cytosolic volume. Archaeal and bacterial cells are generally of comparable size and shape. Both bacteria and archaea reproduce exclusively through asexual reproduction and commonly engage in horizontal gene transfer (an activity also performed by eukaryotes but with far less frequency) [6]. Unlike with eukaryotes, horizontal gene transfer in bacteria and archaea is a dominant evolutionary force [7]. *Archaea* and *Bacteria* each boast an expansive selection of extremophilic organisms and an appropriately broad range of metabolic diversity, relative to the tame species found in *Eukarya*. Prokaryotes share some common features in terms of genome architecture and size. Genome sizes cover a similar range for the two domains,

between just under 500 kilobase pairs (kbp) and just over 13 Mbp, which is narrow and small when compared to the range of eukaryotic genome sizes [8]. The smallest genome discovered thus far belongs to the archaeon *Nanoarchaeum equitans* and is just greater than 490 kbp [9]. Introns, nucleotide sequences within genes which are excised to produce the final mature RNA molecule of a given gene, are prevalent in eukaryotes yet occur rarely within bacteria and archaea. Compared to eukaryotes, prokaryotes have compact and efficient genomes with relatively few intergenic regions of non-coding DNA [10]. Operons, clusters of genes under the control of a single promoter element, are common in archaea and bacteria and allow for the coordinated regulation and production of genes which are often related in function. The prokaryotic grouping, despite being founded on their 'simple' nature relative to eukaryotes, reflects some important characteristics shared between archaea and bacteria.

The archaeal systems of information processing are of a chimeric nature- transcription and translation in *Archaea* are composed of aspects of machinery and processes from *Bacteria* and *Eukarya*. Following, the archaeal system of transcription will be explored. Transcription, the initial step in gene expression, is the production of RNA from DNA and is performed by RNA polymerase (RNAP). Eukaryotes have at least three RNA polymerase systems (fourth and fifth RNAPs can be found in plants), each responsible for the transcription of non-overlapping subsets of genes [11]. Of these eukaryotic RNAPs, the second system (RNAP II) shares the greatest similarity to archaeal RNAP, and thus will be the focus of discussion concerning the comparison of archaeal and eukaryotic transcription within this work. Though a core of catalytic subunits is conserved across all three domains, eukaryotic RNAP II and archaeal RNAP are alike in terms of structure and subunit composition, and both vary greatly from bacterial RNAP [12]. Comparisons of RNAPs across the three domains will be revisited later in this work. RNAP works in conjunction with other transcriptional proteins to initiate promoter-driven transcription. In bacteria, RNAP is chaperoned to a promoter site by one of a variety of sigma factor proteins (**Fig. 1.2, left panel**). The transcription of a subset of genes is controlled by a particular sigma factor and that

sigma factor is the only protein required to assist RNAP with transcription initiation for that subset of genes. Archaea and eukaryotes employ a suite of general transcription factor proteins for any and all genes which prepare a promoter site prior to the binding of RNAP (**Fig. 1.2, center and right panels**). The binding cascade of general transcription factor proteins is similar for archaea and eukaryotes but varies in the number of proteins involved. While eukaryotic initiation requires at least six transcription factor proteins [13, 14], only two transcription factor proteins are required for initiation in *Archaea* [15, 16]. The similarities between the machinery and processes of archaeal and eukaryotic transcription initiation suggest a common origin. Furthermore, the complex process of eukaryotic transcription initiation can be better understood through the study of the relatively simple, yet analogous, archaeal transcription initiation system.

Several elements of the archaeal gene promoter region are analogous to those found in the promoter regions of eukaryal genes (**Fig. 1.3**). The TATA-box motif is a T/A-rich site located ~25 base pairs from the initiator element which loosely resembles the Pribnow box of bacterial promoters [17, 18]. TATA-binding protein (TBP) in both archaea and eukaryotes binds to the eight base pairs of the TATA-box in the minor groove of the DNA as the initial component of the preinitiation complex (PIC), partially unwinding and bending the DNA towards the major groove [19]. A β-recognition element (BRE) is located ~35 base pairs upstream of the initiation site and serves to bind transcription factor B (TFB in archaea, TFIIB in eukaryotes) [20, 21]. The addition of TFB to the TBP-promoter complex ensures the correct orientation of the complete PIC for RNAP to begin transcription [22].

Despite the homology between the general transcription factor proteins of archaea and eukaryotes, a vast majority of archaeal transcription-associated proteins have homology with similarly functioning proteins in bacteria [23]. Most of the proteins involved in transcription regulate specific genes or specific sets of genes; while *Archaea* and *Eukarya* have homologous general transcription proteins, the regulatory transcriptional proteins in *Archaea* often have bacterial homologues. Bacterial

and archaeal transcriptional regulatory proteins often share similar helix-turn-helix (HTH) DNA-binding domains which were likely distributed among *Archaea* and *Bacteria* through horizontal gene transfer events [24]. Many activators and repressors identified in archaea thus far have homologues in bacteria [25], and are similar both in terms of sequence and mechanism [26]. This interaction of bacterial-homologous regulatory factors with eukaryal basic transcription machinery represents a chimeric transcription system within the *Archaea*.

Several unique features set *Archaea* apart from the other two domains. A salient distinguishing feature is the composition of the archaeal cell membrane. While bacteria and eukaryotes have lipids based on ester linkages, archaea have lipids based on ether linkages. The lipids of archaea also feature isoprenoid hydrocarbon chains as opposed to fatty acid chains and a different stereoisomeric configuration, representing a vast departure from the membrane lipids of *Bacteria* and *Eukarya* [27]. These lipids require different reactants and different enzymes for their production, necessitating a unique lipid metabolism in *Archaea* [28].

The domain *Archaea* is host to organisms inhabiting a broad range of environments and exhibiting a suitably broad range of metabolic strategies. Early on, a presumed prerequisite of archaeal ecology was that of extreme habitats [29], and thus a lack of diversity was expected from *Archaea*. However, the discovery of several members of the phylum *Crenarchaeota* as mesophilic has confuted that misconception [30]. For example, members of the order *Methanomicrobiales* have been found as plankton in North American coastal waters, a temperate and oxic environment [31].

Archaea are of major importance to biogeochemical cycling on planet Earth [32]. In the nitrogen cycle, archaea perform several key steps including nitrogen fixation, nitrate respiration, nitrate assimilation, and denitrification. Importantly, these processes are carried out by archaea in environments such as highly saline water and hot springs, where other organisms are unlikely to occur in high abundance to perform these processes [33]. N_2 fixation is performed by archaea from three classes-

Methanobacteria, *Methanococci*, and *Methanomicrobia-* in a broad range of habitats including deep-sea

hydrothermal vents [34]. Archaea of the phylum *Thaumarchaeota* are ubiquitous in soil and marine

environments and carry out ammonia-oxidation [35, 36]. Only a few cultivated archaea, including

Pyrobaculum aerophilum and *Ferroglobus placidus*, perform denitrifying processes [37, 38]. Archaea are

involved in the production and oxidation of sulfidic compounds within the sulfur cycle [39]. The use of

sulfur as an electron acceptor during methane oxidation to produce disulfide (HS_2^-) is a newly discovered

process within the sulfur cycle which is thus known to be performed exclusively by archaea [40].

Archaea play an integral role in the cycling of carbon, uniquely performing several processes

which are necessary for the cycle to function. Archaea from the phyla *Crenarchaeota*, *Thaumarchaeota*,

and *Euryarchaeota* participate in carbon fixation, assimilating oxidized inorganic carbon compounds

such as CO_2 to form simple organics by reduction [41]. Models estimate that archaea of

Thaumarchaeota may provide nearly 1% of the production of oceanic reduced carbon [42].

Thaumarchaeota may even be the main local producers of reduced carbon in mesopelagic and

bathypelagic ocean zones [43], areas of the ocean which may host archaea as the dominant microbial

biomass [44]. Oxidation of methane (an important greenhouse gas which is discussed further below) is

performed by members of *Methanomicrobia*, which may consume a majority of biogenic oceanic

methane before its release to the atmosphere [45].

By and large, methanogenesis is the most impactful aspect of archaeal participation in global

biogeochemical cycling. Methane is a greenhouse gas with a global warming potential 3.7 times that of

CO_2 [46] and which is produced by industrial processes, serpentinization, and microbial metabolism

(with the majority of the Earth's methane production originating from microbes, namely archaea).

Methane-producing archaea (methanogens) are responsible for ~69% of the production of atmospheric

methane (CH_4). Methanogens are a phylogenetically diverse group of archaea from the kingdom

Euryarchaeota. Methanogen metabolism can produce methane from a variety of sources: CO_2 and H_2,

formate, methanol, methylamines, and/or acetate [47]. Globally, approximately 1 billion tons of

methane are formed annually from microbial methanogenesis [48]. Methanogens reside in a wide

variety of habitats: freshwater sediments, swamps, paddy fields, landfills, hydrothermal vents, and the

digestive tracts of many animals. They are generally found in anoxic environments with low

concentrations of sulfate, nitrate, Mn(IV), or Fe(III), as organisms utilizing those electron acceptors

metabolically will outcompete the methanogens [49]. Five orders of methanogens have thus far been

identified: *Methanobacteriales*, *Methanococcales*, *Methanomicrobiales*, *Methanopyrales*, and

Methanosarcinales. Many methanogens can make use of multiple substrates for methanogenesis, but

there has not yet been a methanogen observed which can use all methanogenic substrates. The most

recently branched order is *Methanosarcinales*; methanogens from this order exhibit a broad substrate

spectrum [49].

Methanogenesis is the biological production of methane and is a primitive metabolic pathway,

likely predating the Great Oxidation Event [50]. H_2 concentrations in the early Earth atmosphere were

likely very high [51], making the utilization of H_2 as a metabolic substrate highly likely. Indeed,

hydrogenases have proven to predate the last universal common ancestor (LUCA) organism [52].

Furthermore, many enzymes involved in methanogenic metabolism incorporate iron-sulfur (Fe-S) cluster

cofactors and are thus O_2-labile [53]. For example, heterodisulfide reductase (HDR) is the enzyme

responsible for the reversible reduction of the heterodisulfide (CoM-S-S-CoB) to the two constituent

methanogenic thiol coenzymes, coenzyme B and coenzyme M, a crucial step in the methanogenesis

pathway; HDR incorporates a [4Fe-4S] cluster to perform this reduction [54].

Iron-sulfur clusters are ubiquitous throughout life as enzyme cofactors and are likely of ancient

origins. The clusters incorporate inorganic iron and sulfur often using cysteine residues as ligands (**Fig.

1.4**), though arginine, aspartate, and histidine have been observed to also serve as ligands to coordinate

clusters [55]. Fe-S clusters are often found as components in electron-transport proteins due to the

capacity of the clusters for robust delocalization of electron density over both Fe and S atoms [56]. Yet

Fe-S clusters are versatile and serve a variety of roles as enzyme prosthetic groups: electron transfer,

substrate binding and activation, structural, regulatory, disulfide reduction, or sulfur donation. A varying

number of iron and sulfur atoms may be incorporated into Fe-S cluster groups (e.g. [2Fe-2S] clusters,

[4Fe-4S] clusters, [8Fe-8S] clusters). Proteins from the ferredoxin family utilize [2Fe-2S] clusters, [3Fe-4S]

clusters, or [4Fe-4S] clusters to transfer electrons in metabolic redox reactions in myriad organisms [57].

Nitrogenase, an enzyme involved in the fixation of atmospheric nitrogen (N_2), performs coupled

electron/proton transfer and incorporates Fe-S clusters [58]. An Fe-S cluster is a component of the

active site of acetyl-coenzyme A synthetase, involved in the citric acid cycle of aerobic respiration [59]. A

number of polyferredoxins have been found which exhibit multiple Fe-S clusters in tandem, likely for the

storage of electrons or iron atoms [60]. Endonuclease III, a DNA repair enzyme, uses [4Fe-4S] clusters in

a purely structural role whereby the cluster stabilizes a binding motif for interaction with DNA [61].

Biotin synthase catalyzes the conversion of dethiobiotin to biotin by donating sulfur from its [2Fe-2S]

cluster [55]. The [4Fe-4S] cluster of heterodisulfide reductase is involved in the reduction of thiols in

methanogenic metabolic processes [54].

Fe-S clusters play a variety of roles in the processes of life. In fact, the clusters may have been a

major component in the emergence of life from abiotic conditions. Many variations of biological Fe-S

clusters are capable of assembling spontaneously when local abundances of iron and sulfur are

sufficient [62]. Many Fe-S cluster containing enzymes catalyze redox reactions involving H_2, CO, and N_2,

all of which were likely to have been major components in the atmosphere of early Earth [63]. Iron and

sulfur have been postulated to have been involved in the original catalytic processes resulting in

abiogenesis. Primitive methanogenesis, which heavily involves Fe-S clusters, is a common candidate for

very early metabolism [50, 64]. It has been demonstrated that inorganic iron and sulfur can be involved

in C-C bond formation under primordial conditions [65]. In the very least, iron and sulfur may have

served as the first energy sources for primordial life [66]. The ubiquity and functional versatility of Fe-S clusters make them prime candidates as integral components of biological processes from the very beginning.

Yet Fe-S clusters are not well-suited for the high oxygen content of the Earth's atmosphere. Fe-S clusters are oxygen-labile and exposure to oxygen can cause the collapse and loss of clusters from an enzyme [53]. However, Fe-S clusters are far more amenable to anaerobic conditions; primordial Earth would have been a largely anaerobic environment before oxygenation of the atmosphere by photosynthetic organisms [67]. Thus, it has been postulated that the iron-sulfur clusters which were involved in primordial enzymes were preferentially replaced among aerobes as they evolved. We find many more Fe-S cluster-containing enzymes among extant anaerobes than extant aerobes, often due to the heavy involvement of such enzymes in anaerobic metabolisms, such as methanogenesis. In modern methanogens, many Fe-S cluster-incorporating enzymes necessary for methanogenesis are oxygen sensitive and exposure to oxygen can effect a loss of energy conservation and of methane production [53]. The abundance of Fe-S clusters in an organism is strongly inversely correlated with the oxygen tolerance of that organism [68].

The Great Oxidation Event, then, posed a problem for metabolisms incorporating Fe-S clusters. Cyanobacteria began releasing diatomic oxygen into Earth's atmosphere as a byproduct of photosynthesis approximately 2.7 billion years ago. This production of O_2 eventually led to the significant oxygenation of the Earth's atmosphere approximately 2.3 billion years ago [69]. This allowed for the proliferation of metabolisms exploiting the now-abundant O_2, yet also posed a problem for organisms still engaging in anaerobic metabolisms. Atmospheric oxygen caused widespread precipitation of iron from biologically-available environmental sources by oxidation to its ferric form, making soluble iron a limited resource in many environments. Furthermore, O_2 is effective at taking possession of electrons from reduced enzymes, which both attenuates metabolic activity performed by

9

those enzymes and produces reactive oxygen species- e.g. superoxide (O_2^-) and hydrogen peroxide

(H_2O_2), which can further damage proteins. These reactive oxygen species (ROS) can damage Fe-S cluster

enzymes by the oxidation of the clusters, often causing cluster collapse and the inactivation of the

enzyme [70, 71]. Likewise, hydrogen peroxide can oxidize the thiol groups of neighboring cysteine

residues to disulfide bonds, disrupting protein structures [72]. ROS are also deleterious to DNA, causing

mutations and DNA-protein crosslinks [73]. Anaerobic organisms often have metabolisms which rely

upon highly reduced substrates and enzymes with surface-exposed Fe-S clusters [67], which makes them

vulnerable to oxygen exposure and particularly at a disadvantage in conditions following the Great

Oxidation Event. Understandably, the Event caused a mass extinction of anaerobic organisms. The

remaining anaerobic organisms required methods to sense and manage exposure to oxygen to prevent

damage and loss of energy conservation.

Several methods have developed to provide control and management of ROS. Cellular redox

buffers such as ascorbate, glutathione, alkaloids, carotenoids, and flavonoids serve as antioxidants

within cells to maintain redox state during oxygen exposure [74]. Most organisms have enzymes which

perform the detoxification of ROS- e.g. peroxides, superoxide dismutases, catalase. A variety of

regulatory systems are used to sense oxygen exposure and coordinate a response. A number of redox-

sensitive transcriptional regulator proteins have been discovered, often incorporating Fe-S clusters or

thiol groups as sensing mechanisms. FNR (fumarate and nitrate reduction), a protein found in

Escherichia coli, uses a [4Fe-4S] cluster to control the transcription of genes involved in the adaptation

of cells to growth in O_2-deprived conditions [75]. Under reduced conditions with limited O_2, the

presence of a stable [4Fe-4S] allows for dimerization of FNR molecules; FNR dimers exhibit site-specific

DNA binding to effect the up-regulation of targeted genes. The first redox-responsive archaeal

transcriptional regulator, SurR, was discovered in *Pyrococcus furiosus* [76]. SurR effects a metabolic shift

to the production of H_2S when elemental sulfur is present whereby oxidation of two specific SurR

cysteine residues form a disulfide bridge, inhibiting sequence-specific DNA binding by SurR. The loss of

SurR binding inhibits genes involved in H_2 production and derepresses genes involved in elemental sulfur

metabolism. OxyR exerts regulatory control over an *E. coli* operon which contains genes encoding for

OxyR, catalase, alkyl hydroperoxide reductase, and glutathione reductase- each a component in

oxidative stress recovery. OxyR up-regulates this operon when it is activated, which is achieved by the

formation of a disulfide bond between two cysteine residues due to oxidation by H_2O_2 [77]; In such a

manner, the presence of H_2O_2 is sensed by the state of the cysteine residues which then signals

production of the oxygen response enzymes. In *Methanothermobacter thermautotrophicus*, a similar

system involves MsvR as a redox-sensitive transcriptional regulatory protein which controls an operon of

genes involved in the detoxification of ROS [78]. Such regulatory mechanisms are used to up-regulate

genes involved in recovery during times of oxidative stress.

The correlation of cellular redox state with transcription would be advantageous to an anaerobic

organism. While regulatory proteins can accomplish this in a specific manner, some RNA polymerases

may be capable of redox-informed global transcription regulation. It would be advantageous for an

anaerobic biological system to correlate energy-intensive processes, such as transcription, with the

redox state of a cell so as to conserve energy during times of oxidative stress. RNAP, as mentioned

previously, is the multi-subunit protein complex responsible for the synthesis of RNA from a DNA

template and is essential to cells in all three domains of life. Five universally-conserved core components

of RNAP ($\alpha_2\beta\beta'\omega$) compose bacterial RNAP; eukaryotic and archaeal RNAPs are considerably more

complex and contain these universally-conserved components as well as additional subunits (**Table 1.2**).

Archaeal and eukaryal RNAPs are similar in subunit composition; eukaryotic RNAP II consists of 12

subunits and archaeal RNAPs consist of 12-13 subunits depending on phyla. Subunit D of *Sulfolobus*

solfataricus RNAP has been observed to incorporate an Fe-S cluster (**Fig. 1.4**) [79]. This was the first

instance of an Fe-S cluster involved in an RNAP, and the exact role of the cluster has yet to be

determined. Subunit D of *S. solfataricus* contains three domains. The Rpb3 subunit of eukaryal RNAP II, the AC40 subunits of eukaryal RNAPs I & III, and the α subunit of bacterial RNAP are homologous to archaeal RNAP subunit D (RpoD). Domains 1 and 2 are conserved among RNAPs in all of life, but domain 3 is not found in bacteria [80]. Domain 1 is integral to the dimerization of subunit D with subunit L (RpoL), the first step in the assembly of RNAP [81]. The D-L heterodimer serves as an assembly platform upon which the rest of RNAP is constructed. Domain 2 of subunit D serves to interact with other RNAP subunits during assembly as well as general transcription factor proteins during transcription processes. Domain 3 is capable of hosting up to two [4Fe-4S] cluster binding motifs and can be found in members of *Archaea* and *Eukarya*. Domain 3 resembles a similar Fe-S cluster binding domain in ferredoxin [82]. *S. solfataricus* RpoD domain 3 contains an [4Fe-4S] cluster binding motif yet structural investigations revealed a [3Fe-4S] cluster coordinated with three cysteine residues (C183, C203, and C209), and with a fourth cysteine (C206) likely to serve as another ligand to the cluster *in vivo* [79]. Phylogenetic analyses revealed that there are two potential [4Fe-4S] cluster binding motifs to be found in RpoD/Rpb3/AC40 domain 3 of RNAPs [83]; the subunits can be categorized into six different groups based upon architectural aspects of domain 3 such as the number of complete or partial cluster binding motifs (**Fig. 1.5**). The analogous subunits of some eukaryotic RNAPs have been observed to contain the same cluster binding motif as that found in *S. solfataricus* [79]. Only anaerobic archaea have been observed to exhibit RpoD with both cluster binding motifs. The RpoD/Rpb3/AC40 of members of the same phylum can be placed into different groups based on their domain 3 architecture [83]. The function of [4Fe-4S] clusters within RNAP and their scattered phylogenetic presence remain a mystery. Perhaps the clusters offer the ability to correlate global gene expression with the redox state of the cell. The oxygen-labile clusters may function to sense oxygen, and their presence or absence within RNAP may regulate global transcriptional processes (**Fig 1.6**).

Methanogens offer an opportunity to explore RpoD subunits of diverse domain 3 architecture within a single group of organisms, the metabolism of which is heavily reliant upon Fe-S cluster enzymes [84]. It may be advantageous for organisms to which Fe-S clusters are essential for RNAP to possess Fe-S clusters- the clusters may offer a mechanism to coordinate metabolism with gene expression. However, not all methanogens exhibit RNAPs predicted to bind [4Fe-4S] clusters. The RpoDs of genome-sequenced methanogens have been analyzed: members of *Methanosarcinales* and *Methnomicrobiales* encode RpoD with two cluster motifs, members of *Methanobacteriales* encode RpoD with one or two cluster motifs, the member of *Methanopyrales* encodes an RpoD without cluster motifs, and members of *Methanobacteriales* encode RpoD entirely lacking domain 3 [83]. The architectural diversity of RpoD in methanogens may be attributed to metabolic and ecological differences between the different organisms. Thus, methanogens offer a unique opportunity to investigate the correlation between RpoD cluster configuration and metabolism/environment.

Methanosarcina acetivorans is a methanogen of the order *Methanosarcinales*, the most recently branched order of the methanogen. Only *Methanosarcina* species have demonstrated the ability to engage in all three known pathways of methanogenesis. *M. acetivorans* RpoD is capable of binding two [4Fe-4S] clusters; cluster #1 is coordinated by cysteines C205, C208, C211, and C183; cluster #2 is coordinated by cysteines C173, C176, C179, and C215 [83]. *M. acetivorans* is one of a few archaea which has a robust genetic system [85]; this system can be used to manipulate the cluster binding motifs to investigate the effects of different cluster configurations on RNAP activity and assembly. The organism has a well-investigated metabolism reliant upon Fe-S cluster enzymes and can grow on a variety of substrates, including acetate [86, 87]. This allows for the investigation of the effects of cluster manipulation on metabolism. The diverse & well-documented metabolism, the robust genetic system, and the presence of a full complement of clusters make *M. acetivorans* the ideal candidate to study the relationship between Fe-S clusters and RNAP activity and assembly.

It had previously been postulated that the collapse of the RpoD [4Fe-4S] clusters by oxygen may preclude the *de novo* assembly of RNAP by preventing the heterodimerization of RpoD with RpoL [79]. It has been determined, however, that absence of the clusters in *M. acetivorans* does not, in fact, prevent formation of the D-L heterodimer nor prevent the assembly of all subunits into RNAP [83]. This suggests that rather than affecting RNAP assembly, the clusters serve a regulatory role and their oxygen-lability may affect RNAP conformation in such a way as to regulate global transcriptional processes in some manner. An *in vitro* transcription system would allow for the investigation of the effects of [4Fe-4S] cluster integrity upon RNAP activity.

Other methods of redox-sensitive transcriptional regulation can be investigated in *M. acetivorans*. MsvR is a transcriptional regulatory protein found in the majority of methanogens. *Methanothermobacter thermautotrophicus* MsvR (MtMsvR) regulates its own promoter and an operon encoding genes involved in oxygen stress recovery [78], including $F_{420}H_2$ oxidase, a methanogenic protein demonstrated to reduce O_2 to H_2O [88]. Under reducing conditions, MtMsvR represses this operon by abrogating the promoter site to prevent the binding of general transcription factors. Oxidation of MtMsvR by H_2O_2 yields different DNA binding behavior. This is consistent with MtMsvR being a redox-sensitive transcriptional regulator. *M. acetivorans* also contains an MsvR (MaMsvR), yet MaMsvR has thus far only been observed to bind its own promoter under reducing conditions [89]. Incubation of MaMsvR will prevent DNA binding; this is achieved by the oxidation of key cysteine groups whereupon the conformation of MaMsvR changes, yielding a cessation of DNA binding. If MaMsvR truly functions as a redox-sensitive transcriptional regulator, a physiological reducing partner would be required for cells to restore MaMsvR DNA binding upon the relief of oxidative stress. A potential partner may be *M. acetivorans* thioredoxin 7, which is capable of reducing protein disulfide bonds [90]. MaMsvR may be an important redox-sensitive transcriptional regulator in *M. acetivorans*, yet its reducing partner and the genes which it regulates are unknown.

The anaerobic origins of life have left an indelible mark upon many of life's processes. In many organisms, transcription is regulated by mechanisms which sense the presence of oxygen and ROS. The presence of Fe-S clusters in RNA polymerase may be serving toward that end by globally regulating transcription in correlation with the redox state of the cell. MsvR is a transcriptional regulator which may modify transcription based on redox state in a more specific manner. The investigation of the interaction between redox-sensitive transcriptional components and their coordinated response to oxygen contamination in anaerobic organisms will further our knowledge concerning the evolution of life's information processing systems.

It is the goal of the work detailed herein to further the construction of an *in vitro* transcription system for *M. acetivorans* and to investigate the reducing partner of and the genes targeted by MaMsvR. The presence or absence of Fe-S clusters may affect transcription on the global level by modifying the activity of RNAP at certain or all promoter sites. The establishment of an *in vitro* transcription system will allow for the investigation of the effects of cluster constitution on the activity of RNAP. The purification of *M. acetivorans* TATA-binding protein and transcription factor B is the first step towards constructing such a system. MaMsvR may serve an important role in the regulation of oxygen stress recovery mechanisms in *M. acetivorans*, as has been demonstrated by MsvR of other methanogens. However, little is known about this regulatory system. The investigation of an *M. acetivorans* thioredoxin as a reducing partner to MaMsvR is explored. Additionally, the potential of other MaMsvR target genes is assessed.

Table 1.1. Overview of general aspects of *Archaea* analogous to the other domains.

Process in *Archaea*	Analogous domain
Morphology	*Bacteria*
Metabolism	*Bacteria*
Genome architecture	*Bacteria*
Transcriptional machinery	*Eukarya*
Promoter site architecture	*Eukarya*
Transcriptional regulation	*Bacteria*
Translation	*Bacteria/Eukarya*

Table 1.2. Conservation of archaeal RNAP subunits across the domains. Five core subunits are conserved across all three domains. Other archaeal RNAP subunits are conserved in eukaryal RNAP but not bacterial RNAP.

Subunit of Archaeal RNAP	Analogue in Eukaryal RNAP	Analogue in Bacterial RNAP
A	Rpb1	β'
B	Rpb2	β
D	Rpb3	α^{I}
E	Rpb7	-
F	Rpb4	-
H	Rpb5	-
K	Rpb6	ω
L	Rpb11	α^{II}
N	Rpb10	-
P	Rpb12	-
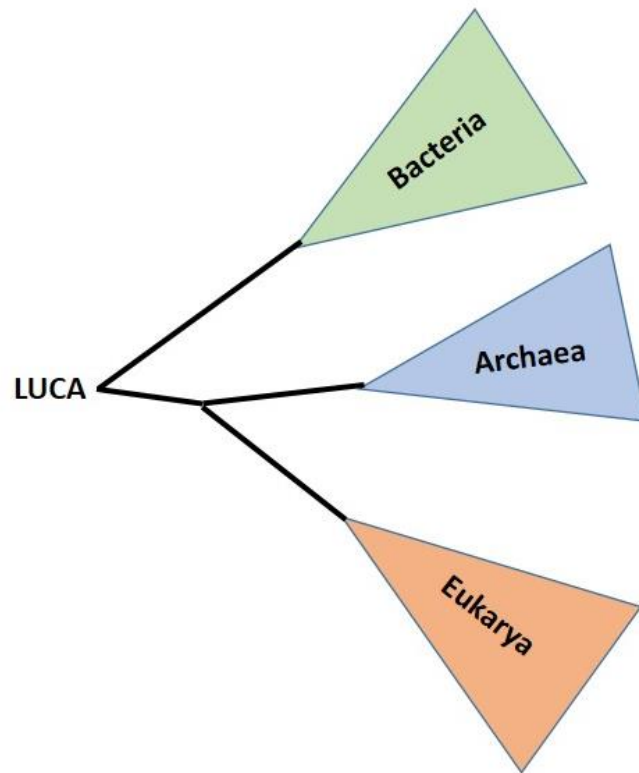
Figure 1.1. The tree of life with respect to the three domains. The tree of life as constructed by 16S rRNA sequencing. The domain *Archaea* shares a more recent common ancestor with the domain *Eukarya* than with the domain *Bacteria*. This precludes the original categorization of *Archaea* and *Bacteria* as one monophyletic group of prokaryotes. LUCA (last universal common ancestor) is the most recent community of organisms from which all extant live has evolved.
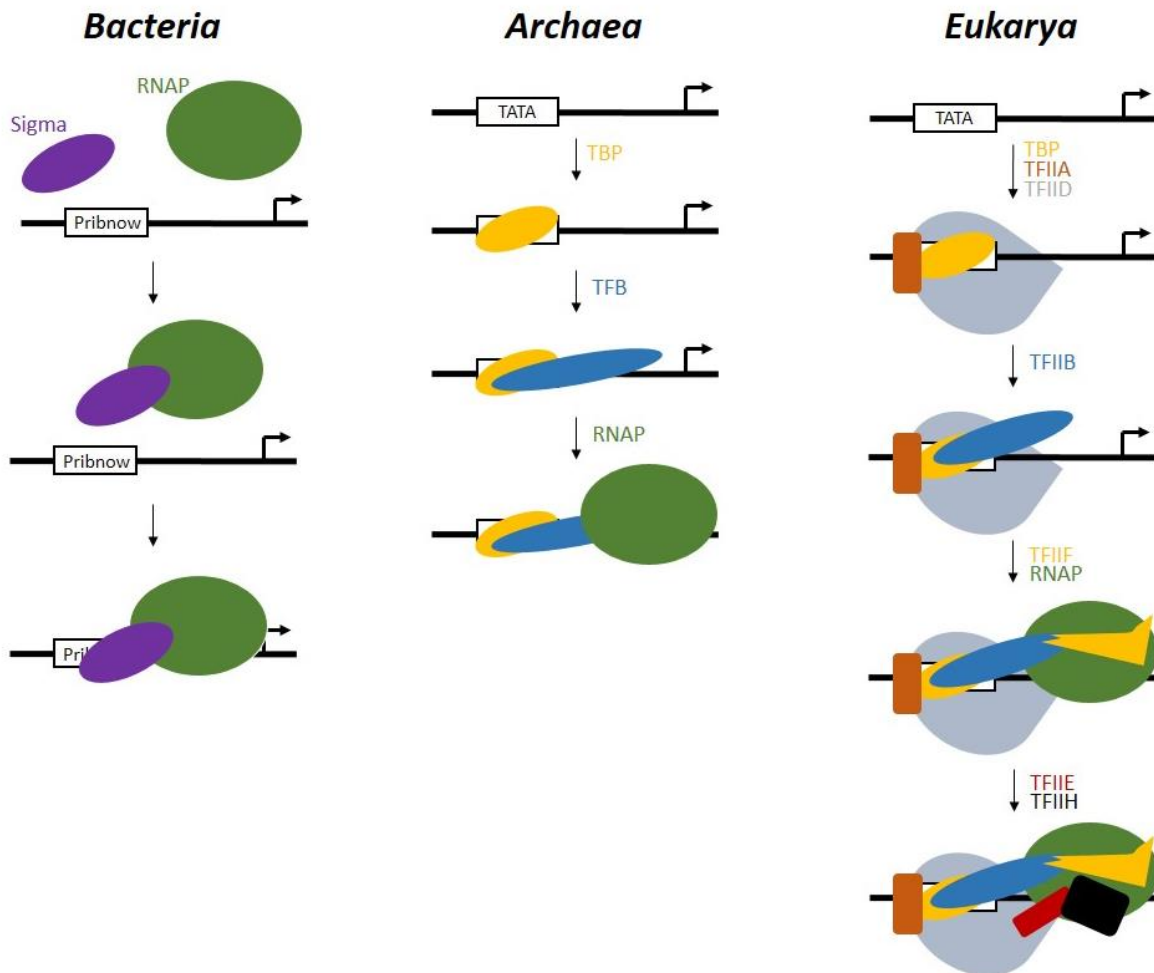
Figure 1.2. Models of transcription initiation for the three domains. Left panel- bacterial transcription initiation. A sigma factor specific to a subset of genes including the particular gene undergoing transcription will bind to RNAP forming the RNAP holoenzyme; the holoenzyme will then bind to the Pribnow box consensus sequence ~10 base pairs from the initiation site and transcription may commence. Center panel- archaeal transcription initiation. Two general transcription factors, TATA-binding protein (TBP) and transcription factor b (TFB) prepare the promoter site for the arrival of RNAP. Right panel- eukaryotic transcription initiation for the RNAP II system. A larger cast of general transcription factors prepare the promoter site for the arrival of RNAP II. Figure adapted from [25].
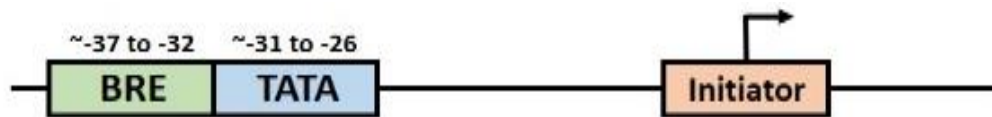
Figure 1.3. Eukaryal/archaeal gene promoter region. The TATA box and β-recognition element (BRE) are located upstream of the initiation site in both eukaryotes and archaea. TATA-binding protein (TBP) binds to the TATA box and transcription factor B (TFB) to the BRE to prepare the promoter region for RNAP binding to begin transcription at the initiator element. Figure adapted from [91].
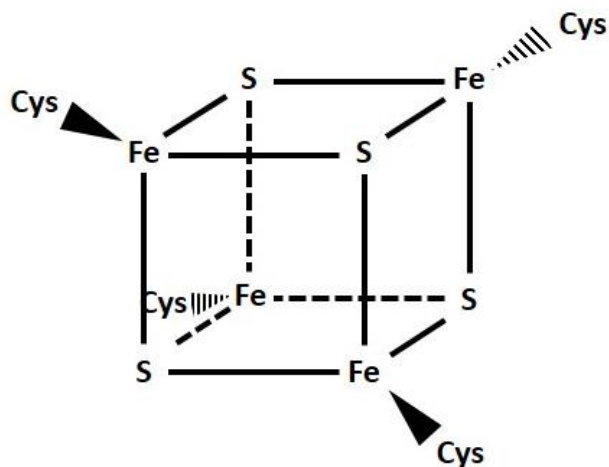
Figure 1.4. Structure of an [4Fe-4S] cluster. Iron-sulfur clusters of [4Fe-4S] type are organized in a cubic structure with an iron or sulfur located at each corner. Each iron is generally coordinated with four sulfur atoms, three of which are the inorganic sulfurs of the cluster, the fourth often being a component of the cysteine residue ligand of the protein to which the cluster is associated.
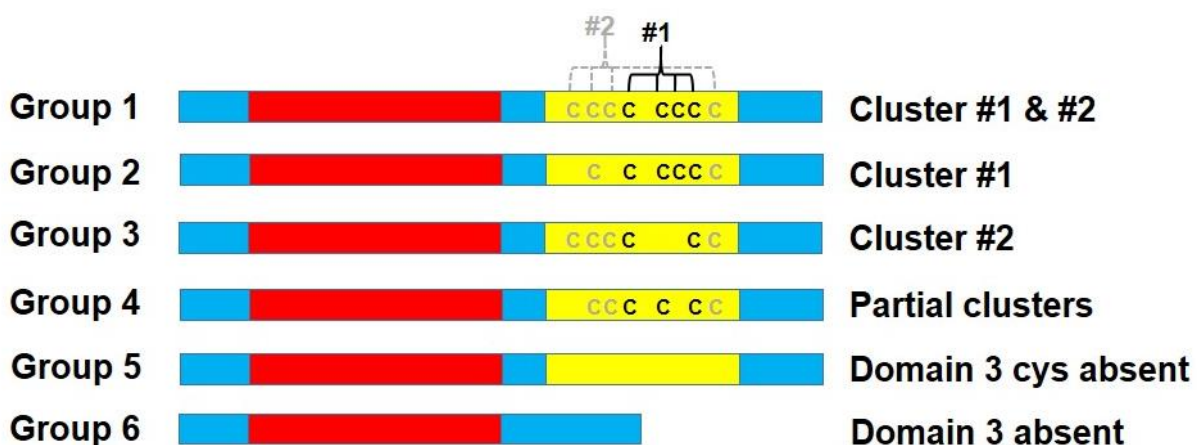
Figure 1.5. Schemata illustrating the diversity of RNAP subunit D domain 3 architecture. Archaeal RNAPs can be grouped by the architecture of their third subunit D (RpoD) domains and the number of whole or partial [4Fe-4S] cluster binding motifs for cluster #1 (cysteine residues in black) and cluster #2 (cysteine residues in grey). Domain 1 of RpoD is represented as blue, domain 2 as red, and domain 3 as yellow. *Sulfolobus solfataricus* RpoD contains a single cluster binding motif, referred to as cluster #1, and is thus a member of group 2. *Methanosarcina acetivorans* RpoD contains binding motifs for both clusters, being a member of group 1. Some eukaryal Rpb3/AC40 have been observed to contain binding motifs for cluster #1. Figure adapted from [83].
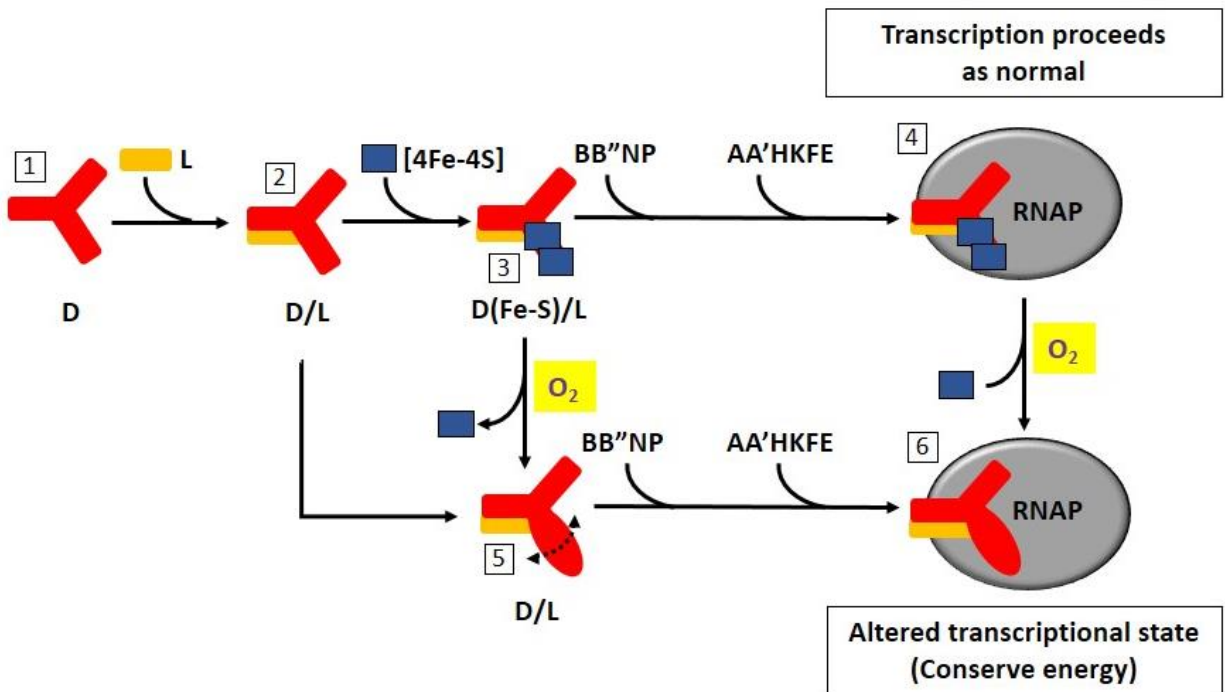
Figure 1.6. Proposed model for Fe-S cluster integration in RNA polymerase. 1- Subunit D (RpoD). 2- D-L heterodimer. 3- D-L heterodimer with Fe-S clusters incorporated. 4- Full RNAP holoenzyme. 5- D-L heterodimer without Fe-S clusters, potentially altering the conformation of the complex. 6- RNAP without Fe-S clusters. The full RNAP holoenzyme and RNAP lacking Fe-S clusters likely effect different transcriptional states. A lack of Fe-S clusters in RNAP may alter the transcriptional activity of the cell by reducing transcription globally to conserve energy.

Chapter Two: Toward the development of an *in vitro* transcription

system in *Methanosarcina acetivorans*

<u>Introduction</u>

Transcription, the initial step in gene expression, is performed in all organisms by the enzyme

RNA polymerase (RNAP). The initiation of transcription is achieved differently in the three domains of

life (**Fig 1.2**). In *Bacteria*, RNAP must first form a holoenzyme complex with one of a group of specificity

proteins known as sigma factors before binding to a promoter sequence and subsequent initiation of

transcription [92]. Bacterial RNAP is composed of five core subunits ($\alpha_2\beta\beta\omega$), with the relevant sigma

factor binding reversibly as a sixth subunit to complete the complex and form the RNAP holoenzyme

[93]. Individual sigma factors endow recognition of a particular promoter site consensus sequence,

effectively regulating transcription by determining the set of genes to which a particular RNAP

holoenzyme will have affinity [94]. In *Escherichia coli*, for example, σ^{70} is the predominant sigma factor

which controls the transcription of genes involved in essential processes [95, 96], but six other sigma

factors are also present and are largely involved in the control of genes which confer adaptation of the

cells to different environmental conditions [97]. Sigma factors interact with the β and β' subunits to bind

to RNAP [98]. Formation of the RNAP holoenzyme reveals DNA binding regions on the complexed sigma

factor. These revealed regions interact with a targeted promoter site at consensus sequence regions

centered at -35 and -10 relative to the transcriptional start site, forming a closed promoter complex

[99]. A conformational change in RNAP following the binding to DNA melts the DNA strand in a region

extending from -10 to the transcriptional start site, forming the open promoter complex [100].

Transcription begins and, when the nascent RNA transcript is elongated to 8-10 nucleotides, sigma

factor is released or partially released as RNAP proceeds with transcription [101]. *Bacteria* utilizes a

specificity protein system wherein transcription initiation is directed to distinct classes of promoters by

sigma factors which confer affinity for their respective RNAP holoenzymes to certain consensus sequences.

Archaea and Eukarya, however, have an altogether different system of transcription initiation. Like Bacteria, Archaea have a single RNAP, whereas eukaryotes have at least three RNAP systems (fourth and fifth RNAPs can be found in plants), each responsible for the transcription of non-overlapping subsets of genes [11]; RNAPs in Archaea and Eukarya are alike in terms of structure and subunit composition [12]. Transcription in both Archaea and Eukarya follow similar procedures in initiation: a suite of general transcription factor proteins prepare a promoter site for the reception of RNAP before transcription commences. The binding cascade of general transcription factor proteins is similar for archaea and eukaryotes but varies in the number of proteins involved. While eukaryotic initiation requires at least six transcription factor proteins [13, 14], only two transcription factor proteins, TATA-binding protein and transcription factor B are required for initiation in Archaea [15, 16].

TATA-binding protein (TBP) recognizes an A- and T-rich element ~25 nucleotides upstream from the transcriptional start site known as the TATA-box [17]. TATA-binding protein is necessary for transcription in archaea and is the most general transcription factor in eukaryotes- whereas some eukaryal transcription factor proteins are specific to their respective RNAP, TBP is required by each eukaryotic RNAP for transcription [102]. TBP is highly conserved and, in eukaryotes, a TBP from one organism can be interchangeable with TBP from another with organisms as diverse as mammals and yeast [15]. Eukaryotes and archaea share near-perfect conservation of the amino acids engaged in contact with DNA [103]. TBP has a saddle-shaped structure endowed by its two symmetrical subdomains [104]. The groove of the saddle serves as the interface with DNA, interacting with the ~8 base pair TATA-box in the minor groove of the DNA. Two pairs of phenylalanine residues contact the DNA in between the first and last nucleotide pairs of the TATA-box, widening the groove [105]. Upon binding, archaeal TBP bends the DNA strands at the TATA-box ~65° [106].

In *Archaea*, TBP prepares the promoter site for binding with transcription factor B (TFB). A purine-rich β-recognition element (BRE) is located ~35 base pairs upstream of the transcriptional start site and serves as the binding site for TFB upon the TBP-DNA complex to complete the preinitiation complex (PIC) [20, 21]. TFB binds to the TBP-DNA complex on the side of the DNA opposite TBP and contacts both the DNA and the c-terminal half of TBP [25]. It has been shown that the BRE is necessary for the correct orientation of the PIC- TFB is capable of binding to the TBP-DNA complex in the opposite direction when a BRE is absent from the promoter [22]. It is the c-terminal domain of TFB which recognizes and binds to the BRE and thus ensures the correct orientation. The n-terminal domain of TFB forms a zinc-ribbon structure which assists in recruiting RNAP to the promoter site [107]. In *Archaea*, the addition of TFB to the TBP-DNA complex completes the PIC and prepares the promoter for reception of RNAP. In eukaryotes, transcription initiation follows a similar cascade involving the addition of transcription factor proteins to the promoter site, but the eukaryal PIC incorporates at least six different proteins [108]. After construction of the PIC, the recruitment of RNAP to the promoter will directly proceed initiation of transcription.

Archaea and *Eukarya* possess RNAPs which are alike in terms of structure and subunit composition [12]. Notably, subunit RpoD of archaeal RNAP and the homologous Rpb3 subunit of eukaryal RNAP II have been observed to incorporate [4Fe-4S] cluster binding motifs in several organisms. The purpose the Fe-S clusters in RNAP is as yet unclear. Potentially, the clusters may allow for RNAP to correlate the cellular redox state of the cell, as detected by the clusters and their integrity, to the transcriptional activity of the cell, as affected by conformational changes precipitated by the lability of the clusters. To investigate the effects of [4Fe-4S] clusters on RNAP activity, an *in vitro* transcription system must be constructed. The methanogen *Methanosarcina acetivorans* will be used as a model organism for its possession of dual RNAP [4Fe-4S] clusters and its position within methanogens,

a group of organisms of diverse environments and possessing of RNAPs with various combinations of whole and partial [4Fe-4S] cluster binding motifs.

A promoter-specific *in vitro* transcription system requires the cellular components necessary for promoter-driven transcription. For *Archaea*, this necessitates the purification of viable TBP, TFB, and RNAP. Promoter-specific *in vitro* transcription systems have been composed for several archaea: *Methanothermus fervidus* [109], *Pyrococcus furiosus* [110], *Sulfolobus shibatae* [111], *Methanosarcina mazeii* [112], etc. Such a system has not yet been developed for *M. acetivorans*. The *M. acetivorans* genome encodes three TBPs; MaTBP1 has been demonstrated to be essential [113], and will thus be the MaTBP used for *in vitro* transcription.

The work described herein details the purification of *M. acetivorans* TBP1 & TFB. His-tagged recombinant MaTBP1 & MaTFB were each cloned into *E. coli* and purified via affinity chromatography. MaTFB was expressed as inclusion bodies, necessitating the denaturing and subsequent renaturing of the protein. Electrophoretic mobility shift assays were performed with MaTBP1 and MaTFB in an attempt to assess DNA-binding functionality of the proteins. The purification of MaTBP1 and MaTFB is the first step towards a *M. acetivorans* promoter-specific *in vitro* transcription system.

Materials & Methods

Cloning. PCR was used to amplify *tbp1* and *tfb* from *M. acetivorans* C2A genomic DNA. The forward primer for the amplification of *tbp1*, NdeTbpF (Integrated DNA Technologies), contained the sequence for an *Nde*I restriction site (5'-GGT GGT CAT ATG AGC GAA TCT AGC ATT AAA ATT G-3'); the reverse primer for *tbp1*, XhoTbpR, contained an *Xho*I restriction site (5'-GGT GGT CTC GAG TTA TAA AAG TCC CAT GTT ATC AAG CTG C-3'); the forward primer for the amplification of *tfb*, NdeTfbF, contained the sequence for an *Nde*I restriction site (5'-GGT GGT CAT ATG GTA GAA GTC GAA AGA GTT CGC TAT-3'); the reverse primer for *tfb*, XhoTfbR (Integrated DNA Technologies), contained an *Xho*I restriction site (5'-

GGT GGT CTC GAG TTA GAG GAT AAT CTC TAT ATC AGT TTC-3'). PCR was executed using Phusion High-Fidelity DNA Polymerase (New England BioLabs) with the following parameters: initial step at 98 $^{\circ}$C for 2 minutes; 30 cycles of: denaturing at 98 $^{\circ}$C for 10 second, annealing at 62 $^{\circ}$C for 10 seconds, elongation at 72 $^{\circ}$C for 15 seconds; final extension at 72 $^{\circ}$C for 5 minutes. The PCR products were digested with *Nde*I and *Xho*I restriction enzymes (New England BioLabs). The digested products were ligated with similarly-digested pET28a plasmid (Novagen) using T4 DNA Ligase (New England BioLabs), producing plasmids pMaTBP1 (with *tbp1*) and pMaTFB (with *tfb*). The ligation mixtures were transformed separately into *Escherichia coli* strain DH5α and colonies containing the plasmid were selected for on LB agar plates containing kanamycin (kan). Plasmids were screened for the presence of insert by restriction digests, verified by sequencing, and transferred to *E. coli* strain Rosetta DE3 for expression.

Expression. MaTBP1 and MaTFB were over-expressed from *E. coli* Rosetta DE3 cells in LB broth supplemented with kan and chloramphenicol (chlor). A number of conditions were attempted to optimize MaTBP1 and MaTFB expression (see Results and Discussion section); following are the conditions used for purifications. Cultures were incubated at 37 $^{\circ}$C, shaking at 240 rpm. Protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) when the cultures reached an OD_{600} of ~0.6. Expression of MaTBP1 proceeded for four hours shaking at 240 rpm at room temperature. Expression of MaTFB proceeded for 16 hours shaking at 240 rpm at 16 $^{\circ}$C. Cells were harvested by centrifugation at 17 500 x *g* for 10 min at 4 $^{\circ}$C and frozen.

MaTBP1 affinity chromatography. MaTBP1 cells were thawed with a few crystals of benzamadine hydrochloride and DNAse I and were resuspended with 3 ml per g of cells in buffer A (50 mM HNa_2PO_4, pH 8, 300 mM NaCl). The cells were lysed by three passages through a French pressure cell at over 110 MPa. The lysate was centrifuged at 35 000 x *g* for 30 minutes at 4 $^{\circ}$C. The supernatant was filtered through a 0.45 μm cellulose acetate syringe filter. Filtered lysate was applied to a pre-equilibrated 1-ml Ni-resin column (Genscript). The column was washed with 5 column volumes (CV) of buffer A and 5 CV

of buffer A containing 10 mM imidazole. The protein was eluted from the column with 2 CV of buffer A

containing 250 mM imidazole. The Ni-eluate was applied to a pre-equilibrated 1-ml heparin column (GE

Healthcare) on a low pressure chromatography system (Biologic) at a flow rate of 0.5 ml per minute. The

column was then washed with 30 CV of buffer B (10 mM H_2NaPO_4, pH 7) followed by 30 CV of buffer B

with a gradient from 0 to 2 M NaCl. Fractions (1 ml) were collected throughout this entire procedure.

Fractions containing MaTBP1 were buffer exchanged into buffer C (50 mM Tris, pH 7.2, 150 mM NaCl,

10% glycerol), concentrated using a Vivacell concentrator (Sigma), and frozen for storage.

MaTFB affinity chromatography. MaTFB cells were thawed with a few crystals of benzamadine

hydrochloride and DNAse I and were resuspended with 3 ml per g of cells in buffer A (50 mM HNa_2PO_4,

pH 8, 300 mM NaCl). The cells were lysed by three passages through a French pressure cell at over 110

MPa. The lysate was centrifuged at 35 000 x g for 30 minutes at 4 ^{o}C. The supernatant was removed and

the pellet resuspended in buffer D (6 M guanidine hydrochloride (GuHCl), pH 7.4, 127 mM NaCl, 2.7 mM

KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10 μM ZnCl, 10 μM $MgCl_2$, 5% glycerol). The insoluble lysate was

incubated at room temperature in buffer D for one hour. During incubation, the mixture was sonicated

for two minutes at twenty minute intervals. The insoluble lysate was applied to a pre-equilibrated 1-ml

Ni-resin column (Genscript). The column was washed with 5 column volumes (CV) of buffer D and 5 CV

of buffer D containing 10 mM imidazole. The protein was eluted from the column with 2 CV of buffer D

containing 250 mM imidazole.

Renaturation of MaTFB by stirred cell concentration. MaTFB Ni-eluate (~2 ml) was placed in a stir cell

concentrator. A series of dilutions to 50 ml and subsequent concentrations to 5 ml using a stirred cell

concentrator (Millipore) was performed; dilutions were performed in the following order by adding the

following buffers to the sample at a rate of one drop every two seconds: buffer E (4 M GuHCl, pH 7.4,

127 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10 μM ZnCl, 10 μM $MgCl_2$, 5% glycerol),

buffer F (2 M GuHCl, pH 7.4, 127 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 10 μM ZnCl,

10 μM MgCl$_2$, 5% glycerol), buffer G (1 M GuHCl, pH 7.4, 127 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 10 μM ZnCl, 10 μM MgCl$_2$, 5% glycerol), buffer H (pH 7.4, 127 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 10 μM ZnCl, 10 μM MgCl$_2$, 5% glycerol), and buffer H again. Precipitated proteins were removed during dilution steps *pro re nata*. The sample was buffer exchanged into buffer J (50 mM Tris, pH 7.2, 150 mM NaCl, 10 μL ZnCl, 10 μL MgCl$_2$, 10% glycerol), concentrated using a Vivacell concentrator, and frozen for storage.

Renaturation of MaTFB by dialysis. MaTFB Ni-resin eluate was loaded into a 3-ml capacity, 3,500 MWCO Slide-A-Lyzer Dialysis Cassette (Thermo) and placed in 200 ml of buffer E and stirred for 2 hours at 10 °C. The cassette containing the sample was transferred to 200 ml of buffer F and stirred for 2 hours at 10 °C. The cassette containing the sample was transferred to 200 ml of buffer K (1.5 M GuHCl, pH 7.4, 127 mM NaCl, 2.7 mM KCl, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 10 μM ZnCl, 10 μM MgCl$_2$, 5% glycerol) and stirred for 2 hours at 10 °C. The cassette containing the sample was transferred to 200 ml of buffer G and stirred for 2 hours at 10 °C. The cassette containing the sample was transferred to 200 ml of buffer H and was incubated overnight while stirring at 10 °C. Protein was buffer exchanged into buffer J, concentrated using a Vivacell concentrator, and frozen for storage.

Electrophoretic Mobility Shift Assays. Procedure adapted from [112]. Complimentary 50-bp oligonucleotides (Integrated DNA technologies) of P$_{mcrB}$, P$_{dnaK}$, P$_{msvRa}$, and P$_{control}$ were annealed to generate DNA probes (see appendix for oligonucleotide sequences). DNA-binding reactions were prepared by incubating 5 μL of each protein (TBP alone or TBP and TFB), 5 μL 1 μM DNA probe, and 35-40 μL of DNA binding reaction master mix (20 mM Tris pH 8, 15 mM MgCl$_2$, 120 mM KCl, 12.5 μg/ml heparin, 5 mM DTT, 10% glycerol) for 20 min at 37 °C. Binding reactions were loaded onto a pre-run 6% polyacrylamide gel in 0.5X TBE buffer and electrophoresed for 75 min at 75 V at 10 °C. Gels were stained using SYBR Gold (Life Technologies) and visualized using a Gel-Doc XR+ system (Bio-Rad Technologies).

Analytical methods. Protein concentrations were determined by the method of Bradford [114] using bovine serum albumin as a standard. Gene sequencing was performed by the University of Arkansas Division of Agriculture DNA Resource Center (Fayetteville, AR) using T7seq (5'-TAA TAC GAC TCA CTA TAG G-3'), a T7 universal sequencing primer, and T7term (5'-GCT AGT TAT TGC TCA GCG G-3'), a universal sequencing primer for the T7 terminator. Mass spectrometry of purified proteins was performed at the University of Arkansas Statewide Mass Spectrometry facility (Fayetteville, AR). SDS-PAGE were performed by standard procedures.

Results & Discussion

Expression and purification of MaTBP1. The insertion of *tbp1* into a pET28a plasmid fused with an N-terminal six-histidine tag (His-tag) under the control of a T7 promoter (producing pMaTBP1) was verified by DNA sequencing. Induction studies were performed to optimize the expression of pMaTBP1 in Rosetta DE3 *E. coli*. Cells were grown at room temperature, shaking at 240 rpm, with 1-ml samples collected at time points of 2 hours after induction, 4 hours after induction, and 16 hours after induction. Time point samples were lysed by sonication, centrifuged to separate soluble and insoluble components, and analyzed by SDS-PAGE. It was determined that 4 hours of expression yielded sufficient MaTBP1 in the soluble form for purification. Ni-resin chromatography was performed, due to the affinity of the His-tag on MaTBP1, but did not yield sufficient purity (**Fig. 2.2A, Lane 9**). Thus, heparin affinity chromatography was included in the purification process, as DNA-binding proteins often have an affinity to heparin due to its similarity to DNA in terms of structure and charge [115]. Though MaTBP1 was found in very early fractions exiting the column (**Fig 2.1**) and likely did not bind strongly to the heparin, the protein was pure after exiting the column (**Fig. 2.2B, Lane 6**). Interestingly, it exited the column in two waves- in fraction 2 and in fraction 3- potentially representing two different species of MaTBP1 with different conformational structures. The identity of MaTBP1 was confirmed by mass spectrometry.

MaTBP1 found in heparin fraction 2 was stored at a concentration of 0.35 mg/ml (15 µM); an amount of ~0.7 mg was obtained. MaTBP1 found in heparin fraction 3 was stored at a concentration of 0.25 mg/ml (11 µM); an amount of ~0.5 mg was obtained.

Expression and purification of MaTFB. The insertion of *tfb* into a pET28a plasmid fused with an N-terminal six-histidine tag (His-tag) under the control of a T7 promoter (producing pMaTFB) was verified by DNA sequencing. Induction studies were performed to optimize the expression of pMaTFB in Rosetta DE3 cells. Cells were grown, shaking at 240 rpm, with 1-ml samples collected at time points of 2 hours after induction, 4 hours after induction, and 16 hours after induction under a number of different conditions, varying the temperature and amount of IPTG used for induction (**Table 2.1**). MaTFB was predominantly expressed in the insoluble form under all conditions attempted. With little success using the Rosetta DE3 expression strain, pMaTFB was transformed into Tuner *E. coli* (Novagen) cells and further induction studies were performed. Tuner cells possess a *lac* permease mutation which allows for a more uniform distribution of IPTG among the cells of a culture to better control induction with respect to IPTG concentration. Induction studies did not yield any conditions which produced an appreciable amount of MaTFB in the soluble form (**Table 2.1**), likely MaTFB was expressed in inclusion bodies under all conditions. The presence of MaTFB in the insoluble cell lysate fraction was confirmed by mass spectrometry. Ultimately, due to the inability to express soluble MaTFB, it was decided to attempt renaturation of the protein in an effort to obtain viable MaTFB. This would require resuspending the inclusion bodies in denaturant and attenuating the concentration of denaturant in solution slowly to refold the proteins. Thus, insoluble cell lysate was resuspended in a denaturant during a 1 hour incubation with frequent sonication to break up the lysate pellet. Denaturation was first attempted in an 8 M urea solution, yet MaTFB would not bind to the Ni-resin column under these conditions. Resuspension of insoluble cell lysate in a 6 M GuHCl solution did allow for protein binding to the Ni-resin column. On-column refolding was attempted by the washing of protein-bound Ni-resin column with

buffers of gradually decreasing concentrations of GuHCl. Protein would not elute from column after renaturing washes. Off-column refolding was then attempted. Protein was eluted from Ni-resin column using a 6 M GuHCl buffer with 250 mM imidazole (see Materials and Methods). Typically, ~7 mg of protein would be present in the Ni-eluate and MaTFB generally composed ~75 % of the total protein content (**Fig. 2.3, Lane 5**). Thus, ~5.25 g of MaTFB would be present in a typical Ni-eluate before renaturation. Renaturation was first achieved by the dilution and subsequent concentration of Ni-eluate with buffers of gradually decreasing concentrations of GuHCl in a stirred cell concentrator. As the concentration of denaturant was reduced, proteins precipitated out of solution and were removed by centrifugation. These proteins included both the contaminating proteins and a majority of the MaTFB. After denaturant had been completely removed, some MaTFB remained in solution (**Fig. 2.3, Lane 6**) and was stored at a concentration of 62 µg/ml (1.6 µM); an amount of ~50 µg was obtained. Stirred cell concentrator refolding was successful once, yet subsequent attempts resulted in all MaTFB crashing out of solution with the other proteins during reduction of the denaturant. As an alternative, dialysis yielded a successful renaturation and isolation of MaTFB (data not shown); MaTFB was stored at a concentration of 0.19 mg/ml (4.9 µM), an amount of ~50 µg was obtained. Both stirred cell concentration and dialysis are methods which have proven fastidious. Each have yielded a purification but once from a number of attempts, and each recovered only ~1 % of MaTFB present in the Ni-eluate.

Functional analysis of purified proteins. In an attempt to assess the functionality of the recombinant proteins, electrophoretic mobility shift assays were performed. The promoters of *mcrB* [116], *msvR* [78], and *dnaK* [112] were chosen due to their capacity as strong promoters and their characterization in the literature; DNA probes of the promoter regions from these genes encompassing the BRE and TATA-box sites were constructed. A control lacking any consensus BRE or TATA-box was chosen from the coding DNA within *msvR*. These DNA probes were incubated with MaTBP1 and with MaTBP1 and MaTFB in attempts to observe protein-DNA complexes. No complexes were observed. Binding conditions and

electrophoresis conditions were manipulated for voltage (120 V, 100 V, 75 V, 50 V), temperature (room temperature and 10°C), and duration (at the various voltages, the proteins were run the entire length of the gel, half length, and quarter length), yet no complexing was observed. As a control to assess EMSA conditions, MaMsvR EMSA trials were conducted (see chapter three). MaMsvR-DNA complexing was observed, confirming correct EMSA conditions for MsvR. Disheartingly, the lack of complexing for MaTBP1 and MaTFB may be due to a lack of functionality of one or both of the proteins. Though, because MaTFB requires MaTBP1 to bind DNA, a dysfunctional MaTBP1 would preclude successful complexing of MaTFB with the DNA in any fashion. Thus, even if EMSA conditions were correct, such a negative result may not necessarily indicate dysfunctionality in MaTFB. More hopefully, EMSA conditions may not yet have been found which were favorable to the complexing of MaTBP1 and MaTFB with the probes or, more likely, conditions may have been such that the complexes formed were unstable such that none or a negligible amount of the complexes survived the entire run of electrophoresis. Thus, functionality of the purified proteins was either negative or inconclusive. A more conclusive measure of protein functionality may be determined from promoter-specific *in vitro* transcription assays with the proteins.

Conclusion. MaTBP1 and MaTFB were produced and purified, yet functional analysis of the proteins was elusive or negative. MaTBP1 and MaTFB were purified with the express purpose of use in an *in vitro* transcription assay with RNAP. A run-off transcription assay would incorporate a linearized plasmid upon which restriction digestion had incised a gene with a strong promoter. Incubation of this linearized plasmid with MaRNAP, MaTBP1, and MaTFB should yield a preinitiation complex upon the promoter. The addition of nucleotides with sufficient incubation time would allow for transcription to initiate and proceed. Gel electrophoresis would then reveal the presence of transcripts of the correct size predetermined by the incised gene if promoter-driven transcription had been performed. In such a way, MaTBP1 and MaTFB could be used in an *in vitro* transcription assay. An example promoter-driven *in*

vitro transcription assay protocol from Thomsen *et al.*, 2001 [112] requires 4.9 pmol of TBP and 1.2 pmol

of TFB for transcription assay reaction mixtures. MaTBP1 and MaTFB purified in this work are both

stored at concentrations which would be sufficient for such an assay. This system could then be used to

investigate the function of [4Fe-4S] clusters on MaRNAP activity by attempting transcription under

oxidizing and reducing conditions. A complete *in vitro* transcription system would also be capable of

looking at other aspects of transcription. For instance, the function of MaMsvR as a redox-sensitive

transcriptional regulator could be further investigated by observing the different effects of oxidized

MaMsvR and reduced MaMsvR on the transcriptional activity of *msvR* and of other potential MaMsvR

target genes.

Table 2.1. Overview of MaTFB induction study conditions. Induction studies were performed with Rosetta DE3 and Tuner cells. Incubation temperatures attempted were 16 ^{o}C and room temperature (25 ^{o}C). Concentrations of IPTG for induction attempted were 0.1 mM IPTG and 0.5 mM IPTG. An "x" indicates that the specific induction conditions yielded no appreciable soluble MaTFB.

	16 ^{o}C	16 ^{o}C	25 ^{o}C	25 ^{o}C
	0.1 mM IPTG	0.5 mM IPTG	0.1 mM IPTG	0.5 mM IPTG
Rosetta DE3	x	x	x	x
Tuner	x	x	x	x
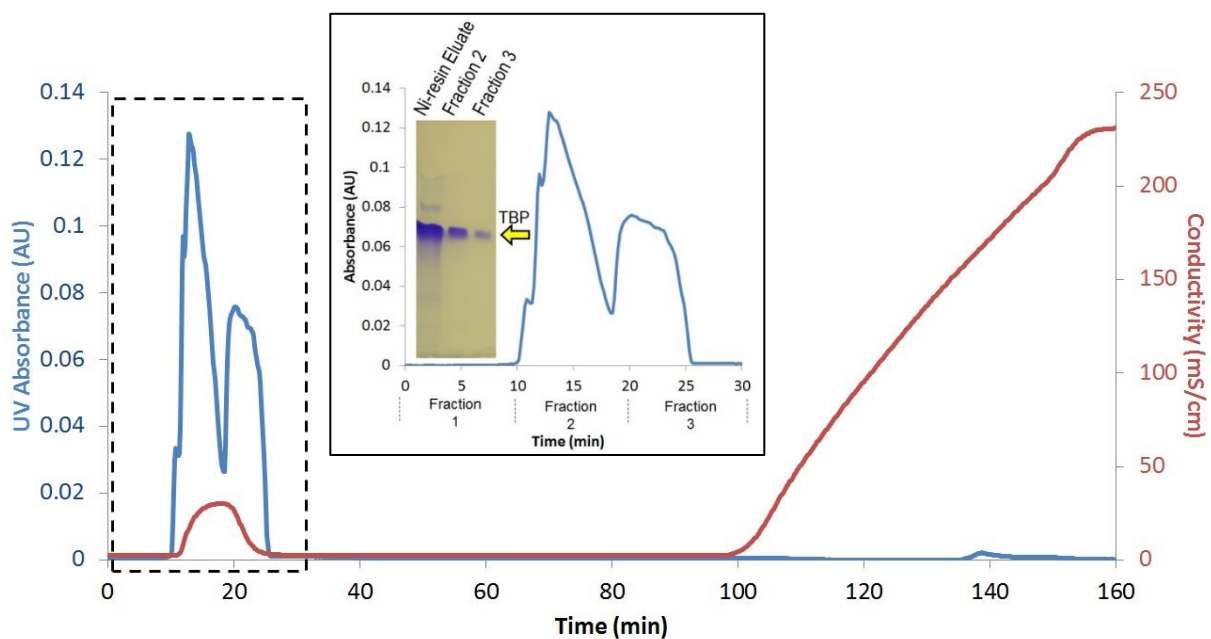
Figure 2.1. Chromatogram of MaTBP1 heparin purification. The Abs$_{280}$ (blue) and conductivity (red) of liquid exiting the LP are illustrated in the chromatogram. UV absorbance peaks indicate protein leaving the column as two distinct species. Inset- chromatogram blown up and SDS-PAGE of fractions 2 and 3 from heparin purification alongside Ni-eluate. Dashed box indicates the portion of the graph depicted in the inset image.
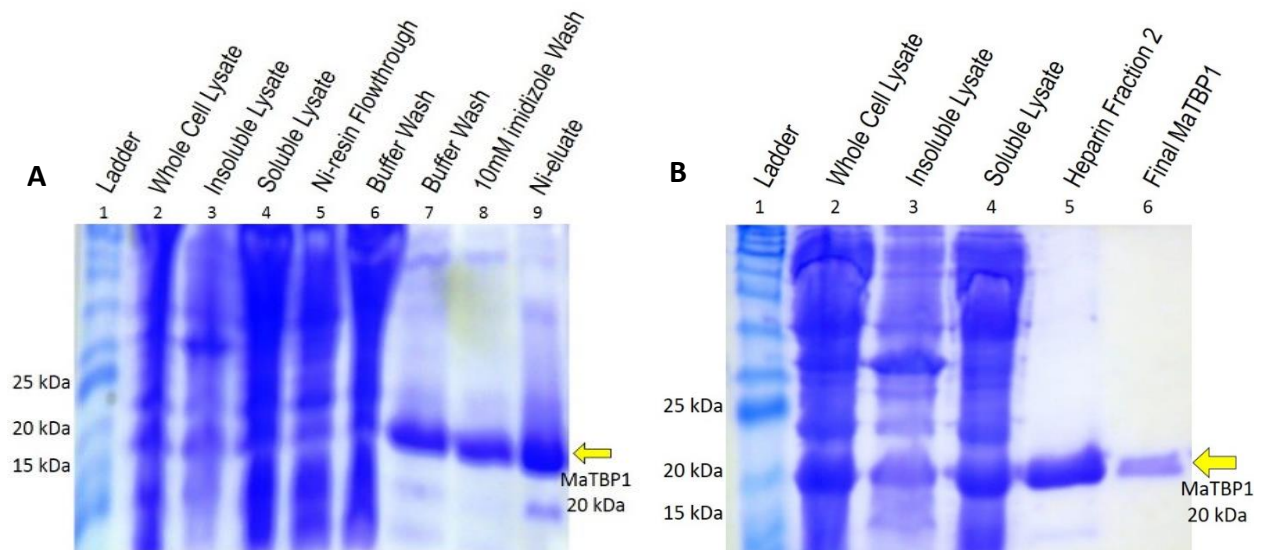
Figure 2.2. SDS-PAGE of MaTBP1 purification. A- SDS-PAGE of the steps from MaTBP1 Ni-resin chromatography. B- SDS-PAGE of select samples from entire MaTBP1 purification procedure. The band representative of MaTBP1 is indicated by a yellow arrow.
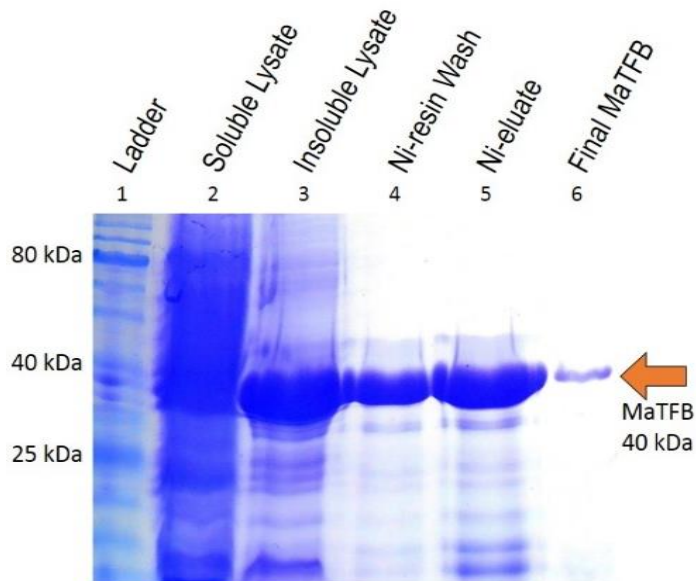
Figure 2.3. SDS-PAGE of MaTFB purification. SDS-PAGE of select samples from entire MaTFB purification procedure involving renaturation by stirred cell concentrator. Cells were lysed and soluble/insoluble components were separated by centrifugation (lanes 2 & 3). Insoluble components were resolubilized in GuHCl and purified by Ni-resin affinity chromatography (lanes 4 & 5). The concentration of GuHCl in Ni-eluate was slowly decreased through dilution and subsequent concentration using a stirred cell concentrator. During attenuation of denaturant, contaminating proteins precipitated and were removed from solution, leaving pure MaTFB (lane 6). The band representative of MaTFB is indicated by an orange arrow.

Chapter Three: The *Methanosarcina acetivorans* thioredoxin system activates DNA binding

of the redox-sensitive transcriptional regulator MsvR

Introduction

Methanogenic metabolism involves a number of Fe-S cluster enzymes which are vulnerable to oxygen [68]. Exposure to oxygen can be fatal to the activity of many methanogenesis enzymes, resulting in a dramatic reduction of energy conservation. Oxygen will not only precipitate the collapse of Fe-S clusters and the deactivation of affected proteins, but will produce reactive oxygen species (ROS), such as superoxide (O_2^-) and hydrogen peroxide (H_2O_2). ROS are deleterious to many cellular structures. For instance, hydrogen peroxide can oxidize the thiol groups of neighboring cysteine residues, producing disulfide bonds and disrupting protein structure [117]. These damaging processes can result in a loss of metabolic activity and eventually lead to cell death, requiring mechanisms to decrease the production of ROS, actively remove ROS, and repair damaged proteins [67]. While many methanogens are capable of surviving oxygen exposure, cellular activity and energy conservation are attenuated during times of oxidative stress. Among the more aerotolerant methanogens are members of the order *Methanosarcinales* [118, 119]. The genomes of sequenced *Methanosarcinales* species encode a large number of putative antioxidant and repair proteins [120]. Mechanisms used by *Methanosarcina* sp. to sense and respond to oxidative stress, however, are not fully understood.

The formation of disulfides in proteins can be used to determine the redox state of a cell in an attempt to monitor the presence of deleterious ROS. Transcriptional regulators, such as OxyR in *Eschericihia coli* and other bacteria, use the formation of disulfide bonds to monitor the presence of H_2O_2 [121]. OxyR possesses cysteine residues with redox-sensing thiols; these thiols are oxidized in the presence of H_2O_2, forming disulfides and modifying the conformation of OxyR. This activation of OxyR increases the expression of H_2O_2 scavengers, Fe-S cluster repair enzymes, and thiol redox buffer

40

systems- oxidative stress recovery measures which combat the presence of the ROS. When the presence

of H_2O_2 has been reduced to acceptable levels, the inactivation of OxyR is accomplished by the reduction

of its redox-sensitive thiols by glutaredoxin 1, with reductant supplied by glutathione/glutathione

reductase and NADPH [122]. Similar transcriptional regulators monitoring H_2O_2 levels have been

identified in eukaryotes [123]. A redox-sensing transcriptional regulator (MsvR) has recently been

discovered in methanogens. *Methanothermobacter thermautotrophicus* MsvR (MtMsvR) regulates the

expression of $F_{420}H_2$ oxidase (FpaA) and itself by redox-sensitive binding to the promoter region of *fpaA*

and *msvR* [78]. Under reducing conditions, MtMsvR serves as a negative regulator to repress expression

of *fpaA* and *msvR*. Oxidation of MsvR results in the induction of *fpaA* and *msvR*.

Similarly, MsvR from *Methanosarcina acetivorans* (MaMsvR), was shown to bind to its own

promoter (P_{msvR}) only under reducing conditions [89]. The C-terminal V4R effector domain of MaMsvR

contains cysteine residues with thiol groups which, when reduced to disulfide(s), have been shown to

abrogate the binding of MaMsvR to the P_{msvR} promoter region. The cysteines within the V4R domain

(C206, C225, C232, and C240) are postulated to function in redox-sensing, whereby thiol-disulfide

exchange causes conformation changes which alter the ability of MaMsvR to bind an inverted repeat

sequence motif ($TTCGN_{7-9}CGAA$) upstream of P_{msvR} [89]; several genes of the *M. acetivorans* genome

host promoter sites located near similar MaMsvR binding motifs, indicating them as candidate genes for

which MaMsvR may serve as a redox-sensitive transcriptional regulator. Using the disulfide-reducing

agent dithiothreitol (DTT), MaMsvR disulfide can be reduced and MaMsvR binding of P_{msvR} can be

restored *in vitro*. However, the physiological reducing system used to restore MaMsvR binding once

oxygen/ROS have been removed is yet unknown. The majority of organisms accomplish disulfide

reduction using thioredoxin and/or glutaredoxin systems [124].

A functional glutaredoxin system is unlikely to be present in methanogens, as they lack

glutathione [125, 126]. However, disulfide reduction in methanogens may be mediated by thioredoxin

systems. Thioredoxins are small (~12 kDa) proteins which possess a CXXC active site motif necessary for disulfide oxidoreductase activity [124]. In the thioredoxin system, thioredoxin (trx) receives reducing equivalents from thioredoxin reductase (TrxR) with NADPH as the electron donor. Recent evidence indicates the majority of methanogens contain thioredoxins [90, 127]. Seven putative Trx homologs (MaTrx1-7) and a single TrxR homolog (MaTrxR) are found in *M. acetivorans*. Recent evidence revealed that *M. acetivorans* contains at least three functional Trxs (MaTrx2, MaTrx6, and Matrx7) and a complete NADPH-dependent thioredoxin system comprised of MaTrxR and MaTrx7 [90].

In the following research, it is shown that the *M. acetivorans* NADPH-dependent thioredoxin system can reduce disulfides in oxidized MaMsvR and restore P*msvR*-binding activity, indicating that the thioredoxin system is the physiological MaMsvR disulfide reducing system. Potential gene MaMsvR binding sites are tested for their redox-sensitive binding with MaMsvR.

Materials & Methods

Protein purification and manipulation. Recombinant MaTrxR and MaTrx7 were expressed in *E. coli* and purified to homogeneity as previously described [90] by the Karr lab. Strep-tagged MaMsvR was expressed in *E. coli* and purified to homogeneity as previously described [89] by Addison McCarver. H_2O_2-oxidized MaMsvR (MaMsvR$_{ox}$) was prepared by incubation of MaMsvR with 100-fold molar excess of H_2O_2 in buffer A (20 mM Tris pH 8, 15 mM $MgCl_2$, 120 mM KCl, 12.5 µg/ml heparin, 10% glycerol) for 30 minutes. Residual H_2O_2 was removed by buffer exchange into buffer A using a NAP5 column (GE Healthcare). DTT-reduced samples of MaMsvR were prepared by incubating 100 µM MaMsvR$_{ox}$ in buffer A containing 10 mM DTT for 20 min at room temperature. Residual DTT was removed using a NAP5 column. The ability of the thioredoxin system to reduce MsMsvR$_{ox}$ was assayed by incubation of 10 µM MaMsvR$_{ox}$ with 1 mM NADPH, 0.5 µMMaTrxR, and 2.5 µM MaTrx7 for 1 hour at 37°C in buffer A. Protein

concentrations were determined by both the Bradford assay and using fluorescence by Qubit protein assay following the manufacturer's instructions (Invitrogen).

Electrophoretic Mobility Shift Assay (EMSA). Complimentary 50-bp oligonucleotides (Integrated DNA technologies) were annealed to generate the P_{msvR} DNA probe for use with MaTrx in EMSAs [89] (see appendix for oligonucleotide sequences). DNA-binding reactions were prepared by incubating 100 nM P_{msvR} with 8 µM MsvR in buffer A for 20 min at $37^{o}C$. Binding reactions were loaded onto a pre-run 6% polyacrylamide gel in 0.5X TBE buffer and electrophoresed for 75 min at 75 V at $10^{o}C$. Gels were stained using SYBR Gold (Life Technologies) and visualized using a Gel-Doc XR+ system (Bio-Rad Technologies). Separately, other promoter regions were assayed. Complimentary 50-bp oligonucleotides (Integrated DNA Technologies) were annealed to generate P_{0502}, P_{0829}, P_{2139}, P_{2689}, P_{3322}, and P_{4164} probes (see appendix for oligonucleotide sequences). Probes were used in EMSAs as described.

Quantitation of thiols in MaMsvR. Aliquots of MaMsvR-containing samples used in EMSAs were analyzed for total thiol content using DTNB [128]. MaMsvR was denatured and thiols quantified by the addition of 10 µL of MaMsvR-containing sample to 90 µL of 6M guanidine-HCl in 100 mM KPO_4, pH 7.8 containing 175 µM DTNB. Samples were incubated anaerobically for 15 min at room temperature and the absorbance at 412 nm was recorded. The number of thiols per MaMsvR monomer was calculated based on the concentration of TNB using $\varepsilon_{412}=13,700$ M^{-1} cm^{-1} [128]. All samples were analyzed in triplicate. The background amount of thiols contributed by the denatured thioredoxin system was determined in samples containing NADPH, MaTrxR, and MaTrx7, but without MaMsvR.

<u>**Results & Discussion**</u>

MaMsvR contains ten cysteine residues, with two located in the DNA-binding domain, four in the V4R domain, and the remaining four located in the linker domain [89]. The cysteines within the V4R domain (C206, C225, C232, and C240) are postulated to function in redox-sensing, whereby thiol-

43

disulfide exchange causes conformation changes which alter the ability of MaMsvR to bind an inverted repeat sequence motif (TTCGN$_{7-9}$CGAA) upstream of P$_{msvR}$. Three of the residues (C206, C232, and C240) are conserved in all MsvR homologs [78]. Wild-type MaMsvR does not bind to P$_{msvR}$ whereas a MaMsvR C206A variant does, indicating that C206 is critical for the redox-sensitive binding of MaMsvR to P$_{msvR}$ [89]. While previous results have revealed that C225 was not involved in redox-sensing, C232 and C240 appear to impact MsvR binding to P$_{msvR}$, yet their precise role is uncles. Thus, C206 is likely, and C232/C240 are possibly, involved in thiol-disulfide formation which serves to control DNA-binding by MaMsvR.

EMSA and thiol quantitation experiments were used to examine the role of thiol-disulfide exchange in controlling DNA binding by MaMsvR. First, MaMsvR was incubated with 100-fold molar excess of H$_2$O$_2$ to generate H$_2$O$_2$-oxidized MaMsvR (MaMsvR$_{ox}$). Quantitation of the thiol content of MaMsvR$_{ox}$ under denaturing conditions revealed that four of the cysteines were not oxidized by H$_2$O$_2$ (**Table 3.1**), indicating some cysteines are inaccessible to H$_2$O$_2$, and likely to not participate in redox-sensing. Importantly, MaMsvR$_{ox}$ was incapable of binding to the P$_{msvR}$ region as revealed by the lack of shift when examined by EMSA (**Fig. 3.1, lane 2**). This result indicates that oxidation of the thiols of six cysteine residues is sufficient to inactivate MaMsvR DNA binding. The subsequent treatment of MaMsvR$_{ox}$ with DTT resulted in detection of approximately nine thiols (**Table 3.1**), consistent with the total number of cysteines present in MaMsvR. Moreover, incubation of MaMsvR$_{ox}$ with DTT restored binding to P$_{msvR}$ (**Fig. 3.1, lane 3**). This result is consistent with H$_2$O$_2$ causing the oxidation of six thiols to disulfides, which causes reversible inactivation MaMsvR binding to P$_{msvR}$. The remaining four thiols are likely buried within the folded protein and are inaccessible to H$_2$O$_2$ or DTT, and therefore do not participate in thiol-disulfide exchange.

Similar experiments were performed to determine if the *M. acetivorans* thioredoxin system could also activate DNA-binding of MaMsvR$_{ox}$. Incubation of MaMsvR$_{ox}$ with NADPH, MaTrxR, and

MaTrx7 (complete thioredoxin system) activated binding of MaMsvR$_{ox}$ to P$_{msvR}$ (**Fig. 3.1, lane 6**). The

thioredoxin system alone did not cause a shift of P$_{msvR}$ in the EMSA (**Fig. 3.1, lane 4**) and NADPH/MaTrxR

in the absence of MaTrx7 also failed to activate binding of MaMsvR$_{ox}$ to P$_{msvR}$ (**Fig. 3.1, lane 5**).

Moreover, incubation of MaMsvR$_{ox}$ with the complete thioredoxin system also resulted in the detection

of ten thiols (**Table 3.1**), consistent with all the H$_2$O$_2$-generated disulfides in MaMsvR being surface

exposed and accessible to reduction by MaTrx7. Taken together these results demonstrate that the *M.*

acetivorans NADPH-dependent thioredoxin system can activate P$_{msvR}$ binding in oxidized MaMsvR and

that MaTrx7 is required for the reduction of disulfides in oxidized MaMsvR. The reduction of MaMsvR

by MaTrx7 is the first evidence of thioredoxin playing a role in the regulation of the activity of a

transcription regulator in a methanogen. The activation of MaMsvR DNA binding by MaTrx7 also

integrates P$_{msvR}$ regulation by MsvR into the physiology of *M. acetivorans*, which supports the future use

of P$_{msvR}$ in engineering oxidant-responsive gene expression strains. For example, it has previously been

demonstrated that increased expression of catalase protects *M. acetivorans* from H$_2$O$_2$ [129].

Additional promoters containing the inverted repeat sequence (TTCGN$_{7-9}$CGAA) MaMsvR

binding site motif were incubated with DTT-reduced MsvR to determine their potential as genes

regulated by MaMsvR. Six genes were selected from the potential MaMsvR binding sites within the *M.*

acetivorans genome [89]. Each of the selected genes encoded for hypothetical/predicted proteins:

MA0502, MA0829, MA2139, MA2689, MA3322, and MA4164. MA0502, MA2139, and MA4164 were

selected because they had been observed to be up-regulated during catalase challenge assays in

unpublished microarray experiments with *M. acetivorans*. MA0829, MA2689, and MA3322 were

selected because their sequences identify with an S-layer protein, an ATP-binding multidrug ABC

transporter, and a flavoprotein oxygenase, respectively. Of the gene promoters investigated, only P$_{3322}$

indicated any interaction with MaMsvR. Naked P$_{3322}$ vanished upon incubation of the probe with

MaMsvR$_{red}$ (**Fig 3.2**). However, a shift indicative of P$_{3322}$-MaMsvR complexing was not observed. This

may have been due to weak stability of the P_{3322}-MaMsvR complex under experimental binding conditions. Nevertheless, *msvR* is the only gene yet proven to be transcriptionally regulated by MaMsvR; these results indicate that MA3322 has promise as another gene for which MaMsvR serves as a redox-sensitive transcriptional regulator, which is interesting as its sequence identifies with a class of proteins which reduce oxygen [130]. The regulation of MA3322 could be further investigated with a promoter-specific *in vitro* transcription system for *M. acetivorans*.

Based upon results from previous studies and the results of this study, the following model is proposed for the regulation of the P_{msvR} binding activity of MaMsvR by thiol-disulfide exchange involving the thioredoxin system (**Fig. 3.3**). Oxidation of critical MaMsvR cysteines to disulfides occurs during exposure of *M. acetivorans* to H_2O_2 and other oxidants. Based on previous studies, C206 plays a crucial role, likely in the formation of intermolecular disulfide between MaMsvR monomers [89]. However, under the conditions tested here, at least six cysteines are involved in H_2O_2-induced disulfide formation, which may generate three intra-molecular, six inter-molecular, or some combination of intra/inter-molecular disulfides. Disulfide formation likely effects a conformational change in MaMsvR, preventing binding to P_{msvR}. The deactivation of MaMsvR allows for RNAP to bind to the promoter site and transcription to proceed. Removal of oxidant and/or an influx of electron donor would allow for the reduction of MaMsvR disulfides by MaTrx7, with reducing equivalents supplied by MaTrxR and NADPH. The *in vitro* results presented here demonstrate that MaTrx7 can specifically reduce disulfides in MaMsvR, but we cannot rule out that the additional MaTrxs or other proteins also participate in the *in vivo* reduction of disulfides in MaMsvR and may do so under different conditions. However, the target specificity and the redox partner(s) of the other MaTrxs is currently unknown [90]. These data link the regulation of MaMsvR to the redox status of *M. acetivorans* and the availability of reducing equivalents (ex. NADPH). The results also reveal that methanogens have oxidant sensing systems which are integrated into metabolism in a manner similar to systems identified in bacteria and eukaryotes.

<u>Conclusions</u>

The reduction of disulfides was required for the restoration of MaMsvR DNA binding and was achieved using the *M. acetivorans* NADPH-dependent thioredoxin system. This is the first evidence of thioredoxin playing a role in the regulation of the activity of a transcriptional regulator in a methanogen and integrates regulation of genes by MsvR into the physiology of *M. acetivorans*. MaMsvR was demonstrated to interact with the promoter region of MA3322, indicating that MA3322 is a likely target for MaMsvR regulation. MA3322 encodes a putative flavoprotein oxidase which, if functional, would involve MaMsvR in the oxidative stress response of *M. acetivorans*. Thus, MaMsvR is likely a functioning component of *M. acetivorans* physiology and may be a participant in the oxidative stress response of the organism.

Table 3.1. Quantitation of MaMsvR thiols.

Sample[a]	Thiols
MaMsvR$_{ox}$	4.0 ± 0.6
MaMsvR$_{ox}$ + DTT	9.0 ± 1.5
MaMsvR$_{ox}$ + NADPH/MaTrxR	4.8 ± 0.1
MaMsvR$_{ox}$ + NADPH/MaTrxR/MaTrx7	9.9 ± 1.0

[a]samples were processed and thiols quantified using DTNB as described in the Materials and Methods

MaMsvR$_{ox}$	-	+	+	-	+	+
DTT	-	-	+	-	-	-
NADPH	-	-	-	+	+	+
MaTrxR	-	-	-	+	+	+
MaTrx7	-	-	-	+	-	+
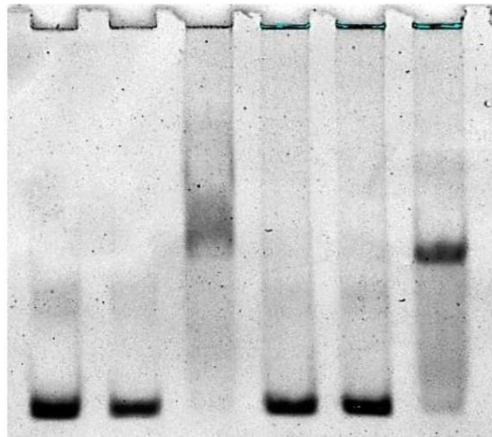
Figure 3.1. Activation of MaMsvR P$_{msvR}$ binding by the *M. acetivorans* thioredoxin system. EMSA performed with P$_{msvR}$ and the addition of the indicated components as described in materials and methods.

P_{msvR}	+	+	−	−	−
P_{3322}	−	−	+	+	+
$MaMsvR_{ox}$	+	−	−	+	−
$MaMsvR_{red}$	−	+	−	−	+
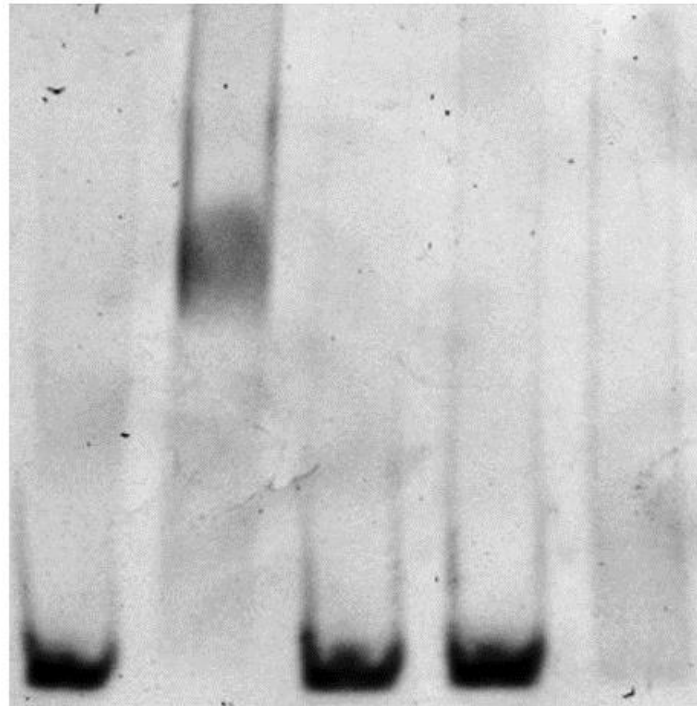
Figure 3.2. Potential binding of P$_{3322}$ with MaMsvR$_{red}$. EMSA performed with P$_{msvR}$ and P$_{3322}$. P$_{msvR}$ exhibits a full shift when incubated with MaMsvR$_{red}$; naked P$_{3322}$ vanishes when incubated with MaMsvR$_{red}$, yet no shift is visible.
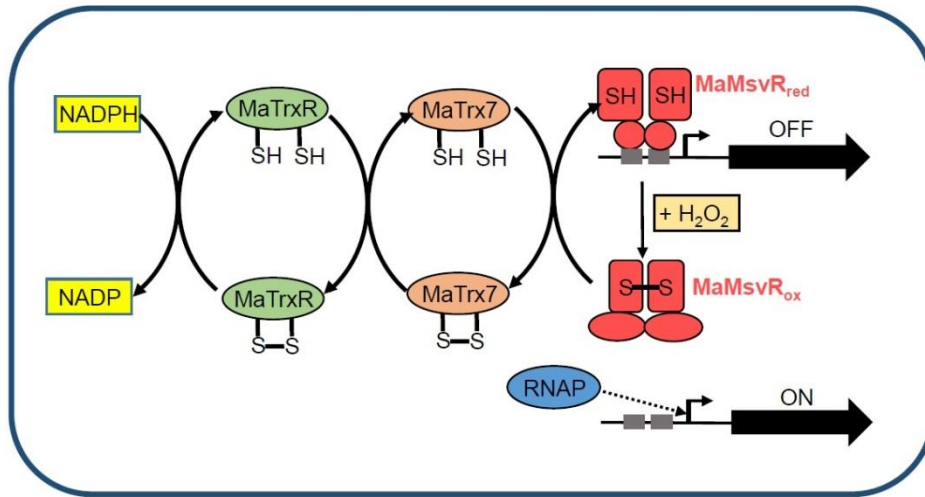
Figure 3.3. Proposed model of MaMsvR activation by the NADPH-dependent MaTrxR-MaTrx7 thioredoxin system in *M. acetivorans*. H_2O_2 causes the oxidation of thiols (SH) to disulfides which inactivates MaMsvR DNA binding, allowing transcription by RNAP. MaTrx7 receives reducing equivalents from NADPH/MaTrxR to reduce the disulfides to thiols and restore MaMsvR DNA binding.

Chapter Four: Conclusions

Methanogenesis is a metabolic pathway which was likely to have developed very early in the history of life. It is heavily reliant upon enzymes which incorporate iron-sulfur clusters and thus is easily disrupted by aerobic conditions. The metabolic pathways of primordial organisms, including methanogenesis, developed in an anoxic environment. Their likely reliance upon components to which oxygen was poisonous posed difficulties when the atmosphere of Earth was oxygenated. Under the aerobic conditions following the Great Oxygenation Event, such organisms required mechanisms to detect and combat oxygen and reactive oxygen species. *Methanosarcina acetivorans* is an extant archaeon engaging in methanogenesis. Two methods of redox-sensitive transcriptional regulation were investigated in *M. acetivorans*. RNA polymerase in *M. acetivorans* incorporates two Fe-S clusters, potentially for the correlation of transcriptional processes with the metabolic state of the cell: if oxygen attenuates methanogenic processes by destroying the Fe-S clusters of metabolic enzymes, the collapse of similar Fe-S clusters in RNAP will effect a different transcriptional state to conserve energy during metabolic stress. TATA-binding protein and transcription factor B were purified with the intent to develop an *in vitro* promoter-dependent transcription system to investigate the effects of Fe-S cluster integrity on RNAP activity. Redox-sensitive transcriptional regulation in *M. acetivorans* can be achieved in a more specific manner by the regulatory protein MsvR. The *M. acetivorans* NADPH-dependent thioredoxin system is capable of reducing MaMsvR disulfides to restore DNA binding. MaMsvR interacts with the promoter regions of its own gene as well as MA3322, a gene encoding a putative flavoprotein oxidase. These results indicate that MaMsvR is a functioning component of *M. acetivorans* physiology and is a likely participant in the oxidative stress response of the organism.

These mechanisms of redox-sensitive transcriptional regulation in *M. acetivorans* are examples of the adaptations which ancient organisms adopted in the transition to an aerobic world. Iron-sulfur clusters and redox-sensitive thiol groups, constructs which make metabolic enzymes vulnerable to oxygen and reactive oxygen species, have been incorporated into enzymes involved in transcription so as to couple transcriptional processes with cellular redox state. This evolutionary trajectory is integral to the history of life on planet Earth and may be informative if examining the evolution of life systems on other worlds.

Literature Cited

1.	Woese, C.R. and G.E. Fox, *Phylogenetic structure of the prokaryotic domain: the primary kingdoms.* Proc Natl Acad Sci U S A, 1977. **74**(11): p. 5088-90.

2.	Sogin, S.J., M.L. Sogin, and C.R. Woese, *Phylogenetic measurement in procaryotes by primary structural characterization.* J Mol Evol, 1972. **1**(2): p. 173-84.

3.	Moore, R.L. and B.J. McCarthy, *Comparative study of ribosomal ribonucleic acid cistrons in enterobacteria and myxobacteria.* J Bacteriol, 1967. **94**(4): p. 1066-74.

4.	Bendich, A.J. and B.J. McCarthy, *Ribosomal RNA homologies among distantly related organisms.* Proc Natl Acad Sci U S A, 1970. **65**(2): p. 349-56.

5.	Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.

6.	Koonin, E.V., K.S. Makarova, and L. Aravind, *Horizontal gene transfer in prokaryotes: quantification and classification.* Annu Rev Microbiol, 2001. **55**: p. 709-42.

7.	Koonin, E.V. and Y.I. Wolf, *Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world.* Nucleic Acids Res, 2008. **36**(21): p. 6688-719.

8.	Bentley, S.D. and J. Parkhill, *Comparative genomic structure of prokaryotes.* Annual Review of Genetics, 2004. **38**: p. 771-792.

9.	Waters, E., et al., *The genome of Nanoarchaeum equitans: Insights into early archaeal evolution and derived parasitism.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(22): p. 12984-12988.

10.	Rogozin, I.B., et al., *Congruent evolution of different classes of non-coding DNA in prokaryotic genomes.* Nucleic Acids Research, 2002. **30**(19): p. 4264-4271.

11.	Werner, F., *Structural evolution of multisubunit RNA polymerases.* Trends Microbiol, 2008. **16**(6): p. 247-50.

12.   Ebright, R.H., *RNA polymerase: structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II.* J Mol Biol, 2000. **304**(5): p. 687-98.

13.   Hahn, S., *Structure and mechanism of the RNA polymerase II transcription machinery.* Nat Struct Mol Biol, 2004. **11**(5): p. 394-403.

14.   Roeder, R.G., *The role of general initiation factors in transcription by RNA polymerase II.* Trends Biochem Sci, 1996. **21**(9): p. 327-35.

15.   Thomm, M., *Archaeal transcription factors and their role in transcription initiation.* FEMS Microbiol Rev, 1996. **18**(2-3): p. 159-71.

16.   Thomas, M.C. and C.M. Chiang, *The general transcription machinery and general cofactors.* Crit Rev Biochem Mol Biol, 2006. **41**(3): p. 105-78.

17.   Reiter, W.D., U. Hudepohl, and W. Zillig, *Mutational analysis of an archaebacterial promoter: essential role of a TATA box for transcription efficiency and start-site selection in vitro.* Proc Natl Acad Sci U S A, 1990. **87**(24): p. 9509-13.

18.   Hausner, W., G. Frey, and M. Thomm, *Control regions of an archaeal gene. A TATA box and an initiator element promote cell-free transcription of the tRNA(Val) gene of Methanococcus vannielii.* J Mol Biol, 1991. **222**(3): p. 495-508.

19.   Rowlands, T., P. Baumann, and S.P. Jackson, *The TATA-binding protein: a general transcription factor in eukaryotes and archaebacteria.* Science, 1994. **264**(5163): p. 1326-9.

20.   Lagrange, T., et al., *New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB.* Genes Dev, 1998. **12**(1): p. 34-44.

21.   Qureshi, S.A. and S.P. Jackson, *Sequence-specific DNA binding by the S. shibatae TFIIB homolog, TFB, and its effect on promoter strength.* Mol Cell, 1998. **1**(3): p. 389-400.

22.   Bell, S.D., et al., *Orientation of the transcription preinitiation complex in archaea.* Proc Natl Acad Sci U S A, 1999. **96**(24): p. 13662-7.

23.   Kyrpides, N.C. and C.A. Ouzounis, *Transcription in archaea.* Proc Natl Acad Sci U S A, 1999. **96**(15): p. 8545-50.

24.   Aravind, L. and E.V. Koonin, *DNA-binding proteins and evolution of transcription regulation in the archaea.* Nucleic Acids Res, 1999. **27**(23): p. 4658-70.

25.   Bell, S.D. and S.P. Jackson, *Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features.* Trends Microbiol, 1998. **6**(6): p. 222-8.

26.   Bell, S.D., *Archaeal transcriptional regulation--variation on a bacterial theme?* Trends Microbiol, 2005. **13**(6): p. 262-5.

27.   Koga, Y. and H. Morii, *Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations.* Microbiol Mol Biol Rev, 2007. **71**(1): p. 97-120.

28.   De Rosa, M., A. Gambacorta, and A. Gliozzi, *Structure, biosynthesis, and physicochemical properties of archaebacterial lipids.* Microbiol Rev, 1986. **50**(1): p. 70-80.

29.   Woese, C.R., L.J. Magrum, and G.E. Fox, *Archaebacteria.* J Mol Evol, 1978. **11**(3): p. 245-51.

30.   DeLong, E.F., *Everything in moderation: archaea as 'non-extremophiles'.* Curr Opin Genet Dev, 1998. **8**(6): p. 649-54.

31.   DeLong, E.F., *Archaea in coastal marine environments.* Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5685-9.

32.   Offre, P., A. Spang, and C. Schleper, *Archaea in biogeochemical cycles.* Annu Rev Microbiol, 2013. **67**: p. 437-57.

33.   Cabello, P., M.D. Roldan, and C. Moreno-Vivian, *Nitrate reduction and the nitrogen cycle in archaea.* Microbiology, 2004. **150**(Pt 11): p. 3527-46.

34.   Mehta, M.P. and J.A. Baross, *Nitrogen fixation at 92 degrees C by a hydrothermal vent archaeon.* Science, 2006. **314**(5806): p. 1783-6.

35.   Spang, A., et al., *Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota.* Trends Microbiol, 2010. **18**(8): p. 331-40.

36.   Tourna, M., et al., *Nitrososphaera viennensis, an ammonia oxidizing archaeon from soil.* Proc Natl Acad Sci U S A, 2011. **108**(20): p. 8420-5.

37.   Volkl, P., et al., *Pyrobaculum aerophilum sp. nov., a novel nitrate-reducing hyperthermophilic archaeum.* Appl Environ Microbiol, 1993. **59**(9): p. 2918-26.

38.   Hafenbradl, D., et al., *Ferroglobus placidus gen. nov., sp. nov., A novel hyperthermophilic archaeum that oxidizes Fe2+ at neutral pH under anoxic conditions.* Arch Microbiol, 1996. **166**(5): p. 308-14.

39.   Canfield, D.E. and R. Raiswell, *The evolution of the sulfur cycle.* American Journal of Science, 1999. **299**(7-9): p. 697-723.

40.   Milucka, J., et al., *Zero-valent sulphur is a key intermediate in marine methane oxidation.* Nature, 2012. **491**(7425): p. 541-+.

41.   Berg, I.A., et al., *Autotrophic carbon fixation in archaea.* Nature Reviews Microbiology, 2010. **8**(6): p. 447-460.

42.   Herndl, G.J., et al., *Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean.* Applied and Environmental Microbiology, 2005. **71**(5): p. 2303-2309.

43.   Hansman, R.L., et al., *The radiocarbon signature of microorganisms in the mesopelagic ocean.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(16): p. 6513-6518.

44.   Karner, M.B., E.F. DeLong, and D.M. Karl, *Archaeal dominance in the mesopelagic zone of the Pacific Ocean.* Nature, 2001. **409**(6819): p. 507-510.

45.   Reeburgh, W.S., *Oceanic methane biogeochemistry.* Chemical Reviews, 2007. **107**(2): p. 486-513.

46.   Lashof, D.A. and D.R. Ahuja, *Relative Contributions of Greenhouse Gas Emissions to Global Warming.* Nature, 1990. **344**(6266): p. 529-531.

47.   Garcia, J.L., B.K.C. Patel, and B. Ollivier, *Taxonomic phylogenetic and ecological diversity of methanogenic Archaea.* Anaerobe, 2000. **6**(4): p. 205-226.

48.   Thauer, R.K., *Biochemistry of methanogenesis: a tribute to Marjory Stephenson.* Microbiology-Uk, 1998. **144**: p. 2377-2406.

49.   Thauer, R.K., et al., *Methanogenic archaea: ecologically relevant differences in energy conservation.* Nature Reviews Microbiology, 2008. **6**(8): p. 579-591.

50.   Ferry, J.G. and C.H. House, *The stepwise evolution of early life driven by energy conservation.* Mol Biol Evol, 2006. **23**(6): p. 1286-92.

51.   Kasting, J.F., *Earths Early Atmosphere.* Science, 1993. **259**(5097): p. 920-926.

52.   Vignais, P.M., B. Billoud, and J. Meyer, *Classification and phylogeny of hydrogenases.* Fems Microbiology Reviews, 2001. **25**(4): p. 455-501.

53.   Jarrell, K.F., *Extreme Oxygen Sensitivity in Methanogenic Archaebacteria.* Bioscience, 1985. **35**(5): p. 298-302.

54.   Hedderich, R., N. Hamann, and M. Bennati, *Heterodisulfide reductase from methanogenic archaea: a new catalytic role for an iron-sulfur cluster.* Biological Chemistry, 2005. **386**(10): p. 961-970.

55.   Berkovitch, F., et al., *Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme.* Science, 2004. **303**(5654): p. 76-79.

56.   Glaser, T., et al., *Ligand K-edge X-ray absorption spectroscopy: A direct probe of ligand-metal covalency.* Accounts of Chemical Research, 2000. **33**(12): p. 859-868.

57.   Bruschi, M. and F. Guerlesquin, *Structure, function and evolution of bacterial ferredoxins.* FEMS Microbiol Rev, 1988. **4**(2): p. 155-75.

58.   Howard, J.B., et al., *Fe:S cluster ligands are the only cysteines required for nitrogenase Fe-protein activities.* J Biol Chem, 1989. **264**(19): p. 11270-4.

59.   Doukov, T.I., et al., *A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase.* Science, 2002. **298**(5593): p. 567-72.

60.   Hedderich, R., et al., *Isolation and Characterization of Polyferredoxin from Methanobacterium-Thermoautotrophicum - the Mvhb Gene-Product of the Methylviologen-Reducing Hydrogenase Operon.* Febs Letters, 1992. **298**(1): p. 65-68.

61.   Kuo, C.F., et al., *Crystallization and crystallographic characterization of the iron-sulfur-containing DNA-repair enzyme endonuclease III from Escherichia coli.* J Mol Biol, 1992. **227**(1): p. 347-51.

62.   Johnson, D.C., et al., *Structure, function, and formation of biological iron-sulfur clusters.* Annu Rev Biochem, 2005. **74**: p. 247-81.

63.   Rees, D.C. and J.B. Howard, *The interface between the biological and inorganic worlds: iron-sulfur metalloclusters.* Science, 2003. **300**(5621): p. 929-31.

64.   Liu, Y., L.L. Beer, and W.B. Whitman, *Methanogens: a window into ancient sulfur metabolism.* Trends Microbiol, 2012. **20**(5): p. 251-8.

65.   Huber, C. and G. Wachtershauser, *Activated acetic acid by carbon fixation on (Fe,Ni)S under primordial conditions.* Science, 1997. **276**(5310): p. 245-7.

66.   Wachtershauser, G., *Pyrite Formation, the 1st Energy-Source for Life - a Hypothesis.* Systematic and Applied Microbiology, 1988. **10**(3): p. 207-210.

67.   Imlay, J.A., *Iron-sulphur clusters and the problem with oxygen.* Mol Microbiol, 2006. **59**(4): p. 1073-82.

68.   Major, T.A., H. Burd, and W.B. Whitman, *Abundance of 4Fe-4S motifs in the genomes of methanogens and other prokaryotes.* Fems Microbiology Letters, 2004. **239**(1): p. 117-123.

69.   Farquhar, J., H.M. Bao, and M. Thiemens, *Atmospheric influence of Earth's earliest sulfur cycle.* Science, 2000. **289**(5480): p. 756-758.

70.   Gardner, P.R. and I. Fridovich, *Superoxide sensitivity of the Escherichia coli aconitase.* J Biol Chem, 1991. **266**(29): p. 19328-33.

71.   Varghese, S., Y. Tang, and J.A. Imlay, *Contrasting sensitivities of Escherichia coli aconitases A and B to oxidation and iron depletion.* J Bacteriol, 2003. **185**(1): p. 221-30.

72.   Paoli, P., et al., *Hydrogen peroxide triggers the formation of a disulfide dimer of muscle acylphosphatase and modifies some functional properties of the enzyme.* J Biol Chem, 2001. **276**(45): p. 41862-9.

73.   Imlay, J.A. and S. Linn, *DNA damage and oxygen radical toxicity.* Science, 1988. **240**(4857): p. 1302-9.

74.   Apel, K. and H. Hirt, *Reactive oxygen species: metabolism, oxidative stress, and signal transduction.* Annu Rev Plant Biol, 2004. **55**: p. 373-99.

75.   Kiley, P.J. and H. Beinert, *Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster.* FEMS Microbiol Rev, 1998. **22**(5): p. 341-52.

76.   Yang, H., et al., *SurR regulates hydrogen production in Pyrococcus furiosus by a sulfur-dependent redox switch.* Molecular Microbiology, 2010. **77**(5): p. 1111-1122.

77.   Lee, C., et al., *Redox regulation of OxyR requires specific disulfide bond formation involving a rapid kinetic reaction path.* Nat Struct Mol Biol, 2004. **11**(12): p. 1179-85.

78.   Karr, E.A., *The methanogen-specific transcription factor MsvR regulates the fpaA-rlp-rub oxidative stress operon adjacent to msvR in Methanothermobacter thermautotrophicus.* J Bacteriol, 2010. **192**(22): p. 5914-22.

79.   Hirata, A., B.J. Klein, and K.S. Murakami, *The X-ray crystal structure of RNA polymerase from Archaea.* Nature, 2008. **451**(7180): p. 851-4.

80.   Zhang, G. and S.A. Darst, *Structure of the Escherichia coli RNA polymerase alpha subunit amino-terminal domain.* Science, 1998. **281**(5374): p. 262-6.

81.   Goede, B., et al., *Protein-protein interactions in the archaeal transcriptional machinery: binding studies of isolated RNA polymerase subunits and transcription factors.* J Biol Chem, 2006. **281**(41): p. 30581-92.

82.   Rodriguez-Monge, L., C.A. Ouzounis, and N.C. Kyrpides, *A ferredoxin-like domain in RNA polymerase 30/40-kDa subunits.* Trends Biochem Sci, 1998. **23**(5): p. 169-70.

83.   Lessner, F.H., et al., *Subunit D of RNA polymerase from Methanosarcina acetivorans contains two oxygen-labile [4Fe-4S] clusters: implications for oxidant-dependent regulation of transcription.* J Biol Chem, 2012. **287**(22): p. 18510-23.

84.   Ferry, J.G., *Enzymology of one-carbon metabolism in methanogenic pathways.* FEMS Microbiol Rev, 1999. **23**(1): p. 13-38.

85.   Rother, M. and W.W. Metcalf, *Genetic technologies for Archaea.* Curr Opin Microbiol, 2005. **8**(6): p. 745-51.

86.   Li, Q., et al., *Electron transport in the pathway of acetate conversion to methane in the marine archaeon Methanosarcina acetivorans.* J Bacteriol, 2006. **188**(2): p. 702-10.

87.   Ferry, J.G. and D.J. Lessner, *Methanogenesis in marine sediments.* Ann N Y Acad Sci, 2008. **1125**: p. 147-57.

88.   Seedorf, H., et al., *F420H2 oxidase (FprA) from Methanobrevibacter arboriphilus, a coenzyme F420-dependent enzyme involved in O2 detoxification.* Arch Microbiol, 2004. **182**(2-3): p. 126-37.

89.   Isom, C.E., et al., *Redox-sensitive DNA binding by homodimeric Methanosarcina acetivorans MsvR is modulated by cysteine residues.* Bmc Microbiology, 2013. **13**.

90.   McCarver, A.C. and D.J. Lessner, *Molecular characterization of the thioredoxin system from Methanosarcina acetivorans.* Febs Journal, 2014. **281**(20): p. 4598-4611.

91.   Butler, J.E. and J.T. Kadonaga, *The RNA polymerase II core promoter: a key component in the regulation of gene expression.* Genes Dev, 2002. **16**(20): p. 2583-92.

92.   Helmann, J.D. and M.J. Chamberlin, *Structure and function of bacterial sigma factors.* Annu Rev Biochem, 1988. **57**: p. 839-72.

93.   Travers, A.A. and Burgessrr, *Cyclic re-use of the RNA polymerase sigma factor.* Nature, 1969. **222**(5193): p. 537-40.

94.   Doi, R.H. and L.F. Wang, *Multiple procaryotic ribonucleic acid polymerase sigma factors.* Microbiol Rev, 1986. **50**(3): p. 227-43.

95.   Burgess, R.R., et al., *Factor stimulating transcription by RNA polymerase.* Nature, 1969. **221**(5175): p. 43-6.

96.   Gruber, T.M. and C.A. Gross, *Multiple sigma subunits and the partitioning of bacterial transcription space.* Annu Rev Microbiol, 2003. **57**: p. 441-66.

97.   Wosten, M.M., *Eubacterial sigma-factors.* FEMS Microbiol Rev, 1998. **22**(3): p. 127-50.

98.   Arthur, T.M. and R.R. Burgess, *Localization of a sigma70 binding site on the N terminus of the Escherichia coli RNA polymerase beta' subunit.* J Biol Chem, 1998. **273**(47): p. 31381-7.

99.   Burgess, R.R. and L. Anthony, *How sigma docks to RNA polymerase and what sigma does.* Curr Opin Microbiol, 2001. **4**(2): p. 126-31.

100.   Craig, M.L., et al., *DNA footprints of the two kinetically significant intermediates in formation of an RNA polymerase-promoter open complex: evidence that interactions with start site and downstream DNA induce sequential conformational changes in polymerase and DNA.* J Mol Biol, 1998. **283**(4): p. 741-56.

101.   Murakami, K.S. and S.A. Darst, *Bacterial RNA polymerases: the wholo story.* Curr Opin Struct Biol, 2003. **13**(1): p. 31-9.

102.   Cormack, B.P. and K. Struhl, *The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells.* Cell, 1992. **69**(4): p. 685-96.

103.   Nikolov, D.B., et al., *Crystal structure of a TFIIB-TBP-TATA-element ternary complex.* Nature, 1995. **377**(6545): p. 119-28.

104.   Ouhammouch, M., W. Hausner, and E.P. Geiduschek, *TBP domain symmetry in basal and activated archaeal transcription.* Molecular Microbiology, 2009. **71**(1): p. 123-131.

105.   Kim, J.L., D.B. Nikolov, and S.K. Burley, *Co-crystal structure of TBP recognizing the minor groove of a TATA element.* Nature, 1993. **365**(6446): p. 520-7.

106.   Kosa, P.F., et al., *The 2.1-angstrom crystal structure of an archaeal preinitiation complex: TATA-box-binding protein/transcription factor (II)B core/TATA-box.* Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(12): p. 6042-6047.

107.   Bell, S.D. and S.P. Jackson, *The role of transcription factor B in transcription initiation and promoter clearance in the archaeon Sulfolobus acidocaldarius.* J Biol Chem, 2000. **275**(17): p. 12934-40.

108.   Zawel, L. and D. Reinberg, *Initiation of transcription by RNA polymerase II: a multi-step process.* Prog Nucleic Acid Res Mol Biol, 1993. **44**: p. 67-108.

109.   Thomm, M., et al., *Transcription in vivo and in vitro of the histone-encoding gene hmfB from the hyperthermophilic archaeon Methanothermus fervidus.* J Bacteriol, 1992. **174**(11): p. 3508-13.

110.   Hethke, C., et al., *A cell-free transcription system for the hyperthermophilic archaeon Pyrococcus furiosus.* Nucleic Acids Res, 1996. **24**(12): p. 2369-76.

111.   Qureshi, S.A., S.D. Bell, and S.P. Jackson, *Factor requirements for transcription in the Archaeon Sulfolobus shibatae.* EMBO J, 1997. **16**(10): p. 2927-36.

112.   Thomsen, J., et al., *The basal transcription factors TBP and TFB from the mesophilic archaeon Methanosarcina mazeii: structure and conformational changes upon interaction with stress-gene promoters.* J Mol Biol, 2001. **309**(3): p. 589-603.

113.   Reichlen, M.J., K.S. Murakami, and J.G. Ferry, *Functional analysis of the three TATA binding protein homologs in Methanosarcina acetivorans.* J Bacteriol, 2010. **192**(6): p. 1511-7.

114.   Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding.* Anal Biochem, 1976. **72**: p. 248-54.

115.   Gadgil, H., L.A. Jurado, and H.W. Jarrett, *DNA affinity chromatography of transcription factors.* Anal Biochem, 2001. **290**(2): p. 147-78.

116.   Guss, A.M., et al., *New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for Methanosarcina species.* Archaea, 2008. **2**(3): p. 193-203.

117.   Winterbourn, C.C. and D. Metodiewa, *Reactivity of biologically important thiol compounds with superoxide and hydrogen peroxide.* Free Radic Biol Med, 1999. **27**(3-4): p. 322-8.

118.   Angel, R., D. Matthies, and R. Conrad, *Activation of methanogenesis in arid biological soil crusts despite the presence of oxygen.* PLoS ONE, 2011. **6**(5): p. e20453.

119.   Kiener, A. and T. Leisinger, *Oxygen sensitivity of methanogenic bacteria.* Syst Appl Microbiol, 1983. **4**(3): p. 305-12.

120.   Maeder, D.L., et al., *The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes.* J Bacteriol, 2006. **188**(22): p. 7922-31.

121.   Zheng, M., F. Aslund, and G. Storz, *Activation of the OxyR transcription factor by reversible disulfide bond formation.* Science, 1998. **279**(5357): p. 1718-1721.

122.   Dubbs, J.M. and S. Mongkolsuk, *Peroxide-Sensing Transcriptional Regulators in Bacteria.* Journal of Bacteriology, 2012. **194**(20): p. 5495-5503.

123.   Cremers, C.M. and U. Jakob, *Oxidant Sensing by Reversible Disulfide Bond Formation.* Journal of Biological Chemistry, 2013. **288**(37): p. 26489-26496.

124.   Meyer, Y., et al., *Thioredoxins and Glutaredoxins: Unifying Elements in Redox Biology.* Annual Review of Genetics, 2009. **43**: p. 335-367.

125.   Fahey, R.C., *Novel thiols of prokaryotes.* Annual Review of Microbiology, 2001. **55**: p. 333-356.

126.   McFarlan, S.C., C.A. Terrell, and H.P. Hogenkamp, *The purification, characterization, and primary structure of a small redox protein from Methanobacterium thermoautotrophicum, an archaebacterium.* J Biol Chem, 1992. **267**(15): p. 10561-9.

127.   Susanti, D., et al., *Thioredoxin targets fundamental processes in a methane-producing archaeon, Methanocaldococcus jannaschii.* Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(7): p. 2608-2613.

128.   Riddles, P.W., R.L. Blakeley, and B. Zerner, *Reassessment of Ellman Reagent.* Methods in Enzymology, 1983. **91**: p. 49-60.

129.   Jennings, M.E., et al., *Expression of a bacterial catalase in a strictly anaerobic methanogen significantly increases tolerance to hydrogen peroxide but not oxygen.* Microbiology, 2014. **160**(Pt 2): p. 270-8.

130.   Gadda, G., *Oxygen activation in flavoprotein oxidases: the importance of being positive.* Biochemistry, 2012. **51**(13): p. 2662-9.

Appendix

Oligonucleotide sequences for EMSA probes

Probe name	Oligo name	Oligonucleotide sequence
P_{mcrB}	RSEMSAmcrB1	5'-TAT CGG AGA ACA CAA AAG ATT TAA GTA CCT TCT AAA CGA ATG AGA TTT CA-3'
P_{mcrB}	RSEMSAmcrB2	5'-AGA AAT CTC ATT CGT TTA GAA GGT ACT TAA ATC TTT TGT GTT CTC CGA TA-3'
P_{dnaK}	RESMSAdnaK1	5'-ATC TGG CTG GAA ATT AAA CTT TAA TTA ATC TAT AAC TAC AAT TTA CAA AC-3'
P_{dnaK}	RESMSAdnaK2	5'-GTT TGT AAA TTG TAG TTA TAG ATT AAT TAA AGT TTA ATT TCCAGC CAG AT-3'
P_{msvRa}	RESMSAmsvR1	5'-GTA GTC ACG CGA ACG TTT TAT ATA TTC AAA ACG CAA CTT ATC TAA AGA TA-3'
P_{msvRa}	RESMSAmsvR2	5'-TAT CTT TAG ATA AGT TGC GTT TTG AAT ATA TAA AAC GTT CGC GTG ACT AC-3'
$P_{control}$	RESMSAControl1	5'-GAG GGA CTG AAG GGT CAT GTA CCC GAC GAG AGA GTC CTC GGA AAA GAC CT-3'
$P_{control}$	RESMSAControl2	5'-AGG TCT TTT CCG AGG ACT CTC TCG TCG GGT ACA TGA CCC TTC AGT CCC TC-3'
P_{msvR}	Pmsvr1	5'-TAT TTC AAA CAT GAT TAT TCG TAG TCA CGC GAA CGT TTT ATA TAT TCA AA-3'
P_{msvR}	Pmsvr2	5'-TTT GAA TAT ATA AAA CGT TCG CGT GAC TAC GAA TAA TCA TGT TTG AAA TA-3'
P_{0502}	ma0502_a	5'-TAG AGT TTA AAA GTA CTT TTC GGG TTA CTC GAA TTT CAT AAA ATA TGC CC-3'
P_{0502}	ma0502_b	5'-GGG CAT ATT TTA TGA AAT TCG AGT AAC CCG AAA AGT ACT TTT AAA CTC TA-3'
P_{0829}	ma0829_a	5'-GAG GGA ATG AAA TCC TCT TCG GCA ATA GAC GAA AAT GGA CAA GTC TGA TT-3'
P_{0829}	ma0829_b	5'-AAT CAG ACT TGT CCA TTT TCG TCT ATT GCC GAA GAG GAT TTC ATT CCC TC-3'
P_{2139}	ma2139_a	5'-TTC TTT ATG AGT CTT CCT TTC GGC ACT TAC GAA CCG AGT CTG GGC TTT AC -3'
P_{2139}	ma2139_b	5'-GTA AAG CCC AGA CTC GGT TCG TAA GTG CCG AAA GGA AGA CTC ATA AAG AA-3'
P_{2689}	ma2689_a	5'-AAA AAC AGG AAG TAA GAA TTC GAA GAA GCC GAA ACC GAA AAA AAG CCG AT-3'
P_{2689}	ma2689_b	5'-ATG GGC TTT TTT TCG GTT TCG GCT TCT TCG AAT TCT TAC TTC CTG TTT TT-3'
P_{3322}	ma3322_a	5'-CAT CTT AAG TTC AAA GGT TTC GAG GTG AAC GAA TGA AAT TAA AAC CAA GC-3'
P_{3322}	ma3322_b	5'-GCT TGG TTT TAA TTT CAT CGT TCC ACC TCG AAA CCT TTG AAC TTA GAT TG-3'
P_{4164}	ma4164_a	5'-ACC TAT CAA TGT TCT CTT TTC GCA CCG CAC GAAGAC AGG ACC GGC CTG GT-3'
P_{4164}	ma4164_b	5'-ACC AGG CCG GTC CTG TCT TCG TGC GGT GCG AAA AGA GAA CAT TGA TAG GT-3'