INSIGHTS INTO SUBSTRATE UTILIZATION AND TRANSCRIPTIONAL REGULATION IN *METHANOSARCINA*

BY

MADELINE M. LÓPEZ MUÑOZ

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Microbiology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor William W. Metcalf, Chair
Associate Professor Rachel J. Whitaker
Emeritus Professor Stephen K. Farrand
Professor Isaac Cann

Abstract

Methanogenic archaea are strictly anaerobic organisms responsible for the production of all biologically produced methane making them key players of the global carbon cycle. These unusual organisms derive their energy from the reduction of one-carbon (C-1) compounds to methane. Among methanogenic archaea, *Methanosarcina* are the most metabolically diverse being able to use H_2/CO_2, acetate, methanol, methylamines and methyl sulfides as substrates for methanogenesis. This dissertation presents biochemical and genetic studies on pyruvate utilization in *Methanosarcina* as a substrate for methanogenesis and the regulation of methylotrophic methanogenesis in *M. acetivorans* in an effort to further delineate the intrinsic metabolic diversity found in *Methanosarcina* species.

Organic compounds such as fatty acids and sugars are rarely substrates for methanogenesis unless they are co-metabolized by syntrophic associations. Previously, an *M. barkeri* mutant that had the capacity to use pyruvate as substrate was isolated. Chapter 2 of this dissertation is focused towards understanding the molecular basis of pyruvate utilization in these cells. Genetic and transcriptomic studies on two cell lines of the mutant strain revealed the presence of two mutations in both strains. These changes in genome sequence resulted in cells that over-expressed pyruvate ferredoxin oxidoreductase (Por) an enzyme responsible for the conversion of pyruvate to acetyl-CoA and lacked pyruvate carboxylase activity, the enzyme responsible for synthesis of oxaloacetate in the cells. Biochemical and transcriptomic experiments revealed that in order to make the central metabolite oxaloacetate this mutant was using an alternative reaction catalyzed by phosphoenolpyruvate carboxylase. Suprisingly, our results revealed that Por is essential in *M. bakeri* suggesting this enzyme to have an alternative function to pyruvate synthesis in the cells.

Chapter three entails the characterization of the regulatory network that controls methylotrophic methanogenesis, the pathway used for growth on methylated

compounds like methanol, methylamines and methyl sulfides. Methylated compounds enter the methanogenesis pathway at the level of methyl-coenzyme M, the synthesis of which is mediated by the substrate specific methyltransferase complex MT1 and MT2. Growth on methanol requires the MT1 complex encoded by the *mtaCB1* operon. Previous work showed that this MT1 complex is among the highest differentially regulated genes in Archaea and that its transcription is regulated by the methanol specific regulators *msrA* and *msrB*. Biochemical and transcriptomic data presented in this dissertation revealed that MsrA and MsrB are DNA binding proteins that activate the transcription of *mtaCB1* by binding to its promoter as heterodimers. Transcriptomic data also revealed that MsrB could act as both repressor and activator of transcription, since its deletion caused an increase in expression of methylamine methyltransferase genes. Complementation of *msrB* deletion strains resulted in restoration of regulation, thus confirming the role of *msrB* in repression of methylamine methyltransferase genes. Finally, Chapter 4 of this dissertation is focused towards understating how *Methanosarcina* cells are able to sense and respond to environmental signals in order to activate the substrate specific methyltransferase genes. Bioinformatic analysis revealed the presence of three putative histidine kinases (MA0863, MA4377, MA4377), a response regulator (MA4376) and an msr-like protein (MA4375) that due to their genomic localization could be involved in the regulation of methyl sulfide methyltransferase (mts) genes. Deletion and transcriptomic studies revealed that the three-histidine kinases may be involved in the regulation of three methyl sulfide methyltransferases and that the single response regulator MA4376 may to be interacting with more than one of these histidine kinases in order to regulate *mts* genes.

Acknowledgements

Throughout my formation as a student and a young scientist several people have serve as role models and inspiration. I would like to first thank my advisor Bill Metcalf. It was my dream as a naïve undergraduate to study the Archaea, you made that dream possible by giving me the chance to be a part of your lab. I want to thank you for your patience and guidance; for allowing me to be an independent thinker and for believing in my capacity to be a good scientists even at times when I did not believed in myself. I want to thank the members of the Metcalf group, for their encouragement and advice and Dr. Wolfe for serving as an inspiration to never give up and to always be enthusiastic about my work.

To my parents, even though they did not get the chance to go to college they always imparted to my siblings and me a strong love for education. I would like to thank them for their constant support and for teaching me that you can never say I can't. To Manuel, you are the best thing that has happened to me. I am glad that science brought us together and I want to thank you because you are not only a good scientist but also an outstanding person. You push me to become not only the best student that I can be but also to be more kind, generous and loving.

We have the family that we are born into but we also have the one that we choose. I want to thank my Champaign family: Ariana, Maria, Angela and Whitney. Of all the things and experiences that I have acquired in these years you are what I treasure the most. Finally, I want to thank my friend Camille that even though she is hundreds of miles away she has always been there for me providing unconditional friendship and love.

Table of Contents

Chapter 1: Introduction

1.1 Methane in the Environment

Methane (CH_4) is the second most abundant greenhouse gas emitted in the United States (EPA, 2010). This greenhouse gas can be formed by abiotic reactions, as in gas flux from the Earth's crust, or as a result of microbial activity (Kietavainen and Purkamo, 2015). Other sources of methane include methane hydrates and sedimentary formations (Kietavainen and Purkamo, 2015). Two main types of geological methane have been described. Thermogenic methane, which is formed by the decomposition of organic matter at high temperatures and pressure and abiotic formation of methane, which is the formation of methane from inorganic compounds without the involvement of living organisms (Kietavainen and Purkamo, 2015). Among these sources, the most common is the digestion of organic matter mediated by anaerobic archaea, which accounted for the formation of 164 MMT CO_2 eq (million metric tons of CO_2 equivalents) in 2013 (EPA, 2010).

Methane concentration has been increasing at an average rate of 0.5-1% per year (EPA, 2010) and atmospheric concentration has increased from 715 (pre-industrial era) to 1774 ppb (Khosa et al, 2011). Globally, over 60% of methane emissions come from anthropogenic sources including agriculture, industry and waste management facilities (Allen, 2016). CH_4 is also emitted by a number of natural sources including wetlands, termites, sediments and volcanoes. Methane is one of the main contributors to the global greenhouse effect and although methane concentration and lifetime in the atmosphere is lower than that of carbon dioxide (CO_2), methane is more efficient at trapping radiation. For this reason it is believed that the impact of CH_4 with regard to climate change is 25 times greater than CO_2 when normalized over a 100-year period (EPA, 2010). Thus, numerous efforts have been implemented to reduce CH_4 emissions, including repurposing CH_4 as a renewable energy source. For instance, methane represents a renewable energy source given that it can replace fossil fuels in heat and power generation. Methane is more efficient and less carbon intensive

than other fossil fuels; its energy content is three times greater than that of hydrogen. As an example, methane rich biogas (biomethane) produced from waste treatment is currently employed to power generators that serve industrial facilities contributing to their power grid (Kumar, 2011). In addition, in some rural areas biomethanation is used for the disposal of domestic waste; biogas generated from this process is utilized in the homes' heating and cooking systems (Kumar, 2011).

Globally, agricultural practices are the number one anthropogenic source of CH_4; in part because ruminant livestock produce a large amount of methane as part of their normal digestive process. In addition, anaerobic decomposition of livestock manure produces significant amounts of methane. To reduce CH_4 emissions from this industry, efforts have been implemented to change the animal's diet as a way to alter the microbial population in the rumen increasing animal performance with concomitant reduction of methane emissions. Current efforts are also focused in the development of a vaccine against methanogenic archaea in the rumen (Martinez-Fernandez et al., 2016, Odongo et al., 2007).

1.2 Phylogeny and Ecology of Methanogenic Archaea

Essentially all biologically produced CH_4 is derived from methanogenic archaea. These strictly anaerobic microorganisms produce methane as the main by-product of their metabolism. Methanogenic archaea are abundant in habitats where electron acceptors such as NO_3^-, SO_4^- and Fe^{3+} are limited, including gastrointestinal tracks of animals, freshwater sediments and hydrothermal vents (Offre et al., 2013). Microbial methanogenesis plays a central role in the global carbon cycle, as it is the terminal step in biomass degradation in many anaerobic environments. Anaerobic biomass degradation is a process that includes the decomposition of complex organic compounds into small organic compounds, as well as acetogenesis and methanogenesis (Fig 1.1). Methanogens obtain all of their energy from methanogenesis and grow using a limited number of substrates including methanol, H_2 / CO_2, methyl-amines, methyl-sulfides and acetate. Most

organic compounds such as sugars and long chain fatty acids are not substrates for methanogenic archaea, yet are known to support methanogenesis via communities of syntrophic organisms. Although not all methanogens are capable of utilizing acetate as a substrate, it is estimated that two-thirds of the total methane produced by these organisms is derived from this compound (Ferry, 2015).

Methanogens are abundant in anoxic habitats; however members of *Methanocella* and *Methanosarcina* have been found in aerated soils and it has been hypothesized that methanogens are universal inhabitants of upland soils (Aschenbach et al., 2013). Methanogenic archaea can be found in a wide range of salinity, from freshwater to hypersaline; in harsh environments including high-altitude cold deserts and in an array of temperatures that can vary between marine sediments at 2°C to hydrothermal vents at 110°C. An adaptations to this wide set of environments is the presence of a thick methanocondroitin layer on freshwater members of the genus *Methanosarcina* that acts as a shield to protect the cells from turgor pressure. Methanogens are often outcompeted by sulfate reducers; iron reducers and denitrifying bacteria in habitats were electron acceptors other than CO_2 are present. However, in the environment methanogens are responsible for keeping H_2 partial pressure low, allowing the conversion of volatile fatty acids and alcohol to acetate, CO_2 and H_2 to be thermodynamically favorable (Fig 1.1). Since CO_2 is a product of the fermentation it is seldom limiting in anaerobic environments, allowing methanogens to thrive in these ecosystems.

Phylogenetic placement of methanogens in the archaeal tree and analysis of their core enzymes suggests that methanogenesis has an early origin within the Euryachaeota and that methanogenesis related genes were lost in the non-methanogenic euryarchaeal lineages (Borrel et al., 2013). In fact, the first five steps of methanogenesis are conserved in some members of the

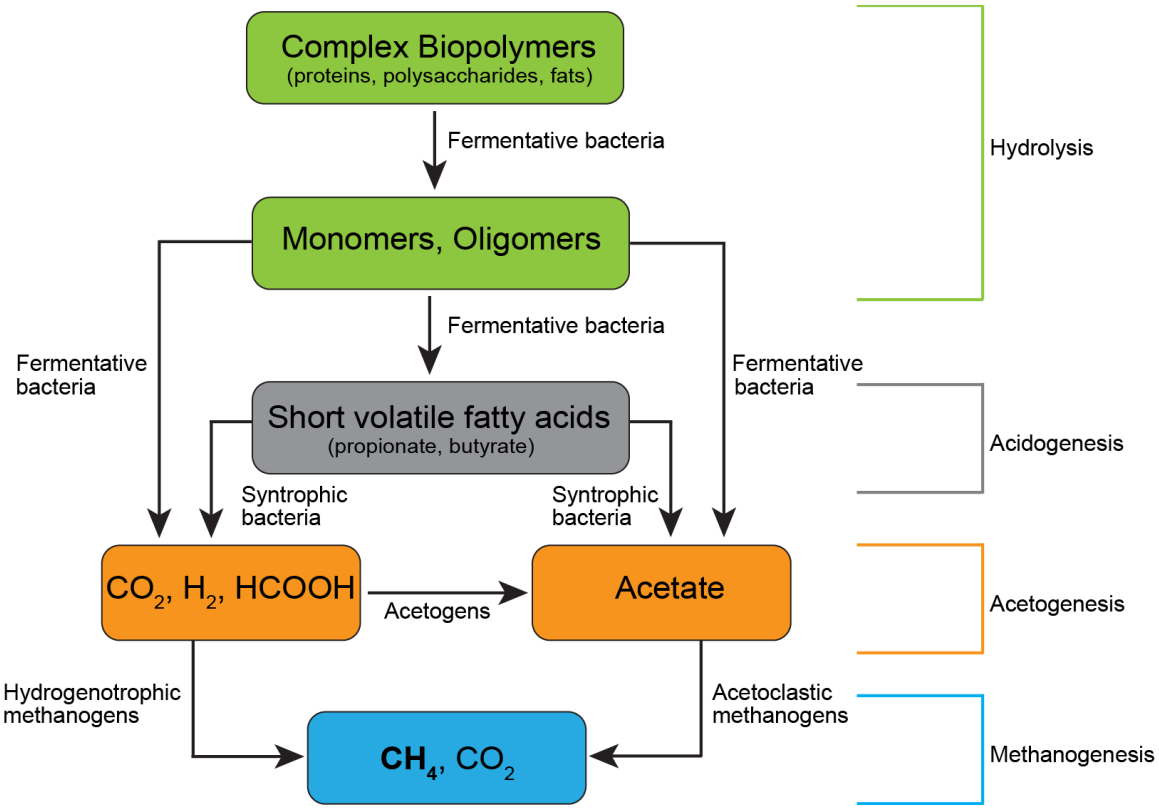

Figure 1.1 Anaerobic food chain for the decomposition of organic matter.

Archaeoglobales, where they are known to be involved in the oxidation of lactate to CO_2. Currently methanogens are classified in seven orders: Methanobacteriales, Methanococcales, Methanopyrales, Methanomicrobiales, Methanocellales, Methanomassiliiococcales and Methanosarcinales. These orders differ in their lipid composition, cell envelope structure and motility. Methanogens are also classified by the substrate they use for methane production: H_2/CO_2 (hydrogenotrophs), acetate (acetotrophs) and methylated compounds (methylotrophs).

Members of Methanopyrales, Methanococcales and Methanobacteriales are capable of producing methane from CO_2 using H_2 or formate (Methanococcales and Methanobacteriales) as electron donors. Cell wall of Methanobacteriales' and Methanopyrales' contain pseudomurein whereas those of the Methanococcales' cell wall is composed of S-layer proteins. Lipids in these three groups include archaeol, caldarchaeol and hydroxyarcaheol. Members of

Methanococcales and Methanopyrales are motile by means of flagella, have been isolated from marine habitats and have members that are thermophiles or hyperthermophiles respectively. Most species of Methanobacteriales are non-motile with the exception of members of *Methanothermus* that are motile via peritricous flagella. Methanobacteriales are distributed in different habitats including freshwater sediments, soil, anaerobic digestors and animal gastrointestinal tracts. Some species within the Methanobacteriales have the ability to use methanol via H_2-dependent methylotrophic methanogenesis. Among them is *Methanosphaera stadtmanae*, an organism that has lost the ability to reduce CO_2 to methane or oxidize methanol to CO_2 although it has kept most of the corresponding enzymes involved in both pathways (Borrel et al., 2013). This observation has led to the hypothesis that specialization for H_2-dependent methylotrophic methanogenesis is an adaptation of the organism to the gut environment (Borrel et al., 2013).

Growth by CO_2 reduction is also found in Methanomicrobiales and Methanocellales, both capable of using H_2, formate or alcohol as electron donors. Members of the Methanomicrobiales have a diverse morphology that includes rods and cocci, have protein or glycoprotein cell walls and cellular lipids that include the diether lipid archaeol and the tetraether lipid caldarchaeol. They are distributed in numerous anaerobic habitats including marine and freshwater sediments, anaerobic digestors, and gastrointestinal tracts. Methanocellales are abundant in upland soils including rice paddy fields, from which the first cultivable member of this order was isolated (Sakai et al., 2011). They are mesophilic, rod shaped and non-motile. Similar to the Methanocellales, Methanomassiliicoccales were first discovered by sequencing of the 16S rRNA and *mcrA* gene from a variety of environmental samples. Members of the order were isolated from intestinal tracts of humans, termites and from anaerobic digesters (Lang et al., 2015). Three Methanomassiliicoccales genomes had been sequenced as of 2015, and each lacked the genes encoding the C_1 pathway for reduction of CO_2 to methyl-CoM, but all have the genes needed for the utilization of other

substrates like methanol and methylamines, suggesting a novel strategy for energy conservation in this order (Lang et al., 2015). In addition, culture independent techniques have recently allowed the identification of two new orders *Candidatus* Verstraetearchaeota and *Candidatus* Methanofastidiosales (Vanwonterghem et al., 2016, Nobu et al., 2016). Genome sequencing revealed these organisms might grow using methylated compounds or acetate, malonate or propionate respectively.

Members of the order Methanosarcinales are the most metabolically versatile species, able to use methylated compounds, acetate and H_2/CO_2 as substrates for methanogenesis. Contrasting with other hydrogenotrophic methanogens, members of Methanosarcinales apparently cannot use formate as an electron donor. With the exception of *Methanosarcina baltica* most members of the Methanosarcinales are non-motile cocci, pseudosarcinae or rods (von Klein et al., 2002). Most cells have a cell wall made of protein and often a heteropolysaccharide external wall composed of a sulfated polymer known as methanochondroitin due to its similarty to chondroitin sulfate. Membrane lipids contain archaeol, hydroxyarchaeol, and caldarchaeol. Members of this order are mesophiles or moderate thermophiles that are present in freshwater and marine water sediments, gastrointestinal tracts, and anaerobic digestors. The order is divided into two families, Methanosarcinaceae and Methanosaetaceae. Members of the latter are represented by only one genus, *Methanosaeta* that is comprised of strains that are specialized acetate utilizers. The strains that are the focus of this study, *Methanosarcina barkeri* Fusaro and *Methanosarcina acetivoras* C2A, are members of the Methanosarcinaceae family. *Methanosarcina barkeri,* a nonhalophilic, mesophilic species, was isolated from a freshwater sample from Lago del Fusaro, Italy (Maeder et al., 2006). Among *Methanosarcina, M. barkeri* is the most metabolically diverse being able to use methylated compounds, acetate and H_2/CO_2 as methanogenesis substrates. *M. acetivorans* C2A, a marine isolate, is capable of using all substrates with the exception of H_2/CO_2.

1.3 Methanogenesis Substrates and Pathways

While methanogens are phylogenetically diverse, they can only use a limited number of substrates for methanogenesis. Substrates for methane production include H_2/CO_2, acetate and methylated compounds, including the recent discovery of growth on methoxylated aromatic compounds (Mayumi et al., 2016). Two-thirds of the biologically generated methane results from acetoclastic methanogenesis, even though acetate utilization is limited to members of *Methanosarcina* and *Methanosaeta* genus. The other fraction is mainly derived from H_2/CO_2 and formate. Other substrates that have been reported to support growth of methanogens include pyruvate (Bock et al., 1994) and carbon monoxide. Methanogenesis from acetate, CO_2 and methylated compounds proceeds via the following pathways: the CO_2 reduction pathway, in which CO_2 is reduced using H_2 or an alternative compound as an electron donor; the methylotrophic pathway, which involves the disproportionation of methylated compounds into CO_2 and CH_4, the methyl respiration pathway in which the methyl group of methylated compounds is reduced to CH_4 and the aceticlastic pathway, wherein the methyl group of acetate is reduced to methane using electrons derived from oxidation of its carbonyl group.

1.3.1 The CO_2 reduction pathway

In hydrogenotrophic methanogenesis, CO_2 is first reduced to a formyl group that is covalently attached to methanofuran (MF), in a reaction catalyzed by formyl-MF dehydrogenase (Fmd) (Fig 1.2). The formyl group is then transferred to the carrier tetrahydromethanopterin (H_4MPT) and is dehydrated to the methenyl form by methenyl-H_4MPT cyclohydrolase (Mch). The methenyl group undergoes a series of reductions first to form a methylene group (methylene-H_4MPT dehydrogenase, Mtd) and finally a methyl group (methylene-H_4MPT reductase, Mer). The methyl group is then transferred to a third carrier, the sulfhydryl containing coenzyme M, by methyl-H_4MPT methyltransferase (Mtr). This is an exergonic reaction that allows the export of Na^+ across the membrane, resulting in the creation of an electrochemical gradient. The methyl group is finally reduced

to CH_4 by methyl-S-CoM reductase (Mcr) using coenzyme B and generating the heterodisulfide CoM-S-S-CoB. This heterodisulfide is reduced to regenerate free thiols by the complex heterodisulfide reductase (Hdr). To perform this reaction the cells employ an electron bifurcation system, in which two pair of electrons enter the methanogenic pathway in a complex formed by Hdr and Fmd. Hdr uses one pair of electrons in the exergonic reduction of CoM-S-S-CoB to regenerate CoM-SH and SH-CoB (Costa and Leigh, 2014). The second pair of electrons is used in the endergonic reaction catalyzed by Fmd, in the reduction of CO_2 to formyl-MF. By coupling the last and first steps of methanogenesis the electron bifurcating system creates a cyclic pathway called the Wolfe cycle (Thauer, 2012).

The cyclic nature of the hydrogenotrophic pathway may explain the limited amount of carbon substrates that hydrogenotrophic methanogens can use. However, several studies have shown that hydrogenotrophic methanogens expand their metabolic diversity by the broad utilization of electron donors other than H_2. The use of formate and secondary alcohols as electron donors has been documented for members of the Methanobacteriales and Methanococcales. The use of formate requires the oxidation of formate to CO_2 using an Hdr associated F_{420} formate dehydrogenase (Costa and Leigh, 2014). An anaplerotic input of electrons is required in the CO_2 reduction step of the pathway. These electrons are provided by a membrane bound hydrogenase that uses H_2 to reduce ferredoxin (Costa and Leigh, 2014). Recently it was shown that CO serves as an electron donor by means of acetyl-CoA synthase/carbon monoxide dehydrogenase system (Costa and Leigh, 2014). Finally, in the same study it was shown that in *Methanococcus maripaludis* overexpression of glyceraldehyde-3-phosphate: ferredoxin oxidoreductase acts as an alternate way of providing the anaplerotic electrons required for hydrogenotrophic growth.

1.3.2 Methylotrophic methanogenesis

During methylotrophic methanogenesis, the methyl group of the C1 compound is transferred to a substrate-specific methyltransferase that subsequently transfers the methyl group to a cognate corrinoid protein. These two proteins together constitute the methyltransferase system 1 (MT1). The methyl group is then transfered to CoM by the methyltransferase system 2 (MT2) (Fig 1.2). As in hydrogenotrophic methanogenesis, methyl-CoM is reduced to CH_4, generating the CoM-S-S-CoB heterodisulfide. The electrons needed for regeneration of the thiol groups are derived by the oxidation of the methyl group to CO_2, running the hydrogenotrophic pathway in reverse. The activation and transfer of each methylated compound requires a unique and substrate-specific set of methyltransferases.

In contrast to hydrogenotrophic methanogens, methylotrophic methanogens do not rely in electron bifurcation. Heterodisulfide reduction in these organisms occurs by a membrane bound complex that allows the formation of an electrochemical gradient. Within methylotrophic methanogens members of the *Methanomassiliicoccales* represent a unique example, because even though they possess all the methanol-methyltranfersase genes, they lack all the genes required for the reduction of CO_2 to the methyl level. Instead, members of this genus like *Candidatus* Methanoplasma termitum have been shown to perform H_2 dependent reduction of methanol or methylamines (Lang et al., 2015). In addition, it has been hypothesized that *Candidatus Methanoplasma termitum* has a Hdr protein that has a dual function, working in electron bifurcation and as a terminal acceptor in redox processes in the cell.

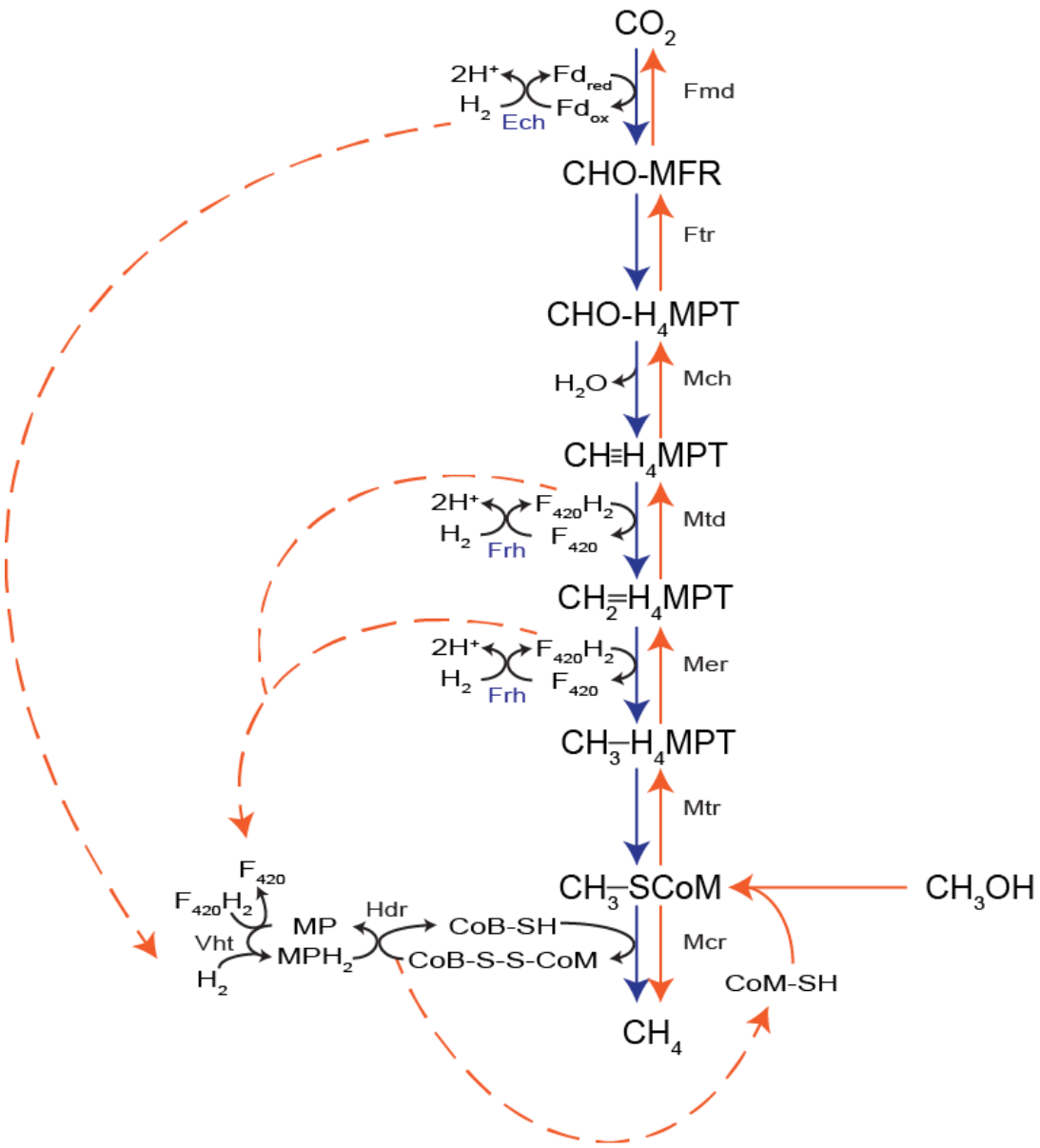

Figure 1.2 Pathways for hydrogenotrophic (blue lines) and methylotrophic methanogenesis (orange lines). Pathways share the reduction of methyl-SCoM and CoB-S-S-CoM. Abbreviations: Fmd, formyl-methanofuran dehydrogenase; FMR, methanofuran; H_4MPT, tetrahydromethanopterin; Ftr, formyl-H_4MPT formyltransferase; Mch, methenyl-H_4MPT cyclohydrolase; Mtd, methylene-H_4MPT dehydrogenase; Mer, methylene-H_4MPT reductase; Mtr, methyl- H_4MPT: CoM methyltransferase; CoB-S-S-CoM, coenzyme M-coenzyme B heterodisulfide; Mcr, methyl-S-CoM reductase; Hdr, heterodisulfide reductase; MP, methanophenezine; Fd, ferredoxin; F_{420}, factor F_{420}; Ech, energy conserving hydrogenase; Vht, methanophenezine dependent hydrogenase; Frh, F_{420} dependent hydrogenase.

1.3.3 The methyl respiration pathway

As in methylotrophic methanogenesis, the methyl group of methylated compounds enters the methanogenesis pathway through methyl-CoM (Fig 1.3). However, in this pathway methylated compounds are not disproportionated to methane and CO_2, but are converted directly into CH_4 using electrons from H_2. This pathway is found in *M. barkeri*.

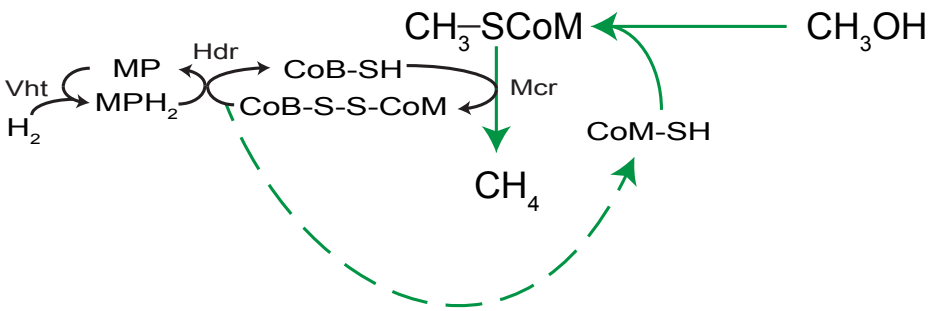

Figure 1.3 Methyl respiration pathway Abbreviations: CoM methyltransferase; CoB-S-S-CoM, coenzyme M-coenzyme B heterodisulfide; Mcr, methyl-S-CoM reductase; Hdr, heterodisulfide reductase; MP, methanophenezine; Vht, methanophenezine dependent hydrogenase.

1.3.4 Acetoclastic methanogenesis

Growth on acetate as a methanogenesis substrate is limited to members of the *Methanosarcina* or *Methanosaeta* genus. The first step of the pathway is the activation of acetate to acetyl-CoA. In *Methanosarcina* this step is done by the acetate kinase and phosphotransacetylase (Ack/Pta) system, using 1 ATP in the process (Fig 1.4). *Methanosaeta* uses the high affinity adenosine monophosphate (AMP) forming acetyl-CoA synthase. Once activated the acetyl group of acetyl-CoA is subsequently cleaved by the carbon monoxide dehydrogenase/acetyl-CoA synthase complex (CODH/ACS) into carbon monoxide and methyl-H_4MPT. Carbon monoxide is oxidized to CO_2 by CODH, providing the electrons needed for reduction of the methyl group to CH_4. Electrons required for the regeneration of the coenzymes are derived in *M. barkeri* Fusaro from the reduced ferredoxin generated by the energy conserving

hydrogenase (Ech) (Fig 1.4). *Methanosarcina acetivorans* lacks hydrogenase genes; in this organism the electrons are derived from Rnf complex.

1.3.5 Non-traditional substrates

Carbon monoxide is not a common substrate for methanogenesis. However, it has been shown that *M. barkeri* and *Methanothermobacter thermautotrophicus* can use CO for growth (Daniels L, 1977). These strains are able to grow by disproportionation of 4 mol CO into 3 mol CO_2 and 1 mol CH_4 (Equation 1.1); producing large amounts of H_2 in the process. The CODH complex catalyzes conversion of CO to CO_2. H_2 is formed by a hydrognease that oxidizes the reduced ferredoxin produced by CODH. H_2 is then used to reduce CO_2 to CH_4. Use of CO by *M. acetivorans* proceeds by a different route that results in the formation of CH_4, formate and acetate (Rother and Metcalf, 2004). This pathway requires the Ack/Pta system for the formation of acetate and generation of ATP via substrate level phosphorylation. Thus, *M. acetivorans* is an organism that is able to use the Ack/Pta pathway to growth either acetogenically or acetotrophically.

Recently, the methoxydotrophic mode of methanogenesis was proposed (Mayumi et al., 2016). This pathway is proposed to couple o-demethylation, CO_2 reduction and acetyl-CoA metabolism to produce methane from methoxylated aromatic compounds (Mayumi et al., 2016). Pyruvate has also been shown to support methanogenic growth in *M. barkeri* Fusaro (Bock et al., 1994). Growth on pyruvate first requires the conversion of pyruvate into acetyl-CoA catalyzed by the anabolic enzyme pyruvate ferredoxin oxidoreductase (Por), generating CO_2 in the process (Fig 1.4). Acetyl-CoA enters the acetoclastic methanogenesis pathway where the acetyl group is subsequently reduced to CH_4, generating 1.25 mol CH_4 and 1.75 mol CO_2 for every mol of pyruvate. A portion of this thesis is focused on further characterizing pyruvate growth in this strain with the ultimate goal to understand pyruvate metabolism in the cells.

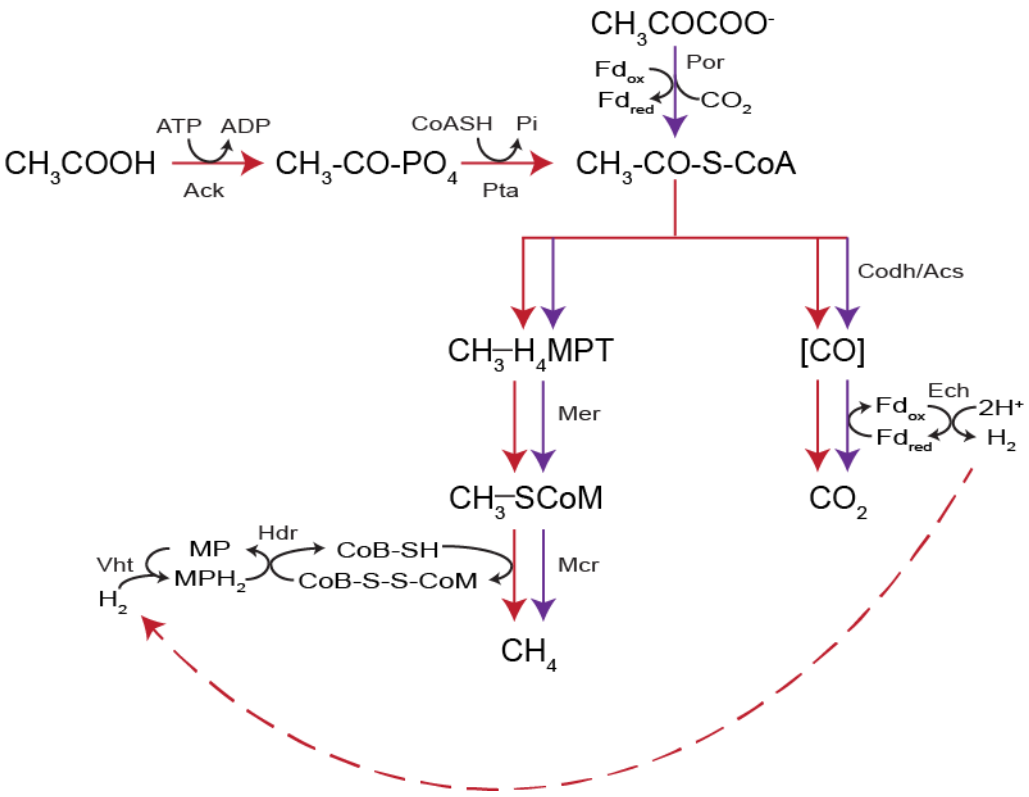

Figure 1.4 Pathways for acetoclastic methanogenesis (red) and pyruvate-derived methanogenesis (purple) in *M. barkeri* Fusaro. Acetate is split into a methyl group and an enzyme bound carbonyl group that is oxidized into CO_2 by Codh/Acs, carbon monoxide dehydrogenase / acetyl-CoA synthase. Pyruvate is decarboxylated by Por, pyruvate ferredoxin oxidoreductase, into acetyl-CoA. Pathways share the enzymes from the oxidation of CO and reduction of CH_3-H4MPT to CH_4. Abbreviations: Ack, acetate kinase; Pta, phosphotransacetylase.

Equation 1.1

$CO + H_2O \longrightarrow CO_2 + H_2$ $\quad \Delta G = -4.8$ kcal/mol of CO

$CO_2 + 4H_2 \longrightarrow CH_4 + 2H_2O$ $\quad \Delta G = -31.3$ kcal/mol of CH_4

1.4 Anabolic Metabolism in Methanogens

The majority of methanogens are autotrophs. Therefore methanogens need to use central metabolic pathways and anaplerotic reactions for the synthesis of precursors for macromolecular biosynthesis. In this section I will discuss the synthesis of two central precursors for biosynthesis; acetyl-CoA and pyruvate, and the link between anabolic and catabolic reactions in methanogenic archaea.

1.4.1 Autotrophic acetyl-CoA biosynthesis

Acetate is a key intermediate in the metabolism of methanogens, contributing 60% of the cell carbon formed during heterotrophic growth of *M. barkeri* on methanol, acetate or H_2/CO_2 (Weimer and Zeikus, 1979). Acetate serves as precursor for the synthesis of pyruvate and amino acids including alanine, aspartate and glutamate and it is the major product of autotrophic CO_2 fixation (Weimer and Zeikus, 1979). Several pathways for CO_2 fixation have been described in Archaea, all of which have in common the synthesis of acetyl-CoA from CO_2 (Reviewed in (Berg et al., 2010)).

Autotrophic methanogens assimilate CO_2 using the reductive acetyl-CoA (Wood-Ljungdahl) pathway. This is a non-cyclic pathway that results in the fixation of two molecules of CO_2 into acetyl-CoA (Fig 1.5). One molecule of CO_2 is reduced to the formyl level and bound to the carrier formyl-methanofuran; this group is subsequently reduced to a methyl group bound to H_4MPT using the enzymes of the CO_2 reduction pathway. The second CO_2 molecule is reduced to CO and remains bound to the CODH/ACS complex. CODH/ACS accepts the methyl group bound to the H_4MPT carrier and combines it with CO to ultimately form and release acetyl-CoA.

1.4.2 Pyruvate metabolism in *Methanosarcina spp.*

Pyruvate plays a central role in biosynthetic and catabolic reactions of all three domains of life. In *Methanosarcina spp*. it is a precursor for the synthesis of alanine and tricarboxylic acid cycle (TCA) intermediates including oxaloacetate and α-ketoglutarate. The TCA cycle provides intermediates for the synthesis of amino acids and tetrapyrroles. In most organisms pyruvate is derived from glycolysis via pyruvate kinase. Methanogens do not have the ability to grow on complex compounds, therefore they are required to synthesize pyruvate from acetyl-CoA using the enzyme pyruvate ferredoxin oxidoreductase (Por). Pyruvate dehydrogenase and pyruvate formate lyase are absent in methanogens suggesting that Por provides the only mechanism for pyruvate/acetyl-CoA

conversion in this group of organisms. However, in most anaerobic microorganisms Por catalyzes the oxidative decarboxylation of pyruvate into acetyl-CoA and CO_2. In anaerobic acetogenic bacteria it acts as a link between glycolysis and the Wood-Ljungdahl pathway. But in *Methanosarcina* Por catalyzes the reverse reaction in order to replenish pyruvate in the cells and serve as a link between reactions from the Wood-Ljungdahl pathway and the incomplete oxidative branch of the TCA cycle (Fig 1.5). Por from *M. barkeri* is a heterotetramer, oxygen sensitive and is able to catalyze the oxidation of α-ketobutyrate (Bock et al., 1997). As will be shown in Chapter 2 of this thesis, Por is essential for the growth of M. *barkeri* Fusaro since, rescue with pyruvate could not be achieved. This result suggests an additional function of Por in the cells other than replenishing pyruvate/acetyl-CoA.

In addition to pyruvate, oxaloacetate is an essential physiological component of methanogens (Mukhopadhyay et al., 2001). Methanogenic archaea have two enzymes capable of catalyzing the synthesis of oxaloacetate; pyruvate carboxylase (Pyc) and phosphoenolpyruvate carboxylase (Ppc). Pyc is ubiquitous in all methanogens however, members from the Methanosarcinales and *Methanothermobacter thermautotrophicus* also posses phosphoenolpyruvate carboxylase (Fig 1.5). The pyruvate carboxylase operon from *M. barkeri* Fusaro has a unique arrangement that encodes the two subunits of the biotin-dependent enzyme and its putative biotinylating enzyme (Bpl). This biotinylating enzyme gene is homologous to the *bpl* gene present in *E. coli,* which encodes a bifunctional protein that also acts as a repressor of the biotin operon (BirA) and Bpl subunit genes. This gene is homologous to the *bpl* gene present in *E. coli,* which is bifunctional with a repressor of the biotin operon (BirA) and Bpl subunits. Ppc activity is absent in *M. barkeri* extracts (Weimer and Zeikus, 1979). However, in Chapter 2 we show that Pyc mutants are viable without supplementation of oxaloacetate. This finding suggests that Ppc is expressed and is able to replenish oxaloacetate in the cells. Finally, *M. barkeri* also encodes an operon for pyruvate phosphate dikinase, an ATP-dependent enzyme

responsible for converting pyruvate into phosphoenolpyurvate (PEP), thus providing the PEP necessary for the Ppc activity.

1.4.3 Other biosynthetic reactions

Amino acid biosynthesis in methanogens resembles the pathways found in Bacteria and Eukarya. C^{13} labeling studies in *M. barkeri* using labeled acetate demonstrated that alanine, aspartate and glutamate are derived from pyruvate and oxaloacetate. Synthesis of TCA intermediates from OAA provides the precursors for more complex amino acids like lysine. Pyruvate acts as a precursor for the synthesis of α-ketobutyrate and hence isoleucine. Synthesis of pyrolysine, the 22^{nd} genetically encoded amino acid, proceeds from two molecules of lysine (Gaston et al., 2011). In methanogens and other organisms growing on carbon sources different from hexoses, gluconeogenesis is an essential pathway for the synthesis of biosynthetic intermediates and polysaccharides. A recent study in *M. acetivorans* C2A showed that cells expressed all the enzymes required for gluconeogenesis and glycogen synthesis in a growth phase-dependent manner (Santiago-Martinez et al., 2016). *M. acetivorans* is able to synthesize glycogen in conditions where methanol was the carbon source. When the carbon source is depleted, cells are able to run the pathway in reverse (glycolytic direction) to synthesize acetyl-CoA via Por (Santiago-Martinez et al., 2016). Acetyl-CoA is then converted to CH_4 providing an ATP supply to the cells.

1.5 Transcription in Archaea

Chapters 3 and 4 of this thesis focus on the transcriptional regulation of methylotrophic methanogenesis thus; in the next sections I will introduce the topic of transcription in Archaea. Archaeal gene expression processes have features corresponding to bacteria and eukaryotes. For example, archaea have basal transcription machinery that resembles that present in eukaryotes and

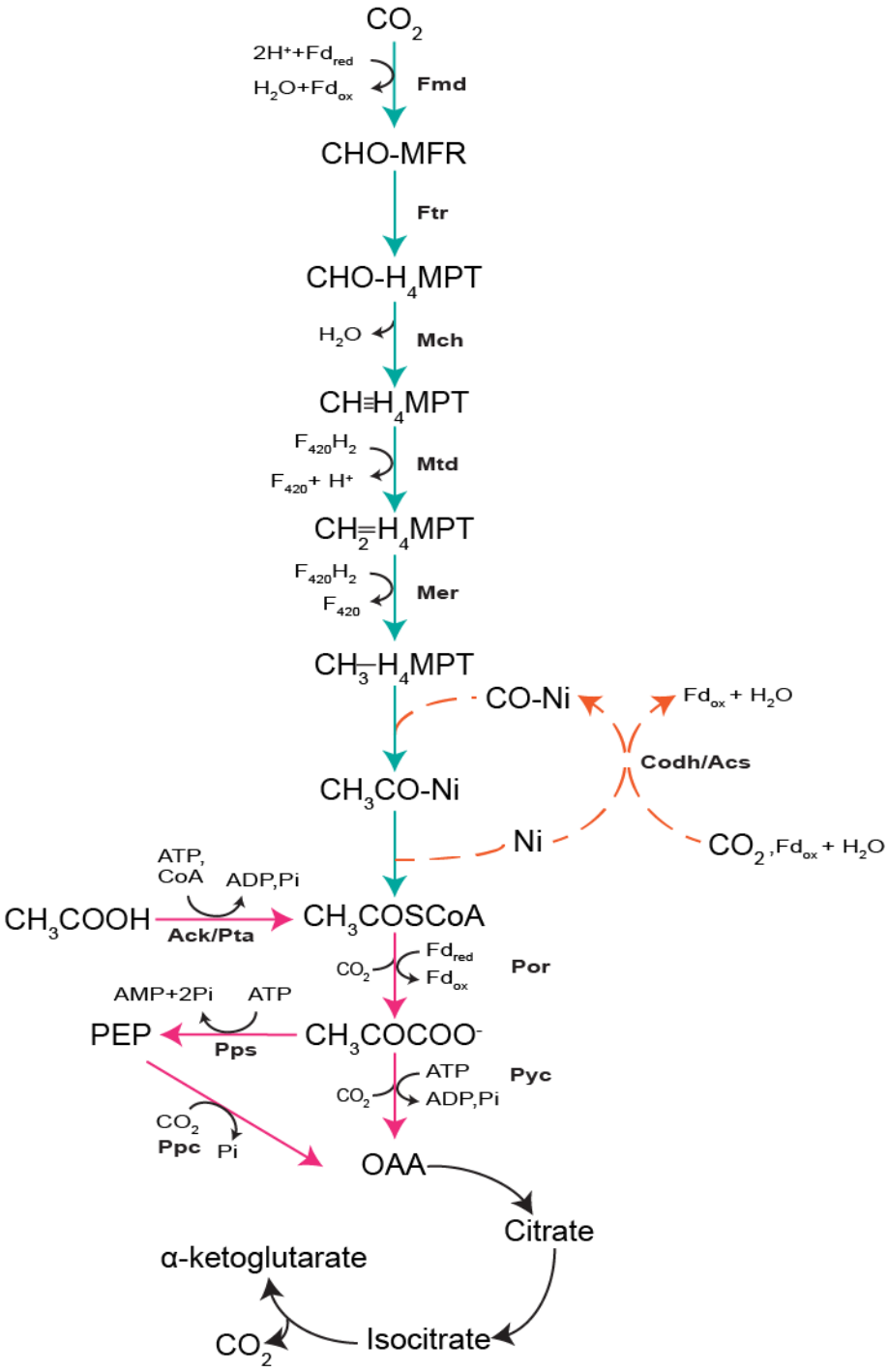

Figure 1.5 Anabolic reactions in *Methanosarcina spp*. In the reductive acetyl-CoA pathway (green) two molecules of CO_2 are used to synthesize acetyl-CoA (abbreviations found on Figure 1.2 legend). Reactions linked to pyruvate metabolism (pink arrows) provide the precursors for the synthesis of tricarboxilic acid cycle intermediates. Abbreviations: Cdh/Acs, carbon monoxide dehydrogenase / acetyl-CoA synthase; Por, pyruvate ferredoxin oxidoreductase; Pyc, pyruvate carboxylase; Ack/Pta, acetate kinase /phosphotransacetylase; Pps, phosphoenolpyruvate synthase; Ppc, phosphoenolpyruvate carboxylase.

some of their members have genomes that are organized by eukaryotic-like histones. However, archaeal gene organization and gene-specific transcription regulation are similar to those present in bacteria.

1.5.1 Transcription basal machinery

The archaeal transcription machinery shares a remarkable similarity to that of eukaryotes. The archaeal RNA polymerase is a 12-subunit polymerase that shares similar composition, structure and mechanism to the eukaryal RNA polymerase II (Ao et al., 2013). The canonical promoter of members of the Archaea consists of two well-conserved elements: the TATA box (A-T rich region 25-27bp upstream of the transcription start site) and the B recognition-element (purine rich region) (BRE) (Fig 1.6). Two main transcription factors have been described for transcription initiation in Archaea: the TATA box binding protein (TBP) and the transcription factor B (TFB) (Isom et al., 2013). Binding of TBP to the promoter results in a recruitment cascade that triggers the binding of the TFB protein to BRE. The DNA-TBP-TFB complex recruits the transcription factor E (TFE) and the RNA polymerase to initiate transcription. TFE interacts with the nontemplate strand and aids in stabilizing open complex formation (Ao et al., 2013). Transcription factor B on the other hand, facilitates the loading of the DNA into the RNA polymerase catalytic site (Lassak et al., 2013). A third promoter element has been described in *Sulfolobus solfataricus,* known as the initiator element (Inr) (Ao et al., 2013). Inr consensus sequence includes the transcription start site and was shown to enhance promoter strength.

1.5.2 Transcription regulation in methanogens

Two forms of transcription regulation have been described in archaea: one mediated by histone-like proteins and chromatin organization, and a second one mediated by small bacterial-like transcription factors.

DNA condensation and organization is essential in all three domains of life to facilitate processes like transcription, replication and DNA repair. In bacteria,

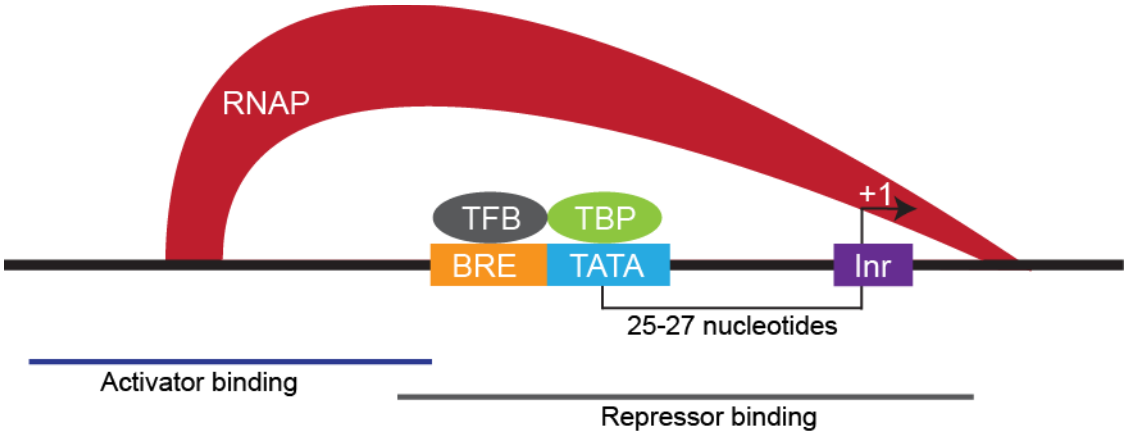

Figure 1.6 Transcription basal machinery and promoter architecture in Archaea (modified from Karr, 2014). Abbreviations: RNAP, RNA polymerase; TFB, transcription factor B; BRE, B recognition element; TBP, TATA box binding protein; TATA, TATA box; Inr, initiator element; +1, transcription start site.

chromosomes are organized into supercoiled domains using nucleoid-associated proteins (Peeters et al., 2015). Members of the Archaea, including methanogenic archaea, organize their genomes using histone-like proteins similar to the ones present in eukaryotes. Histone octamers wrap DNA around them forming structures known as nucleosomes. In eukaryotic models it has been shown that the position of the nucleosomes in the genome plays an important role in gene regulation. Moreover, histone modifications are an important mechanism to regulate gene accessibility [Reviewed in (Peeters et al., 2015)]. Several studies have shown a link between chromatin organization and transcription regulation in archaea. Archaeal histones do not have the sites for posttranslational modifications that allow eukaryal histones to have regulatory functions. Therefore, it seems unlikely that these proteins regulate transcription initiation in Archaea (Pereira et al., 1997). Nevertheless histone binding is decreased in regions that are highly transcribed like the *mcr* gene in *Methanothermobacter thermautotrophicus,* suggesting a relationship between position of nucleosomes and levels of gene expression (Pereira et al., 1997). *Methanococcus voltae* represents another example of gene regulation mediated by histone-like proteins. Deletion of chromatin gene *htsA* in this strain resulted in the overexpression of genes related to autotrophic reactions in *M. voltae*, suggesting once more a relationship between chromatin organization and gene regulation in Archaea.

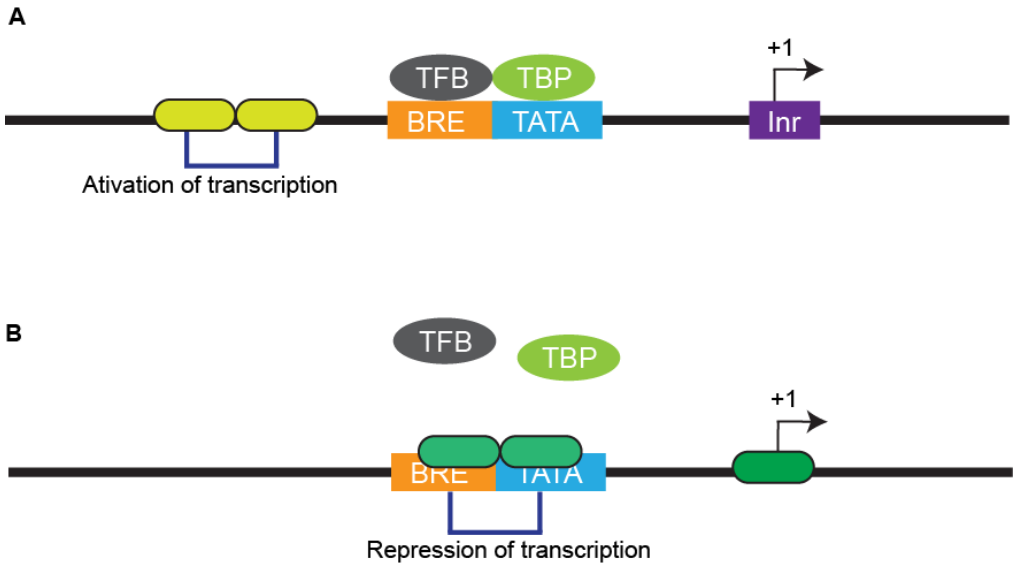

Figure 1.7 A. Transcription factor acting as an activator. Binding of transcription factors aids in the recruitment of TFB and TBP into promoter region, helping in the recruitment of the RNA polymerase and to initiate transcription. **B.** Transcription factor acting as a repressor of transcription. Binding of transcription factor to promoter region inhibits binding of TFB and TBP preventing recruitment of RNA polymerase and transcription initiation.

In archaea, gene regulation as a physiological response is mediated by small bacterial-like transcription factors (TF). Transcription factors are defined by their ability to bind a specific DNA sequence and modulate gene expression. They can act as repressors binding to a promoter and preventing recruitment of the RNA polymerase (Fig 1.7) or as activators aiding in the recruitment and stabilization of the polymerase and other transcription factors. Most of archaeal transcription regulators are similar to the ones present in bacteria having a helix-turn-helix motif, but they differ from bacterial regulators in their sensory domains. Several families of regulators have been described for the archaea, three of which have been characterized in methanogens.

The transcription regulator MsvR (methanogen specific V4R domain-containing regulator) was first characterized in *M. thermautotrophicus (Karr, 2010).* The N terminus domain contains a helix-turn-helix DNA binding domain and the C terminus has five-cysteine containing domain VR4 (Fig 1.8). MsvR is a redox sensing transcriptional regulator that binds to its own promoter and the promoter of the *fpaA-rlp-rub* operon (involved in oxidative stress response) in this strain.

Similar to *M. thermautotrophicus* MsvR, *M. acetivorans* MsvR (MaMsvR) binds to its own promoter in a redox dependent manner. However, it does not seem to be involved in regulation of oxidative stress response genes (Isom et al., 2013). Gel filtration experiments revealed that MaMsvR acts as a dimer and mutational analysis showed that the five-cysteine residues are important for redox sensing and gene regulation. The second regulator described in methanogenic archaea

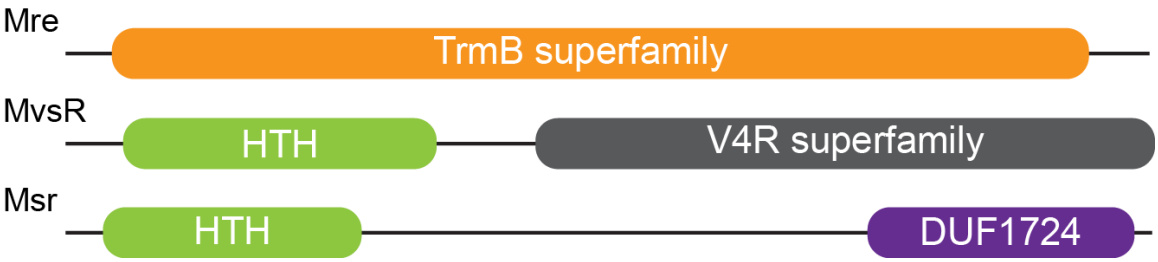

Figure 1.8 Protein domains of the three transcriptional regulators described in *Methanosarcina*. Abbreviations: TrmB, sugar metabolism regulator; HTH, helix-turn-helix domain; V4R, predicted small molecular binding domain; DUF1724, domain of unknown function present only in archaea.

is Mre, a regulator of *Methanosarcina* energy-converting metabolism (Reichlen et al., 2012). Mre has an amino-terminal DNA binding domain and lacks the sensory domain typically found in bacterial transcription factors (Fig 1.8). It was shown that *mre* is highly expressed during growth on acetate and that its gene product is able to bind the promoters and activate the transcription of genes related to acetate metabolism. In addition, it was shown that Mre could also act as a repressor of methylotrophic methanogenesis genes. This dual function of archaeal transcription factors as both activators and repressor has been reported in other archaeal regulators including methanol specific regulators (Msr) present in *M. acetivorans* and which are further studied in this thesis.

The methyltransferase genes (refer to Fig 1.2) of *Methanosarcina* are among the most highly regulated genes in all Archaea (Bose and Metcalf, 2008). Studies on the factors that allow this stringent regulation led to the identification of the third kind of transcription factors in *Methanosarcina,* the Msr proteins. Six Msr regulators have been described by the Metcalf lab (Bose and Metcalf, 2008) all of which have a N-terminal helix-turn-helix domain and a domain of unidentified

function (DUF1724) found only in Archaea (Fig 1.8). Mrs proteins form a heterodimer that binds to the promoter of the methyltransferase gene to either repress or activate its transcription (unpublished data). Part of the work in this thesis focuses on elucidating the sensory mechanisms that trigger Msr mediated regulation.

1.5.3 Two-component systems in methanogenic archaea

Even though multiple transcription factors have been described in Archaea, the mechanisms that trigger the regulatory responses remain understudied. Two component (TCS) systems are one way by which microorganisms sense and respond to their environments. TCS are typically composed of a sensory histidine kinase (HK) and its cognate response regulator (RR) (Fig 1.9). HK proteins are phosphorylated at a conserved histidine residue as a result of an environmental signal. The phosphoryl group is then transferred to a conserved aspartate residue in the cognate response regulator, which in turn regulates gene expression in the cell (Najnin et al., 2016). Multiple sensory domains have been identified in HKs including CHASE, PAS and HAMP, enabling HKs to sense and respond to a wide variety of stimuli. Characteristic RRs have a receiver domain (REC), which is linked to an effector domain that mediates DNA or ligand binding (Li et al., 2014). However, this effector domain is rarely present in RRs annotated in Archaea suggesting that archaeal response regulators act in a different way than those of bacteria, possibly by modulating the activity of other transcriptional regulators (Wuichet et al., 2010).

Only two TCS have been biochemically characterized in Archaea both in members of the methanogens. The first TCS characterized is present in *Methanosaeta hurundinacea*, where phosphorylation assays showed that the HK FilI was able to autophosphorylate, followed by transfer of the phosphoryl moiety to two independent response regulators FilR1 and FilR2 (Li et al., 2014). In addition, it was shown that the phosphorylation state of the RRs affected the expression of genes related to acetoclastic methanogenesis in *M. hurundinacea*.

Recently a temperature sensitive TCS was characterized in the psychrophilic archaeon *Methanococcoides burtonii* (Najnin et al., 2016). *Methanococcoides burtonii* histidine kinase LtrK has a HisKA, HATPase and the sensory domain

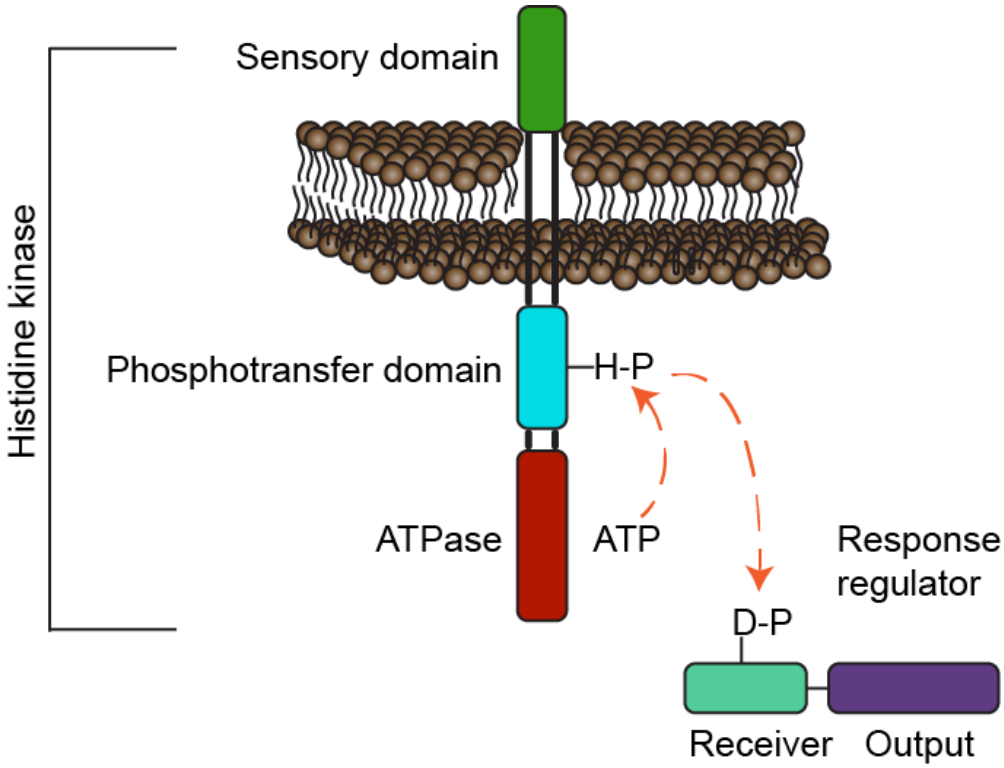

Figure 1.9 Schematic of a two-component system (adapted from Laub, 2007). The extracellular domain senses a stimuli that triggers a conformational change in the cytoplasmic region if the HK. HK autophosphorylates and transfers the phosphoryl group to the receiver domain of the response regulator. The response regulator undergoes a conformational change that allows it to do gene regulation.

CHASE, while the response regulator LtrR has both a receiver and an effector domain. Kinase and phosphatase activities on wild-type and mutant proteins showed the ability of LtrK to autophosphorylate and transfer the phosphoryl group to LtrR in a temperature-dependent manner. Two component systems remain vastly understudied in the Archaea. Chapter 4 of this thesis focuses on the study of three histidine kinases systems and their effect in the control of genes related to methylotrophic methanogenesis in *M. acetivorans* C2A, expanding our knowledge of gene regulation in Archaea.

1.6 References

1. ALLEN, G. 2016. Biogeochemistry: Rebalancing the global methane budget. *Nature,* 538**,** 46-48.

2. AO, X., LI, Y., WANG, F., FENG, M., LIN, Y., ZHAO, S., LIANG, Y. & PENG, N. 2013. The Sulfolobus initiator element is an important contributor to promoter strength. *J Bacteriol,* 195**,** 5216-22.

3. ASCHENBACH, K., CONRAD, R., REHAKOVA, K., DOLEZAL, J., JANATKOVA, K. & ANGEL, R. 2013. Methanogens at the top of the world: occurrence and potential activity of methanogens in newly deglaciated soils in high-altitude cold deserts in the Western Himalayas. *Front Microbiol,* 4**,** 359.

4. BERG, I. A., KOCKELKORN, D., RAMOS-VERA, W. H., SAY, R. F., ZARZYCKI, J., HUGLER, M., ALBER, B. E. & FUCHS, G. 2010. Autotrophic carbon fixation in archaea. *Nat Rev Microbiol,* 8**,** 447-60.

5. BOCK, A. K., PRIEGER-KRAFT, A. & SCHONHEIT, P. 1994. Pyruvate- a novel substrate for growth and methane formation in *Methanosarcina barkeri Arch Microbiol* 161**,** 33-46.

6. BOCK, A. K., SCHONHEIT, P. & TEIXEIRA, M. 1997. The iron-sulfur centers of the pyruvate:ferredoxin oxidoreductase from Methanosarcina barkeri (Fusaro). *FEBS Lett,* 414**,** 209-12.

7. BORREL, G., O'TOOLE, P. W., HARRIS, H. M., PEYRET, P., BRUGERE, J. F. & GRIBALDO, S. 2013. Phylogenomic data support a seventh order of Methylotrophic methanogens and provide insights into the evolution of Methanogenesis. *Genome Biol Evol,* 5**,** 1769-80.

8. BOSE, A. & METCALF, W. W. 2008. Distinct regulators control the expression of methanol methyltransferase isozymes in Methanosarcina acetivorans C2A. *Mol Microbiol,* 67**,** 649-61.

9. COSTA, K. C. & LEIGH, J. A. 2014. Metabolic versatility in methanogens. *Curr Opin Biotechnol,* 29**,** 70-5.

10. DANIELS L, F. G., THAUER R.K., ZEIKUS J..G. 1977. Carbon Monoxide Oxidation by Methanogenic Bacteria *Journal of Bacteriology* 132**,** 118-126.

11. EPA 2010. Methane and Nitrous Oxide Emissions from Natural Sources. Washington, DC, USA: U.S. Environmental Protection Agency.

12. FERRY, J. G. 2015. Acetate Metabolism in Anaerobes from the Domain Archaea. *Life (Basel),* 5**,** 1454-71.

13. GASTON, M. A., ZHANG, L., GREEN-CHURCH, K. B. & KRZYCKI, J. A. 2011. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature,* 471**,** 647-50.

14. ISOM, C. E., TURNER, J. L., LESSNER, D. J. & KARR, E. A. 2013. Redox-sensitive DNA binding by homodimeric Methanosarcina acetivorans MsvR is modulated by cysteine residues. *BMC Microbiol,* 13**,** 163.

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

16. KIETAVAINEN, R. & PURKAMO, L. 2015. The origin, source, and cycling of methane in deep crystalline rock biosphere. *Front Microbiol,* 6**,** 725.

17. KUMAR, V. K., FERRY, J.G., MARANAS, C.D. 2011. Metabolic reconstruction of the archaeon methanogen Methanosarcina acetivorans *BMC Systems Biology,* 5.

18. LANG, K., SCHULDES, J., KLINGL, A., POEHLEIN, A., DANIEL, R. & BRUNEA, A. 2015. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of "Candidatus methanoplasma termitum". *Appl Environ Microbiol,* 81**,** 1338-52.

19. LASSAK, K., PEETERS, E., WROBEL, S. & ALBERS, S. V. 2013. The one-component system ArnR: a membrane-bound activator of the crenarchaeal archaellum. *Mol Microbiol,* 88**,** 125-39.

20. LI, J., ZHENG, X., GUO, X., QI, L. & DONG, X. 2014. Characterization of an archaeal two-component system that regulates methanogenesis in Methanosaeta harundinacea. *PLoS One,* 9**,** e95502.

21. MAEDER, D. L., ANDERSON, I., BRETTIN, T. S., BRUCE, D. C., GILNA, P., HAN, C. S., LAPIDUS, A., METCALF, W. W., SAUNDERS, E., TAPIA, R. & SOWERS, K. R. 2006. The Methanosarcina barkeri genome: comparative analysis with Methanosarcina acetivorans and Methanosarcina mazei reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol,* 188**,** 7922-31.

22. MARTINEZ-FERNANDEZ, G., DENMAN, S. E., YANG, C., CHEUNG, J., MITSUMORI, M. & MCSWEENEY, C. S. 2016. Methane Inhibition Alters the Microbial Community, Hydrogen Flow, and Fermentation Response in the Rumen of Cattle. *Front Microbiol,* 7**,** 1122.

23. MAYUMI, D., MOCHIMARU, H., TAMAKI, H., YAMAMOTO, K., YOSHIOKA, H., SUZUKI, Y., KAMAGATA, Y. & SAKATA, S. 2016. Methane production from coal by a single methanogen. *Science,* 354**,** 222-225.

24. MUKHOPADHYAY, B., PURWANTINI, E., KREDER, C. L. & WOLFE, R. S. 2001. Oxaloacetate synthesis in the methanarchaeon Methanosarcina barkeri: pyruvate carboxylase genes and a putative Escherichia coli-type bifunctional biotin protein ligase gene (bpl/birA) exhibit a unique organization. *J Bacteriol,* 183**,** 3804-10.

25. NAJNIN, T., SIDDIQUI, K. S., TAHA, ELKAID, N., KORNFELD, G., CURMI, P. M. & CAVICCHIOLI, R. 2016. Characterization of a temperature-responsive two component regulatory system from the Antarctic archaeon, Methanococcoides burtonii. *Sci Rep,* 6**,** 24278.

26. NOBU, M. K., NARIHIRO, T., KURODA, K., MEI, R. & LIU, W. T. 2016. Chasing the elusive Euryarchaeota class WSA2: genomes reveal a uniquely fastidious methyl-reducing methanogen. *ISME J,* 10**,** 2478-87.

27. ODONGO, N. E., OR-RASHID, M. M., BAGG, R., VESSIE, G., DICK, P., KEBREAB, E., FRANCE, J. & MCBRIDE, B. W. 2007. Long-term effects of feeding monensin on milk fatty acid composition in lactating dairy cows. *J Dairy Sci,* 90**,** 5126-33.

28. OFFRE, P., SPANG, A. & SCHLEPER, C. 2013. Archaea in biogeochemical cycles. *Annu Rev Microbiol,* 67**,** 437-57.

29. PEETERS, E., DRIESSEN, R. P., WERNER, F. & DAME, R. T. 2015. The interplay between nucleoid organization and transcription in archaeal genomes. *Nat Rev Microbiol,* 13**,** 333-41.

30. PEREIRA, S. L., GRAYLING, R. A., LURZ, R. & REEVE, J. N. 1997. Archaeal nucleosomes. *Proc Natl Acad Sci U S A,* 94**,** 12633-7.

31. REICHLEN, M. J., VEPACHEDU, V. R., MURAKAMI, K. S. & FERRY, J. G. 2012. MreA functions in the global regulation of methanogenic pathways in Methanosarcina acetivorans. *MBio,* 3**,** e00189-12.

32. ROTHER, M. & METCALF, W. W. 2004. Anaerobic growth of Methanosarcina acetivorans C2A on carbon monoxide: an unusual way of life for a methanogenic archaeon. *Proc Natl Acad Sci U S A,* 101**,** 16929-34.

33. SAKAI, S., TAKAKI, Y., SHIMAMURA, S., SEKINE, M., TAJIMA, T., KOSUGI, H., ICHIKAWA, N., TASUMI, E., HIRAKI, A. T., SHIMIZU, A., KATO, Y., NISHIKO, R., MORI, K., FUJITA, N., IMACHI, H. & TAKAI, K. 2011. Genome sequence of a mesophilic hydrogenotrophic methanogen Methanocella paludicola, the first cultivated representative of the order Methanocellales. *PLoS One,* 6**,** e22898.

34. SANTIAGO-MARTINEZ, M. G., ENCALADA, R., LIRA-SILVA, E., PINEDA, E., GALLARDO-PEREZ, J. C., REYES-GARCIA, M. A., SAAVEDRA, E., MORENO-SANCHEZ, R., MARIN-HERNANDEZ, A. & JASSO-CHAVEZ, R. 2016. The nutritional status of Methanosarcina acetivorans regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes. *FEBS J,* 283**,** 1979-99.

35. THAUER, R. K. 2012. The Wolfe cycle comes full circle. *Proc Natl Acad Sci U S A,* 109**,** 15084-5.

36. VANWONTERGHEM, I., EVANS, P. N., PARKS, D. H., JENSEN, P. D., WOODCROFT, B. J., HUGENHOLTZ, P. & TYSON, G. W. 2016. Methylotrophic methanogenesis discovered in the archaeal phylum Verstraetearchaeota. *Nat Microbiol,* 1**,** 16170.

37. VON KLEIN, D., ARAB, H., VOLKER, H. & THOMM, M. 2002. Methanosarcina baltica, sp. nov., a novel methanogen isolated from the Gotland Deep of the Baltic Sea. *Extremophiles,* 6**,** 103-10.

38. WEIMER, P. J. & ZEIKUS, J. G. 1979. Acetate assimilation pathway of Methanosarcina barkeri. *J Bacteriol,* 137**,** 332-9.

39. WUICHET, K., CANTWELL, B. J. & ZHULIN, I. B. 2010. Evolution and phyletic distribution of two-component signal transduction systems. *Curr Opin Microbiol,* 13**,** 219-25.

Chapter 2: Genetic, Genomic and Transcriptomic Studies of Pyruvate Metabolism in *Methanosarcina barkeri* Fusaro[*]

2.1 Abstract

Pyruvate, a central intermediate in the carbon fixation pathway of methanogenic archaea, is rarely used as an energy source by these organisms. The sole exception to this rule is a genetically-uncharacterized *Methanosarcina barkeri* mutant capable of using pyruvate as a sole energy and carbon source (Pyr^+ phenotype). Here, we show that the Pyr^+ mutant is able to metabolize pyruvate by overexpressing pyruvate ferredoxin oxidoreductase (*por*) and mutating genes involved in central carbon metabolism. Genomic analysis showed that the Pyr^+ strain has two mutations localized to Mbar_A1588, the biotin protein ligase subunit of the pyruvate carboxylase (*pyc*) operon, and Mbar_A2165, a putative transcriptional regulator. Mutants expressing the Mbar_A1588 mutation showed no growth defect when compared to WT, yet the strains lacked *pyc* activity. Re-creation of the Mbar_A2165 mutation resulted in a two-fold increase of Por activity and gene expression, thereby suggesting a role in *por* transcriptional regulation. Further transcriptomic analysis revealed that Pyr^+ strains also overexpress the gene encoding phosphoenolpyruvate carboxylase, indicating the presence of a previously uncharacterized route for synthesizing oxaloacetate in *M. barkeri* an explaining the unimpaired growth in the absence of Pyc. Surprisingly, stringent repression of the *por* operon was lethal, even when the media were supplemented with pyruvate and/or casamino acids, suggesting that *por* plays an unidentified essential function in *M barkeri*.

[*] Parts of this chapter are reprinted with permission from López Muñoz, M., Shönheit, P., and Metcalf, W.W. (2015) Genetic, genomic and transcriptomic studies of pyruvate metabolism in *Methanosarcina barkeri* Fusaro. *J. Bacteriol.* 197, 3592–3600. Copyright (2015) American Society for Microbiology.

2.2 Introduction

A diverse group of strictly anaerobic archaea are responsible for virtually all biologically produced methane, making them key players in the global carbon cycle (Thauer et al., 2008). These unusual microorganisms grow using a very limited number of substrates, obtaining all of their energy from the methanogenic process. While most methanogens use only H_2/CO_2 or formate as growth substrates, members of the *Methanosarcinales* show some metabolic diversity, with many species using H_2/CO_2, various one-carbon (C-1) compounds (*e.g.* methanol, methylamines, and methyl-sulfides) and acetate as growth substrates. A few methanogenic species can use longer chain alcohols as energy sources, but these molecules are not assimilated. Instead they are oxidized to provide the reducing equivalents needed for reduction of CO_2 to methane, then excreted (Widdel, 1986). Larger organic compounds, such as sugars and fatty acids, are almost never substrates for methanogenic archaea, although they are co-metabolized via syntrophic microbial communities (McInerney et al., 2009). The sole exception to this rule is *Methanosarcina barkeri,* which has been reported to grow on pyruvic acid.

Two reports of methanogenic growth on pyruvate appeared in the early 1990s. In the first, a pyruvate-utilizing (Pyr^+) mutant of *M. barkeri* Fusaro, was obtained after prolonged incubation (8-12 weeks) of methanol-grown cells inoculated into a medium containing pyruvate as the sole substrate (Bock et al., 1994). Although initial growth was rather slow, repeated passages in pyruvate medium gave rise to a strain with a doubling time of *ca.* 24 hrs. The conclusion that growth on pyruvate was enabled by mutation, rather than by adaptation, was based on the observation that the long lag needed for switching from methanol to pyruvate was lost in this strain. Characterization of the mutant suggested that growth involved a pathway in which pyruvate was oxidized to acetyl-CoA and CO_2 using the enzyme pyruvate:ferredoxin oxidoreductase (Por), with subsequent conversion of acetyl-CoA to CO_2 and CH_4 by the standard aceticlastic pathway for methanogenesis (Bock et al., 1994, Bock and Schonheit, 1995). The reducing

equivalents produced by pyruvate oxidation were also used to reduce CO_2 to CH_4, probably using H_2 as an intermediate. Interestingly, the mutant displayed limited non-methanogenic growth when methanogenesis was inhibited by addition of bromoethanesulfionic acid (BES); however, growth ceased after *ca.* eight generations (Bock and Schonheit, 1995). Thus, the notion that all methanogens are obligate methanogens remains intact. Rajagopal and LeGall also reported growth of *M. barkeri* on pyruvate (Rajagopal and Legall, 1994), but using different strains (227, MS and UBS). The phenotype of these strains when grown on pyruvate was substantially different than the one observed in the Pyr^+ mutant of *M. barkeri* Fusaro. *M. barkeri* 227, MS and UBS grew substantially slower (doubling time ~150 hr vs 25 hr), with lower levels of methane production (~0.7 mol CH_4/mol pyruvate vs ~1.25 mol CH_4/mol pyruvate) and significantly lower growth yields (~1.7 g/mol CH_4 vs ~14 g/mol CH_4), relative to the *M. barkeri* Fusaro mutant. Whether growth of these strains on pyruvate required mutation or adaptation was not addressed.

In addition to serving as an energy source, pyruvate plays a critical role as a biosynthetic intermediate for numerous anabolic pathways in methanogens (reviewed in (Simpson and Whitman, 1993)). Based on biochemical data, *M. barkeri* synthesizes pyruvate by reductive carboxylation of Ac-CoA mediated by Por, which is a freely reversible enzyme (Furdui and Ragsdale, 2000). Pyruvate then enters the gluconeogenic pathway via phosphoenolpyruvate (PEP) and the branched tricarboxylic cycle via oxaloacetate (OAA). The former is produced via the enzyme PEP synthase (Weimer and Zeikus, 1979), whereas the latter is produced via the biotin-dependent enzyme pyruvate carboxylase (Pyc, (Mukhopadhyay et al., 2001)).

The genetic basis for the Pyr^+ phenotype of the *M. barkeri* Fusaro mutant has yet to be established. In this chapter, I explored the molecular, genetic and biochemical bases for growth of the *M. barkeri* Fusaro Pyr^+ mutant on pyruvate. The Pyr^+ mutant is able to metabolize pyruvate by overexpressing pyruvate ferredoxin oxidoreductase (*por*) and mutating genes involved in central carbon

metabolism. Genomic analysis showed that the Pyr^+ strain has two mutations localized to Mbar_A1588, the biotin protein ligase subunit of the pyruvate carboxylase (*pyc*) operon, and Mbar_A2165, a putative transcriptional regulator. Mutants expressing the Mbar_A1588 mutation showed no growth defect when compared to WT, yet the strains lacked *pyc* activity. Re-creation of the Mbar_A2165 mutation resulted in a two-fold increase of Por activity and gene expression, thereby suggesting a role in *por* transcriptional regulation. Further transcriptomic analysis revealed that Pyr^+ strains also overexpress the gene encoding phosphoenolpyruvate carboxylase, indicating the presence of a previously uncharacterized route for synthesizing oxaloacetate in *M. barkeri* an explaining the unimpaired growth in the absence of Pyc. Surprisingly, stringent repression of the *por* operon was lethal, even when the media were supplemented with pyruvate and/or casamino acids, suggesting that *por* plays an unidentified essential function in *M barkeri*. Collectively, the work presented here reveals a complex interaction between anabolic and catabolic pathways involving pyruvate metabolism in *Methanosarcina barkeri* Fusaro.

2.3 Materials and Methods

2.3.1 Strains, media and growth conditions

Methanosarcina strains used in the study are shown in Table 2.1. All *Methanosarcina* strains were grown under strictly anaerobic conditions. Derivatives of *M.* barkeri WWM85 (Guss et al., 2008) were grown in single cell morphology at 37°C in high-salt (HS) medium (Metcalf et al., 1997). *M. barkeri* DSM804, Pyr-1 and Pyr-2 were grown in aggregated cell morphology at 37°C in imidazole medium as previously described (4). Growth substrates provided were 125 mM methanol, 120 mM acetate, 100 mM sodium pyruvate, or a mixture of methanol (125 mM) and acetate (120 mM); methanol (125 mM) and pyruvate (100 mM); methanol (125 mM), acetate (120 mM) plus pyruvate (100 mM); or methanol (125 mM) with H_2/CO_2 (80:20) at 200 kPa over ambient pressure. Substrates were sterilized by filtration using a 0.2 µm filter. *M. barkeri* Pyr-1 was independently maintained after transfer to the Metcalf lab in January, 2004, *M.*

barkeri Pyr-2 has been maintained in the Schönheit laboratory since its isolation (Bock et al., 1994). *M. barkeri* DSM804 was received from DSMZ in May 2011. Growth on media solidified with 1.5% agar was as described previously (William Metcalf, 1996). All plating manipulations were carried out under strictly anaerobic conditions in an anaerobic incubator as described (William Metcalf, 1998). Puromycin (CalBiochem, San Diego, CA) was added from sterile, anaerobic stocks at a final concentration of 2 μg/ml for selection of *Methanosarcina* strains carrying puromycin transacetylase gene (*pac*). The purine analog 8-aza-2,6-diaminopurine (Sigma, St. Louis, MO) was added from sterile, anaerobic stocks at a final concentration of 20 μg/ml for selection against the *hpt* gene. Tetracycline was added to a final concentration of 100 μg/mL to induce expression of genes driven by the pMcrB(tetO1) promoter. Unless otherwise stated 10% inoculum was used as a starting culture for pyruvate-utilizing strains

Table 2.1 *Methanosarcina* strains used in this study.

Strain	Genotype	Source
Methanosarcina barkeri DSM804	wild type	DSMZ[a]
Methanosarcina barkeri Pyr-1	Pyr^+, see Table 2	lab stock
Methanosarcina barkeri Pyr-2	Pyr^+, see Table 2	(Bock et al., 1994)
Methanosarcina barkeri WWM85	*Dhpt::ϕC31int-attP*	(Guss et al., 2008)
Methanosarcina barkeri WWM818	*Dhpt:: ϕC31int-attP, pycC1*	this study
Methanosarcina barkeri WWM940	*Dhpt:: ϕC31int-attP, pycC1, Mbar_A2165(G59R)*	this study
Methanosarcina barkeri WWM941	*Dhpt:: ϕC31int-attP, pMcrB(tetO1)::porCDAB*	this study
Methanosarcina barkeri WWM945	*Dhpt:: ϕC31int-attP, pMcrB(tetO1)::porCDAB,pycC1*	this study

[a]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures.

2.3.2 Construction and verification of mutant strains

M. barkeri strains WWM818 and WWM940 were constructed by markerless genetic exchange via liposome-mediated transformation (Metcalf et al., 1997, Pritchett et al., 2004b). WWM818 was made from WWM85 using pML09. WWM940 was made from WWM818 using pML34. WWM941 was made by gene replacement (Zhang et al., 2002) after transformation of WWM85 to puromycin resistance with pML44. All plasmids were confirmed by DNA sequencing at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois. All mutations were verified by PCR and DNA hybridization experiments

(data not shown). DNA hybridizations were performed using the DIG System (Roche, Mannheim, Germany) as recommended using MagnaGraph Nylon transfer membranes (Micron Separations Inc, Westborough, MA). Standard methods were used to isolate and manipulate plasmid DNA from *E. coli* strains (Ausubel, 1992). For plasmid descriptions and strain genotypes refer to Table 2.2 and Table 2.1 respectively.

2.3.3 Genome sequencing and mapping analysis

Genomic DNA from *M. barkeri* was isolated as described in (Paolo Boccazzi, 2000) and sequenced on the Illumina HiSeq2000 platform at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois. DNA libraries were prepared with Illumina TruSeq DNAseq Sample Prep kit (Illumina, San Diego, CA); quantified by qPCR and sequenced for 100 cycles using a TruSeq SBS sequencing kit

Table 2.2 Plasmids used in this study.

Plasmids	Description	Reference
pMP44	Ampicillin resistant vector used to mutate cells using the markerless exchange method in *Methanosarcina barkeri*	(Pritchett et al., 2004a)
pML09	NotI/SpeI digested 2kb PCR product surrounding the mutation on Mbar_A1588 (using primers Mbar_A1588F and Mbar_A1599R) cloned into NotI/SpeI digested pMP44	This study
pML34	NotI/SpeI digested 2kb PCR product surrounding the mutation on Mbar_A2165 (using primers Mbar_A2165F and mbar_A2165 R) cloned into pMP44 NotI/SpeI digested pMP44	This study
pGK050	Kanamycin resistant vector used for testing gene essentiality using PmcrB (tetO1)	(adam m guss, 2008)
pML42	ApaI/NcoI digested 1.5kb *por* upstream PCR product (using primers por_up_F and por_up_R) cloned into ApaI/NcoI digested pGK50	This study
pML44	NdeI/SpeI digested 1.5kb *por* downstream PCR product (using primers por_gene_F and por_gene_R) cloned into NdeI/SpeI digested pML42	This study

(Illumina, San Diego, CA). Reads were trimmed (quality: 0.001, maximum: 100, minimum: 99) and mapped to the *M. barkeri* reference genome of (NC_007333, NC_007349) using CLC genomics workbench (CLC Genomics Workbench 7.0.3; CLCbio, Aarhus, Denmark). Parameters for mapping were mismatch: 2, insertion: 3, deletion: 3, length: 0.8, similarity: 0.95, mapped randomly. The same program was used to identify point mutations, insertions, and deletions (parameters for single nucleotide polymorphism detection max coverage: 10, variant count: 2, variant probability: 90, parameters for insertions-deletions p-value: 0.0001, mismatch: 3, minimum number of reads: 2). A variant ratio of 0.9-1 and a frequency of 90-100% were used as selection criteria for the mutations. All mutations identified by this method were confirmed by Sanger sequencing of PCR products amplified from the appropriate strains. Genome sequencing data are deposited in GenBank SRP059541.

2.3.4 Transcriptomic analyses (RNA-seq)

Strains were adapted to specific growth substrates for at least 30 generations prior to RNA isolation. RNA was extracted from triplicate cultures grown to mid exponential phase using a modification of the ZR Fungal/Bacterial RNA MiniPrep kit (Zymo Research Irvine, CA). Briefly, 10 mL cultures were quickly chilled in a dry ice/ethanol bath to stop RNA synthesis. Cultures were pelleted by centrifugation for 10 min at 4 °C, then resuspended in 800 µL TRIzol Reagent (ThermoFisher, Grand Island, NY). Total RNA was extracted following manufacturer's instructions. Ribosomal RNA was depleted using subtractive hybridization employing biotinylated rRNA probes as previously described (Fu and Metcalf, 2015), except that probes were generated using *M. barkeri* rDNA templates.

RNA sequencing was performed at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois using strand specific libraries prepared with the Illumina ScriptSeq version 2 (Epicentre Biotechnologies, Madison, WI). Bar-coded libraries were pooled and sequenced for 101 cycles on a HiSeq2000 using the TruSeq SBS sequencing kit version 3 (Illumina, San Diego, CA). Reads

were analyzed using the CLC genomics platform (version 7.5); reads trimmed based on quality (quality: 0.001, ambiguous: 2, discard min:30, max: 100). To remove bias caused by remaining stable RNAs, the trimmed sequences were mapped to *M. barkeri* rRNA and tRNA sequences using the following parameters similarity: 0.9, length: 0.85, mismatch: 2, insertion: 3, deletion: 3. The remaining reads were used for the RNAseq analysis (similarity: 0.9, length: 0.85, mis:2, ind:3, del:3, global alignment, maximum number or reads per read: 10). Expression values were calculated as RPKM (Reads Per Kilobase per Million reads). Reads were normalized and statistical analysis was performed using the empirical analysis of digital gene expression (Robinson and Smyth, 2008). To determine differential gene expression, reads from each mutant strain were compared against reads of the WT strain when grown in the same substrate. Genes showing fold expression changes greater than 2, or less than -2, with P-values ≤ 0.05 were considered to be differentially expressed. The raw and processed RNA-seq data have been deposited in the GSE70370.

2.3.5 Phenotypic characterization

Growth rates were determined by monitoring optical density of three or more independent cultures at 600 nm using a Bausch and Lomb Spectronic 21. Gene essentiality tests were conducted using a modified procedure described previously (Buan et al., 2011). Briefly, 10 mL cultures were grown to saturation in HS-methanol broth with tetracycline (100 µg/ml), spun down and washed three times with plain HS medium, then resuspended in 5 mL HS medium. Aliquots of 10 µl of a ten-fold dilution series was then spotted onto agar-solidified HS methanol media containing some, or all, of the following: tetracycline (100 µg/ml), pyruvate (100mM), acetate (100mM), 0.01% casamino acid, 0.01% yeast extract.

2.3.6 Determination of enzymatic activities

Pyruvate carboxylase was assayed from cell extracts by coupling the formation of oxaloacetate to malate dehydrogenase (Mukhopadhyay et al., 2001). NADH oxidation by malate dehydrogenase was measured spectrophotometrically (ε_{340} = 6.22 $mM^{-1}cm^{-1}$) at 37°C. Crude cell extracts were prepared from mid exponential

phase cultures (OD_{600nm} 0.5-0.6) by spinning down 10 mL cultures and resuspending in 500 µL of lysis buffer (3M KCl, 50mM Tris-HCl (pH8), 5% inositol, 5mM $MgCl_2$, 2mM DTT, 0.5 mM phenylmethylsulfonyl fluoride). The reaction mixture consisted of 100 mM Tris-HCl (pH8), 4 mM $MgCl_2$, 250mM KCl, 20 mM $KHCO_3$, 4 mM ATP, 0.2 mM NADH, 2 U malic dehydrogenase (Sigma Aldrich, St. Louis, MO) 20mM sodium pyruvate and 20 µL crude cell extract. Reactions were started by addition of 20 µL of sodium pyruvate into 1 mL of pre-warmed assay mixture. Preparation of cell extract and enzyme assays was performed under strictly anaerobic conditions. Phosphoenolpyruvate carboxylase activity was also measured by coupling its reaction to malate dehydrogenase. Crude cell extracts were prepared as described for pyruvate carboxylase activity. The reaction mixture contained 100 mM Tris-HCl (pH8), 4 mM $MgCl_2$, 250 mM KCl, 20 mM $KHCO_3$, 4 mM ATP, 0.2 mM NADH, 2 U malic dehydrogenase, 2mM phosphoenolpyruvic acid monosodium salt and 20 µL crude cell extracts. Reactions were started by addition of phosphoenolpyruvate to 1 mL of pre-warmed assay mixture. Preparation of cell extracts and assays was performed under strictly anaerobic conditions.

Pyruvate ferredoxin oxidoreductase was assayed by monitoring the CoA- and pyruvate-dependent reduction of benzyl viologen (ε_{578} = 8.65 $mM^{-1}cm^{-1}$) (Bock et al., 1996). Crude cell extracts were prepared as previously described using 50 mM MOPS (2 mM DTT, pH 7) as lysis buffer. The assay was performed under strictly anaerobic conditions. The reaction mixture consisted of 2 mM benzyl viologen, 0.1 mM CoA, 10 mM sodium pyruvate, and 10 µL crude cell extract in a total volume of 1 mL. Reactions were started by addition of CoA to pre-warmed assay mixture. Protein concentration for all assays was determined by the Bradford method (Coomassie protein reagent, Sigma Aldrich, St. Louis, MO) using BSA as a standard.

2.4 Results

2.4.1 Identification of mutations associated to the *Pyr*$^+$ phenotype

To identify the mutations responsible for the pyruvate–utilizing phenotype of the *M. barkeri* Fusaro mutant, the genome of two cell lines (Pyr-1 and Pyr-2) that had been independently maintained by periodic serial transfer over the course of a decade was sequenced. Additionally the genome of the parental strain (DSM804) from the DSMZ stock collection, which represents the closest available relative of the Pyr$^+$ mutants, was sequenced as well.

Eighteen differences, including sixteen single nucleotide polymorphisms (SNPs) and two relatively large deletions, were observed between the three newly sequenced strains and the published genome sequence (Table 2.4). It should be noted that the published *M. barkeri* genome sequence was obtained from a cell line that had also been independently maintained by serial passage in a high-salt medium to promote growth in the single-celled morphology (Sowers et al., 1993). Thus, numerous differences between the four strains were anticipated. Comparison of shared sequence differences allowed reconstruction of the most likely mutagenic history of the four strains (Fig. 2.1). Thirteen SNPs and both of the deletions are shared between Pyr-1, Pyr-2 and DSM804. These mutations clearly occurred after separation of the cell lines leading to the genome-sequenced strain and the remaining three strains. Because the DSM804 is Pyr$^-$, these mutations do not by themselves confer the ability to utilize pyruvate. A similar argument can be made for a mutation that is found only in Pyr-1, because Pyr-1 and Pyr-2 grow equally well on pyruvate. The two remaining mutations must have arisen in the ancestor of the Pyr$^+$ lines, suggesting that one or both are responsible for the ability to use pyruvate. These putative Pyr$^+$ mutations include a frameshift in the *pycC* gene (designated the *pycC1* allele) and a glycine to arginine mutation (G59R) in the winged helix-turn-helix motif of a putative transcriptional regulator encoded by the Mbar_A2165 locus (Fig. 2.2).

Table 2.3 Primers used in this study.

Primer	Sequence	Added sites
Primers used for plasmid construction		
Mbar_A1588F	AAAAAAAAAAAAAGCGGCCGCATACAGGTTACAC	NotI
Mbar_A1588R	AAAAAAAAAAAAAACTAGTTGTCACCAGAAGCACGGC	SpeI
Mbar_A2165F	GGCGCGCCGCGGCCGCGACTTTCCTACATTCGGATAG	NotI
Mbar_A2165R	GGCGCGCCACTAGTCGCAGCCGAGCTCAGTGCCC	SpeI
por_up_F	GGCGCGCCGGGCCCTCTTTTCCCTAATCTGC	ApaI
por_up_R	GGCGCGCCCCATGGCAGGTATCCTGTTCTTT	NcoI
por_gene_F	GGCGCGCCACTAGTGATTTCTCCGTCTGCCAT	SpeI
por_gene_R	GGCGCGCCCATATGTGAAGGAAATCAGAATACACGG	NdeI
Primers used for resequencing of mutations		
Mbar_A0055F	TTACTGCAAAGAATCTTCC	none
Mbar_A0055R	TCCATCACATACAAAAATTTG	none
Mbar_A0540F	TTGTGAAAGACTCGTTTATAGC	none
Mbar_A0540R	TTTCCTGAAAGCTTATTCTG	none
Mbar_A1111F	GCACTAGTTTATATTACAGACC	none
Mbar_A1111R	GTTTCTCGCTAAAGATCG	none
Mbar_A1119F	GAAGAGGGACTTGGAGAG	none
Mbar_A1119R	AGCAAAGGATATTTTAGTTG	none
Mbar_A1236F	TGAGTATAAAAACACCTGCCAG	none
Mbar_A1236R	GAACTTCTAGAATATAGCTG	none
Mbar_A1407F	CATAAAACATCTGGATGATG	none
Mbar_A1407R	CACTCATTATTTTTATTCGTTC	none
Mbar_A1501F	ATGGAAAAGGAAGTACTGG	none
Mbar_A1501R	CTATAGAGACTTCTCTCCGG	none
Mbar_A1588F	TGAAGTGCAGAAAGGAGG	none
Mbar_A1588R	CTGGGGTTACTCCTACAG	none
Mbar_A1758F	GCTGACAAACTCGCCAAG	none
Mbar_A1758R	CAATTCTCTCGATGGCTG	none
Mbar_A2165F	ATTAATTATGGTAAATATATGCG	none
Mbar_A2165R	TATCAGGCGGTTTACAGG	none
Mbar_A3037F	GGCAAAGTACCCTTTGCGCAATATCTC	none
Mbar_A3037R	AGTCAGGAATTAATGCCAGCTCTGC	none
Mbar_A3249F	CCAATGAGTCGTTATCAGG	none
Mbar_A3249R	AAGAGCCAGAGAATTCGA	none
Mbar_A3466F	TTACCTTCCTGTTGCCTC	none
Mbar_A3466R	TCCGAGTCTGTAGCTATACT	none
Mbar_A3480F	TTAAAGTGCTCACCGGGA	none
Mbar_A3480R	ATAAATCGGATCCTTGAAAGCC	none
Mbar_A3644F	GGTAGAGGAGAGATTTAGGA	none
Mbar_A3644R	CTGTTATTCAAAGAATATCATG	none
Mbar_A3692F	CCACAAGACCTTCCCTACC	none
Mbar_A3692R	CATGCTTGCAGTTAGGGC	none
Mbar_A1226F	ACTACGGGAGCGCAGAGG	none
Mbar_A1226R	CGAACTTACTGCGATTCCTG	none
Mbar_A3656F	CAAGAAATTCCTTAACCGATGACCAATG	none
Mbar_A3656R	CTGGATCGTGTTTAATTTGATTTAGATGTA	none

Table 2.4 DNA sequence polymorphisms in *M. barkeri* Fusaro genome sequences[a].

Region[b]/ Locus tag	Annotation of altered gene	Observed polymorphism[c]	AA change[d]	Presence/absence[b]		
				Pyr-1	Pyr-2	DSM804
Mbar_A0055 (bp 67,649)	CAAX protease self-immunity	T=>C	Leu => Ser	+	+	+
Mbar_A0540 (bp 642,250)	Oligosaccharide amylase	A=>G	Arg => Gly	+	+	+
Intergenic (bp 1,366,984)	NA	T=>C	NA	+	+	+
Mbar_A1119 (bp 1,377,047)	UDP-glucose 6-dehydrogenase	ΔT	Ile => Phe	+	+	+
Mbar_A1236 (bp 1,533,125)	Hypothetical protein	T=>C	Thr => Ala	+	+	+
Mbar_A1407 (bp 1,736,118)	Transposase	T=>C	Val => Ala	+	+	+
Intergenic (bp 1,857,490)	NA	Ins:AT	NA	+	+	+
Mbar_A1588 (bp 1,958,354)	**Biotin-(acetyl-CoA carboxylase) ligase**	**ΔA**	**STOP codon**	**+**	**+**	**-**
Mbar_A1758 (bp 2,157,232)	Predicted S-layer protein	G=>A	Arg => Gln	+	+	+
Mbar_A2165 (bp 2,739,641)	**Predicted transcriptional regulator**	**G=>A**	**Gly => Arg**	**+**	**+**	**-**
Intergenic (bp 3,911,323)	NA	ΔA	A	+	-	-
Mbar_A3249 (bp 4,169,592)	Histidine kinase-like ATPases	ΔA	Frameshift	+	+	+
Intergenic (bp 4,445,293)	NA	A=>G	NA	+	+	+
Intergenic (bp 4,462,290)	NA	C=>T	NA	+	+	+
Mbar_A3644 (bp 4,695,769)	Flagellar assembly protein H	G=>A	Arg => Cys	+	+	+
Mbar_A3692 (bp 4,753,372)	DNA-directed RNA polymerase subunit A	T=>C	Ile => Met	+	+	+
Mbar_A1226 (bp 1,514,571-1,514,834)	Winged helix-turn-helix transcription repressor DNA-binding	Δ264bp	NA	+	+	+
Mbar_A3656 (bp 4,713,304-4,715,008)	Transposase	Δ1705bp	NA	+	+	+

[a]Putative Pyr$^+$ mutations are shown in bold.
[b]Numbering and sequence differences are relative to GenBank Accession number CP000099.1.
[c]Presence (+) or absence (-) of the indicated polymorphism is shown. [d]NA, Not applicable, change is not within a coding sequence.

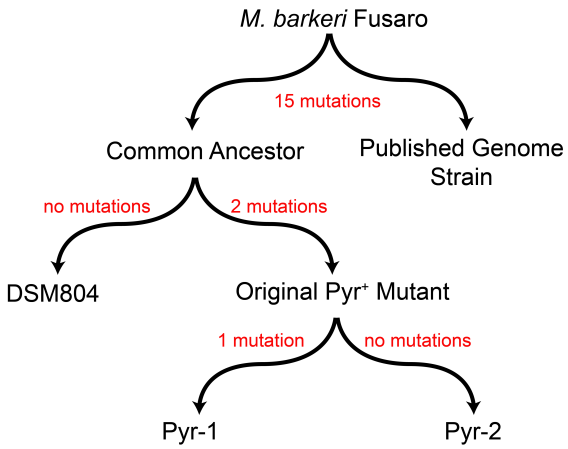

Figure 2.1 Reconstruction of the most likely genetic history of the four sequenced *Methanosarcina barkeri* Fusaro genomes based on shared sequence differences.

2.4.2 Pyruvate carboxylase activity is abolished by the *pycC1* mutation

In *M. barkeri*, OAA is synthesized by pyruvate carboxylase (Pyc), a two-subunit, biotin-dependent enzyme encoded by the *pycBAC-slp* operon. The α and β subunits of Pyc are encoded by *pycA* and *pycB*, while *pycC* encodes a bifunctional protein that covalently attaches the essential biotin cofactor to PycB, while also serving as a transcriptional repressor for biotin biosynthetic genes (Mukhopadhyay et al., 2001). The frameshift mutation identified occurs early in the biotin protein-ligase domain (Fig. 2.2). Thus, it is likely that the mutation results in a non-functional Pyc. To validate this hypothesis, a *pycC1* mutant in the strain background derived from the genome-sequenced strain (WWM85) was constructed and characterized. As expected, the resulting strain (WWM818) lacked Pyc activity; however, growth of the *pycC1* mutant was similar to that of its isogenic parent (Fig. 2.3).

The viability of the *pycC1* mutant shows that *M. barkeri* has an alternate pathway for OAA biosynthesis. Many organisms, including some methanogenic archaea, use the enzyme phosphoenolpyruvate carboxylase (Ppc) for this purpose (Patel et al., 2004). Although previous efforts failed to detect Ppc activity in *M. barkeri* cell extracts (Weimer and Zeikus, 1979), a homolog of the archaeal *ppc* gene is found in the *M. barkeri* genome sequence (locus tag Mbar_A2632). Based on the

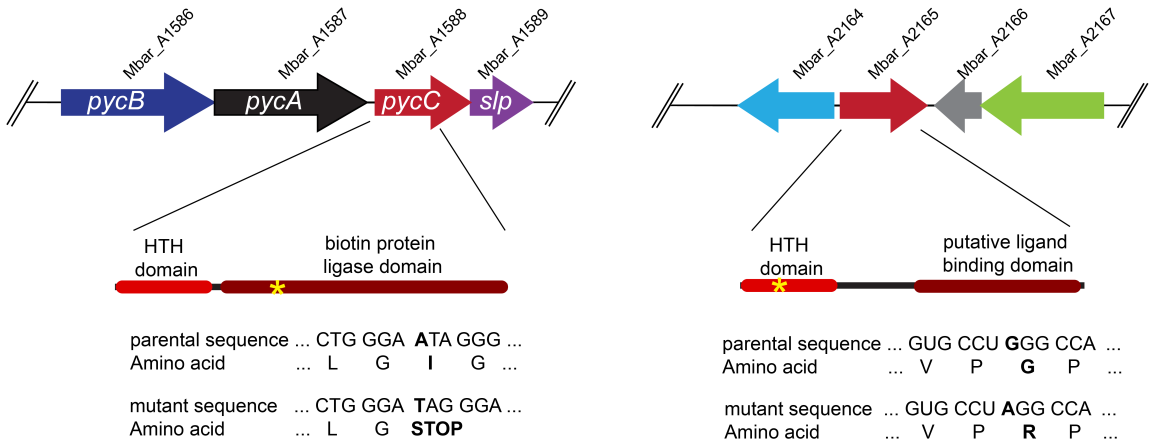

Figure 2.2 Graphic depiction of the putative Pyr$^+$ mutations. A frameshift mutation, designated *pycC1*, in the Mbar_A1588 locus introduces a STOP codon in the biotin protein ligase domain of the PycC protein. The mutated allele retains the potential to produce a truncated protein encompassing the DNA-binding domain of PycC. The second mutation results in a Gly to Arg change in the helix-turn-helix domain of the putative DNA-binding transcriptional regulator encoded by the Mbar_A2165 locus. The locus tags for each gene are shown above the arrows. Mbar_A2164 encodes a putative GTP binding protein, Mbar_A2166 encodes a putative phosphoribosyl-ATP-pyrophosphatase and Mbar_A2167 encodes a putative 3-isopropylmalate dehydratase. Mutations positions are shown by the yellow asterisk.

idea that Ppc activity might be expressed in Pyr$^+$ strains, I attempted, unsuccessfully, to assay this activity in crude extracts of the *pycC1* mutant. Thus, the route of OAA synthesis in the mutant could not be biochemically established.

2.4.3 Global transcriptomic profiling of Pyr$^+$ mutants

Because both of the putative Pyr$^+$ mutations occur in genes encoding transcriptional regulators, it seems likely that altered gene expression plays a role in the ability of the mutants to use pyruvate as a growth substrate. To test this hypothesis, we performed global transcriptional profiling of the mutant and parental strains via RNA sequencing (RNA-Seq). Because the parental strain is incapable of growth on pyruvate, RNA was isolated after growth on methanol with and without added pyruvate and, because pyruvate is thought to be metabolized via acetate, after growth on methanol with pyruvate and acetate, methanol with acetate, and acetate alone. Pairwise comparisons of mRNA abundance were then made between each of the three strains for each of the growth conditions. The full dataset, including parsed tables showing only

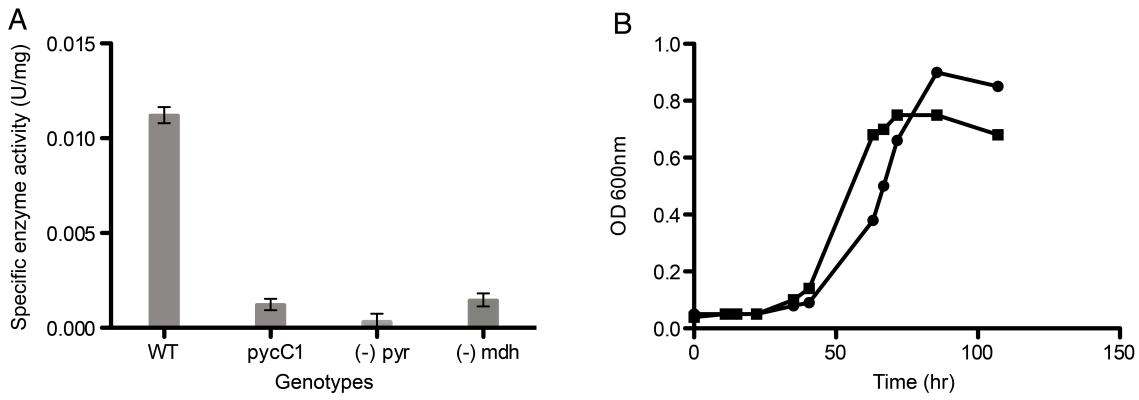

Figure 2.3 Characterization of the *pycC1* mutant. *Panel A:* Pyc activity was measured in a coupled reaction using malate dehydrogenase (Mdh) to reduce enzymatically-produced oxaloacetate to malate with concomitant oxidation of NADH. Activity for the complete reaction using extracts from the parental strain (WWM85) and the *pycC1* mutant (WWM818) are shown, as are control reactions using *pycC1* cell extract, but lacking the pyruvate substrate, (-) Pyr, and coupling enzyme, (-) Mdh. Error bars represent the standard deviation of three biological replicates. *Panel B:* Growth of the parental strain (WWM85) and the *pycC1* mutant (WWM818) on HS-medium supplemented with methanol as substrate. Error bars represent the standard deviation of three biological replicates.

significant differences and those shared in common between both mutant cell lines, has been submitted to NCBI Gene Expression Omnibus (GEO accession number GSE70370).

Although numerous differences in mRNA abundance were observed, we reasoned that differential expression would need to be seen in both Pyr-1 and Pyr-2 in order to be relevant to the Pyr^+ phenotype. Relatively few genes met this criterion, Table 3. Most notable among these are the genes of the *por* operon, which were expressed at significantly higher levels in the Pyr^+ mutants. This phenotype was especially pronounced when the cells were grown in the presence of pyruvate or acetate. This substrate-specific phenotype can be explained by the observation that DSM804 down-regulates expression of *por* when pyruvate or acetate is present, whereas Pyr-1 and Pyr-2 do not (refer to RNAseq complete data set GSE70370). This regulatory phenotype exaggerates the differences seen between the mutants and their parent in media with pyruvate or acetate and minimizes the difference seen in media without these additions. Two other genes displayed a similar behavior: Mbar_A2632, which

encodes the putative Ppc enzyme mentioned above, and Mbar_A2011, which encodes an S-layer protein.

Genes involved in nitrogen metabolism also showed substantially different expression in the Pyr^+ mutants, although no clear pattern could be observed (Table 2.5). Thus, some nitrogen assimilation genes were up-regulated, with others down-regulated. Moreover, the expression pattern was reversed for many of these genes on methanol media containing either pyruvate or acetate to media with either methanol or acetate alone. Currently, there is no satisfactory explanation for this observation. Surprisingly, the expression of genes predicted to be regulated by PycC (Rodionov et al., 2002), including the *pyc* and *bioY* operons, is not significantly changed in the *pycC1* mutants under most of the conditions examined.

2.4.4 Mbar_A2165 is involved in regulation of Por

To examine the role of the Mbar_A2165(G59R) allele in pyruvate metabolism, we attempted to recreate the mutation in the WWM85 strain background, yet despite numerous attempts, were unable to obtain this mutant. However, we were able to construct a double mutant (WWM940) containing both the *pycC1* mutation and the G59R change in Mbar_A2165. Taken together, these data suggest that the Mbar_A2165(G59R) mutation is lethal in the presence of a wild-type *pycC* allele. To assess the regulatory phenotype of this mutation, a transcriptional profiling of the *pycC1* single mutant, the *pycC1*/Mbar_A2165(G59R) double mutant and their isogenic parent, was conducted (refer to RNAseq complete data set, GSE70370). Comparison of the *pycC1* mutant with the parent revealed only fourteen differentially expressed genes, none of which have obvious roles in pyruvate metabolism. In contrast, comparison of the double mutant to the parent revealed 34 differentially expressed genes, including *porA* (+2.3 fold), *porB* (+2.2-fold) and *ppc* (Mbar_A2632, +2.5-fold). Comparison of the single mutant to the double showed 78 regulated genes including the *por* operon (*porA*, +2.9 fold; *porB*, +3.1-fold, *porC*, +3.1-fold; *porD*, +2.75-fold), but not *ppc*. Consistent with the transcriptional data, Por activity was more than two-fold higher when the

Mbar_A2165 (G59R) mutation was present (specific activities: WWM85, 0.59 ± 0.06; WWM940, 0.50 ± 0.04; WWM941, 1.28 ± 0.22).

2.4.5 Pyruvate ferredoxin oxidoreductase is an essential gene in *M. barkeri* Fusaro

To probe the role of Por activity in pyruvate metabolism, we constructed a strain that expresses *porCDAB* from a tetracycline-regulated promoter (WWM941). Consistent with Por being required for pyruvate biosynthesis (Weimer and Zeikus, 1979), the strain grew well on methanol medium in the presence of the inducer tetracycline, but failed to grow in its absence, (Fig. 2.4, panels A & B). Similar results were obtained in liquid medium when the strain was grown on acetate, methanol with H_2/CO_2 and methanol/acetate mixture.

However, unlike a previously characterized *M. barkeri ech* mutant, which is a pyruvate auxotroph (Meuer et al., 2002), this strain failed to grow in the absence of inducer in media supplemented with various combinations of pyruvate, acetate, casamino acids and yeast extract (Fig 2.4, panels C through H). The failure of supplementation to restore growth of the *por*-depleted strain indicates that Por, or one of its subunits, plays an additional essential role(s) beyond pyruvate biosynthesis.

2.4.6 Suppressors of the Por phenotype

To further asses pyruvate metabolism in *M. barkeri*, we constructed a strain that expresses porCDAB from a tetracycline controlled promoter in a background were the *pyc* operon has the *pycC1* mutation (WWM945). Unlike strain WWM941 this strain was able to grow in the absence of the inducer tetracycline (Fig. 2.5). Since no pyruvate or amino acids were supplemented in the medium these results suggests the presence of an alternative pathway to replenish these constituents in the cell. To identify genes that might be involved in these functions we isolated suppressors of the Por^- phenotype by plating WWM941 in the absence of the inducer tetracycline.

Table 2.5 Select mRNA abundance changes in Pyr-1 and Pyr-2 relative to DSM804[a].

		MeOH/Pyr		MeOH/Pyr/Ac		MeOH/Ac		MeOH		Acetate	
		Pyr-1	Pyr-2	Pyr-1	Pyr-2	Pyr-1	Pyr-2	Pyr-1	Pyr-2	Pyr-1	Pyr-2
Locus Tag	Annotation	Fold Change[b]	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change
Pyruvate induced genes											
Mbar_A0999	*por* operon	**6.8**	**5.1**	**4.9**	2.6	**7.8**	**4.2**	2.4	1.2	1.2	-1.3
Mbar_A1000	*por* operon	**7.7**	**4.3**	**4.3**	2.5	**5.8**	3.0	**4.2**	1.2	1.1	-1.6
Mbar_A1001	*por* operon	**11.4**	**7.1**	**7.1**	3.8	**9.1**	4.1	4.1	1.1	1.4	-1.4
Mbar_A1002	*por* operon	**15.3**	**7.9**	**8.1**	**4.5**	**9.0**	3.9	3.6	1.1	1.2	-1.5
Mbar_A2011	S-layer protein	**34.2**	**19.7**	**8.6**	4.8	**82.7**	**39.5**	-1.1	-3.5	2.0	1.3
Mbar_A2632	pep carboxylase	**4.4**	**3.7**	**2.9**	1.6	**4.37**	**2.9**	**4.0**	1.9	-1.2	-1.4
Putative PycC-regulated genes											
Mbar_A0583	*bioY* operon	1.1	1.1	-1.3	-1.3	-1.4	-1.1	1.5	1.8	1.2	1.2
Mbar_A0584	*bioY* operon	-1.0	-1.1	1.1	1.1	-1.2	-1.4	1.7	1.7	1.1	-1.1
Mbar_A0585	*bioY* operon	1.6	1.3	1.2	1.1	1.5	1.3	1.8	1.5	-1.1	-1.1
Mbar_A0586	*bioY* operon	**2.3**	1.9	1.7	1.1	**2.5**	1.8	**3.6**	1.6	1.8	1.4
Mbar_A1586	*pyc* operon	1.5	1.2	1.0	1.1	1.5	1.0	**2.8**	1.3	1.7	1.2
Mbar_A1587	*pyc* operon	**3.8**	1.9	1.8	1.7	**3.3**	2.0	2.0	-1.1	1.2	-1.1
Mbar_A1588	*pyc* operon	1.6	1.3	1.2	1.2	2.4	1.7	1.2	-1.3	1.3	1.1
Mbar_A1589	*pyc* operon	1.9	1.2	1.2	1.1	1.3	1.0	1.3	-1.5	1.2	-1.0
Nitrogen fixation genes											
Mbar_A0165	Mo-Nitrogenase operon	2.4	1.2	2.3	1.7	**2.8**	2.0	-2.1	-1.8	-1.2	-1.2
Mbar_A0166	Mo-Nitrogenase operon	**3.9**	1.7	**4.1**	2.8	**3.5**	2.4	**-3.0**	-2.4	-1.6	-1.6
Mbar_A0167	Mo-Nitrogenase operon	3.5	1.4	**5.4**	3.6	**5.8**	3.9	**-4.8**	-3.6	-3.1	-3.4
Mbar_A0168	Mo-Nitrogenase operon	4.2	1.7	**6.0**	3.9	**6.2**	4.2	-4.2	-3.3	-3.4	-3.2
Mbar_A0169	Mo-Nitrogenase operon	**5.3**	1.9	**6.4**	4.1	**5.3**	**3.5**	**-3.7**	-2.9	-3.5	-2.9
Mbar_A0170	Mo-Nitrogenase operon	4.1	1.8	**5.9**	3.9	**8.8**	**6.0**	**-7.1**	**-5.7**	-1.9	-1.6
Mbar_A0171	Mo-Nitrogenase operon	**6.7**	2.7	**8.9**	**5.6**	-1.2	5.5	-4.9	-4.5	-4.3	-4.5

[a]The full datasets for each condition, including parsed tables showing only significant differences, are presented in GSE70370 .

[b]Fold change of expression values of mutant strains compared to WT *M. barkeri* grown in the same substrate. Statistically significant values (P<0.05) are shown in bold.

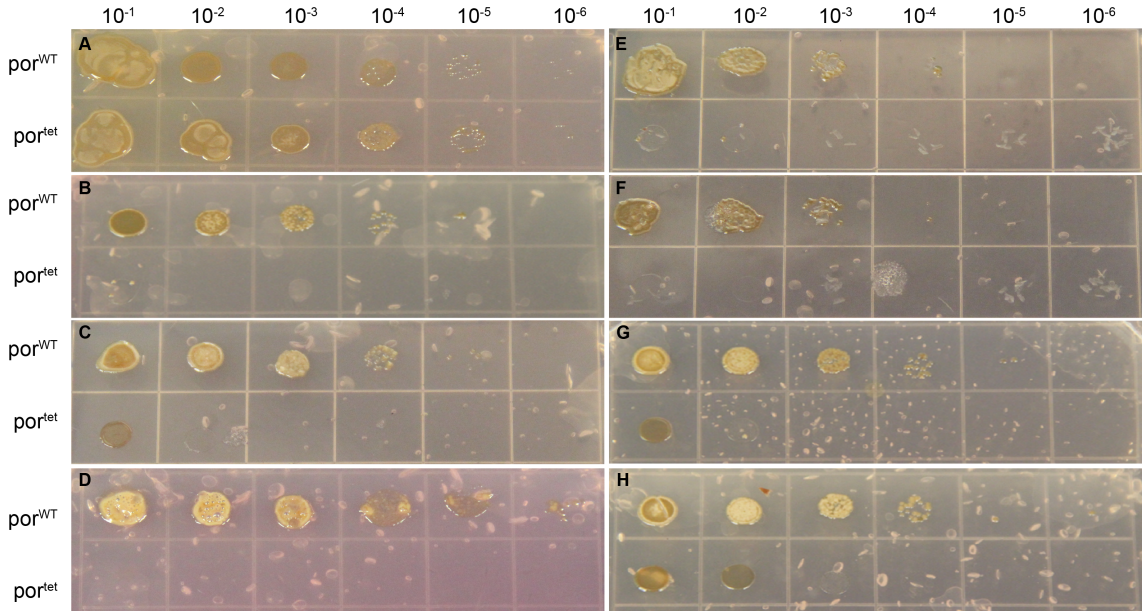

Figure 2.4 The *por* operon is essential for viability in *M. barkeri*. The parental strain, which expresses *por* from its normal promoter (*por*$^{+}$, WWM85) and a derivative that expresses *por* from a tetracycline-dependent promoter (*por*^{tet}, WWM941) were pre-grown in HS-methanol broth with tetracycline, then washed, serially diluted and plated on HS-methanol agar in the presence (A) or absence (B) of tetracycline. Cultures were also plated on HS-methanol agar, without tetracycline, supplemented with 100 mM pyruvate (C), 10 mM pyruvate and 100 mM acetate (D), 0.1% casamino acids (E), 0.1% yeast extract (F), 100mM pyruvate, 0.01% case amino acids, 0.01% yeast extract (G) or 100 mM pyruvate, 0.1% casamino acids, 0.1% yeast extract (H).

After three weeks of incubation nine suppressors strains were obtained. Eight of the isolated strains (clones 2-9) showed a mutation in the promoter region of the tetracycline binding cite (Figure 2.6). The genome of the remaining strain was sequenced. A total of six mutations were identified including a mutation in a potential transcriptional regulator (Table 2.5). Since most of he mutations are present in genes with unknown functions recreations of the mutations will be necessary to determine which role if any they play in the suppressor phenotype.

2.4.7 Attempts to recreate the Pyr^{+} phenotype

Based on the results presented above, it was expected that either the *pycC1* or Mbar_A2165 (G59R) mutation, or both together, would be required for growth on pyruvate. Unfortunately, a direct test of this hypothesis is experimentally difficult due to the morphological properties of *Methanosarcina* cells and the unusual physiology of the Pyr-1 and Pyr-2 cell lines. Most *Methanosarcina* species grow

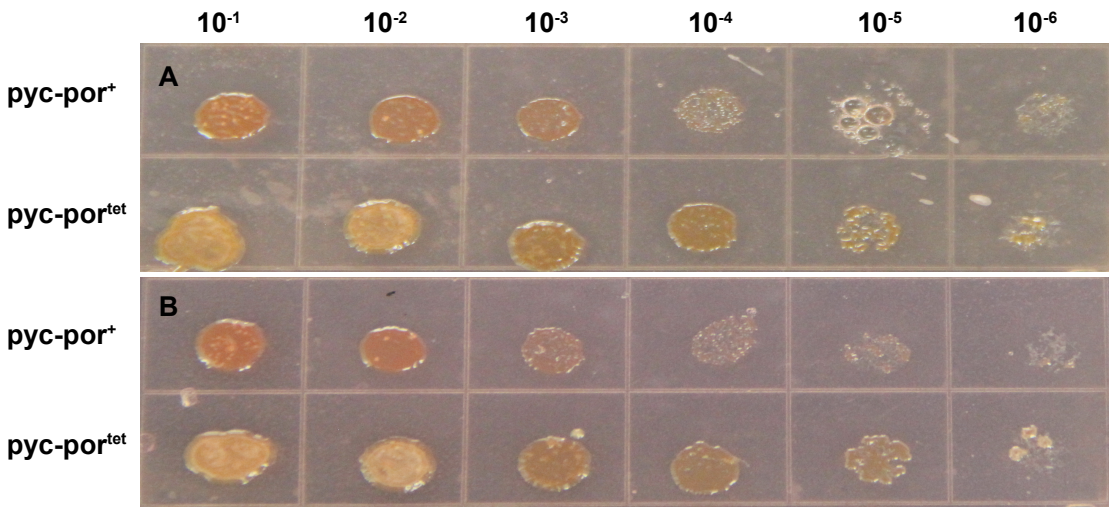

Figure 2.5 The *por* operon is essential for viability in *M. barkeri*. The parental strain, which expresses *por* from its normal promoter (*pycC1-por*$^{+}$, WWM818) and a derivative that expresses *por* from a tetracycline-dependent promoter (*pycC1-por*^{tet}, WWM945) were pre-grown in HS-methanol broth with tetracycline, then washed, serially diluted and plated on HS-methanol agar in the presence (A) or absence (B) of tetracycline.

in large cellular aggregates that are encased in a thick polysaccharide cell wall. This cell wall presents a significant barrier to uptake of foreign DNA and despite numerous attempts, to our knowledge no one has yet succeeded in transformation of aggregated *Methanosarcina* cells. However, prolonged cultivation of aggregated cells in media of high osmotic strength with high concentrations of divalent cations results in conversion to a highly transformable, single celled morphology that lacks the polysaccharide cell wall (Sowers et al., 1993, Metcalf et al., 1997). Thus, the WWM85 cell line used in this study grows as single cells and is genetically tractable, while the DSMZ804, Pyr-1 and Pyr-2 strains grow as aggregates and are not transformable.

As described in the preceding sections, we were able to recreate a *pycC1* single mutant, and a *pycC1*, Mbar_A2165(G59R) double mutant in the WWM85 strain background. Because the single-celled strains lyse in the low-salt medium used for cultivation of DSM804, Pyr-1 and Pyr-2, we characterized the growth phenotype of these strains in high-salt (HS) pyruvate medium. Surprisingly, the aggregated Pyr-1 and Pyr-2 strains failed to grow on pyruvate in HS medium. Pyr-1 and Pyr-2 grow well in HS media with other substrates, suggesting that the

medium lacks some component of the low-salt medium that is required for growth on pyruvate, or that growth on pyruvate is blocked by some component of the high-salt medium. To examine this, numerous variants of HS medium that included the components of low salt medium at a range of concentrations were tested. We also tried variants with decreasing concentrations of the components of the high-salt medium; however, no change to the concentrations of NaCl (400 mM) or $MgCl_2$ (50 mM) was made, as these are required to prevent lysis. Although some of these media allowed limited pyruvate-dependent growth of Pyr-1 and Pyr-2, none would support growth upon serial transfer. We also attempted to develop a high osmotic strength, low salt medium using sucrose and $MgCl_2$, but obtained similar results. Consequently, we were unable to assess whether the *pycC1* and/or Mbar_A2165(G59R) alleles are necessary and sufficient for growth on pyruvate.

Table 2.6 DNA sequence polymorphisms in *M. barkeri* Por$^-$ suppressor.

Locus tag	Annotation	Protein prediction	Polymorphism	Amino acid change
Mbar_A0824/Mbar_A0825	Hypothetical protein/ Universal stress protein	Hypothetical protein / Universal stress protein	T=>C	Intergenic
Mbar_A0824/Mbar_A0825	Bacterioferritin/5S ribosomal RNA	Bacterioferritin/5S ribosomal RNA	Δ1 bp	Intergenic
Mbar_A1672	Hypothetical protein	Carbohydrate domain family	G=>T	Gly-Val
Mbar_A3185	Hypothetical protein	Archaeal ribulose 1,5-bisphosphate synthetase, Rossmann like domain	+G	
Mbar_A3314	Hypothetical protein	Carboxypeptidase regulatory domain/HTH domain	Δ2 bp	Frame shift
Mbar_A3681/Mbar_A3682	Hypothetical protein/ Hypothetical protein	Hypothetical protein/ Hypothetical protein	G=>A	Intergenic

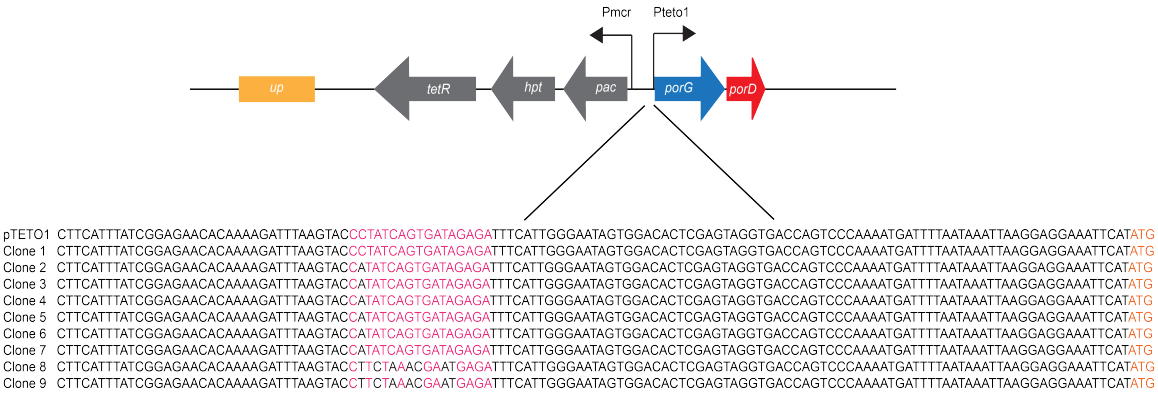

Figure 2.6 Suppressors of Por$^-$ phenotype. Promoter region of nine clones and parental strain were sequence. Annotated ATG site in orange, TetR binding region in red, mutated bases within TetR binding site in black.

2.5 Discussion

Previous studies showed that wild-type strains of *M. barkeri* Fusaro are incapable of growth on pyruvate, but that it is possible to select mutants with the ability to do so (Bock et al., 1994). Our analysis of these mutants suggests that the Pyr$^+$ phenotype depends on the *pycC1* and/or the Mbar_A2165(G59R) allele, because these mutations represent the only shared sequence differences between the Pyr$^+$ strains and their isogenic Pyr$^-$ parent. Unfortunately, we were unable to rigorously establish whether one or both of the alleles are required due to the physiological and morphological constraints of our current experimental system. Thus, the possibility that additional polymorphisms between the DSM804 and WWM85 cell lines contribute to the Pyr$^+$ phenotype can not be excluded. Nevertheless, the data presented in this chapter strongly suggest that both mutations are required for the phenotype because it appears that the Mbar_A2165(G59R) mutation is lethal in the absence of *pycC1*.

The data show that the ability to grow on pyruvate requires two substantial changes in central metabolism (Fig. 2.5). First, instead of using the Pyc-catalyzed carboxylation of pyruvate to form OAA, the strains use a two-step pathway involving conversion of pyruvate to PEP using the enzyme phosphoenolpyruvate synthase (Pps), followed by conversion of PEP to OAA via the enzyme Ppc. Although demonstration of Ppc activity in extracts of the mutants was not possible, it is known that the archaeal enzyme can be unstable

(Patel et al., 2004). Given that OAA biosynthesis is essential and that the *ppc* gene is upregulated in the mutants, it is highly likely that Ppc is functionally expressed in the Pyr-1 and Pyr-2 cell lines. The second key change found in the Pyr^+ strains is a substantially higher level of Por activity. Transcriptomic analyses of recreated mutants show that the Mbar_A2165(G59R) mutation is responsible for increased expression of both Por and Ppc. Our data do not address whether this regulation is direct or indirect, but the phenotype is probably caused by the inability to repress expression of *por* and *ppc* in response to pyruvate and/or acetate, a regulatory phenotype seen in DSM804, but not in Pyr-1 and Pyr-2. The non-conservative, glycine to arginine change in winged, helix-turn-helix motif of the Mbar_A2165-encoded protein could easily prevent DNA-binding to cognate promoters, thereby abrogating repression.

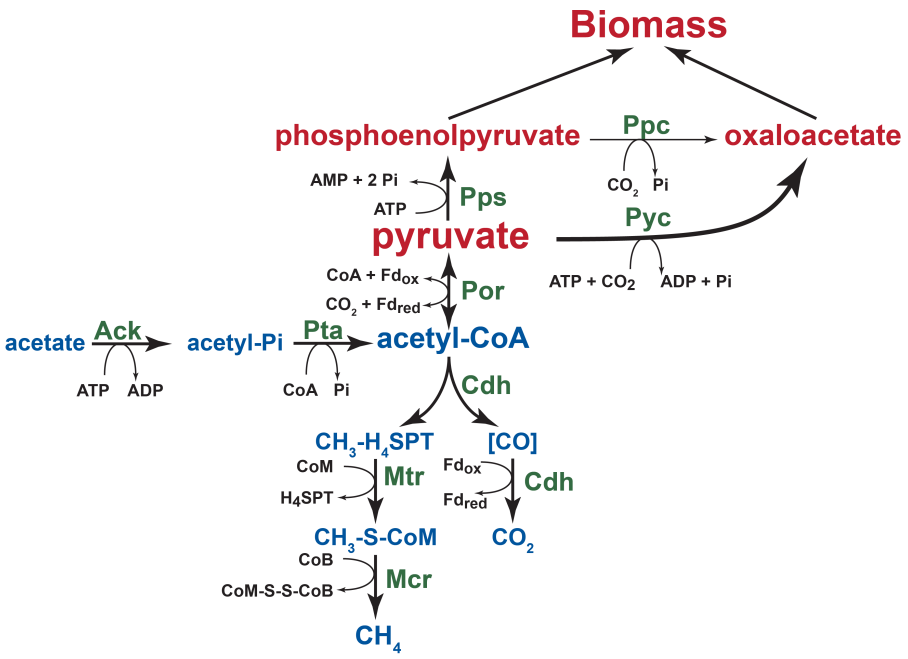

Figure 2.7 Proposed anabolic and catabolic metabolism of pyruvate in *Methanosarcina barkeri* Fusaro. The native aceticlastic pathway, which is used for catabolism of pyruvate, is shown in blue, the anabolic pathway for production of biomass is shown in red. The enzymes proposed to catalzye specific steps are shown in green: Mcr, methy-CoM reductase; Mtr, CoM:H4MPT methytransferase; Cdh, carbon monoxide dehydrogenase; Por, pyruvate:ferredoxin oxidoreductase; Pyc pyruvate carboxylase; Ppc, PEP carboxylase; and Pps, PEP synthase. Overexpression of Por and Ppc in the Pyr+ mutants, coupled with loss of Pyc, is sufficient to allow growth on pyruvate as the sole methanogenic substrate.

These physiological differences suggest a model for growth of the Pyr^+ mutant. The key feature of this model is the inability of the mutant to down-regulate Por in response to pyruvate. High Por activity in the presence of exogenous pyruvate would thus allow the biosynthetic enzyme to take on a catabolic role, oxidizing pyruvate to acetyl-CoA and CO_2. Subsequent conversion of acetyl-CoA to methane could then occur via the standard aceticlastic pathway as previously suggested. The requirement for the *pycC1* mutation has two possible explanations in this model. In the first, *pycC1* is required because the Mbar_A2165(G59R) mutation that allows Por overexpression is lethal in the presence of a functional Pyc. This mutation simultaneously increases expression of Ppc and speculate that the loss of Pyc is required to prevent futile cycling between OAA and pyruvate when both enzymes are present, as seen in other organisms (Flores and Gancedo, 1997). Alternatively, the loss of Pyc may direct pyruvate into the oxidative pathway by blocking its direct conversion to OAA. In this scenario, the kinetic parameters of the two-step pathway would favor channeling of pyruvate into the oxidative pathway, rather than the biosynthetic pathway.

Perhaps the most surprising result was the observation that *porCDAB* expression is essential in *M. barkeri,* even when the media are supplemented with pyruvate. In many anaerobic microorganisms Por plays a catabolic role similar to that observed in the *M. barkeri* mutants. Por-deficient mutants have been made in several of these organisms and, although the mutants lose the ability to utilize specific substrates, they are viable and generally healthy (Kenta Nohara, 2014, McNeely et al., 2011). In contrast, autotrophic methanogens typically use Por as a biosynthetic enzyme to produce pyruvate by reductive carboxylation of acetyl-CoA (Simpson and Whitman, 1993). Based on the fact that the mutants grow on pyruvate, and that pyruvate supplementation can rescue an *M. barkeri ech* mutant (Meuer et al., 2002), this phenotype is not caused by an inability to take up the substrate. Thus, Por clearly has an additional role beyond synthesis of pyruvate. Perhaps, this additional function is also biosynthetic, consistent with

the nature of the encoded gene, and because supplementation did allow slight growth of the Por-depleted cells (see Fig. 2.4, panel C through H). Biochemical characterization of *M. barkeri* Por showed that the enzyme has a rather narrow substrate range using only pyruvate and 2-oxo-butyrate (in the oxidative direction) (Bock et al., 1996). Thus, it is conceivable that the enzyme is required for production of either propionyl-CoA or 2-oxo-butryate, according to *eq1*.

eq1 propionyl-CoA + CO_2 + reduced-ferredoxin ⇔ 2-oxo-butyrate + oxidized ferredoxin

Por enzymes from other sources have additional activities, including hydrogenase and pyruvate decarboxylase (Eram et al., 2014, Ragsdale, 1996), so it is possible that one of these activities is required for viability. It is also possible that one of the proteins in the *porCDAB* operon acts as a subunit for another required enzyme complex. Candidates include the small ferredoxin encoded by *porD* or one of the catalytic subunits, which have been shown to be interchangeable in other 2-oxo acid:ferrredoxin oxidoreductases (Kenta Nohara, 2014, Arnulf Kletzin, 1996). If, as suspected, a biosynthetic block causes the lethal phenotype, then the required intermediate is either not present in yeast extract or casamino acids, or the cell does not take it up. In cells were Pyc was no longer active Por was not essential even in the absence of pyruvate. The connection between Pyc and Por activities remain undetermined however we can speculate that it is possible that absence of both Por and Pyc induces the expression of an alternative biosynthesintetic pathway for the synthesis of pyruvate and other intermediates in the cell.

Finally, our data shed some light on the role of *pycC* in biotin-dependent gene regulation. In *E. coli,* mutations in the biotin-protein ligase domain of the homologous protein affect regulation of biotin biosynthetic genes (Chakravartty and Cronan, 2012). *M. barkeri* does not encode genes for biotin biosynthesis however, *in silico* analysis has revealed the presence of two PycC binding sites in the genome, one presumably regulating the biotin transporter *bioY* and a

second one upstream the *pyc* operon (Rodionov et al., 2002). Our transcriptomic analysis did not reveal consistent and statistically significant changes in the expression of these genes across substrates. Thus, *M. barkeri* PycC does not regulate expression of genes related to biotin metabolism, at least to levels that can be detected by the RNA-seq methodology used here.

Taken together, our results show that pyruvate-utilization in *M. barkeri* is an intricate process that balances the needs of anabolic and catabolic pathways. Moreover, the finding of novel anaplerotic reactions and the essentiality of Por suggest that further analyses of central metabolism in *Methanosarcina* are warranted.

2.6 References

1. ADAM M GUSS, M. R., JUN KAI ZHANG, GARGI KULKARNI AND WILLIAM METCALF 2008. new methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for Methanosarcina species *Archaea,* 2**,** 193-203.

2. ARNULF KLETZIN, M. W. W. A. 1996. Molecular and phylogenetic characterization of pyruvate and 2-ketoisovalerate ferredoxin oxidoreductases from *Pyrococcus furiosus* and pyruvate ferredoxin oxidoreductase from *Thermotoga maritima*. *J Bacteriol,* 178**,** 248-257.

3. AUSUBEL, F. M. 1992. Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology*,* New York, NY, Greene Pub. Associates ; Wiley.

4. BOCK, A. K., KUNOW, J., GLASEMACHER, J. & SCHONHEIT, P. 1996. Catalytic properties, molecular composition and sequence alignments of pyruvate: ferredoxin oxidoreductase from the methanogenic archaeon *Methanosarcina barkeri* (strain Fusaro). *Eur J Biochem,* 237**,** 35-44.

5. BOCK, A. K., PRIEGER-KRAFT, A. & SCHONHEIT, P. 1994. Pyruvate- a novel substrate for growth and methane formation in *Methanosarcina barkeri Arch Microbiol* 161**,** 33-46.

6. BOCK, A. K. & SCHONHEIT, P. 1995. Growth of *Methanosarcina barkeri* (Fusaro) under nonmethanogenic conditions by fermentation of pyruvate to acetate - ATP synthesis via the mechanism of substrate level phosphorylation. *J Bacteriol* 177**,** 2002-2007.

7. BUAN, N., KULKARNI, G. & METCALF, W. 2011. Genetic methods for methanosarcina species. *Methods Enzymol,* 494**,** 23-42.

8. CHAKRAVARTTY, V. & CRONAN, J. E. 2012. Altered regulation of *Escherichia coli* biotin biosynthesis in BirA superrepressor mutant strains. *J Bacteriol,* 194**,** 1113-26.

9. ERAM, M. S., ODUARAN, E. & MA, K. 2014. The bifunctional pyruvate decarboxylase/pyruvate ferredoxin oxidoreductase from *Thermococcus guaymasensis*. *Archaea,* 2014**,** 349379.

10. FLORES, C. L. & GANCEDO, C. 1997. Expression of PEP carboxylase from *Escherichia col*i complements the phenotypic effects of pyruvate carboxylase mutations in Saccharomyces cerevisiae. *FEBS Lett,* 412**,** 531-4.

11. FU, H. & METCALF, W. W. 2015. Genetic basis for metabolism of methylated sulfur compounds in Methanosarcina species. *J Bacteriol,* 197**,** 1515-24.

12. FURDUI, C. & RAGSDALE, S. W. 2000. The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood-Ljungdahl pathway. *J Biol Chem,* 275**,** 28494-9.

13. GUSS, A. M., ROTHER, M., ZHANG, J. K., KULKARNI, G. & METCALF, W. W. 2008. New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for Methanosarcina species. *Archaea,* 2**,** 193-203.

14. KENTA NOHARA, I. O., SATOSHI NAKAMURA, TADAYUKI IMANAKA, TOSHIAKI FUKUI 2014. Genetic examination and mass balance analysis of pyruvate/amino acid oxidation pathways in the hyperthermophilic Archaeon *Thermococcus kodakarensis*. *J Bacteriol,* 196**,** 3831-3839.

15. MCINERNEY, M. J., SIEBER, J. R. & GUNSALUS, R. P. 2009. Syntrophy in anaerobic global carbon cycles. *Curr Opin Biotechnol,* 20**,** 623-32.

16. MCNEELY, K., XU, Y., ANANYEV, G., BENNETTE, N., BRYANT, D. A. & DISMUKES, G. C. 2011. Synechococcus sp. strain PCC 7002 nifJ mutant lacking pyruvate:ferredoxin oxidoreductase. *Appl Environ Microbiol,* 77**,** 2435-44.

17. METCALF, W. W., ZHANG, J. K., APOLINARIO, E., SOWERS, K. R. & WOLFE, R. S. 1997. A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc Natl Acad Sci U S A,* 94**,** 2626-31.

18. MEUER, J., KUETTNER, H. C., ZHANG, J. K., HEDDERICH, R. & METCALF, W. W. 2002. Genetic analysis of the archaeon *Methanosarcina barkeri* Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proc Natl Acad Sci U S A,* 99**,** 5632-7.

19. MUKHOPADHYAY, B., PURWANTINI, E., KREDER, C. L. & WOLFE, R. S. 2001. Oxaloacetate synthesis in the methanarchaeon *Methanosarcina barkeri*: pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional biotin protein ligase gene (bpl/birA) exhibit a unique organization. *J Bacteriol,* 183**,** 3804-10.

20. PAOLO BOCCAZZI, J. K. Z., WILLIAM METCALF 2000. Generation of dominant selectable markers for resistance to pseudomonic acid by cloning and mutagenesis of the ileS gene from teh archaeon *Methanosarcina barkeri* Fusaro *Journal of Bacteriology,* 182**,** 2611-2618.

21. PATEL, H. M., KRASZEWSKI, J. L. & MUKHOPADHYAY, B. 2004. The phosphoenolpyruvate carboxylase from *Methanothermobacter thermautotrophicus* has a novel structure. *J Bacteriol,* 186**,** 5129-37.

22. PRITCHETT, M. A., ZHANG, J. K. & METCALF, W. W. 2004a. Development of a Markerless Genetic Exchange Method for *Methanosarcina acetivorans* C2A and Its Use in Construction of New Genetic Tools for Methanogenic Archaea. *Applied and Environmental Microbiology,* 70**,** 1425-1433.

23. PRITCHETT, M. A., ZHANG, J. K. & METCALF, W. W. 2004b. Development of a markerless genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. *Appl Environ Microbiol,* 70**,** 1425-33.

24. RAGSDALE, S. M. A. S. W. 1996. Unleashing Hydrogenase Activity in Carbon Monoxide Dehydrogenase/Acetyl-CoA Synthase and Pyruvate:Ferredoxin Oxidoreductase *Biochemistry,* 35**,** 15814-15821.

25. RAJAGOPAL, B. S. & LEGALL, J. 1994. Pyruvate as a substrate for growth and methanogenesis for *Methanosarcina barkeri*. *Curr Microbiol,* 28**,** 307-311.

26. ROBINSON, M. D. & SMYTH, G. K. 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics,* 9**,** 321-32.

27. RODIONOV, D. A., MIRONOV, A. A. & GELFAND, M. S. 2002. Conservation of the biotin regulon and the BirA regulatory signal in Eubacteria and Archaea. *Genome Res,* 12**,** 1507-16.

28. SIMPSON, P. G. & WHITMAN, W. B. 1993. Anabolic pathways in methanogens. *In:* FERRY, J. G. (ed.) *Methanogenesis.* New York, NY: Chapman & Hall.

29. SOWERS, K. R., BOONE, J. E. & GUNSALUS, R. P. 1993. Disaggregation of *Methanosarcina* spp. and growth as single cells at elevated osmolarity. *Appl Environ Microbiol,* 59**,** 3832-3839.

30. THAUER, R. K., KASTER, A. K., SEEDORF, H., BUCKEL, W. & HEDDERICH, R. 2008. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat Rev Microbiol,* 6**,** 579-91.

31. WEIMER, P. J. & ZEIKUS, J. G. 1979. Acetate assimilation pathway of Methanosarcina barkeri. *J Bacteriol,* 137**,** 332-9.

32. WIDDEL, F. 1986. Growth of methanogenic bacteria in pure culture with 2-propanol and other alcohols as hydrogen donors. *Appl Environ Microbiol,* 51**,** 1056-62.

33. WILLIAM METCALF, J. K. Z., XUN SHI, RAPLH S. WOLFE 1996. Molecular, Genetic, and Biochemical Characterization of the serC Gene of Methanosarcina barkeri Fusaro. *J Bacteriol,* 178**,** 5797-5802.

34. WILLIAM METCALF, J. K. Z., XUN SHI, RAPLH S. WOLFE 1998. An Anaerobic, Intrachmber Incubator for Growth of Methanosarcina spp. on Methanol-Containing Solid Media. *J Bacteriol,* 64**,** 768-770.

35. ZHANG, J. K., WHITE, A. K., KUETTNER, H. C., BOCCAZZI, P. & METCALF, W. W. 2002. Directed mutagenesis and plasmid-based complementation in the methanogenic archaeon *Methanosarcina acetivorans* C2A demonstrated by genetic analysis of proline biosynthesis. *J Bacteriol,* 184**,** 1449-54.

Chapter 3: An Archaeal-specific Family of DNA-binding Proteins Mediates Transcriptional Regulation of C-1 Metabolism in *Methanosarcina acetivorans* [1]

3.1 Abstract

The *mtaCB1* operon, which encodes one of three methanol methyltransferase isozymes in the methanogenic archaeon *Methanosarcina acetivorans,* is one of the most highly regulated transcriptional units known in Archaea. Previous studies showed that the homologous *msrA* and *msrB* genes are required for transcription of *mtaCB1*, suggesting that the products of these genes act as transcriptional activators; however, direct interaction of these proteins with the target promoter was not shown. To address this, we purified the MsrA and MsrB proteins after overexpression in *Escherichia coli*. The purified proteins from a heterodimer in solution and specifically bind to the *mtaCB1* promoter, as shown by electrophoretic mobility shift assays. Global transcriptional profiling of *msrA* and *msrB* mutants via RNA-seq revealed *mtaCB1* to be the sole positive target of the MsrA/B dimer. Mutants lacking MsrB, but not those lacking MsrA, showed constitutive expression of genes involved in methylamine catabolism during growth on methanol, a condition where these genes are normally repressed. Thus, MsrB is required for both positive and negative regulation of one carbon (C-1) metabolism, although the mechanism of the two processes is clearly distinct. Transcriptional and translational fusions, as well as RNAseq data, showed that the *msr* genes themselves are constitutively expressed and not subject to auto-regulation. Bioinformatics analyses show that MsrA and MsrB represent a novel, archaeal-specific family of DNA binding proteins that is particularly abundant in the Methanosarcinaceae family.

[1] This chapter presents work contributed by three authors: Arpita Bose (bioinformatics analysis of msr proteins, EMSA experiments to determine binding specificity), Petra Kohler (RNAseq experiments and size exclusion chromatography) and Madeline López Muñoz (RNAseq data analysis, qRT-PCR, determination of binding specificity by Footprint assay, EMSA experiment to confirm metal requirement and Western blot experiments)

3.2 Introduction

Members of the *Methanosarcinales*, including *Methanosarcina acetivorans*, are the most metabolically diverse methanogens, being able to use H_2/CO_2, acetate and a variety of methylated compounds as substrates for methanogenesis (Zinder, 1993). In these organisms, reduction of C-1 compounds such as methanol, trimethylamine (TMA), dimethylamine (DMA) and monomethylamine (MMA) proceeds via the methylotrophic pathway of methanogenesis, in which substrates enter the methanogenenic pathway by transferring the methyl group to coenzyme M. Transfer of the methyl group occurs via two consecutive reactions catalyzed by heterodimeric substrate specific methyltransferase 1 complexes (MT1) and non-specific methyltransferase 2 enzymes (MT2) (Keltjens and Vogels) (van der Meijden et al., 1983).

The methanol-specific MT1 enzyme is composed of MtaC (corrinoid protein component) and MtaB (the methyltransferase component) (Sauer et al., 1997) encoded by the operon *mtaCB* (Sauer et al., 1997, Thauer, 1998). Interestingly, *M. acetivorans* C2A, *M. barkeri* Fusaro and *M. mazei* Gö1 each have three operons encoding methanol-specific MT1s, designated *mtaCB1, mtaCB2,* and *mtaCB3*. (Galagan et al., 2002, Maeder et al., 2006, Deppenmeier et al., 2002, Pritchett and Metcalf, 2005). Genetic analysis of these operons in *M. acetivorans* C2A showed that these genes have discrete roles during growth on methanol, with different expression patterns during exponential and stationary phases of the culture cycle (Pritchett and Metcalf, 2005). Moreover, methanol is the strongest inducer of the *mtaCB1* and *mtaCB2* operons causing 100 and 575-fold inductions respectively, one of the highest induction levels observed in Archaea to date (Bose et al., 2006).

Previously, we showed that differential expression of the *mtaCB1* operon is mediated by discrete promoter elements and by the trans acting methanol specific regulators MsrA and MsrB (Bose and Metcalf, 2008). The *msrA* and *msrB* genes are located directly upstream of the *mtaCB1* operon and deletion of

either causes a decrease in the expression of *mtaCB1,* suggesting that they act as activators of *mtaCB1* transcription (Bose and Metcalf, 2008). MsrA and MsrB form part of a larger family of transcriptional regulators that have been shown to control expression of the methyltransferases needed for utilization of the C-1 substrates methanol, methanethiol, dimethyl sulfide and methylmercaptopropionate (Bose et al., 2009, Fu and Metcalf, 2015). However, Msr family proteins are not known to be involved in the regulation of methylamine methyltransferases. Although MsrA and MsrB proteins are predicted to possess HTH DNA-binding domains, experimental demonstration of promoter binding has yet to be reported. Thus, the molecular mechanism by which the Msr proteins mediated regulation remains unproven. Moreover, the means by which C-1 substrates are sensed to modulate regulation are also elusive.

Here, we report a series of transcriptomic and biochemical experiments that address the function and mechanism of action of the Msr proteins. Our results show that MsrA and MsrB comprise an archaeal-specific family of regulators that bind in a sequence specific manner to the *mtaCB1* promoter. In addition, our results show that *msrB* acts to repress several methylamine methyltransferases during growth on methanol and acetate, providing the first evidence for regulation of this group of C-1 substrates by an Msr family member. These results shed significant light into the complex regulatory network that controls methylotrophic metabolism in *Methanosarcina spp.*

3.3 Materials and Methods

3.3.1 Strains, media, and growth conditions

Standard conditions were used for growth of *E. coli* strains (Wanner, 1986). DH5α/λ-*pir* (Miller and Mekalanos, 1988) was used as the host for all *pir*-dependent replicons. DH10B (Stratagene, La Jolla, CA) was used for all other plasmid replicons. *Methanosarcina acetivorans* C2A (DSM 2834) (Sowers et al., 1984) was from laboratory stocks. *Methanosarcina* were grown in single cell morphology (Sowers et al., 1993) at 37°C in high salt medium containing either

125 mM methanol (HS-MeOH), 50 mM trimethylamine (HS-TMA) or 120 mM acetate (HS-Ac). Growth of *M. acetivorans* on media solidified with 1.5 % agar was as described previously (Boccazzi et al., 2000). All plating manipulations were carried out under strictly anaerobic conditions in an anaerobic glove box. Solid media plates were incubated in an intrachamber anaerobic incubator as described (Metcalf et al., 1998). Puromycin (CalBiochem, San Diego, CA) was added from sterile, anaerobic stocks at a final concentration of 2 μg/ml for selection of *Methanosarcina* strains carrying the puromycin transacetylase gene cassette (*pac*) (Metcalf et al., 1997, Gernhardt et al., 1990). The purine analog 8-aza-2,6-diaminopurine (8ADP) (Sigma-Aldrich, St. Louis, MO) was added from sterile, anaerobic stocks at a final concentration of 20 μg/ml for selection against strains expressing the *hpt* gene.

3.3.2 DNA methods

Standard methods were used throughout for isolation and manipulation of plasmid DNA from *E. coli* (Ausubel et al., 1992). DNA sequences were determined from double-stranded templates at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois.

3.3.3 Construction of mutant strains

Construction of Δ*msrA* and Δ*msrB* was previously described (Bose and Metcalf, 2008). For construction of complementation strains or translational and transcriptional fusions all plasmids were integrated on the *M. acetivorans* C2A chromosome using site-specific recombination between the ϕC31 *attB* site on the plasmid with the ϕC31 *attP* using liposome mediated transformation as described previously (Guss et al., 2008). Strains with a chromosomal tag of *msrA* or *msrB* were constructed using the CRISPR-Cas-9 system as described in (Nayak, 2017). Plasmids used for strain constructions are described in Table 3.1. Segments encoding the sgRNA and 1kb homology region were synthetized as double-stranded DNA fragments (gBlocks) by Integrated DNA Technologies (Coralville, IA) and introduced into *AscI*-cut pDN201 using the NEBuilder HiFi DNA assembly

kit following the manufacturer's instructions (New England Biolabs, Ipswich,MA). sgRNA sequences were designed using the CRISPR site finding tool in Genieious version R9 (Kearse et al., 2012). Standard methods were used for construction of plasmids. pJK41 and all derivatives do not replicate in *Methanosarcina*. The Gateway® BP Clonase® II enzyme mix (Invitrogen, Carlsbad, CA) was used to mediate site-specific recombination between pJK027A derivatives and pAH55 derivatives following manufacturer's instructions. All plasmids and strains were verified by PCR, DNA sequencing and/or restriction digest. For primers, plasmids and strain genotypes refer to Tables 3.2, 3.3 and 3.4 respectively.

3.3.4 Purification of His_6-Msr proteins using Ni-NTA affinity chromatography

E. coli Rosetta™ (DE3) (Novagen, Gibbstown, NJ) cells carrying the appropriate overexpression plasmid were grown in LB broth with 25 μg/ml chloramphenicol and 50 μg/ml ampicillin to mid-exponential phase at 37°C. The cells were then incubated on ice for 15 minutes followed by induction with 1 mM IPTG and growth at 18^{o}C for 24 hours. The cells were then pelleted at 4000 X *g*, the supernatant was discarded and cell pellets were frozen at -80 ^{o}C until use. The cell pellets from 250 ml cultures carrying overexpressed His_6-Msr proteins were resuspended in 5 ml ice-cold 50 mM NaH_2PO_4 (pH 8.0) buffer containing 0.5 M NaCl (1X Native binding buffer) with 1X Complete, EDTA-free, Protease Inhibitor Cocktail (Roche, Indianapolis, IN). Cells were broken by passage two times in a French pressure cell (20,000 psi) and cell debris was removed by centrifugation at 13,000 X *g* to pellet cell debris. The resultant supernatant was passed over a Ni-NTA agarose (2 ml bed volume) (Qiagen, Valencia, CA) gravity flow column pre-equilibrated with 10 bed volumes of 1X Native binding buffer.

Contaminating *E. coli* proteins were washed off the column in a stepwise fashion with buffers identical to 1X Native binding buffer, but with varying concentrations of imidazole (20, 50, 100 and 200 mM; 10 bed volumes of each of the buffers

was used for washes). His_6-Msr proteins were collected from the column after washing with 1 ml of buffers containing 250 or 500 mM imidazole. His_6-Msr after affinity chromatography was purified to apparent homogeneity, as ascertained by SDS-PAGE followed by Coomassie Brilliant Blue staining. High-purity His_6-Msr was desalted using Zeba™ Desalt spin columns (Pierce, Rockford, IL) as per the manufacturer's guidelines, replacing the original buffer with 50 mM HEPES (pH 7.2), 200 mM NaCl, containing 20% glycerol, 1 mM dithiothreitol (DTT) and 1 mM $MgCl_2$.

3.3.5 Analytical gel filtration chromatography

The oligomerization state of the His_6-tagged Msr proteins in solution was determined via size exclusion chromatography (SEC) using an Åkta fast protein liquid chromatography system (GE Healthcare Life Sciences) at 4°C with a constant flow-rate of 0.7 ml/min. Protein samples were prepared to a total of 250 µg in Buffer C (50 mM NaH_2PO_4 pH 7.4, 120 mM NaCl, 5% glycerol, filtered and degassed). Freshly purifiedHis_6-MsrA and His_6-MsrB were combined in a 1 to 1, a 2 to 1, and a 1 to 2 ratio. After 15 min of incubation at room temperature the samples were loaded onto a HiLoad 16/600 Superdex 75 prep grade Gel Filtration Column (GE Healthcare Life Sciences), equilibrated with Buffer C. Proteins were eluted with 120 ml Buffer C while monitoring the absorbance at 280 nm. The molecular mass of monomeric and oligomerized Msr protein was estimated using a standard curve constructed from the respective logarithmic molecular weights of standard proteins: Vitamin B12 (1.3 kDa), Myoglobin (17 kDa), Ovalbumin (44 kDa), γ–Globulin (158 kDa) and Thyroglobulin (670 kDa) (Biorad) plotted against their elution volume. The protein contents of the samples before SEC, and of the peak fractions were visualized via SDS-PAGE and Coomassie staining.

Table 3.1 gBlock sequences (underline homology regions for assembly; red: sgRNA sequence; blue: 3X-FLAG sequence; **bold**: ATG and mutations of gsRNA sequence).

Name	Sequence
msrA gsRNA	AACAACATCAGTCACCTAAAAAGAGAAAACGAATTACACGATCACTAATTTTAA ATTTTATATATGTTGACTGAGATTGCAAATTTGAACATTGAAATTTTTTACCCGC TTGATCTGAATAATGACATTGTTCAAAAAAAGTACAAATGATAAAAAAGAAAGC TTCTCAAAAAACAGTAAAGAAGTTCTCCCCAAAATCACCTCAAAAATTCAGAGC TCTATTATCAGAAAAAGCGAGCTTAAAAAATTCAAAGGAAAGATACCCCTCTG CACCCTCAAAATTTTAGACCCTGTGTTGACCTGTAAAAATCAGGAAAAAATTTC CGTCGGTTATGGTATATGTGATGATTTCCCCTAATTATGCTGTAAGCACATGT**A** **GAGACCTCCTCTTACTGCT**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGCCCTCA GTTCTCTTTTTCTTTTTCTTAAACTTCACGCACTGCACTTTTGTCCTCACTTTTT TCATGCCGTCAGATTAACTACTTTTTCTATCCTTGAAATCAGCGGCTTTTCAGC CCTCATGTAGGCGCGCCGGCGATCGCGGCCGCTTAATTAA
msrA Homology	TTTTTTCGAAGTTTAAACCTGCAGGCGCGCCAAAAATTTCAATGTTCAAAATCT CAGTCAACATATATAAAATTTAAAATTAGTGATCGTGTAATTCGTTTTCTCTTTT TAGGTGACTGATGTTGTTTTTCCTTCTATCTTGCAACTCTCTTCAACTGTTAAA GTAAACTCAATTAGATTATCTCGCGTTTTAAATTATTCATCCAAATTCTTTTATA TAAATCCATATAACGGTAATCTCTAATATATGGCGCAAACTCTTCCTCCTATTA GGGTCGTATATGGCCCGAGTGTAACTCAAATCCTTATTCGCCCACATGTAAAC ATTTTTTCCTTGCCGCATAATGCAGATTATCCTATATTTAAAGAAAAAAGAGAT AGGAGTGGTTTGA**ATGCAA**GACTACAAGGACCACGACGGTGACTACAAGGA CCACGACATCGACTACAAAGACGATGACGACAAGTCTGAGTTAATAGATGTA GTATTTCGTTCTCAGAAAAGA**AGAGACCTCCTCTTACTCCTGGG**AGAGGAAC CCCGCACAATGGAAGATATTAAAGTCCTCCTTGATGTTTCTCCCACGGCTATC TTGCCCCAGATCAAGAGGCTTACGGACAGTAACCTTGTTATTCAGAAAAATGG CAGCTATGAATTGACAGATATGGGAGAGCAGGTCTTCAAAAAAGCCAGGGCC CTTGTTGATGTCCTTACCCTGGTTGAAAAAGATAACTACTGGATCGAACACGA CCTGGGGGGAATTCCCCAGTACCTACTTGATAAGATCGGGGAAATTAAGGAC TGCAATCTGGTTAAAGCTGATCCCAGCCAGATTTTCGAGCCGAATACCGAGCT TCTGGAATATTTCGCTTCTTCCCGCTACCTTATGGTATTTTCATCTTTCTACAG GCCGGAGTTCCTGCCCCTCTATTCCAAGCTAGGAAGGCTTGAGTCAGAGGTC TCCCTTATTTTCACGGAATCGGTACTCGAAAAGTTCCTGTATAACTATGAAAAG AAAATCAGAAGGCTTCCAACAACATCAG

Table 3.2 Primers used for cloning in this study.

Primer	Sequence (restriction sites underlined and mutated bases lower case)	Added sites
MA0459ndeIfor	GGCGCGCCCATATGCAATCTGAGTTAATAGA	*Asc*I, *Nde*I
MA0459bamHIrev	GGCGCGCCGGATCCTCAGTGGAGGCTTAAGAGCAG	*Asc*I, *Bam*HI
MA0460ndeIfor	GGCGCGCCCATATGAGTATGGAACTGCAG	*Asc*I, *Nde*I
MA0460xhoIrev	GGCGCGCCCTCGAGTCAGATTTCCTTTATCTCTTC	*Asc*I, *Xho*I
MA0460pET-Duet1for	GGCGCGCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATCATATGAGATGGAACTGCAGTT AATAGATACT	*Asc*I, *Nde*I, *Xba*I
MA0460pET-Duet1rev	GGCGCGCCCTCGAGGCTAGCTCAGATTTCCTTTATCTCTTC	*Asc*I, *Xho*I
MA0459NdeIfor	GGCGCGCCCATATGCAATCTGAGTTAATAGATGTAG	*Asc*I, *Nde*I
MA0459HindIIIrev	GGCGCGCCAAGCTTTCAGTGGAGGCTTAAGAGCAGAGATTC	*Asc*I, *Hind*III
MA0460NdeIfor	GGCGCGCCCATATGAGTATGGAACTGCAGTTAATAG	*Asc*I, *Nde*I
MA0460HindIIIrev	GGCGCGCCAAGCTTTCAGATTTCCTTTATCTCTTC	*Asc*I, *Hind*III
MA0459Ntermhisfor	GGCGCGCCCATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC AGCCACATGCAATCTGAGTTAATAG	*Asc*I, *Nde*I
MA0460Ntermhisfor	GGCGCGCCCATATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC AGCCACATGAGTATGGAACTGCAGT	*Asc*I, *Nde*I
MA0459CtermHisrev	GGCGCGCCAAGCTTTCAATGATGATGATGATGATGGTGGAGGCTTAAGAGCAGAGATTC	*Asc*I, *Hind*III
MA0460CtermHisrev	GGCGCGCCAAGCTTTCAATGATGATGATGATGATGGATTTCCTTTATCTCTTC	*Asc*I, *Hind*III
MA0459uidArev	GGCGCGCCCTGCGGCTTCTTCTGGCGTCATTGCCT	*Asc*I
MA0459uidAfor	AACCACTCCTATCTCTTTTTTCTTTAA	none
MA0460uidArev	GGCGCGCCATATTAAAGTCCTCCTTGATGTTTC	*Asc*I
MA0460uidAfor	GTTCGGCCTCCTTCCGGTAGGAG	none
P0459pTC1for	GGATCCGGCGCGCCTCTGCGGCTTCTTCTGGCGTCATTG	*Asc*I, *Bam*HI
P045930aarev	GGCGCGCCGGATCCTTCCATTGTGCGGGGTTCCTCTCCCAGC	*Asc*I, *Bam*HI
P04601aarev	GGCGCGCCGGATCCCATTCTGTTCGGCCTCCTTCC	*Asc*I, *Bam*HI
P0460pTC1for	GGATCCGGCGCGCCATATTAAAGTCCTCCTTGATGTTTC	*Asc*I, *Bam*HI
CB1forSphI	GGCGCGCCGCATGCGTCCGTAAGCCTCTTGATCTGG	*Asc*I, *Sph*I
uidArevEcoRI	GGCGCGCCGAATTCTCATTGTTTGCCTCCCTGCTGCG	*Asc*I, *EcoR*I
CB1rev(-SphI)	CTCATGCCTGCATaCGGACATGTGCTTAC	none
CB2forPstI	GGCGCGCCCTGCAGTGCAACTGGAAGACAGGAAGC	*Asc*I, *Pst*I
CB3forPstI	GGCGCGCCCTGCAGTAGCAGTAACCTCTGGCAGGATCAACCAGAATTG	*Asc*I, *Pst*I
CB3rev(-PstI)	GCAGAGACTCCTaCAGATGACGAGGG	none
uidArevPstI	GGCGCGCCCTGCAGTCATTGTTTGCCTCCCTGCTGCG	*Asc*I, *Pst*I
CB1L	GGCGCGCCACTAGTCCGTAAGCCTCTTGATCTGG	*Asc*I, *Spe*I
uidArev	GGCGCGCCGCATGCTCATTGTTTGCCTCCCTGC	*Asc*I, *Sph*I
msrAF	GGCGCGCCAACATATGCAATCTGAGTTAATAGATGTAGT	*AscI, NdeI*
msrAR	TTGGCGCGCCAAGCTTGGGGAAAAGCCTCAAAACTG	*AscI, HindIII*
msrBF	GGCGCGCCAACATATGAGTATGGAACTGCAGTTAATAGATACTATTT	*AscI, NdeI*
msrBR	TTGGCGCGCCAAGCTTAGAGCGGATCATTGAGGAGT	*AscI, HindIII*

Table 3.3 Plasmids used in the study.

Plasmid	Description and/or construction	Reference
pET-15b	Ampicillin resistant *E. coli* 6X histidine-tag fusion expression vector	Novagen
pET-11a	Ampicillin resistant *E. coli* protein expression vector	Novagen
pET-Duet1	Ampicillin resistant *E. coli* protein expression vector	Novagen
pAH55	Kanamycin resistant CRIM plasmid with *oriR* and I *attP* site.	(Haldimann and Wanner, 2001)
pAMG82	Vector for construction of translational fusions to the *uidA* gene of *E. coli*	(Guss et al., 2008)
pTC1	Vector for construction of transcriptional fusions to the *uidA* gene of *E. coli. Asc*I/*Bam*HI cloning site, two stop codons, and the *mcrB* ribosome binding site (RBS) upstream of the *uidA* gene; contains *pac* gene (puromycin resistance), *attP* site, and fC31 integrase for insertion of fusions into *M. acetivorans* chromosome.	This study
pJK027A	ΦC31-*attB* vector with P*mcrB*(tetO1) promoter fusion to the *uidA* gene of *E. coli*	(Guss et al., 2008)
pJK026A	ΦC31-*attB* vector with P*mcrB* promoter fusion to the *uidA* gene of *E. coli*	(Guss et al., 2008)
pAMG40	Kanamycin resistant pC2A derivative, which serves as an *E. coli-Methanosarcina* shuttle vector for fosmid retrofitting using the *attP* site.	(Guss et al., 2008)
pDN201	Plasmid used for expression of the CRISPR-Cas9 system in *Methanosarcina acetivorans*	(Reference)
pAB70	*Nde*I/*Bam*HI-digested MA0459 PCR product (using primers MA0459ndeIfor and MA0459bamHIrev) was cloned into *Nde*I/*Bam*HI-digested pET-15b.	This study
pAB71	*Asc*I/*Hind*III-digested pAB98 subcloned into *Asc*I/*Hind*III-digested pET-Duet1.	This study
pAB74	*Nde*I/*Bam*HI-digested MA0459 PCR product (using primers MA0459ndeIfor and MA0459bamHIrev) was cloned into *Nde*I/*Bam*HI-digested pET-11b.	This study
pAB75	*Nde*I/*Bam*HI-digested MA0460 PCR product (using primers MA0460ndeIfor and MA0460xhoIrev) was cloned into *Nde*I/*Bam*HI-digested pET-11b.	This study
pAB78	*Xba*I/*Xho*I-digested MA0460 PCR product (using primers MA0460pET-Duet1for and MA0460pET-Duet1rev) was cloned into *Xba*I/*Xho*I-digested pET-Duet1.	This study
pAB92	*Nde*I/*Hind*III-digested MA0459 PCR product (using primers MA0459NdeIfor and MA0459HindIIIrev) was cloned into *Nde*I/*Hind*III-digested pJK027A.	This study
pAB93	*Nde*I/*Hind*III-digested MA0460 PCR product (using primers MA0460NdeIfor and MA0460HindIIIrev) was cloned into *Nde*I/*Hind*III-digested pJK027A.	This study
pAB64	*Asc*I-digested 1000 bp upstream region of MA0459 PCR product (using primers MA0459uidArev and MA0459uidAfor) was treated with T4 kinase and cloned into *Ecl*136II/*Asc*I-digested pAMG82	This study
pAB65	*Asc*I-digested 1000 bp upstream region of MA0460 PCR product (using primers MA0460uidArev and MA0460uidAfor) was treated with T4 kinase and cloned into *Ecl*136II/*Asc*I-digested pAMG82	This study
pAB83	*Asc*I/*Bam*HI-digested 1000 bp upstream region including the 30 codons into the open reading frame of MA0459 PCR product (using primers P045930aarev and P0459pTC1for) was cloned into *Asc*I/*Bam*HI-digested pTC1	This study
pAB84	*Asc*I/*Bam*HI-digested 1000 bp upstream region including the ATG start codon of MA0460 PCR product (using primers P04601aarev and P0460pTC1for) was cloned into *Asc*I/*Bam*HI-digested pTC1	This study

Table 3.3 Continued

pAB150	Gateway® BP Clonase™ was used as described in materials and methods to mediate site-specific recombination between the l *attP* site on pAMG40 with the l *attB* site on pAB92.	This study
pAB151	Gateway® BP Clonase™ was used as described in materials and methods to mediate site-specific recombination between the l *attP* site on pAMG40 with the l *attB* site on pAB93.	This study
pPK062	*NdeI /HindIII* digested *msrB* PCR product cloned into *NdeI/HindIII* digested pJK026A	This study
pML72	Gblocks for sgRNA and 1kb *msrB* homology region introduced into pDN201 linearized with *AscI* by NEBuilder HiFi DNA assembly kit	This study
pML79	Gblocks for sgRNA and 1kb *msrA* homology region introduced into pDN201 linearized with *AscI* by NEBuilder HiFi DNA assembly kit	This study
pML73	Gateway® BP Clonase™ was used as described in materials and methods to mediate site-specific recombination between the l *attP* site on pAMG40 with the l *attB* site on pML72.	This study
pML80	Gateway® BP Clonase™ was used as described in materials and methods to mediate site-specific recombination between the l *attP* site on pAMG40 with the l *attB* site on pML79.	This study

Table 3.4 *Methanosarcina acetivorans* strains used in this study.

Strain	Genotype or description	Source or Reference
C2A	Wild type (DSM 2834)	(Guss et al., 2008)
WWM82	Δ*hpt*:: ΦC31-*int-attP*	(Bose and Metcalf, 2007)
WWM60	Δ*hpt*::PmcrB-*tetR*	(Bose and Metcalf, 2007)
WWM330	Δ*hpt*::ΦC31-*int-attP* Δ*msrA*	(Bose and Metcalf, 2007)
WWM331	Δ*hpt*:: ΦC31-*int-attP* Δ*msrB*	(Guss et al., 2008)
WWM401	Δ*hpt*:: ΦC31-*int-attR* P*mtaC1*::*uidA* P*mcrB*(tetO1)-*msrA-attL* Δ*msrA*	This study
WWM404	Δ*hpt*:: ΦC31-*int-attR* P*mtaC1*::*uidA* P*mcrB*(tetO1)-*msrB-attL* Δ*msrB*	This study
WWM416	Δ*hpt*:: ΦC31-*int-attR* P*msrA*::*uidA -attL*	This study
WWM417	Δ*hpt*:: ΦC31-*int-attR* P*msrB*::*uidA -attL*	This study
WWM422	Δ*hpt*:: ΦC31-*int-attR* P*msrA*(30aa)::*uidA -attL*	This study
WWM423	Δ*hpt*:: ΦC31-*int-attR* P*msrB*(1aa)::*uidA -attL*	This study
WMM934	Δ*hpt*:: ΦC31-*int-attR* P*mcrB::msrB-attL*	This study
WWM1000	Δ*hpt*::PmcrB-*tetR::3xFLAGmsrB*	This study
WWM1116	Δ*hpt*::PmcrB-*tetR::3xFLAGmsrA*	This study

3.3.6 Preparation of DNA substrates for gel-shift assays

The DNA substrates for gel-shift assays were amplified by PCR with primers shown in Table 3.5 using *M. acetivorans* C2A chromosomal DNA as template. For some substrates ssDNA oligos were annealed as follows; oligos were dissolved in STE Buffer (10 mM Tris pH 8.0, 50 mM NaCl, 1 mM EDTA). Each oligo was dissolved at high concentration (1 - 10 OD 260_{nm} units / 100 uL). The two strands were mixed together in equal molar amounts. The mixture was boiled and cooled gradually to room temperature. The PCR products or annealed oligos were 5′ radiolabeled with γ-[^{32}P]-ATP and T4 polynucleotide kinase, and then separated by electrophoresis on a native 5 - 20% polyacrylamide-1X TBE gels. The gel was exposed to film for two minutes to determine the location of the labeled DNA fragment. The area of the gel containing the DNA fragment was cut out and the DNA was eluted in low salt buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl, 1 mM EDTA) overnight at room temperature. The DNA was passed over DE-52 columns and eluted using high salt buffer (10 mM Tris-HCl pH 8, 1 M NaCl, 1 mM EDTA). Following ethanol precipitation, the DNA pellet was resuspended in Tris-EDTA buffer pH 8.0. The DNA concentration was determined by measuring absorbance at 260 nm.

3.3.7 Electrophoretic mobility shift assays (EMSA)

Two methods were used for EMSA experiments. EMSA reactions performed using radiolabeled substrates proceeded as follows. Binding reactions were prepared with the indicated concentrations of DNA in 50 mM HEPES (pH 7.2), 50 mM NaCl, containing 20% glycerol, 1 mM dithiothreitol (DTT), 1 mM $MgCl_2$, 1 mM EDTA, 0.25 mg/ml BSA and 0.075 mg/ml of herring sperm DNA. Appropriate amounts of the purified protein or crude extract to be tested were then added and the reactions were allowed to equilibrate at room temperature for 15-20 min. The equilibrated reactions were loaded on a 5% polyacrylamide 0.5X TBE gel that was pre-run at 200V for 30 - 60 min, and run at 200 V in 0.5X TBE, until the DNA had migrated through most of the gel. The gel was then transferred to Whatman filter paper and exposed to phosphorimager screens. The images were obtained

using a Fujifilm FLA-3000 phosphorimager. A non-radioactive EMSA assay was used for some experiments. Reactions were prepared with the following concentrations 50 mM HEPES (pH 7.2), 50 mM NaCl, 1 mM DTT, 10% glycerol, 1 pmol 6FAM (Fluorescein)-labeled DNA and 100 µM $MgCl_2$ or 100 µM $MgCl_2$ and 100 µM $CaCl_2$. Proteins used were incubated with 1 mM EDTA prior before concentrating and desalting as indicated in section 3.3.6. MsrAB was added to the reactions in the concentration indicated in Figure 3.2. Reactions were incubated at RT for 30 minutes and run in 7.5% Mini-PROTEAN® TGX Precast Protein Gels (Bio-Rad Laboratories, Hercules, CA). Gels were run at 150 V using 1X TBE as running buffer. DNA bands were visualized by staining in 50 ml TBE with 1X SBR® Safe DNA Gel Stain for 30 min (ThermoFisher, Grand Island, NY).

3.3.8 DNAse I footprinting

Binding reactions were set up at room temperature with 50 mM HEPES (pH 7.2), 50 mM NaCl, 1 mM DTT, 100 µM $MgCl_2$, 100 µM $CaCl_2$, 10% glycerol, 1 pmol fluorescein-labeled DNA, incorporated in the PCR primers used to prepare substrates, and 1000 pmol MsrA-MsrB. Reactions were incubated for 30 min. The complexes were digested with 50 mU DNAse I (New England Biolabs, Ipswich, MA) at 37°C for five minutes. Reactions were stopped by addition of one volume of 0.5M EDTA and cleaned using the DNA Clean and Concentrator kit from Zymo Research following manufacturer's instructions (Zymo Research Irvine, CA) eluting in 20 µl of water. Fragments were submitted for fragment analysis at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois. Genemapper Software version 3.7 (Applied Biosystems Foser City, CA) was used to analyze and map the fragment data.

3.3.9 Transcriptomic analyses (RNA-seq)

RNA was extracted from cultures grown to mid exponential phase using the Direct-zol RNA MiniPrep kit (Zymo Research Irvine, CA) following the manufacturer's instructions. Ribosomal RNA was depleted using subtractive hybridization employing biotinylated rRNA probes as previously described (Fu

and Metcalf, 2015). The quality of the RNA was verified using a Agilent Bioanalizer. RNA sequencing was performed at the Keck Center for Comparative and Functional Genomics, University of Illinois using strand specific libraries prepared with the Illumina ScriptSeq kit version 2 (Epicentre Biotechnologies, Madison, WI). Bar-coded libraries were pooled, quantified by qPCR and sequenced for 101 cycles on a HiSeq2000 using the TruSeq SBS sequencing kit version 3 (Illumina, San Diego, CA). Reads were analyzed using the CLC genomics platform (version 7.5) and reads were trimmed based on quality (quality: 0.001, ambiguous: 2, discard min:30, max: 100). To remove bias caused by remaining stable RNAs, the trimmed sequences were mapped to *M. barkeri* rRNA and tRNA sequences using the following parameters similarity: 0.9, length: 0.85, mismatch: 2, insertion: 3, deletion: 3. The remaining reads were used for the RNAseq analysis (similarity: 0.9, length: 0.85, mis:2, ind:3, del:3, global alignment, maximum number or reads per read: 10). Expression values were calculated as RPKM (Reads Per Kilobase per Million reads). Reads were normalized and statistical analysis was performed using the empirical analysis of digital gene expression (Robinson and Smyth, 2008). To determine differential gene expression, reads from each mutant strain were compared against reads of WT and each mutant strain. Genes showing fold expression changes greater than 4, or less than -4, with P-values ≤ 0.01 were considered to be differentially expressed. The raw and processed RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) under the accession number (submission pending).

3.3.10 qRT-PCR measurements

RNA was extracted from biological cultures grown in triplicates to mid exponential phase using Direct-zol RNA MiniPrep kit (Zymo Research Irvine, CA) following manufacturer's instructions. DNA was removed from RNA samples by digestion with DNaseI (New England Biolabs, Ipswich, MA). The quality of the RNA was verified using Agilent Bioanalizer at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois. qRT-PCR reactions

Table 3.5 Primers used to generate substrates for gel-shift assays.

Primer	Sequence
mtaCB1revmain	TCTAAACCTCCATTTAGATCAAACAACAAAAAGGTC
mtaCB1for100	GGCGCGCCACTTTACAAAGGCAAAACCGG
mtaCB1for200	GGCGCGCCGTTTCGGTAACAGGAGAAAGAAATTCT
mtaCB1for300	GGCGCGCCTAGACCCTGTGTTGACCTGTAAAAATC
mtaCB1for400	GGCGCGCCCCTCAAAAATTCAGAGCTCTATTATC
mtaCB1for500	GGCGCGCCTTTACCCGCTTGATCTGAATAATG
mtaCB1for600	GGCGCGCCAACAACATCAGTCACCTAAAAAGAG
mtaCB1rev300	GATTTTTACAGGTCAACACAGGGTCTA
mtaCB1rev400	GATAATAGAGCTCTGAATTTTTGAGG
mtaCB1for 700	GGCGCGCCCCGTTATATGGATTTATATAAAAGAAT
mtaCB1for650	CTAATTGAGTTTACTTTAACAG
mtaCB1for550	TTAAATTTTATATATGTTGACTG
mtaCB1rev500	AAATTTCAATGTTCAAATTTGC
mtaCB1rev450	TCATTTGTACTTTTTTTGAACAATGTC
mtaCB1rev400	GATAATAGAGCTCTGAATTTTTGAGG
mtaCB1rev350	ATCTTTCCTTTGAATTTTTTAAGC
mtaCB1rev300	GATTTTTACAGGTCAACACAGGGTCTA
cb1BS(3) topstrand	TAAAGAAGTTCTCCCCAAAATCACCTCAAAAATTCAGAGCTCTATTATCAG AAAAAGCGA
cb1BS(3) botstrand	TCGCTTTTTCTGATAATAGAGCTCTGAATTTTTGAGGTGATTTTGGGGAGA ACTTCTTTA
PmcrBfor	TTCATTTATCGGAGAACACAAAAGATTTAAG
PmcrBrev	TTTGCTCCCTCCGATTATTCCCAATG
mtaCB1_F300	TTGACTGAGATTGCAAATTTGAACA
mtaCB1_R270	ATGTGCTTACAGCATAATTAGGGGA

Table 3.6 qRT-PCR primers used in this study.

Primer	Sequence
qRT_rpoA1F	GGCAGAGACGAGACAGG
qRT_rpoA1R	GGTTCAGGATGGAACCTCT
qRT_mtaBF	CACGCAGTTCTCAGACCAGA
qRT_mtaBR	GATGAACATTGCGGTGTTTG
qRT_mttBF	GACTCCGTGGAACAGGACAT
qRT_mttBR	GTTAGTGCTTTGCGGTGTTT

were performed with 200 ng of total RNA, using Mastercycler ep realplex (Eppendorf, Hamburg, Germany) and the SuperScript® III One-Step RT-PCR System with Platinum® Taq (Thermo Fisher Scientific, Waltham, MA) per manufacturer's instructions. Changes in gene expression were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) using *rpoA1* as the standard. For primers used refer to Table 3.6.

3.3.11 Western blots

Cell lysates from mid exponential phase cultures grown on methanol or TMA were prepared as follows; 5 ml of cell cultures were centrifuged and resuspended in 500 µl of lysis buffer (50 mM Tris HCl, 3 M KCl and 5 mM $MgCl_2$). Protein concentration was determined by the Bradford method (Coomassie protein reagent, Sigma Aldrich, St. Louis, MO) using BSA as a standard. Increasing amounts of protein from cell lysates were electrophoretically separated using SDS-PAGE (4-20%), transferred onto PVDF membranes and blocked with 5% (w/v) milk in TBST buffer (150 mM NaCl, 5 mM Tris base and 0.05% Tween 20) for 30 min at room temperature. Membranes were incubated with 3 µl monoclonal anti-FLAG M2 murine antibody (Sigma-Aldrich, St. Louis, MO) and incubated over night at 4°C. Membranes were washed three times and incubated with 1 µl of Goat anti-Mouse IgG (H+L) Secondary Antibody, HRP conjugated (Thermo Scientific, Waltham, MA) for an hour. Reactions were detected by chemiluminescence using SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Scientific, Waltham, MA).

3.3.12 *In silico* methods

Protein domain predictions were made using InterProScan (Mulder et al., 2007), HTH program (Dodd and Egan, 1990) and BLASTP (Altschul et al., 1990). Precise length of the domains were determined using CDD (Marchler-Bauer et al., 2007). Orthologs were identified using the Integrated Microbial Genomes System (Markowitz et al., 2006). CLUSTALW was used to determine sequence identity and to construct amino acid sequence alignments (Chenna et al., 2003). Start codon prediction for open reading frames was performed using GeneMark (Lukashin and Borodovsky, 1998). A rooted neighbor-joining tree was generated by the DRAWTREE program (http://workbench.sdsc.edu) (Thompson et al., 1994, Higgins et al., 1992, Felsenstein, 1989).

3.4 Results

3.4.1 Bioinformatics analyses suggest that MsrA and MsrB proteins are members of an archaeal specific family of DNA-binding proteins

We previously reported that the MsrA and MsrB proteins from *M. acetivorans* C2A have orthologs in *M. mazei* Gö1 and *M. barkeri* Fusaro (Bose and Metcalf, 2008). Careful pair-wise comparison of these proteins showed that the homologous proteins share 74-96% amino acid sequence identity between *M. acetivorans*, *M. barkeri* Fusaro and *M. mazei* Gö1, while paralogs share only 30-50% identity (Table 3.7). Both MsrA and MsrB have a helix-turn-helix (HTH) DNA-binding motif that shares a weak similarity to the HTH motif of the ArsR family of regulators. However, in contrast to the ArsR proteins that only encode a HTH domain, the MsrA and MsrB proteins also have a C-terminal domain of unknown function (DUF1724) that is found solely in Archaea (Fig 3.1, panel A) (Busenlehner et al., 2003). Significantly, the Msr proteins lack the canonical ArsR family metal binding sites (data not shown), thus, they are unlikely to function as metal-repsonsive transcriptional regulators. Homology detection and structure prediction by Hidden Markov Model using HHpred (Soding et al., 2005)predicted that both MsrA and MsrB are structural homologs of TrmBL2 and TrmB, an archaeal chromatin protein and a global transcriptional regulator from Pyrococcus *furiosus* respectively. Additional analysis using I-TASSER (Yang et al., 2015) suggested the presence of a metal binding pocket in the C-terminus of the proteins, suggesting that although they lack ArsR-like conserved residues metals might still play a role in MsrA/B function.

Similarity searches against the non-redundant protein database reveal that MsrA and MsrB are members of a much larger protein family found solely in Archaea, which we have designated the Msr family (Fig 3.1). Msr proteins occur exclusively in the members of the domain Archaea including members of the Halobacteria, Methanobacteria, Archaeoglobi and Methanomicrobia. Although many archaeal species encode one or two Msr family members, there is a

Table 3.7 Percent amino acid identity of the Msr proteins in *M. acetivorans* C2A, *M. barkeri* Fusaro and *M. mazei* Gö1[a].

	Ma MsrA	Mm MsrA	Mb MsrA	Ma MsrB	Mm MsrB	Mb MsrB	Ma MsrC	Mm MsrC	Mb MsrC	Ma MsrD	Mm MsrD	Mb MsrD	Ma MsrE	Mm MsrE	Mb MsrE
Ma MsrA	-	83	77	31	34	33	31	36	29	41	40	42	35	34	32
Mm MsrA		-	74	29	30	29	27	30	25	36	35	36	31	28	31
Mb MsrA			-	34	34	35	31	34	28	39	38	40	35	33	35
Ma MsrB				-	84	79	40	41	38	36	35	36	48	48	46
Mm MsrB					-	82	39	42	37	40	38	39	50	50	48
Mb MsrB						-	37	42	38	38	36	38	48	48	47

[a].CLUSTALW was used for calculations (Thompson et al., 1994).
[b.] Ma: *M. acetivorans* C2A; Mb: *M. barkeri* Fusaro; Mm: *M. mazei* Gö1.

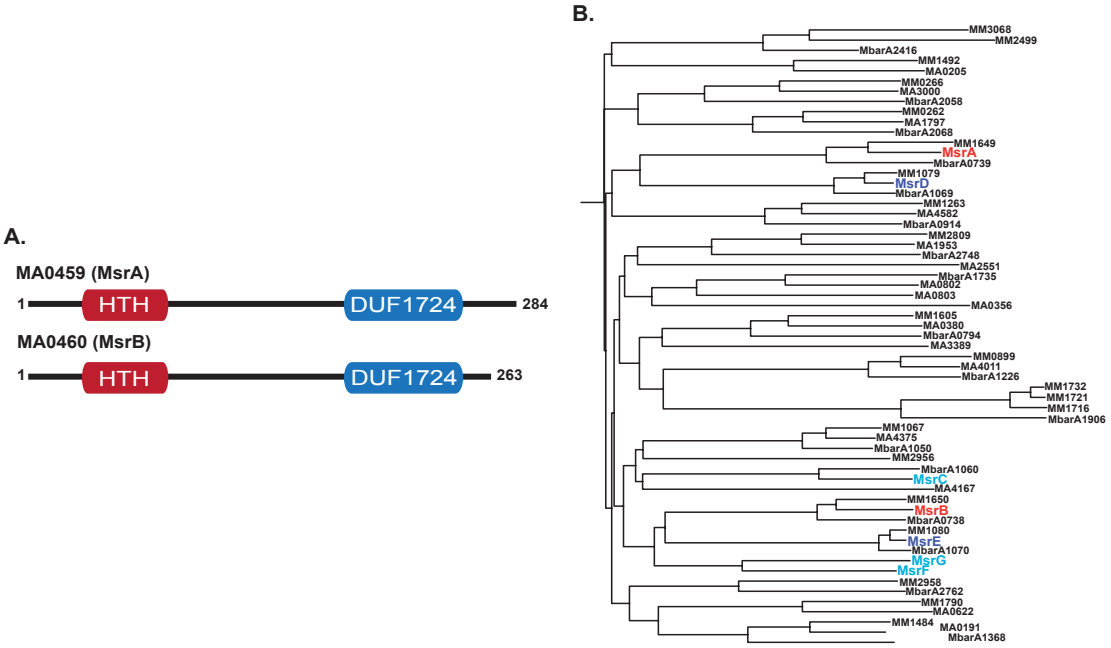

Figure 3.1 Bioinformatic and phylogenetic analysis of MsrA and MsrB Panel A: Domain structure of the MsrA and MsrB proteins. Abbreviations: HTH, helix-turn-helix domain; DUF1724, domain of unknown function present only in archaea. Panel B: Neighborhood-joining three of *Methanosarcina* Msr proteins. MM, *M. mazei*, MA, *M. acetivorans* and Mbar, *M. barkeri*. Red: MsrA and MsrB; Blue: MsrD and MsrE; Aqua MsrF, MsrG and MsrC .

substantial expansion in the Family Methanosarcinaceae, with at least 25 Msr-like proteins in *M. acetivorans*, 16 in *M. barkeri* and 20 in *M. mazei* Gö1. To date 8 members of this family in *M. acetivorans*. have been studied, all of which are involved in transcriptional regulation of C-1 metabolism (Bose and Metcalf, 2008, Bose et al., 2009, Fu and Metcalf, 2015). These include the methanol methyltransferase regulators, MsrA, MsrB, MsrD and MsrE, the dimethylsulfide methyltransferase regulators, MsrC, MsrF and MsrG (Bose et al., 2009, Bose and Metcalf, 2008) and the methylmercaptopropionate regulator MsrH (Fu and Metcalf, 2015). Several subgroups can be recognized within the family, with MsrA and MsrB falling into distinct clades (Fig 3.1, panel B). These subgroups are likely to reflect functional differences between the members. Although sequence analysis suggests that these proteins are DNA-binding proteins, the mechanism by which they regulate transcription has yet to be established.

3.4.2 Promoter binding is sequence specific and requires both MsrA and MsrB

To test the hypothesis that MrsA/B mediate transcriptional activation by direct binding to *mtaCB1*, we purified recombinant versions of the proteins after expression in *E. coli* and investigated their DNA binding activities by EMSA assays. When used alone, neither MsrA nor MsrB was able to bind the putative promoter region; however, when the two proteins were combined a pronounced gel-shift was observed, suggesting that an MsrA/MsrB complex is required for DNA binding (Fig 3.2 A). Binding to the labeled fragment could be eliminated by addition of a large excess of unlabeled DNA containing the *mtaCB1* promoter, whereas addition of a large excess of a non-target sequence did not (Fig 3.2 B). Thus, binding of the MsrA/B complex is sequence-specific.

Subsequent EMSA experiments using truncated promoter fragments narrowed the MsrA/B binding region to a 134 bp fragment, beginning 81 bp upstream of mapped transcription start site (Bose and Metcalf, 2008). Thus, more distal, or more proximal, promoter fragments, were incapable of binding the recombinant

proteins. However, within this region, a series of non-overlapping fragments showed MsrA/B binding, suggesting that multiple binding sites are present in the *mtaCB1* promoter (Fig 3.3 panel A). The observation of larger band-shifts due to increasing levels of recombinant protein is consistent with this interpretation (Fig 3.2 B).

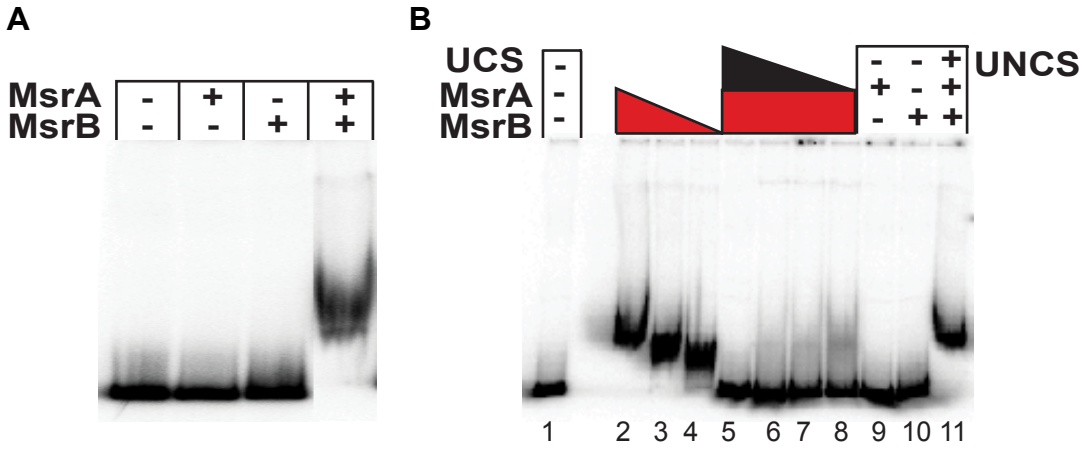

Figure 3.2 Binding specificity of MsrA/B to the mtaCB1 promoter. Panel A: DNA substrate encompassing the *mtaCB1* promoter was used to test the ability of MsrA and MsrB to bind to its promoter (360 bp to 544 bp upstream of the ATG start codon of the *mtaC1* gene). 0.2 fmol of labeled substrate was reacted with 30 pmol of recombinant MsrA and MsrB and incubated in the presence (+) or absence (-) of the proteins. Panel B: Unless otherwise specified reactions included 0.2 fmol of DNA and 30 pmol of MsrA/B. Fragment K, 334-544 (Fig 3.3 panels A and B) was used as substrate. Lane 1: DNA only; Lanes 2-4: 30 pmol, 15 pmol and 7.5 pmol of MsrA/B respectively; Lanes 5-8 reactions have MsrA/B, DNA probe and unlabeled competitive substrate (UNC) in the following concentrations: 10 fmol, 5 fmol, 2.5 fmol and 1.25 fmol of DNA; Lane 9: MsrA and DNA probe; Lane 10: MsrB and DNA probe; Lane 11: MsrA/B, DNA probe and 1 pmol of unlabeled non-competitive substrate.

3.4.3 MsrA/B footprint

To determine the precise MsrA/B binding sequence, we performed a DNAseI footprint experiment using a probe that included the regions from 270 to 578 from the annotated ATG of *mtaCB1*. Protection in the presence of MsrA/B was detected over a region of 62 bp from positions 318 to 380 (Fig 3.3 and 3.4). An area of hypersensitivity was detected in position 314.

3.4.4 The Msr proteins bind the *mtaCB1* promoter as a heterodimer

Because MsrA/B bind as a complex, we wanted to examine the oligomeric structure of the proteins in solution. To do this, we combined the two proteins in

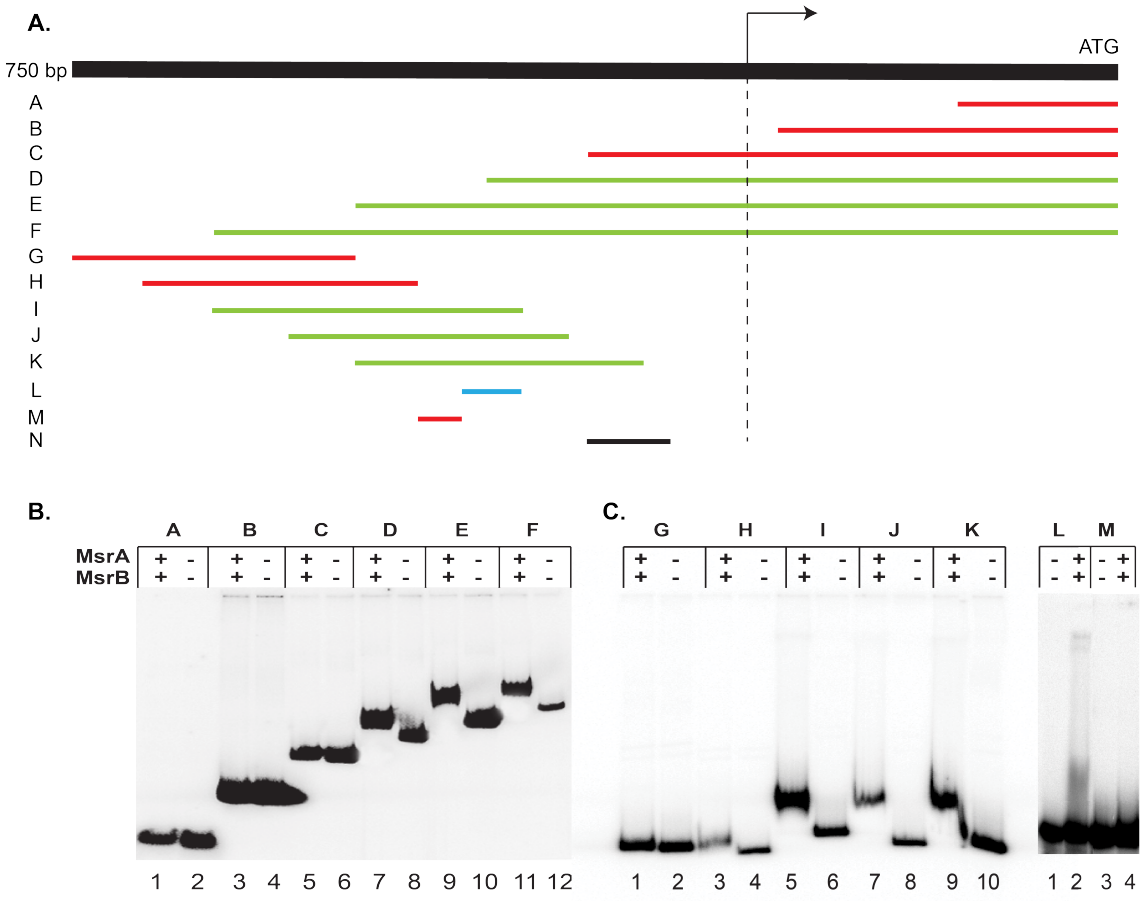

Figure 3.3 Determination of the binding sites of the MsrA and MsrB to their target promoter *mtaC*. Panel A. Schematic depiction of promoter truncations used as probes for EMSA experiments. **Green**: fragments tested positive for binding, **Red**: fragments tested negative for binding, **Blue**: 38 bp fragment tested positive for binding, **Black**: footprint. Fragment position with respect to annotated ATG: A, 109; B, 230; C, 360; D, 444; E, 544; F, 644; G, 545-751; H, 495-694; I, 418-594; J, 386-594; K, 334-544; L, 419-456; M, 457-494. Panel B: EMSA with truncated promoter regions. Reactions were performed in the presence (+) or absence (-) of MsrA/B. Panel C: EMSA with truncated promoter regions. Unless specified otherwise 0.1-0.3 fmol of DNA was incubated in the presence (+) or absence (-) of 30 pmol MsrA/B. Reactions on lanes 11 and 12 had 21 fmol DNA probe and 100 pmol protein. Reactions on lanes 13 and 14 had 26 fmol of DNA and 100 pmol of protein.

various ratios and performed size exclusion chromatography (SEC). When combined in a 1:1 ratio the complex eluted as a single peak consistent with expected size for a heterodimeric complex (Fig 3.5). Two peaks were observed when mixtures containing an excess of either MsrA or MsrB were analyzed, one eluting in the same fraction as the heterodimer, the other eluting in fractions consistent with the expected size of the excess monomer (Fig 3.5). SDS/PAGE analysis of the various fractions is consistent with this interpretation. Thus, MsrA

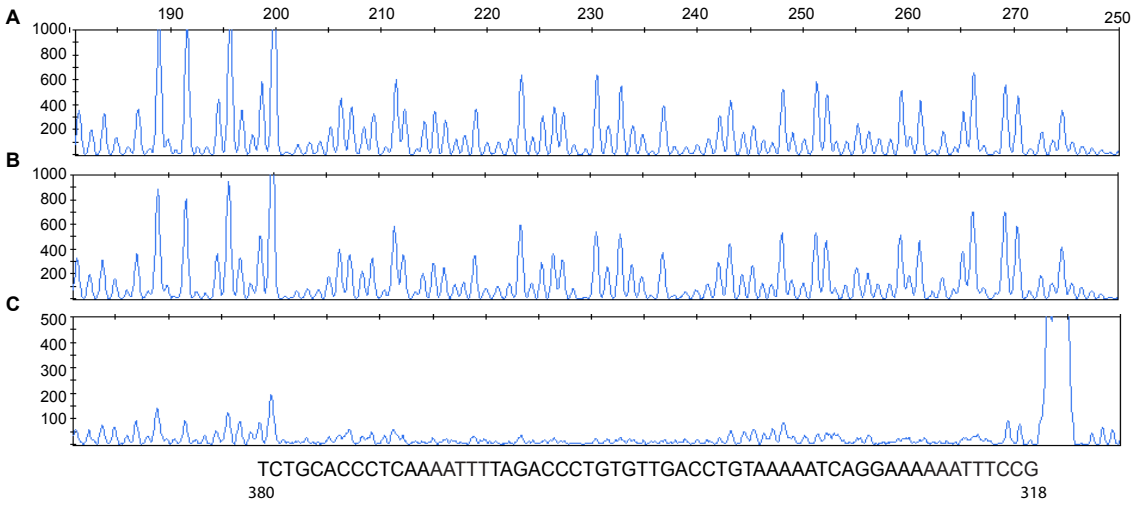

Figure 3.4 DNAse I footprint of *mtaCB1* promoter region. The x-axis represents length of the fragments. The y-axis represents fluorescence intensity. Panel A represents DNA only reaction, panel B represents BSA control and panel C represents a reaction containing 1000 pmol MsrAB. Protected region and its position regarding translation start site is presented on the figure.

and MsrB form a heterodimer in solution, which most likely represents the minimal DNA-binding species; however, given the observation of multiple binding sites within the promoter, it is possible that higher order structures are required for transcriptional activation.

3.4.5 Binding of MsrAB is enhanced by addition of metals

Although MsrA and MsrB do not have the conserved metal binding residues conserved in the ArsR family of regulators, analysis of predicted protein structure and function suggests the presence of a metal binding pocket in the C-terminus of the proteins. To test if addition of metals improves MsrA/B binding we performed an EMSA experiment in the presence and absence of calcium. Addition of 100 μM of $CaCl_2$ in the presence of equimolar amount of $MgCl_2$ caused a stronger shift in binding when compared to reactions performed in the presence of $MgCl_2$ alone (Fig 3.6). These results suggest that MsrAB might require the presence of a metal to bind and regulate transcription of methyltransferases. Addition of Mn also improves binding of MsrA/B to *mtaCB1* promoter. More experiments are needed to determine which metal is bound to the proteins and the role they play in Msr function.

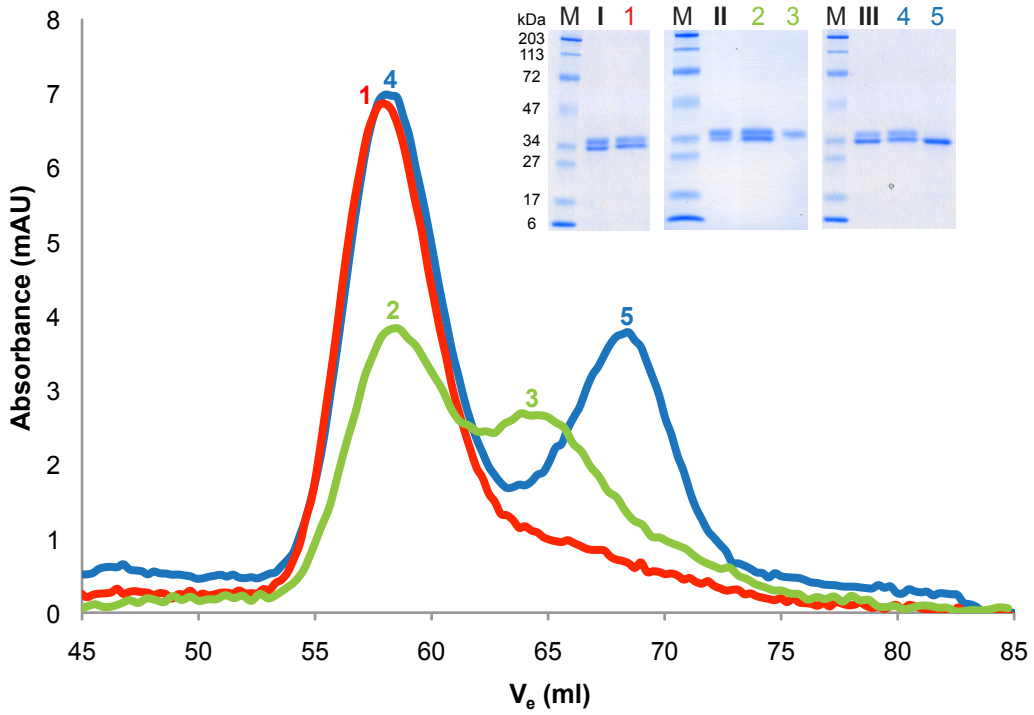

Figure 3.5 Oligomeric state of the MsrAB proteins. Analytical size exclusion chromatography (SEC) of MsrA and MsrB proteins. Red, 1:1 ratio of MsrA and MsrB; 1, dimer MsrAB (elution 58 ml). Green, 2:1 ratio of MsrA and MsrB, 2, MsrAB dimer (elution 58 ml); 3, MsrA monomer (elution at 64 ml). Blue, 1:2 ratio of MsrA and MsrB. 4, MsrAB (elution 58 ml); 5, MsrB (elution 68 ml). Inset shows an SDS/PAGE of MsrA and MsrB fractions collected during SEC run. M, marker; I, II, III MsrAB before SEC.

3.4.6 Regulation of *msrA* and *msrB* genes

The observation that recombinant MsrA/B proteins bind their target promoter suggests that substrate-specific gene expression is mediated at a different level, perhaps by regulating expression of the *msrA* and *msrB* genes. Therefore, we assessed the expression of these genes using RNAseq data. Wild-type *M. acetivorans* showed similar *msrA* or *msrB* mRNA levels when cells were grown on methanol or TMA (Table 3.9). Our data also showed that *msrA* expression was unaffected by deletion of *msrB* and that *msrB* transcripts are unaffected by deletion of *msrA*, suggesting that msr proteins do not regulate their own expression. Since no transcriptional level regulation was evident in our results, we decided to test if the MsrA and MsrB proteins were regulated post-transcriptionally by protein degradation. To test this we constructed strains that

expressed the chromosomal copy of either gene with a 3x-FLAG tag on the N-terminus. qPCR analysis of the strains showed the chromosomal tag did not affect the regulation of the *mtaCB1* operon mediated by the Msr proteins (data not shown). Western blot analysis of the strains grown on methanol or TMA revealed that the proteins seem to be expressed at a similar level in cells grown on both substrates (Fig 3.7).

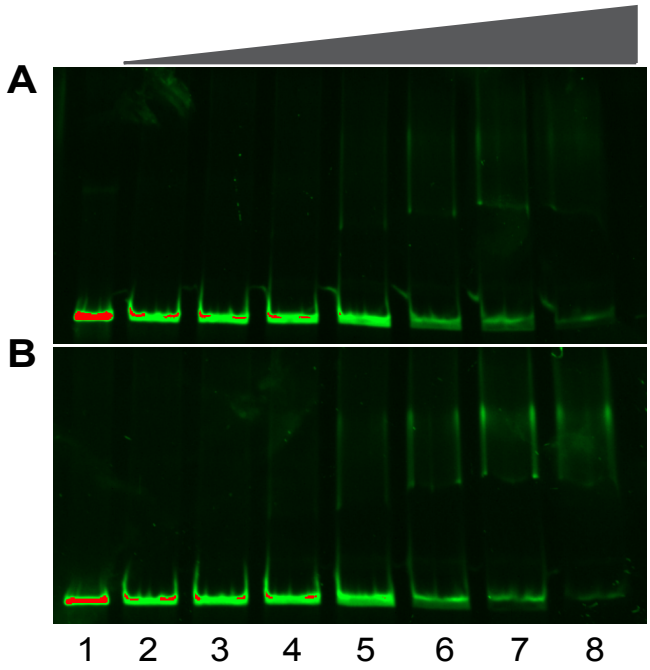

Figure 3.6 EMSA exeperiment in the abscence (panel A) or presence (panel B) of $CaCl_2$
Lane 1: no protein, lanes 2-8 had 8, 16, 32, 64, 128, 256 and 512 pmol of protein respectively.

3.4.7 Global transcriptomic profiling of *ΔmsrA* and *ΔmsrB* strains

To further investigate the regulatory functions of MsrA and MsrB we compared the transcriptional profile of wild-type *M. acetivorans* to that of *msrA* and *msrB* mutants after growth in a variety of media (Table 3.8). Deletion of *msrA* and *msrB* resulted in dramatically lower levels of the *mtaCB1* transcript during growth on methanol, while no difference was observed in other media. This result is consistent with previously obtained reporter gene assays (Bose and Metcalf, 2008). Interestingly, *mtaCB1* was the sole transcript whose expression level was specifically affected by deletion of *msrA*. In contrast, deletion of *msrB* alone

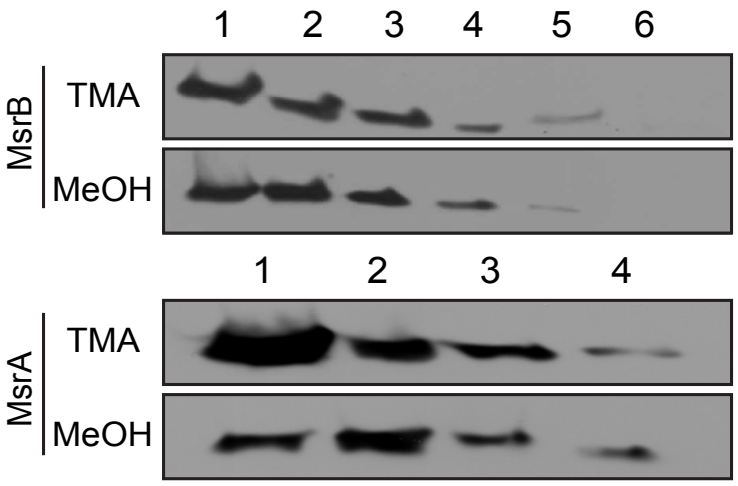

Figure 3.7 Western blot analysis of 3X-FLAG tag MsrA and MsrB. Cell lysates of strains carrying a chromosomally tagged 3X-FLAG tag msrA and msrB were separated by SDS-PAGE, transferred onto nitrocellulose membrane and probe with anit-FLAG and anti-mouse antibodies. 25 µg (lane 1), 12.5 µg (lane 2), 6.25 µg (lane 3), 3.12 µg (lane 4), 1.56 µg (lane 5) and 0.78 µg (lane 6) of total protein from cells lysates of methanol and TMA grown cells.

had more widespread effects. Accordingly, *msrB* mutants showed significantly higher levels of transcripts of genes known to be involved in growth on monomethylamine (*mttP*, *mtmB1*, *mtmC1*, *mtbA*), dimethylamine (*mtbC1*, *mttP*, *mtbB*) or trimethylamine (*mttb*, *mttC*) when the cells were grown on methanol or acetate. The increase in abundance was still apparent when the mRNA levels between the two deletion strains was compared, suggesting that the role of MsrB in regulation of methylamine related genes is independent from MsrA and vice versa. To show that this affect was MsrB-specific, we complemented the deletion mutant with a copy of the gene expressed under the constitutive promoter *pmcrB (*WWM934*)*. qPCR analysis of methanol-grown Δ*msrB* and WWM934 cells showed a decrease in mRNA abundance in the complemented strains when compared to WT (fold change regulation: *ΔmsrB:mttB1,* +750; WWM934:*mttB1,* +22; *ΔmsrB:mtmB1,* +32; WWM934:*mtmB1,* +1; *ΔmsrB:*mtaB1, -92; WWM934:*mtaB1*, +4).

3.5 Discussion

The data presented in this Chapter are fully consistent with a model in which the positive transcriptional effects of MsrA and MsrB are mediated by binding to the *mtaCB1*promoter. Because transcript and protein levels for MsrA and MsrB do not change significantly in the presence of different substrates and addition of methanol to EMSA experiments did not enhanced binding, it is likely that in the presence of methanol MsrA/B might be subjected to conformational changes due to post-translational modifications that allow them to dimerize and bind the *mtaCB1* promoter presumably to aid in the recruitment and stabilization of the RNA polymerase.

This idea is supported by our EMSA experiments, which show that an MsrA/B complex binds the mtaCB1 promoter in a sequence specific manner. EMSA assays showed that a region between 419-456 from the ATG start codon of the *mtaCB1* operon is sufficient for binding of MsrA/B. However, binding to this region was weak suggesting that surrounding regions might contribute to binding. In addition, it is important to mention that although we saw a strong footprint in the region of 318 to 380 the entire promoter fragment used as probe was protected in comparison to the no protein and BSA controls. This suggests that similar to TrmBL2 (Maruyama et al., 2011, Wierer et al., 2016), MsrAB might bind to multiple sites of its cognate promoter *in vitro*. Moreover, tools to predict and analyze protein structures (HHphred) have identified MsrA and MsrB as structural homologs to the TrmBL2 protein from *P. furiosus* thus supporting our hypothesis of TrmBL2-like behavior in Msr proteins. The location of the MsrAB binding site with reference to the TSS, TATA-box and BRE shows that there is a substantial distance between these two regions in comparison to that seen for other archaeal regulators like LysM, Ptr-2 and GvpE (Bauer et al., 2008, Brinkman et al., 2002, Ouhammouch et al., 2003, Ouhammouch and Geiduschek, 2005, Ouhammouch et al., 2005). This, in addition to the observation of multiple binding sites suggests that DNA looping might be a way

that MsrA/B contact the pre initiation complex and aids in the recruitment and stabilization of RNA polymerase.

Contrary to the previously described methanogen regulators, MsvR and MreA, which bind their respective promoters as homodimers, the Msr proteins bind to their cognate promoter as heterodimers. Although common in eukaryotes, binding of target promoters by heterodimeric proteins is rarely seen in bacterial and archaeal regulators. Presumably, binding as heterodimers expands the regulatory properties of these proteins since it would allow them to pair up with multiple partners to regulate multiple genes. Only a few examples have been described in literature including the FlhDC complex, which activates expression of flagella synthesis in *E. coli* (Wang et al., 2006), the RcsB response regulator responsible for the activation of motility, biofilm and stress response genes in *E. coli* (Salscheider et al., 2014) and the CPE1446-CPE1447 complex in *C. perfringens* which activates the transcription of toxin synthesis genes in this organism (Obana and Nakamura, 2011). Some similarities exist between these systems and the MsrA/B complex including the expression of the genes in an operon and the proximity of the regulators to one of their target operons. However, the domain structure of the proteins is different since Msr proteins lack a canonical receiver domain (Fig 3.1) and as most of the archaeal transcription regulators it posses a sensing domain unique to the archaeal domain of life. This suggests a novel mechanism of regulation by the Msr proteins.

Our experiments showed that expression of *msrA* and *msrB* is constitutive in all conditions tested (Table 3.9). Western blot analysis from N-terminus FLAG-tagged cells grown on methanol or TMA revealed that not only is transcript made but also protein stability remains intact even in conditions were MsrA/B are not expected to regulate methyltransferase expression. Therefore it is probable that activity of these proteins is regulated by covalent modifications or by direct interactions with other members of the regulatory cascade. However, there is a caveat in our experiment design since it has been shown that some Archaea

proteins have degradation signals in the N-terminus domain. Hence, there is the possibility that proteins are stable in our experiment because the N-terminus FLAG-tag is preventing protein degradation. Bioinformatic analysis of the MsrA and MsrB amino acid sequence did not revealed any degradation signal in the N-terminus of the proteins. However, we are constructing strains that have C-terminus FLAG tag MsrA and MsrB to test for this possibility.

Global transcriptional profiling of *ΔmsrA* and *ΔmsrB* strains revealed that in addition to activating the transcription of the *mtaCB1* operon, MsrB represses the methylamine methyltransferase genes when the cells are grown on methanol or acetate (Table 3.8). Thus, MsrB has a dual function in transcription regulation. This dual regulatory function has also been proposed for the *M. acetivorans* regulator MreA (Reichlen et al., 2012). Since only *ΔmsrB* and not *ΔmsrA* strains showed an over-expression of methylamine related genes (Table 3.8) it is unlikely that MsrA and MsrB are forming a heterodimer to regulate these operons. Instead it is possible that MsrB is forming a homodimer like the *M. acetivorans* regulator MaMsvR (Isom et al., 2013). Another possibility is that similar to the *E. coli* RcsB response regulator, which has been shown to pair with different transcription regulators (Salscheider et al., 2014), MsrB might be pairing up with a different Msr protein to act as a repressor. However, further EMSA and pull down experiments discussed in Chapter 5 of this thesis would be necessary to asses binding of MsrB to the methylamines promoters and to identify other binding partners of MsrB.

Although the DNA binding domain of the Msr family regulators shares a weak similarity to the bacterial ArsR family of regulators, numerous differences suggest that the Msr proteins are functionally and evolutionarily distinct. These differences include: i, the capacity of Msr proteins to act both as repressors and activators of transcription while ArsR regulators only act as repressors of transcription; ii, contrary to ArsR regulators MsrA/B function as heterodimers and iii, unlike ArsR regulators Msr do not regulate their own expression. Moreover,

Msr proteins do not contain the conserved metal binding residues found in other ArsR-family members (Busenlehner et al., 2003). Based on these observations we propose that the Msr proteins are not part of the ArsR family of regulators but rather form a novel family of DNA binding proteins exclusive to the domain Archaea.

Despite our increased understanding of the mechanism of Msr action, numerous critical questions remain. In particular, our data shed no light on the mechanism of methanol sensing by *M. acetivorans* that ultimately leads to the expression of the methanol specific methyltransferases. Addition of methanol to gel-shift experiments did not affect the ability of these proteins to bind the *mtaCB1* (data not shown), thus these proteins are not themselves the methanol sensors which strongly suggest that these proteins respond to the presence of methanol by an unknown mechanism. In addition, EMSA experiments showed that addition of metals improve binding of MsrAB to the *mtaCB1* promoter. However, the role that metals have in DNA binding remains undetermined. Finally, to see binding our EMSA experiments required a large protein to DNA ratio but the binding to the promoter region was specific. These results suggest that there is a missing component to the system that we hypothesize could be in the form of a post-translational modification or in a required third component of the system. In Chapter 5 of this thesis I propose experiments to identify the potential missing component.

In conclusion, this study solidifies the basic scheme of the methanol regulatory network in *M. acetivorans* C2A. Although, the details of this network have not been worked out completely, a number of important clues were revealed from this study. We anticipate that information from future studies combined with present knowledge will shed light on the physiological significance of complex metabolic regulation in this environmentally significant group of organisms.

Table 3.8 Select mRNA abundance changes of ΔmsrA and ΔmsrB strains[a] (CLC genomics).

		ΔmsrA			ΔmsrB			ΔmsrB vs ΔmsrA		
		MeOH	TMA	Ac	MeOH	TMA	Ac	MeOH	TMA	Ac
Locus Tag	Annotation	Fold Change[b]	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change	Fold Change
MA0143	*mttP*	1	2	-1	**5**	-1	**3**	**5**	-1	**3**
MA0144	*mtmB1*	1	**5**	-1	**44**	1	**5**	**42**	**-5**	**6**
MA0145	*mtmC1*	1	**6**	-1	**52**	1	**6**	**41**	**-5**	**6**
MA0146	*mtbA*	-1	**4**	-1	**7**	1	**2**	**8**	**-4**	2
MA0455	*mtaB1*	**-56**	1	-1	**-56**	-2	-1	-1	-2	-1
MA0456	*mtaC1*	**-46**	2	-1	**-41**	-1	-1	1	-2	-1
MA0527	*mtbC*	-1	3	-1	**52**	-2	**25**	**60**	**-5**	**30**
MA0528	*mttB*	-1	**5**	-1	**76**	-1	**28**	**82**	**-6**	**32**
MA0529	*mttC*	1	**7**	-1	**120**	-1	**28**	**108**	**-6**	**29**
MA0530	*transmembrane protein MttP*	1	3	-1	**27**	-2	**10**	**26**	**-5**	**10**
MA0531	*predicted protein*	1	**5**	-1	**39**	-1	**11**	**29**	**-5**	**12**
MA0532	*mtbB*	1	**4**	-1	**44**	-1	**9**	**43**	**-5**	**10**

[a]The full datasets for each condition, including parsed tables showing only significant differences, are deposited in (submission pending).
[b]Fold change of expression values of mutant strains compared to WT *M. acetivorans* grown in the same substrate. ΔmsrB vs ΔmsrA columns present fold change of expression values of ΔmsrB strains when compared to ΔmsrA grown in the same substrate. Statistically significant values (P<0.01) based on EDGE test are highlighted in bold.

Table 3.9 Expression of *msrA* and *mrsB*.

		Normalized expression values					
		WT		ΔmsrA		ΔmsrB	
Locus Tag	Annotation	MeOH	TMA	MeOH	TMA	MeOH	TMA
MA0459	*msrA*	403	331	1	17	479	584
MA0460	*msrB*	166	215	229	130	1	2

3.6 References

1. ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. & LIPMAN, D. J. 1990. Basic local alignment search tool. *J Mol Biol,* 215**,** 403-10.

2. AUSUBEL, F. M., BRENT, R., KINGSTON, R. E., MOORE, D. D., SEIDMAN, J. G., SMITH, J. A. & STRUHL, K. 1992. *Current protocols in molecular biology,* New York, John Wiley & Sons.

3. BAUER, M., MARSCHAUS, L., REUFF, M., BESCHE, V., SARTORIUS-NEEF, S. & PFEIFER, F. 2008. Overlapping activator sequences determined for two oppositely oriented promoters in halophilic Archaea. *Nucleic Acids Res,* 36**,** 598-606.

4. BOCCAZZI, P., ZHANG, J. K. & METCALF, W. W. 2000. Generation of dominant selectable markers for resistance to pseudomonic acid by cloning and mutagenesis of the *ileS* gene from the archaeon *Methanosarcina barkeri* fusaro. *J Bacteriol,* 182**,** 2611-8.

5. BOSE, A., KULKARNI, G. & METCALF, W. W. 2009. Regulation of putative methyl-sulphide methyltransferases in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 74**,** 227-38.

6. BOSE, A. & METCALF, W. W. 2007. Distinct regulators control the expression of methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 67**,** 649-661.

7. BOSE, A. & METCALF, W. W. 2008. Distinct regulators control the expression of methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 67**,** 649-61.

8. BOSE, A., PRITCHETT, M. A., ROTHER, M. & METCALF, W. W. 2006. Differential Regulation of the Three methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *J Bacteriol,* 188**,** 7274-83.

9. BRINKMAN, A. B., BELL, S. D., LEBBINK, R. J., DE VOS, W. M. & VAN DER OOST, J. 2002. The *Sulfolobus solfataricus* Lrp-like protein LysM regulates lysine biosynthesis in response to lysine availability. *J Biol Chem,* 277**,** 29537-49.

10. BUSENLEHNER, L. S., PENNELLA, M. A. & GIEDROC, D. P. 2003. The SmtB/ArsR family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev,* 27**,** 131-43.

11. CHENNA, R., SUGAWARA, H., KOIKE, T., LOPEZ, R., GIBSON, T. J., HIGGINS, D. G. & THOMPSON, J. D. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res,* 31**,** 3497-500.

12. DEPPENMEIER, U., JOHANN, A., HARTSCH, T., MERKL, R., SCHMITZ, R. A., MARTINEZ-ARIAS, R., HENNE, A., WIEZER, A., BAUMER, S., JACOBI, C., BRUGGEMANN, H., LIENARD, T., CHRISTMANN, A., BOMEKE, M., STECKEL, S., BHATTACHARYYA, A., LYKIDIS, A., OVERBEEK, R., KLENK, H. P., GUNSALUS, R. P., FRITZ, H. J. & GOTTSCHALK, G. 2002. The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J Mol Microbiol Biotechnol,* 4**,** 453-61.

13. DODD, I. B. & EGAN, J. B. 1990. Improved detection of helix-turn-helix DNA-binding motifs in protein sequences. *Nucleic Acids Res,* 18**,** 5019-26.

14. FELSENSTEIN, J. 1989. Phylogeny Inference Package (Version 3.2). *Cladistics* 5**,** 164-166.

15. FU, H. & METCALF, W. W. 2015. Genetic basis for metabolism of methylated sulfur compounds in Methanosarcina species. *J Bacteriol,* 197**,** 1515-24.

16. GALAGAN, J. E., NUSBAUM, C., ROY, A., ENDRIZZI, M. G., MACDONALD, P., FITZHUGH, W., CALVO, S., ENGELS, R., SMIRNOV, S., ATNOOR, D., BROWN, A., ALLEN, N., NAYLOR, J., STANGE-THOMANN, N., DEARELLANO, K., JOHNSON, R., LINTON, L., MCEWAN, P., MCKERNAN, K., TALAMAS, J., TIRRELL, A., YE, W., ZIMMER, A., BARBER, R. D., CANN, I., GRAHAM, D. E., GRAHAME, D. A., GUSS, A. M., HEDDERICH, R., INGRAM-SMITH, C., KUETTNER, H. C., KRZYCKI, J. A., LEIGH, J. A., LI, W., LIU, J., MUKHOPADHYAY, B., REEVE, J. N., SMITH, K., SPRINGER, T. A., UMAYAM, L. A., WHITE, O., WHITE, R. H., CONWAY DE MACARIO, E., FERRY, J. G., JARRELL, K. F., JING, H., MACARIO, A. J., PAULSEN, I., PRITCHETT, M., SOWERS, K. R., SWANSON, R. V., ZINDER, S. H., LANDER, E., METCALF, W. W. & BIRREN, B. 2002. The genome of M. acetivorans reveals extensive metabolic and physiological diversity. *Genome Res,* 12**,** 532-42.

17. GERNHARDT, P., POSSOT, O., FOGLINO, M., SIBOLD, L. & KLEIN, A. 1990. Construction of an integration vector for use in the archaebacterium *Methanococcus voltae* and expression of a eubacterial resistance gene. *Mol Gen Genet,* 221**,** 273-9.

18. GUSS, A. M., ROTHER, M., ZHANG, J. K., KULKARNI, G. & METCALF, W. W. 2008. New methods for tightly regulated gene expression and highly efficient insertion of foreign genes for *Methanosarcina* species. *Archaea,* 2**,** 193-203.

19. HALDIMANN, A. & WANNER, B. L. 2001. Conditional-replication, integration, excision, and retrieval plasmid-host systems for gene structure-function studies of bacteria. *J Bacteriol,* 183**,** 6384-93.

20. HIGGINS, D. G., BLEASBY, A. J. & FUCHS, R. 1992. CLUSTAL V: improved software for multiple sequence alignment. *Comput Appl Biosci,* 8**,** 189-91.

21. ISOM, C. E., TURNER, J. L., LESSNER, D. J. & KARR, E. A. 2013. Redox-sensitive DNA binding by homodimeric *Methanosarcina acetivorans* MsvR is modulated by cysteine residues. *Bmc Microbiology,* 13.

22. KEARSE, M., MOIR, R., WILSON, A., STONES-HAVAS, S., CHEUNG, M., STURROCK, S., BUXTON, S., COOPER, A., MARKOWITZ, S., DURAN, C., THIERER, T., ASHTON, B., MEINTJES, P. & DRUMMOND, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28**,** 1647-9.

23. KELTJENS, J. T. & VOGELS, G. D. Conversion of methanol and methylamines to methane and carbon dioxide. *In:* FERRY, J. G. (ed.) *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics.* New York: Chapman and Hall.

24. LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

25. LUKASHIN, A. V. & BORODOVSKY, M. 1998. GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res,* 26**,** 1107-15.

26. MAEDER, D. L., ANDERSON, I., BRETTIN, T. S., BRUCE, D. C., GILNA, P., HAN, C. S., LAPIDUS, A., METCALF, W. W., SAUNDERS, E., TAPIA, R. & SOWERS, K. R. 2006. The *Methanosarcina barkeri* genome: comparative analysis with *Methanosarcina acetivorans* and *Methanosarcina mazei* reveals extensive rearrangement within methanosarcinal genomes. *J Bacteriol,* 188**,** 7922-31.

27. MARCHLER-BAUER, A., ANDERSON, J. B., DERBYSHIRE, M. K., DEWEESE-SCOTT, C., GONZALES, N. R., GWADZ, M., HAO, L., HE, S., HURWITZ, D. I., JACKSON, J. D., KE, Z., KRYLOV, D., LANCZYCKI, C. J., LIEBERT, C. A., LIU, C., LU, F., LU, S., MARCHLER, G. H., MULLOKANDOV, M., SONG, J. S., THANKI, N., YAMASHITA, R. A., YIN, J. J., ZHANG, D. & BRYANT, S. H. 2007. CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res,* 35**,** D237-40.

28. MARKOWITZ, V. M., KORZENIEWSKI, F., PALANIAPPAN, K., SZETO, E., WERNER, G., PADKI, A., ZHAO, X., DUBCHAK, I., HUGENHOLTZ, P., ANDERSON, I., LYKIDIS, A., MAVROMATIS, K., IVANOVA, N. & KYRPIDES, N. C. 2006. The integrated microbial genomes (IMG) system. *Nucleic Acids Res,* 34**,** D344-8.

29. MARUYAMA, H., SHIN, M., ODA, T., MATSUMI, R., OHNIWA, R. L., ITOH, T., SHIRAHIGE, K., IMANAKA, T., ATOMI, H., YOSHIMURA, S. H. & TAKEYASU, K. 2011. Histone and TK0471/TrmBL2 form a novel heterogeneous genome architecture in the hyperthermophilic archaeon Thermococcus kodakarensis. *Mol Biol Cell,* 22**,** 386-98.

30. METCALF, W. W., ZHANG, J. K., APOLINARIO, E., SOWERS, K. R. & WOLFE, R. S. 1997. A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc Natl Acad Sci U S A,* 94**,** 2626-31.

31. METCALF, W. W., ZHANG, J. K. & WOLFE, R. S. 1998. An anaerobic, intrachamber incubator for growth of *Methanosarcina* spp. on methanol-containing solid media. *Appl Environ Microbiol,* 64**,** 768-70.

32. MILLER, V. L. & MEKALANOS, J. J. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol,* 170**,** 2575-83.

33. MULDER, N. J., APWEILER, R., ATTWOOD, T. K., BAIROCH, A., BATEMAN, A., BINNS, D., BORK, P., BUILLARD, V., CERUTTI, L., COPLEY, R., COURCELLE, E., DAS, U., DAUGHERTY, L., DIBLEY, M., FINN, R., FLEISCHMANN, W., GOUGH, J., HAFT, D., HULO, N., HUNTER, S., KAHN, D., KANAPIN, A., KEJARIWAL, A., LABARGA, A., LANGENDIJK-GENEVAUX, P. S., LONSDALE, D., LOPEZ, R., LETUNIC, I., MADERA, M., MASLEN, J., MCANULLA, C., MCDOWALL, J., MISTRY, J., MITCHELL, A., NIKOLSKAYA, A. N., ORCHARD, S., ORENGO, C., PETRYSZAK, R., SELENGUT, J. D., SIGRIST, C. J., THOMAS, P. D., VALENTIN, F., WILSON, D., WU, C. H. & YEATS, C. 2007. New developments in the InterPro database. *Nucleic Acids Res,* 35**,** D224-8.

34. NAYAK, W. W. M. D. N. 2017 Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans Proc Natl Acad Sci U S A*.

35. OBANA, N. & NAKAMURA, K. 2011. A novel toxin regulator, the CPE1446-CPE1447 protein heteromeric complex, controls toxin genes in Clostridium perfringens. *J Bacteriol,* 193**,** 4417-24.

36. OUHAMMOUCH, M., DEWHURST, R. E., HAUSNER, W., THOMM, M. & GEIDUSCHEK, E. P. 2003. Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proc Natl Acad Sci U S A,* 100**,** 5097-102.

37. OUHAMMOUCH, M. & GEIDUSCHEK, E. P. 2005. An expanding family of archaeal transcriptional activators. *Proc Natl Acad Sci U S A,* 102**,** 15423-8.

38. OUHAMMOUCH, M., LANGHAM, G. E., HAUSNER, W., SIMPSON, A. J., EL-SAYED, N. M. & GEIDUSCHEK, E. P. 2005. Promoter architecture and response to a positive regulator of archaeal transcription. *Mol Microbiol,* 56**,** 625-37.

39. PRITCHETT, M. A. & METCALF, W. W. 2005. Genetic, physiological and biochemical characterization of multiple methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 56**,** 1183-94.

40. REICHLEN, M. J., VEPACHEDU, V. R., MURAKAMI, K. S. & FERRY, J. G. 2012. MreA functions in the global regulation of methanogenic pathways in Methanosarcina acetivorans. *MBio,* 3**,** e00189-12.

41. ROBINSON, M. D. & SMYTH, G. K. 2008. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics,* 9**,** 321-32.

42. SALSCHEIDER, S. L., JAHN, A. & SCHNETZ, K. 2014. Transcriptional regulation by BglJ-RcsB, a pleiotropic heteromeric activator in *Escherichia coli*. *Nucleic Acids Res,* 42**,** 2999-3008.

43. SAUER, K., HARMS, U. & THAUER, R. K. 1997. Methanol:coenzyme M methyltransferase from *Methanosarcina barkeri*. Purification, properties and encoding genes of the corrinoid protein MT1. *Eur J Biochem,* 243**,** 670-7.

44. SODING, J., BIEGERT, A. & LUPAS, A. N. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res,* 33**,** W244-8.

45. SOWERS, K. R., BARON, S. F. & FERRY, J. G. 1984. *Methanosarcina acetivorans* sp. nov., an Acetotrophic Methane-Producing Bacterium Isolated from Marine Sediments. *Appl Environ Microbiol,* 47**,** 971-978.

46. SOWERS, K. R., BOONE, J. E. & GUNSALUS, R. P. 1993. Disaggregation of *Methanosarcina* spp. and growth as single cells at elevated osmolarity. *Appl Environ Microbiol,* 59**,** 3832-3839.

47. THAUER, R. K. 1998. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology,* 144 (Pt 9)**,** 2377-406.

48. THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res,* 22**,** 4673-80.

49. VAN DER MEIJDEN, P., HEYTHUYSEN, H. J., POUWELS, A., HOUWEN, F., VAN DER DRIFT, C. & VOGELS, G. D. 1983. Methyltransferases involved in methanol conversion by *Methanosarcina barkeri*. *Arch Microbiol,* 134**,** 238-42.

50. WANG, S., FLEMING, R. T., WESTBROOK, E. M., MATSUMURA, P. & MCKAY, D. B. 2006. Structure of the *Escherichia coli* FlhDC complex, a prokaryotic heteromeric regulator of transcription. *J Mol Biol,* 355**,** 798-808.

51. WANNER, B. L. 1986. Novel regulatory mutants of the phosphate regulon in *Escherichia coli* K-12. *J Mol Biol,* 191**,** 39-58.

52. WIERER, S., DALDROP, P., UD DIN AHMAD, M., BOOS, W., DRESCHER, M., WELTE, W. & SEIDEL, R. 2016. TrmBL2 from Pyrococcus furiosus Interacts Both with Double-Stranded and Single-Stranded DNA. *PLoS One,* 11**,** e0156098.

53. YANG, J., YAN, R., ROY, A., XU, D., POISSON, J. & ZHANG, Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Methods,* 12**,** 7-8.

54. ZINDER, S. H. 1993. Physiological ecology of methanogens. *In:* FERRY, J. G. (ed.) *Methanogenesis: Ecology, Physiology, Biochemistry and Genetics.* New York: Chapman and Hall.

Chapter 4: Characterization of Two-Component Regulatory Systems in *Methanosarcina acetivorans* C2A

4.1 Abstract

The Methanosarcinales are the most metabolically diverse methanogens using a variety of substrates, including the one-carbon (C-1) compounds methanol, methylamine and methyl sulfides. Use of these C-1 substrates requires the expression of highly specific methyltransferases, which are only expressed in the presence of their cognate substrates. The genes encoding these methyltransfereases are among the most regulated genes in archaea; however, the mechanism(s) by which *Methanosarcina* cells are able to sense and respond to the presence/absence of these substrates has yet to be established. Analysis of the genomic region surrounding the methyl sulfide methyltransferase genes *mtsD*, *mtsF* and *mtsH* allowed the identification of three histidine kinases (MA0863, MA4377, MA4561), one response regulator (MA4376) and an Msr-like protein (MA4375) in close proximity to the *mts* genes. In this chapter we examine the transcriptional regulation of *mtsD* and *mtsF* by the mentioned sensory transduction systems. To test this we performed global transcriptional profiling of strains that lacked the putative two-component systems in the presence of trimethylamine, methanol or a mixture of methanol and dimethylsulfide. Transcriptomic studies showed that three putative histidine kinases (MA0863, MA4377 and MA4561) act as negative regulators of *mts* genes in the presence of methanol. In addition, we have identified a putative response regulator (MA4376) that mediates the regulation of *mts* genes by interacting with two independent histidine kinases (MA0863 and MA4377). Finally, we show that an additional *msr* like protein is involved in the regulation of *mts* genes.

4.2 Introduction

Two-component signal transduction systems are commonly used by bacteria to sense and respond to environmental changes including nutrients, pH, cellular redox changes and temperature (Ashby, 2006). Signal transduction is accomplished by the phosphorylation of an aspartate residue in a response regulator (RR) usually mediated by a membrane bound histidine kinase (HK). Typically, upon sensing environmental stimuli the histidine kinase is autophosphorylated at a histide residue. The phosphoryl group is then transferred to an aspartate residue in a response regulator; this triggers a conformational change in the RR that activates the output domain of the regulator allowing it to respond to the environmental stimuli. Histidine kinases can be membrane bound or cytoplasmic and consist of a conserved C terminus HATPase domain coupled with an array of sensory domains including PAS, CHASE4 and GAF. Response regulators can be coupled to an effector domain or can be found as stand alone receiver domains.

Two-component systems are also found in many genomes within the Euryarchaeota (Salvado et al., 2015). However, they are poorly studied. Genome sequencing of *M. acetivorans* identified the presence of 50 histidine kinases (HK) (Galagan et al., 2002). In addition, 18 response regulator receiver domains were identified, seven present within histidine kinases and eleven single domain proteins of which only one is big enough to have both a receiver and an effector domain (Galagan et al., 2002). Thus far, only one system has been characterized in this strain, the methyl sulfide methyltransferase associated sensor MsmS (Molitor et al., 2013). MsmS has been described as a heme-based redox sensor that modulates its autophosphorylation activity through the redox state of its heme cofactor (Molitor et al., 2013). Deletion of MsmS resulted in strains that constitutively expresses the methyl sulfide methyltransferase MtsF, indicating that MsmS negatively regulates expression of the enzyme. The effect of MsmS on the expression of other methyl sulfide methyltransferases was not examined.

Based on their genomic localization two additional histidine kinases (MA0863 and MA4377) and a response regulator (MA4376) might be involved in the regulation of the methyl sulfide methyltransferases *mtsD* and *mstF*. MA0863 was originally annotated as two genes MA0863 and MA0864; however, a previous work revealed that MA0863 encodes a pyrrolysine (Pyl) containing histidine kinase that might be involved in regulation of *mts* genes (O'Donoghue et al., 2014). Here we present an array transcriptomic experiments that show that the histidine kinases MA0863, MA4377 and MA4561 are involved in the regulation of all three methyl sulfide methyltransferases. Moreover, the response regulator MA4376 seems to be mediating the regulation of the *mts* gene by interacting with more than one histidine kinase.

4.3 Materials and Methods

4.3.1 Strains, media and growth conditions

Methanosarcina strains used in the study are shown in Table 4.1. All *Methanosarcina* strains were grown under strictly anaerobic conditions. Derivatives of *M. acetivorans* WWM73 and WWM60 (Guss et al., 2008a) were grown at 37°C in high-salt (HS) medium (Metcalf et al., 1997). Growth substrates provided were 125 mM methanol, 120 mM acetate, 50 mM trimethylamine (TMA) and a mixture of 125 mM methanol with 5 mM dimethyl sulfate (DMS). Substrates were sterilized by filtration using a 0.2 µm filter. Growth on media solidified with 1.5% agar was as described previously (William Metcalf, 1996, Metcalf et al., 1997). All plating manipulations were carried out under strictly anaerobic conditions in an anaerobic incubator as described (William Metcalf, 1998). Puromycin (CalBiochem, San Diego, CA) was added from sterile, anaerobic stocks at a final concentration of 2 mg/ml for selection of *Methanosarcina* strains carrying puromycin transacetylase gene (*pac*). The purine analog 8-aza-2,6-diaminopurine (Sigma, St. Louis, MO) was added from sterile, anaerobic stocks at a final concentration of 20 mg/ml for selection against the *hpt* gene.

4.3.2 Construction and verification of mutant strains

M. acetivorans strains WWM586, WWM587, WWM926 and WWM975 were constructed by markerless genetic exchange via liposome-mediated transformation (Metcalf et al., 1997, Pritchett et al., 2004). Construction of the plasmid used for making WWM975 was made with a variation of the golden gate method (Liang et al., 2014). WWM977 was constructed using the CRISPR-Cas-9 system as described in (Nayak, 2017). DNA fragments encoding the sgRNA were synthetized as double-stranded DNA fragments (gBlocks) from Integrated DNA Technologies (Coralville, IA) and introduced into pDN201 backbone linearized with *AscI* using the NEBuilder HiFi DNA assembly kit following manufacturer's instructions (New England Biolabs, Ipswich,MA). The sgRNA sequence was designed using the CRISPR site finding tool in Genieious version R9 (Kearse et al., 2012). For construction of complementation strains, plasmids were integrated on the *M. acetivorans* C2A chromosome using site-specific recombination between the ϕC31 *attB* site on the plasmid and the ϕC31 *attP* site on the *M. acetivorans* chromosome using liposome mediated transformation as described previously (Guss et al., 2008b). All plasmids were confirmed via restriction endonuclease digestion and DNA sequencing. Sequencing was performed at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois. Chromosomal mutations were verified by PCR and DNA hybridization experiments (data not shown). DNA hybridizations were performed using the DIG System (Roche, Mannheim, Germany) as recommended using MagnaGraph Nylon transfer membranes (Micron Separations Inc, Westborough, MA). Standard methods were used to isolate and manipulate plasmid DNA from *E. coli* strains (Ausubel, 1992). For strain genotypes, plasmid descriptions, primers and gBlocks refer to Tables 4.1, 4.2, .4.3 and 4.4 respectively.

4.3.3 Transcriptomic analysis (RNAseq)

Strains were adapted to specific growth substrates for at least 30 generations prior to RNA isolation. RNA was extracted from triplicate cultures grown to

Table 4.1 *Methanosarcina* strains used in this study.

Strain	Genotype	Source
WWM60	Δ*hpt*::PmcrB-*tetR*	(Guss et al., 2008a)
WWM82	Δ*hpt*:: PmcrB-ϕC31*int*-attP	(Guss et al., 2008a)
WWM586	Δ*hpt*:: PmcrB-ϕC31*int*-attP ΔMA4376	this study
WWM587	Δ*hpt*:: PmcrB-ϕC31*int*-attP ΔMA4377	this study
WWM926	Δ*hpt*:: PmcrB-ϕC31*int*-attP ΔMA4561	this study
WWM975	Δ*hpt*:: PmcrB-ϕC31*int*-attP ΔMA4375	this study
WWM977	Δ*hpt*::PmcrB-*tetR* ΔMA0863	this study
WWM995	Δ*hpt*::PmcrB-tetR P*mcrB*(tetO4)-MA0863-Pyl-Lys ΔMA0863	this study
WWM1003	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4377-attL ΔMA4377	this study
WWM1004	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4376-attL ΔMA4376	this study
WWM1005	Δ*hpt*::PmcrB-tetR P*mcrB*(tetO4)-MA0863 ΔMA0863	this study
WWM1006	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4375-attL ΔMA4375	this study
WWM1007	Δ*hpt*::PmcrB-tetR P*mcrB*(tetO4)-MA4561 ΔMA4561	this study
WWM1009	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4377-R2-Asp-Glu-attL ΔMA4377	this study
WWM1010	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4377-His-Ala-attL ΔMA4377	this study
WWM1011	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4377-R1-Asp-Ala-attL ΔMA4377	this study
WWM1012	Δ*hpt*:: ΦC31-*int*-attR PmcrB(tetO4)-MA4377-R2-Asp-Ala-attL ΔMA4377	this study

mid-exponential phase using Direct-zol RNA MiniPrep kit (Zymo Research Irvine, CA) following manufacturer's instructions. Ribosomal RNA was depleted using subtractive hybridization employing biotinylated rRNA probes as previously described (Fu and Metcalf, 2015). The quality of the RNA was verified using Agilent Bioanalyzer. RNA sequencing was performed at the W. M. Keck Center for Comparative and Functional Genomics, University of Illinois using strand specific libraries prepared with the Illumina TruSeq Stranded mRNAseq Sample Prep kit (Madison, WI). Bar-coded libraries were pooled, quantified by qPCR and sequenced for 101 cycles on a HiSeq4000 using the HiSeq 4000 sequencing kit version 1 (Illumina, San Diego, CA). Reads were trimmed using TrimmomaticSE (-threads 23, ILLUMINACLIP: 2:30:10, LEADING:28, TRAILING:28, MINLEN:30). Trimmed reads were mapped using the Rockhopper version 2.0.3 (Tjaden, 2015, McClure et al., 2013) mapping to a reference genome with the following parameters allowed mismatches: 0.15 and minimum seed length: 0.33.

Table 4.2 Plasmids used in this study.

Plasmids	Description	Reference
pJK026A	ϕC31-attB vector with PmcrB promoter fusion to uidA	(Guss et al., 2008a)
pJK029A	ϕC31-attB vector with PmcrB(tetO4) promoter fusion to uidA	(Guss et al., 2008a)
pMP44	Ampicillin resistant vector used to mutate cells using the markerless exchange method	(Pritchett et al., 2004)
pDN201	Plasmid used for expression of the CRISPR-Cas9 system in *Methanosarcina acetivorans*	(reference)
pJPK1	NotI/SpeI digested 2kb homology region surrounding MA4376 (using primers MA4376-del-up–R and MA4376-del-Dn-F) cloned into NotI/SpeI digested pMP44	This study
pJPK2	NotI/SpeI digested 2kb homology region surrounding MA4377 (using primers MA4377-del-up–R and MA4376-del-Dn-F) cloned into NotI/SpeI digested pMP44	This study
pML52	Plasmid for deletion of MA4375 assembled using golden gate assembly using primers MA4375_upF, MA4375_upR, MA4375_dF, MA4375_dR, pUC19_R2, pach-hpt_R3	This study
pML55	*Nde*I/*Hind*III-digested MA4561 PCR product (using primers MA4376_start and MA4376_comp_R was cloned into *Nde*I/*Hind*III-digested pJK029A	This study
pML57	*Nde*I/*Hind*III-digested MA4561 PCR product (using primers MA4561_comp_F and MA4561_comp_R) was cloned into *Nde*I/*Hind*III-digested pJK029A	This study
pML58	*Nde*I/*Hind*III-digested MA0863 PCR product (using primers MA0863F and MA0863R) was cloned into *Nde*I/*Hind*III-digested pJK029A	This study
pML59	*Nde*I/*Nsi*I-digested MA4377 (receiver domain 1 Asp-Ala) PCR product (using primers MA4377_start_gib and MA4377_comp_R_gib) was cloned into *Nde*I/*Hind*III-digested pJK029A assembled using Gibson assembly	This study
pML60	*Nde*I/*Nsi*I-digested MA4377 (receiver domain 2 Asp-Glu) PCR product (using primers MA4377_start_gib and MA4377_comp_R_gib) was cloned into *Nde*I/*Hind*III-digested pJK029A assembled using Gibson assembly	This study
pML61	*Nde*I/*Nsi*I-digested MA4377 (receiver domain 1 Asp-Glu) PCR product (using primers MA4377_start_gib and MA4377_comp_R_gib) was cloned into *Nde*I/*Hind*III-digested pJK029A assembled using Gibson assembly	This study
pML62	*Nde*I/*Nsi*I-digested MA4377 (receiver domain 2 Asp-Ala) PCR product (using primers MA4377_start_gib and MA4377_comp_R_gib) was cloned into *Nde*I/*Hind*III-digested pJK029A assembled using Gibson assembly	This study
pML63	2 kb PCR product of homology region of MA0863 (MA0864_CRISPR_PmeI and MA0863_CRISPR_AscI primers) amplified from pML53 was introduced into pDN201 linearized with *PmeI*/*AscI*	This study
pML64	gBlocks for sgRNA was introduced into pML63 linearized with *AscI* by NEB Gibson Cloning Assembly Kit	This study
pML65	pML64 retrofitted into pAMG40	This study
pML69	*Nde*I/*Hind*III-digested MA4375 PCR product (using primers MA4375_start and MA4375_comp_R) was cloned into *Nde*I/*Hind*III-digested pJK029A	This study
pML70	*Nde*I/*Hind*III-digested MA0863 PCR product (Pyl residue mutated to Lys) (using primers MA0863F and MA0863R) was cloned into *Nde*I/*Hind*III-digested pJK029A	This study

Table 4.3 Primers used in the study.

Primer	Sequence (restriction sites underlined and mutated bases lower case)	Added sites
MA4376-del-up–F	TTATTCTCCCAGATATTTATCGATCAGGGGTTCGACAATAAGAATTTCTTTCATCTTCAC	
MA4376-del-up–R	GGCGCGCCACTAGTAAGAAAATTGTAAACTTACACA	*AscI,SpeI*
MA4376-del-Dn–F	GGCGCGCCGCGGCCGCAGATGGAATGAAATGAAAATA	*AscI,NotI*
MA4376-del-Dn–R	GTGAAGATGAAAGAAATTCTTATTGTCGAACCCCTGATCGATAAATATCTGGGAGAATAA	
MA4377-del-up–F	TTATCTTCGACAATAAGAATTTCTTTCATCAATAACAAGAATTTTTCTACTCACATTCAT	
MA4377-del-up–R	GGCGCGCCACTAGTTTTCTTCTGTTATTTTGCCGTA	*AscI,SpeI*
MA4377-del-Dn–F	GGCGCGCCGCGGCCGCACGGTCAGGGTCTTTGAGGAAAACTACGATTACGTGATT	*AscI,NotI*
MA4377-del-Dn–R	ATGAATGTGAGTAGAAAAATTCTTGTTATTGATGAAAGAAATTCTTATTGTCGAAGATAA	
MA4375_upF,	GGCGCGCCCACCTGCAAAACGAGAGGCTCTGGGACCGGTTT	
MA4375_upR	GGCGCGCCCACCTGCAAAATGGCATAGAAACGCCAGGTCAAT	
MA4375_dF	GGCGCGCCCACCTGCAAAAGCCAGCTCTTTTTATATTATAAAGATATGTC	
MA4375_dR	GGCGCGCCCACCTGCAAAAAGGGGCAACTTATTCCTTTACAAAACC	
pUC19_R2	GGCGCGCCCACCTGCAAAACCCTAGGTGGCACTTTTCGG	
pac-hpt_R3	GGCGCGCCCACCTGCAAAACTCGACAGGAACACTTAACGGC	
MA4561_comp_F	ATGATTTTAATAAATTAAGGAGGAAATTCATATGATAGGTGTTGACATGAAACTGG	
MA4561_comp_R	TATAGCATACATTATACGAAGTTATCAAGATTCCAGTTTCATTCCAGCTCACG	
MA0863F	GTGGCCAGTGCCAAGCTTGCATGCCTGCATTCAACATTTACAGTAGTTCAAAATTGGTG	
MA0863R	GAGTCGTCGCCACCAATCCCCATAGATTTTATCCAGGATATTATCGAGGTCAA	

Table 4.3 Continued

MA4377_start_gib	ATGATTTTAATAAATTAAGGAGGAAATTCATATGAATGTGAGTAGAAAAATTCTTG	
MA4377_comp_R_gib	TATAGATACATTATACGAAGTTATCAAGAAGCTTCAGAAGGTCCAGGATCAATTC	
MA4375_start	GGCGCGCCCATATGAAAATATCACTTATTGACCTGGC	*AscI, NdeI*
MA4375_comp_R	GGCGCGCCAAGCTTGCCGCCTCTGTCTTAAACATGT	*AscI, HindIII*
MA4377R1_D-A_F	TGTTATCACTCTGgcaGTTTTACTTCCCGA	
MA4377R1_D-A_R	TCGGGAAGTAAAACtgcCAGAGTGATAACA	
MA4377R1_D-E_F	TGTTATCACTCTGgaaGTTTTACTTCCCGA	
MA4377R1_D-E_R	TCGGGAAGTAAAACttcCAGAGTGATAACA	
MA4377R2_D-A_F	GATATCCTCATCCTCgcaCTGCTGATGCCG	
MA4377R2_D-A_R	CGGCATCAGCAGtgcGAGGATGAGGATATC	
MA4377R2_D-E_F	GATATCCTCATCCTCgaaCTGCTGATGCCG	
MA4377R2_D-E_R	CGGCATCAGCAGttcGAGGATGAGGATATC	
MA0863_PylF	GGGTATGAAGAATGCACCaagGAGTACCGGGTCCTGACA	
MA0863_PylR	TGTCAGGACCCGGTACTCcttGGTGCATTCTTCATACCC	
MA4377_H-A_F	AATATGAGCGCAGAACTGCGG	
MA4377_H-A_R	CCGCAGTTCTGCGCTCATATT	
MA4376_start	GGCGCGCCCATATGAAGATGAAAGAAATTCTTATTGTCG	*AscI, NdeI*
MA4376_comp_R	GGCGCGCCAAGCTTGAAAAATGTCTTTCCGACCGCA	*AscI, HindIII*

Table 4.3 Continued

MA0864_CRISPR_PmeI	TATATATAGTTTAAACAAGCCCCCTGAAGATACTAATGAAT	*PmeI*
MA0863_CRISPR_AscI	TATATATAGGCGCGCCGCTGCATCTACCGGGTCGTGTAGAGATT	*AscI*

Table 4.4 gBlocks used in this study[a]

gBlock	Sequence
MA0863 gRNA 1	AATCTCTACACGACCCGGTAGATGCAGCGGCGCGCCAACAACATCAGTCACCTAAAAAGAGAAAACGAATTACACGATCACTAATTTTAAATTTTA TATATGTTGACTGAGATTGCAAATTTGAACATTGAAATTTTTTACCCGCTTGATCTGAATAATGACATTGTTCAAAAAAAGTACAAATGATAAAAAAG AAAGCTTCTCAAAAAACAGTAAAGAAGTTCTCCCCAAAATCACCTCAAAAATTCAGAGCTCTATTATCAGAAAAAGCGAGCTTAAAAAATTCAAAG GAAAGATACCCCTCTGCACCCTCAAAATTTTAGACCCTGTGTTGACCTGTAAAAATCAGGAAAAAATTTCCGTCGGTTATGGTATATGTGATGATT TCCCCTAATTATGCTGTAAGCACATGT**GAAGTTGAATTCACATTGAG**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA CTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGCCCTCAGTTCTCTTTTTCTTTTTCTTAAACTTCACGCACTGCACTTTTGTCCTCACTTTTTTCAT GCCGTCAGATTAACTACTTTTTCTATCCTTGAAATCAGCGGCTTTTCAGCCCTCATGTAGGCGCGCCGGCGATCGCGGCCGCTTAATTAA
MA0863 gRNA 2	TCAGCCCTCATGTAGTTTAAACCTGCAGGCGCGCCAACAACATCAGTCACCTAAAAAGAGAAAACGAATTACACGATCACTAATTTTAAATTTTAT ATATGTTGACTGAGATTGCAAATTTGAACATTGAAATTTTTTACCCGCTTGATCTGAATAATGACATTGTTCAAAAAAAGTACAAATGATAAAAAAGA AAGCTTCTCAAAAAACAGTAAAGAAGTTCTCCCCAAAATCACCTCAAAAATTCAGAGCTCTATTATCAGAAAAAGCGAGCTTAAAAAATTCAAAGG AAAGATACCCCTCTGCACCCTCAAAATTTTAGACCCTGTGTTGACCTGTAAAAATCAGGAAAAAATTTCCGTCGGTTATGGTATATGTGATGATTT CCCCTAATTATGCTGTAAGCACATGT**CGGGCACAACCATGCATGTA**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAA CTTGAAAAAGTGGCACCGAGTCGGTGCTTTTGCCCTCAGTTCTCTTTTTCTTTTTCTTAAACTTCACGCACTGCACTTTTGTCCTCACTTTTTTCAT GCCGTCAGATTAACTACTTTTTCTATCCTTGAAATCAGCGGCTTTTCAGCCCTCATGTAGGCGCGCCGGCGATCGCGGCCGCTTAATTAA

[a] guide RNA sequences in bold

Expression values were calculated as RPKM (Reads Per Kilobase per Million reads) using the upper quartile of gene expression, differential expression was determined using Rockhopper (McClure et al., 2013) comparing the expression of mutant strains against WT. Genes showing fold expression changes greater than 4, or less than -4, with q-values ≤ 0.01 were considered to be differentially expressed.

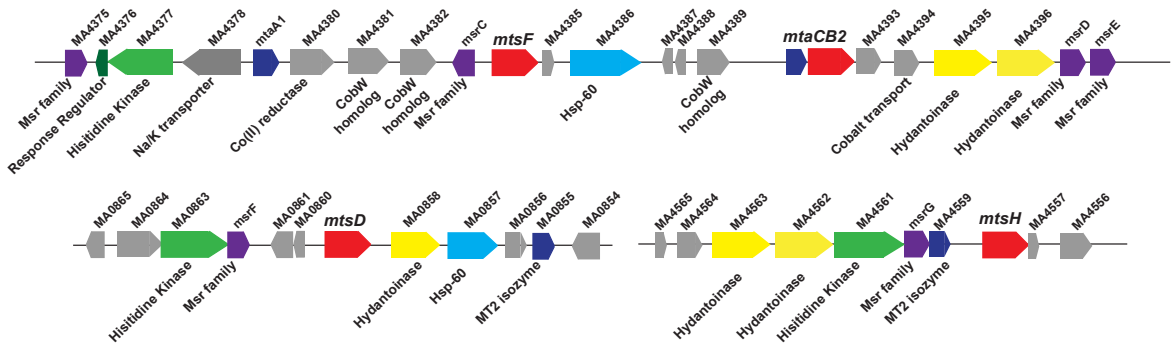

Figure 4.1 Schematic representation genomic region of the mtsD, mtsF and mtsH genes. Methyltransferase genes: red; Histidine kinase: green; Msr family genes: purple; Hsp60: light blue and hydantoinase: yellow.

4.4 Results

4.4.1 Bioinformatics analyses of putative histidine kinase and response regulators in *M. acetivorans* C2A

A of total of three histidine kinases (MA4377, MA4561 and MA0863), one response regulator (MA4376) and an uncharacterized Msr-like family protein (MA4375) were identified in the vicinity of the methylsulfide methyltransferases *mstD*, *mtsF* and *mtsH* (Fig 4.1). Of the three HKs analyzed only one (MA4377) has the characteristic cytoplasmic HisKA and HATPase domains (Fig 4.2). MA4377 also has a predicted extracellular sensory domain CHASE4, a membrane spanning Hamp domain and two intracellular receiver domains (Fig 4.2) all of which are common in prototypical membrane bound histidine kinases. Based on its domain composition, MA4377 is likely part of sensory transduction system that senses extracellular signals. In addition, MA4377 seems to be forming a two-component system with the putative response regulator MA4376. MA4376 encodes the only stand-alone receiver domain containing-protein in the

analyzed region. MA4375 on the other hand has a domain distribution similar to the methanol specific regulators described in Chapter 3; it contains a helix-turn-helix domain and the archaeal- exclusive DUF1724 domain.

Six domains were identified in the previously described MsmS, including three PAS domains, two GAF sensory domains and a HATPase domain (Fig 4.2). Previous proteomic analysis of WT *M. acetivorans* and a strain in which the pyrrolysine tRNA was deleted showed that MA0863 and MA0864 are likely not individual proteins, but rather a single protein that has an in frame pyrrolysine residue (O'Donoghue et al., 2014). This conclusion was due to the fact that MS/MS analysis identified MA0863 peptides beyond the UAG codon, which indicated the presence of Pyl in the protein (O'Donoghue et al., 2014). Analysis of the full-length gene predicted MA0863 to have an identical domain composition to MsmS (Fig 4.2). The absence of extracellular and intramembranous domains suggests that these two regulators are likely to be sensing and responding to intracellular conditions, in fact it has been shown that MsmS serves as an intracellular redox sensor (Molitor et al., 2013).

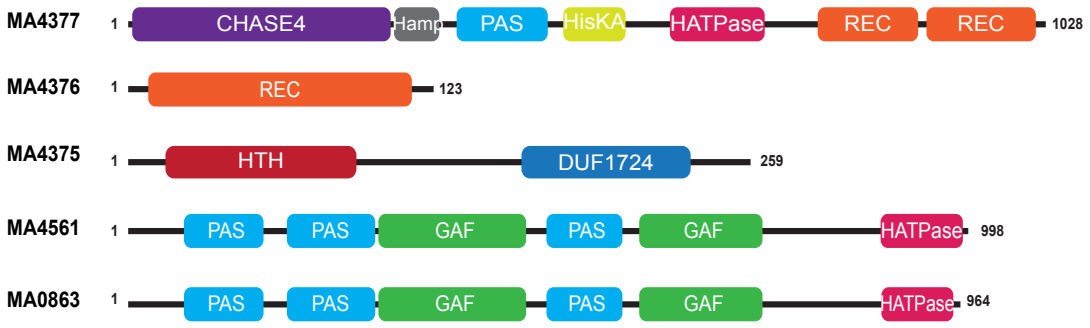

Figure 4.2 Schematic representation of the domain structures of HKs and RRs in this study. Abbreviations: CHASE4, sensory extracellular domain; Hamp, sensory domain, PAS, small molecule binding domain; HisKA, phospho-acceptor domain, HATPase; Rec, receiver domain, GAF, sensory domain; HTH, helix-turn-helix domain; DUF1724, domain of unknown function present only in archaea.

4.4.2 Regulation of methyl sulfide methyltransferases by the putative two-component system MA4376-MA4377

Based on their proximity to previously characterized methanol and methylsulfide-specific methyltransferases, we hypothesized that the putative two-component system encoded by MA4376-MA4377 might be involved in C-1 substrate responsive gene regulation. To test this, strains that carried deletions of these genes were constructed and characterized by global transcriptional profiling using RNAseq. Strains carrying the deletion of the putative response regulator MA4376 did not show differential regulation of any methyltransferase gene when grown on methanol + DMS. However, when grown on methanol ΔMA4376 strains showed an increase of mRNA abundance of *mstD* and *mtsF* methyltransferase genes (Table 4.5). In addition, the ΔMA4376 mutant overexpressed the MA0857 gene, which encodes a putative Hsp60 protein encoded in the vicinity of *mtsD*. Further, when grown on TMA, ΔMA4376 showed a decrease in mRNA abundance of *mtsD.* In contrast, the ΔMA4377 mutant showed an increase in mRNA abundance of *mtsF* when cells where grown on methanol or TMA. This mutation also caused a 23-fold increase in the expression of MA4386; another putative Hsp60 protein encoded in the vicinity of mtsF, and caused a decrease in the abundance of the MA4375 transcript, which encodes a putative Msr-family protein (Table 4.6). Taken together these results suggest that the putative two-component system MA4376-MA4377 is involved in the regulation of expression of *mtsF* when cells are using methanol as substrate and in the regulation of *mtsD* expression when cells are grown on TMA. To our surprise, neither MA4376 nor MA4377 affected the expression of *mtaCB2* or any other methanol specific methyltransferase, regardless of the growth substrate. Thus, MA4376-MA4377 are not involved in the regulation of methanol specific methyltransferases.

Because the Msr-family protein encoded by MA4375 is regulated by MA4376, we suspected that it might be a mediator in the signal transduction cascade. Thus we also constructed and characterized a strain with a deletion of this gene. Based on global transcriptional profiling, the ΔMA4375 mutation does not affect

any of the methyltransferase gene when the cells are grown on methanol + DMS. However, the mutant did exhibit an increase in *mtsD* mRNA when cells were grown on methanol and an increase of *mtsF* mRNA when cells were grown on TMA (Table 4.5).

4.4.3 Global transcriptional profiling of mutants lacking histidine-kinases encoded by MA0863 and MA4561

MA4561 was previously described as the methyl sulfide methyltransferase-associated sensor after deletion of MA4561 caused the constitutive synthesis of MtsF. However, the role of MA4561 in regulation of other genes was not assessed. To test this, we performed RNAseq experiments using an MA4561 mutant grown on a variety of media. Deletion of MA4561 caused a decrease of mRNA abundance of *mtsD* when cells were grown on TMA. Meanwhile deletion of MA4561 resulted in an increase in expression of *mtsF* and the putative Hsp60-encoded by MA4386. No significant regulation of methylotrophic methanogenesis genes was observed in ΔMA4561 strains grown on methanol or methanol+DMS.

Similar, RNAseq experiments showed that deletion of MA0863 caused an increase of mRNA abundance of *mtsD* when cells were grown on methanol or a combination of methanol and DMS (Table 4.5). Deletion of MA0863 also affected the expression of other genes present in the vicinity of the predicted histidine kinase, including the up-regulation of a hydantoinase encoded by MA0858 in methanol+DMS grown cells and the increase in mRNA abundance of the Hsp60 MA0857 when cells were grown on methanol (Table 4.6). Finally, during growth on TMA, ΔMA0863 strains showed an increase in expression of *mtsF*.

4.4.4 Differentially regulated antisense RNA

The Rockhopper software used to analyze our transcriptomic data also identified putative antisense RNAs whose expression is affected by the deletion of the HKs and RR in this study. Interestingly some of these differentially regulated predicted antisense RNAs are encoded within regions of the methyl sulfide methyltransferases or in genes that are in close proximity to them. For example,

deletion of the HK MA0863 resulted in a 53-fold increase in the abundance of an antisense RNA encoded within the *msrD* gene when cells were grown in a mixture of methanol and DMS. Another example is the over expression of several antisense RNAs within *mtsH* that are present when either MA4376 or MA4377 are deleted in TMA grown cells. More experiments are needed to determine what role if any these putative regulatory RNA have on this signal transduction system. A list of predicted antisense RNA is presented in Table 4.7.

4.5 Discussion

Work presented in Chapter 3 characterized the regulation of methyltransferase gene expression mediated by members if the Msr family of proteins, revealing that Msr proteins mediate gene regulation by directly binding the promoter of their target genes. However the sensory transduction pathway that triggers the regulation of methylatransferase genes remains unidentified. In this chapter, I identified a predicted two-component system (MA4376-MA4377) that we hypothesized would be involved in the regulation of the expression of the methyltransferases *mtsD* and *mtsF*. Our experiments support this model, but also showed that the regulatory network is more complex, including two additional HKs (MA0863 and MA4561) and a previously uncharacterized msr protein (MA4375).

Previous work identified the multisensory domain histidine kinase MA4561 as a heme-based sensor that senses and responds in changes to the cellular redox state (Molitor et al., 2013). This study also showed that mutants lacking MA4561 express MtsF constitutively during growth on methanol (Molitor et al., 2013). In contrast, our results showed that deletion of MA4561 did not affect the expression of *mtsF* in methanol grown cells. Similar differences in regulation of *mtsF* expression have been observed between mRNA levels and translational fusion data from methanol grown cells (Bose et al., 2009).Taken together these results suggest that members of the Mts system are regulated at both transcriptional and translational levels. One possibility is that MA4561 could

mediate *mtsF* regulation by affecting mRNA translation rather than expression, like the *Pseudomona aeruginosa* GacS/GacA system whose phosphorylation state affects the transcription of small RNAs that bind to specific mRNA targets inducing their degradation (Goodman et al., 2009). Although highly speculative, this scenario is consistent with our observation of antisense RNA encoded within *mts* genes. Further experiments are necessary to address this possibility.

Our *in silico* analysis showed that the pyrrolysine-containing MA0863 protein, encodes a putative histidine kinase with a domain distribution similar to that of MA4561. Previous work showed that deletion of $tRNA^{Pyl}$ caused cells to have a higher expression of MtsD, MtsF and MtsH (O'Donoghue et al., 2014). We suspect that the inability to translate the pyl codon in MA0863 is actually responsible for this phenotype. Consistent with this prediction, transcriptomic analysis of ΔMA0863 strains showed an increase in mRNA abundance of *mtsD* and *mtsF*. Experiments from our group have shown that the msr proteins MsrF

Table 4.5 mRNA abundance changes of methyl sulfide methyltransferases.

			Fold change Methanol[a]		
Gene	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**	**ΔMA0863**	**ΔMA4561**
mtsD	**10**	**14**	1	**13**	1
mtsF	3	**14**	**12**	**4**	1
			Fold change Methanol + DMS[a]		
Gene	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**	**ΔMA0863**	**ΔMA4561**
mtsD	1	2	**4**	**9**	1
mtsF	1	1	1	2	1
			Fold change TMA[a]		
Gene	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**	**ΔMA0863**	**ΔMA4561**
mtsD	1	**-8**	**-8**	2	**-5**
mtsF	**12**	2	**6**	**11**	**8**

[a]Fold change of expression values of mutant strains compared against WT *M. acetivorans* grown in the same substrate. Statistically significant values (q<0.01) are shown in shaded boxes.

and MsrC are the direct transcriptional regulators of *mtsD* and *mtsF* respectively. MA0863 is localized directly upstream of *msrF* so it is possible that MA0863 and *msrF* act as a two component system to mediate regulation of *mtsD* (Fig 4.1). To control transcription of *mtsF* it is possible that MA0863 is mediating regulation

through *msrC*. Complementation experiments need to be performed to confirm the role of MA0863 in the regulation of Mts proteins and to determine what role if any the pyrrolysine residue plays in MA0863 function.

Methanosarcina acetivorans encodes a total of 50 HKs and only 18 RRs. From the response regulators 11 form orphan proteins in the cell most of them without any DNA binding domain. Since there is not a 1:1 pairing of HKs and RRs, this suggests that each RR might be receiving signals from multiple histide kinases. Also, the lack of a DNA binding domain in these response regulators suggests that instead of directly binding to their target promoters they act by interacting with other DNA binding proteins possibly the previously characterized Msr proteins (Chapter 3). Our results are consistent with this model. For example, when cells are grown on methanol deletion of the RR MA4376 causes a 14-fold increase in mRNA abundance of *mtsD*. A similar increase is observed in ΔMA0863 (HK) and ΔMA4375 (msr) strains. Therefore it is possible that during growth on methanol MA0863 interacts with the RR MA4376 that in turn affects the function of the msr protein MA4375 to turn off the transcription of *mtsD*. Deletion of MA4376 also causes an increase in abundance of *mstF* in methanol. However, in this case a similar increase in mRNA abundance is present when MA4377 but not MA0863 is deleted in the strains (Table 4.5). These observations support the idea that in *Methanosarcina* a single RR can interact with multiple HKs. Once again these proposed pathways are highly speculative and further experiments are needed to confirm them.

In the presence of methanol all the characterized systems are acting as negative regulators of *mts* expression, this could be the reason why we do not see any regulatory effect upon deletion of these genes when cells are grown in the presence of DMS. Since expression of *mtsD*, *mtsF* and *mtsH* is essential for growth in the presence of DMS, it is possible that during growth on this substrate the HK kinase and RR studied have a phosphorylation state that prevents them from repressing *mts* genes expression. Additional experiments using mutated

versions of MA4377 and MA0863 (Table 4.1) would allow us to identify the effect that the different phosphorylation states of the histidine kinase has on the expression of the *mts* genes. No regulation pattern can be established from the TMA transcriptomic data, since deletion of all five genes affected the expression of *mstF* (Table 4.5). More experiments are necessary to understand the regulation of *mts* genes in TMA-grown cells. Other genes that have been shown to be regulated by four of the putative regulators in cells grown on TMA are the putative Hsp60 genes MA0857 and MA4377 (Table 4.6 and 4.7). MA0857 and MA4377 are homologs to the archaeal thermosome protein, a chaperonin that assists in the folding of nascent or denatured proteins (Klunker et al., 2003). Here we have proposed the Mts system is regulated at both transcriptional and translational levels. Hence, it is possible that these putative Hsp60 proteins are not responding to stress-induced situations in the cell but rather are necessary to aid in the regulation and folding of the Mts protein, explaining why they are regulated by the same system that controls *mts* gene expression.

Taken together, our results show that a complex regulatory network that includes multiple histidine kinases, crosstalk with response regulators and msr DNA binding proteins mediates expression of *mts* genes in *M. acetivorans*. Further experiments are necessary to determine autophosphorylation capacity of the HKs, how their phosphorylation states affect regulation of the *mts* genes and to determine the role other proteins like Hsp60 play in this regulatory cascade.

Table 4.6 Select mRNA abundance changes in ΔMA4375, ΔMA4376 and ΔMA4377.

		Fold change Methanol[a]			Fold change Methanol + DMS[a]			Fold change TMA[a]		
Locus tag	**Gene**	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**	**ΔMA4375**	**ΔMA4376**	**ΔMA4377**
MA0857	Hsp60	2	**4**	1	2	1	1	**92**	-2	**24**
MA0858	Hyd	2	2	1	2	1	2	1	-2	-2
MA4375	Msr	**-3**	1	**-4**	**-4**	1	**-8**	1	**-10**	-2
MA4386	Hsp60	2	1	1	2	1	1	**89**	-3	**23**

[a]Fold change of expression values of mutant strains compared against WT *M. acetivorans* grown in the same substrate. Statistically significant values (q≤0.01) are shown in shaded boxes.

Table 4.7 Select mRNA abundance changes in ΔMA0863 and ΔMA4561.

		Fold change Methanol[a]		Fold change Methanol + DMS[a]		Fold change TMA[a]	
Locus tag	**Gene**	**ΔMA0863**	**ΔMA4561**	**ΔMA0863**	**ΔMA4561**	**ΔMA0863**	**ΔMA4561**
MA0857	Hsp60	**5**	1	2	1	**83**	**66**
MA0858	Hyd	**3**	1	**6**	1	2	-2
MA4386	Hsp60	**4**	1	2	1	**81**	**72**

[a]Fold change of expression values of mutant strains compared against WT *M. acetivorans* grown in the same substrate. Statistically significant values (q≤0.01) are shown in shaded boxes.

Table 4.8 Select mRNA abundance changes of selected predicted RNA.

Fold change Methanol					
ΔMA4375		**ΔMA4376**		**ΔMA4377**	
Gene	**Fold change**	**Gene**	**Fold change**	**Gene**	**Fold change**
antisense: MA4384	76	antisense: MA0858/ MA0859	219	antisense: MA4558	13
antisense: MA4558	49	antisense: MA4558	39		

Fold change Methanol + DMS					
ΔMA4375		**ΔMA4377**		**ΔMA0863**	
Gene	**Fold change**	**Gene**	**Fold change**	**Gene**	**Fold change**
antisense: MA4385 MA4384	-5	antisense: MA0859	14	antisense: MA0859	53

Fold change TMA					
ΔMA4375		**ΔMA4376**		**ΔMA4377**	
Gene	**Fold change**	**Gene**	**Fold change**	**Gene**	**Fold change**
antisense: MA4558	17	antisense: MA4558	-5	antisense: MA4558	-4
		antisense: MA0859	-160	antisense: MA0859	-183

4.6 References

1. ASHBY, M. K. 2006. Distribution, structure and diversity of "bacterial" genes encoding two-component proteins in the Euryarchaeota. *Archaea,* 2**,** 11-30.

2. AUSUBEL, F. M. 1992. Short protocols in molecular biology : a compendium of methods from Current protocols in molecular biology*,* New York, NY, Greene Pub. Associates ; Wiley.

3. BOSE, A., KULKARNI, G. & METCALF, W. W. 2009. Regulation of putative methyl-sulphide methyltransferases in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 74**,** 227-38.

4. FU, H. & METCALF, W. W. 2015. Genetic basis for metabolism of methylated sulfur compounds in Methanosarcina species. *J Bacteriol,* 197**,** 1515-24.

5. GALAGAN, J. E., NUSBAUM, C., ROY, A., ENDRIZZI, M. G., MACDONALD, P., FITZHUGH, W., CALVO, S., ENGELS, R., SMIRNOV, S., ATNOOR, D., BROWN, A., ALLEN, N., NAYLOR, J., STANGE-THOMANN, N., DEARELLANO, K., JOHNSON, R., LINTON, L., MCEWAN, P., MCKERNAN, K., TALAMAS, J., TIRRELL, A., YE, W., ZIMMER, A., BARBER, R. D., CANN, I., GRAHAM, D. E., GRAHAME, D. A., GUSS, A. M., HEDDERICH, R., INGRAM-SMITH, C., KUETTNER, H. C., KRZYCKI, J. A., LEIGH, J. A., LI, W., LIU, J., MUKHOPADHYAY, B., REEVE, J. N., SMITH, K., SPRINGER, T. A., UMAYAM, L. A., WHITE, O., WHITE, R. H., CONWAY DE MACARIO, E., FERRY, J. G., JARRELL, K. F., JING, H., MACARIO, A. J., PAULSEN, I., PRITCHETT, M., SOWERS, K. R., SWANSON, R. V., ZINDER, S. H., LANDER, E., METCALF, W. W. & BIRREN, B. 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res,* 12**,** 532-42.

6. GOODMAN, A. L., MERIGHI, M., HYODO, M., VENTRE, I., FILLOUX, A. & LORY, S. 2009. Direct interaction between sensor kinase proteins mediates acute and chronic disease phenotypes in a bacterial pathogen. *Genes Dev,* 23**,** 249-59.

7. GUSS, A. M., ROTHER, M., ZHANG, J. K., KULKARNI, G. & METCALF, W. W. 2008a. New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for Methanosarcina species. *Archaea,* 2**,** 193-203.

8. GUSS, A. M., ROTHER, M., ZHANG, J. K., KULKARNI, G. & METCALF, W. W. 2008b. New methods for tightly regulated gene expression and highly efficient insertion of foreign genes for *Methanosarcina* species. *Archaea,* 2**,** 193-203.

9. KEARSE, M., MOIR, R., WILSON, A., STONES-HAVAS, S., CHEUNG, M., STURROCK, S., BUXTON, S., COOPER, A., MARKOWITZ, S., DURAN, C., THIERER, T., ASHTON, B., MEINTJES, P. & DRUMMOND, A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics,* 28**,** 1647-9.

10. KLUNKER, D., HAAS, B., HIRTREITER, A., FIGUEIREDO, L., NAYLOR, D. J., PFEIFER, G., MULLER, V., DEPPENMEIER, U., GOTTSCHALK, G., HARTL, F. U. & HAYER-HARTL, M. 2003. Coexistence of group I and group II chaperonins in the archaeon *Methanosarcina mazei*. *J Biol Chem,* 278**,** 33256-67.

11. LIANG, J., CHAO, R., ABIL, Z., BAO, Z. & ZHAO, H. 2014. FairyTALE: a high-throughput TAL effector synthesis platform. *ACS Synth Biol,* 3**,** 67-73.

12. MCCLURE, R., BALASUBRAMANIAN, D., SUN, Y., BOBROVSKYY, M., SUMBY, P., GENCO, C. A., VANDERPOOL, C. K. & TJADEN, B. 2013. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res,* 41**,** e140.

13. METCALF, W. W., ZHANG, J. K., APOLINARIO, E., SOWERS, K. R. & WOLFE, R. S. 1997. A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors. *Proc Natl Acad Sci U S A,* 94**,** 2626-31.

14. MOLITOR, B., STASSEN, M., MODI, A., EL-MASHTOLY, S. F., LAURICH, C., LUBITZ, W., DAWSON, J. H., ROTHER, M. & FRANKENBERG-DINKEL, N. 2013. A heme-based redox sensor in the methanogenic archaeon *Methanosarcina acetivorans*. *J Biol Chem,* 288**,** 18458-72.

15. NAYAK, W. W. M. D. N. 2017 Cas9-mediated genome editing in the methanogenic archaeon *Methanosarcina acetivorans Proc Natl Acad Sci U S A*.

16. O'DONOGHUE, P., PRAT, L., KUCKLICK, M., SCHAFER, J. G., RIEDEL, K., RINEHART, J., SOLL, D. & HEINEMANN, I. U. 2014. Reducing the genetic code induces massive rearrangement of the proteome. *Proc Natl Acad Sci U S A,* 111**,** 17206-11.

17. PRITCHETT, M. A., ZHANG, J. K. & METCALF, W. W. 2004. Development of a markerless genetic exchange method for *Methanosarcina acetivorans* C2A and its use in construction of new genetic tools for methanogenic archaea. *Appl Environ Microbiol,* 70**,** 1425-33.

18. SALVADO, B., VILAPRINYO, E., SORRIBAS, A. & ALVES, R. 2015. A survey of HK, HPt, and RR domains and their organization in two-component systems and phosphorelay proteins of organisms with fully sequenced genomes. *PeerJ,* 3**,** e1183.

19. TJADEN, B. 2015. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol,* 16**,** 1.

20. WILLIAM METCALF, J. K. Z., XUN SHI, RAPLH S. WOLFE 1996. Molecular, Genetic, and Biochemical Characterization of the serC Gene of *Methanosarcina barkeri* Fusaro. *J Bacteriol,* 178**,** 5797-5802.

21. WILLIAM METCALF, J. K. Z., XUN SHI, RAPLH S. WOLFE 1998. An Anaerobic, Intrachmber Incubator for Growth of Methanosarcina spp. on Methanol-Containing Solid Media. *J Bacteriol,* 64**,** 768-770.

Chapter 5: Conclusions

5.1 Summary

5.1.1 PEP carboxylase is an alternative route of OAA biosynthesis in *M. barkeri*

Methanosarcina barkeri genome encodes two oxaloacetate (OAA) generating enzymes: pyruvate (Pyc) and phosphoenolpyruvate (Ppc) carboxylase. Until this work it was believed that OAA synthesis in *M. barkeri* proceeded exclusively from Pyc. However, in Chapter 2 we demonstrated that introduction of a stop codon in the *pyc* operon resulted in cells that lacked Pyc activity, but showed no growth phenotype, suggesting an alternative route for synthetizing OAA in the cells. Previously reported in vitro biochemical data show that *M. barkeri* cells lacked Ppc activity even in the absence of Pyc (Mukhopadhyay et al., 2001), but RNAseq analysis showed an overexpression of the *ppc* operon in cells that lacked Pyc activity. Our results showed that *M. barkeri* is capable of expressing two independent enzymes to synthetize the essential physiological component OAA.

5.1.2 Por is an essential gene in *M. barkeri*

Pyruvate ferredoxin oxidoreductase (Por) catalyzes the reversible carboxylation of acetyl-CoA to pyruvate, a step that is crucial for carbon fixation in methanogens. Experiments using cells that expressed *por* under a tetracycline inducible promoter revealed that *por* is essential in *M. barkeri*. Moreover, cells could not be rescued by the addition of pyruvate, casamino acids, yeast extract or acetate. Taken together these results suggest that in *Methanosarcina spp*. Por has an additional role beyond synthesis of pyruvate. Some possibilities include the expression of additional activities like hydrogenase or pyruvate carboxylase both observed in other organisms (Eram et al., 2014, Ragsdale, 1996) and the utilization of one of the *porCDAB* subunits by another enzymatic complex.

5.1.3 A proposed pathway for pyruvate utilization in the *M. barkeri* Fusaro pyruvate-utilizing strain

Experiments presented in Chapter 2 allowed the proposal of a model for pyruvate utilization in mutant *M. barkeri* strains. In this model pyruvate utilization requires the overexpression of pyruvate ferredoxin oxidoreductase and phosphoenolpyruvate carboxylase mediated by the potential regulator encoded in the Mbar_A2165 locus. The overexpression of *ppc* and *por* is coupled to the loss of activity of pyruvate carboxylase. These changes in central carbon metabolism allow the channeling of pyruvate into the oxidative pathway, which allows the growth of cells on pyruvate using enzymes involved in aceticlastic methanogenesis (Fig. 2.5).

5.1.4 MsrAB bind to the *mtaCB1* promoter as heterodimers

Previous work had identified the methanol specific regulators MsrA and MsrB as the factors or proteins responsible for regulating the transcription of the methanol specific methyltransferase *mtaCB1*(Bose and Metcalf, 2008). Here, we describe a series of biochemical experiments that show that MsrAB regulates expression of *mtaCB1* by directly binding to its promoter region. Moreover, we have shown that MsrA and MsrB form a heterodimeric complex in a 1:1 ratio and that formation of this complex is required for binding since neither protein can bind the promoter region by itself. Heterodimer formation potentially allows these proteins to expand their regulatory capacities by binding to multiple binding partners to regulate different genes. To the best of our knowledge, this is the first report of heterodimeric regulators in methanogenic archaea.

5.1.5 MsrB acts as both activator and repressor of transcription

Experiments in Chapter 3 showed that in addition to activating expression of methanol methyltransferases, MsrB seems to be mediating the repression of methylamine methyltransferase genes in the presence of methanol or acetate. A comparison of mRNA abundance in Δ*msrA* and Δ*msrB* strains revealed that the role of *msrB* in regulation of *mtt*, *mtm* and *mtb* genes is independent from *msrA*.

Finally, complementation experiments corroborated the role of *msrB* in the regulation of methylamine methyltransferases. These results suggest that similar to the previously described *Methanosarcina* regulator MreA (Reichlen et al., 2012), MsrB acts as both co-activator and a repressor of transcription in the cells. The mechanisms by which this DNA binding protein is able to identify the presence or absence of a substrate to bind and regulate methyltransferase transcription remain undetermined.

5.1.6 Regulation of methyl sulfide methyltransferases by putative signal transduction systems

Methanosarcina acetivorans C2A has three methyl sulfide methyltransferases: *mtsD*, *mtsF* and *mtsH*. Previous experiments have shown that regulation of their transcription is mediated by the DNA binding proteins MsrF, MsrC and MsrG respectively (Bose et al., 2009). The Mts proteins have been shown to be undetectable in the presence of methanol (Bose et al., 2009, Oelgeschlager and Rother, 2009). Here, we have shown that regulation of the expression of *mts* genes in the presence of methanol involves the expression of at least two independent histidine kinases (MA0863 and MA4377), a response regulator (MA4376) and a fourth *msr* protein (MA4375). Each of these putative signal transduction systems seems to be acting as a negative regulators of *mts* genes, because deletion of any resulted in an increase in the levels of mRNA of either *mtsD* or *mtsF* (Table 4.5). The work presented in this dissertation provides insight into the complexity of regulatory network that controls expression of methyltransferases in *M. acetivorans*. Future experiments using complemented strains and phosphorylation assays will further characterize the signal transduction pathway that governs *mts* expression.

5.2 Future Directions

5.2.1 Determination of the relationship between Pyc and Por in pyruvate metabolism

Experiments presented in Chapter 2 showed that Por was no longer essential in strains that lacked Pyc activity. In addition, lack of Pyc activity seems to be related to *por* overexpression in Pyr^+ strains. These results suggest that there is a different route to synthetize pyruvate and OAA that is induced in the absence of both Pyc and Por. To test for this possibility we could use transcriptomic experiments by RNAseq on strains that lack both activities. This would help identify the genes that are differentially regulated in these strains and determine what other genes might be involved in the synthesis of pyruvate in the absence of *por* and *pyc*. Since *por* is an essential gene in *M. barkeri* deletion of it was not possible. However, if lack of Pyc activity were indeed required to abolish *por* essentiality then construction of a double deletion mutant would be possible, thus confirming the relationship between *por* and *pyc.* In addition, there is the possibility that in the cells Pyc and Por are working in close proximity channeling substrates resulting in their rapid diffusion through the complexes. To explore this possibility we could perform a pull down experiment using affinity tagged proteins to determine if there is complex formation.

5.2.2 Determination of the role the identified mutations have on suppression of Por^- essentiality

Genome sequencing of a single Por^- suppressor strain revealed six mutations that included point mutations, deletions and insertions both in genes and intergenic regions of the mutant genome (Table 2.6). All six mutations are present in regions of hypothetical proteins thus, to identify what role if any they play in pyruvate metabolism the mutations would need to be recreated in a WT background. Special interest relies on the mutation present in Mbar_A3314 (deletion of 2 bp that causes a frame shift) since it is present in a gene that has a predicted helix-turn-helix domain.

5.2.3 Regulation of methylamine methyltransferases by MsrB

In Chapter 3 we presented data that supports the idea that the methanol specific regulator MsrB is acting as both repressor and activator the methyltransferase genes. We have shown that in order to activate transcription of methanol methyltransferases genes MsrB must form a complex with MsrA. Repression of the expression of methylamine methyltransferases genes by MsrB remains understudied; MsrB could be regulating these genes by binding to the promoters by itself or by partnering with other Msr proteins. To test these ideas two experiments could be perform. The first would be to assess the binding of MsrB by itself to the promoter regions of the methylamine methyltransferases genes by EMSA experiments. This would allow us to identify if MsrB is binding alone to repress gene expression. If MsrB by itself is unable to bind the promoter region, a second experiment could be perform to identify potential binding partners of MsrB. In Chapter 3 we have used FLAG tag proteins to perform western blots. Since we know that this tag is not disrupting MsrB function, and that MsrB concentration in the cell is estimated to be low, we could make strains that have MsrB tagged with a FLAG/streptavidin tag to perform a tandem purification of the protein. Samples could then be sent for mass spectrometry analysis to identify which other proteins if any appear to be bound to MsrB.

5.2.4 Identification of potential posttranslational modifications and metals bound to MsrA and MsrB

The EMSA experiments shown in Chapter 3 demonstrated that addition of metals enhanced biding of MsrAB to its target promoter. Moreover, the three-dimensional prediction of the protein structures identified a putative Zn binding site in both MsrA and MsrB. To identify which metal is bound to the proteins we could purify proteins from the native host (using the tandem approach suggested above) and submit them for Inductively coupled plasma mass spectrometry (ICP-MS). Our experiments have shown that *msrA* and *msrB* are always transcribed and translated regardless of the substrate in which the strains are grown on, so

regulation of their activity must be mediated by another route. To test the possibility that post-translational modifications are mediating MsrA and MsrB regulation we could purify the proteins from the native host when cells are grown with methanol, TMA and acetate and perform mass spectrometry to identify and compare any covalent modification of the two proteins in different substrates. This would allow us to predict how Msr proteins sense and respond to substrate availability in the cells and might hint at a possible role for the domain of unidentified function DUF1724 at the C-terminal end of the proteins.

5.2.5 Characterization of two component systems in *M. acetivorans* C2A

In Chapter 4 we described the effect that multiple histidine kinases (HK) and response regulators (RR) have in the expression of methyl sulfide methyltransferase genes (*mts*). To confirm the effect that deletions of the putative HKs and RR have in the expression of the Mts system we have complemented the deletion strains with WT copies of the HKs and RR genes (Table 4.1). qPCR analysis of *mts* gene expression in the complemented strains will allow us to confirm if in fact their transcription is controlled by these putative two-component systems. In addition, we have complemented ΔMA4377 with copies of the gene with key histidine and aspartate residues mutated to either glutamate or alanine to mimic phosphorylation states of the histidine kinase (Table 4.1). We expect that the His-Ala change would mimic a histidine kinase that is always de-phosphorylated or whose function is always off. Changing Asp residues to Glu in the other hand would mimic a system in which the receiver domain of the HK is always phosphorylated and potentially could result in the constitutive expression of the genes regulated by the system. Epistasis analysis using strains containing mutated His and Asp residues would allow the characterization of the signal transduction pathway that MA4377 regulates. We have also identified a pyrrolysine containing HK (MA0863), to investigate the role of the Pyl residue in MA0863 function; we mutated the residue to lysine (Table 4.1). To test the effect this mutation has on the cells we could preform a qPCR probing for the expression of the *mts* genes. Finally we would like to assay the kinase and

phosphatase activities of the putative HKs and their ability to transfer phosphate groups to the proposed response regulator MA4376.

5.3 References

1. BOSE, A., KULKARNI, G. & METCALF, W. W. 2009. Regulation of putative methyl-sulphide methyltransferases in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 74**,** 227-38.

2. BOSE, A. & METCALF, W. W. 2008. Distinct regulators control the expression of methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Mol Microbiol,* 67**,** 649-61.

3. ERAM, M. S., ODUARAN, E. & MA, K. 2014. The bifunctional pyruvate decarboxylase/pyruvate ferredoxin oxidoreductase from *Thermococcus guaymasensis*. *Archaea,* 2014**,** 349379.

4. MUKHOPADHYAY, B., PURWANTINI, E., KREDER, C. L. & WOLFE, R. S. 2001. Oxaloacetate synthesis in the methanarchaeon *Methanosarcina barkeri*: pyruvate carboxylase genes and a putative *Escherichia coli*-type bifunctional biotin protein ligase gene (bpl/birA) exhibit a unique organization. *J Bacteriol,* 183**,** 3804-10.

5. OELGESCHLAGER, E. & ROTHER, M. 2009. In vivo role of three fused corrinoid/methyl transfer proteins in *Methanosarcina acetivorans*. *Mol Microbiol,* 72**,** 1260-72.

6. RAGSDALE, S. M. A. S. W. 1996. Unleashing hydrogenase activity in carbon monoxide dehydrogenase/acetyl-CoA synthase and pyruvate:ferredoxin oxidoreductase *Biochemistry,* 35**,** 15814-15821.

7. REICHLEN, M. J., VEPACHEDU, V. R., MURAKAMI, K. S. & FERRY, J. G. 2012. MreA functions in the global regulation of methanogenic pathways in *Methanosarcina acetivorans*. *MBio,* 3**,** e00189-12.