UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SUBSTITUTIONAL ANALYSIS OF CYSTEINE RESIDUES AND PRELIMINARY REGULON ANALYSIS OF THE REDOX-SENSING TRANSCRIPTION REGULATOR MSVR IN *METHANOSARCINA ACETIVORANS*

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

KYLIE A. GILCREST
Norman, Oklahoma
2016

SUBSTITUTIONAL ANALYSIS OF CYSTEINE RESIDUES AND PRELIMINARY REGULON ANALYSIS OF THE REDOX-SENSING TRANSCRIPTION REGULATOR MSVR IN *METHANOSARCINA ACETIVORANS*

A THESIS APPROVED FOR THE
DEPARTMENT OF MICROBIOLOGY AND PLANT BIOLOGY

BY

_____
Dr. Elizabeth A. Karr, Chair

_____
Dr. Amy V. Callaghan

_____
Dr. Tyrrell Conway

Dedicated to

My loving parents and my best friends, David and Erin, for all their support

Acknowledgements

Research reported in this thesis was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20GM103640. I would like to acknowledge my committee chair, Dr. Elizabeth Karr, for intellectual contributions to this thesis and for mentoring me through graduate school. I'd also like to thank lab members Dr. Cat Isom, Dr. Vy Trinh, Kristen Shelton and Melissa Chanderban for general contributions and for teaching me many of the techniques utilized to conduct research reported in this thesis.

Table of Contents

List of Tables

List of Figures

Abstract

Methane produced from methanogenesis in archaea may be a promising source of renewable energy as methanogens are the primary producers of biogenic methane. However, methanogens are anaerobic and experience transient oxygen exposure in their environment, leading to degradation of proteins involved in methanogenesis pathways. Elucidating the oxidative stress response mechanisms in methanogens may lead to the development of more oxygen-tolerant strains with robust methanogenesis capabilities in bioreactors for production of methane as a renewable energy source. MsvR is a redox-sensitive transcription regulator exclusively found in methanogens. It was initially identified in *M. thermautotrophicus* as a regulator of the *fpaA-rlp-rub* operon, which is implicated in the oxidative stress response. The MsvR homolog in *M. acetivorans* (MaMsvR) was found to have ten cysteine residues, three of which are conserved in MsvR from *M. thermautotrophicus*. This lead to the substitutional analysis of each cysteine in MaMsvR that implicated seven cysteines (C63, C89, C148, C175, C206, C232 and C240) in redox sensing and/or conformational change affecting MaMsvR binding to its promoter (Ma P_{msvR}). RNA-seq was utilized for preliminary analysis of the MaMsvR regulon and it was found that MaMsvR directly or indirectly regulates genes encoding function in methanogenesis, signal transduction, transcription regulation and various other cellular functions. The data suggests MaMsvR is involved in regulatory hierarchies that modulate central metabolism and that regulation may be redox-dependent.

Chapter 1: Introduction

Climate change is one of the most important yet challenging issues facing humanity today. Greenhouse gases absorb heat emitted from the earth and trap it in the atmosphere. Since the industrial revolution, the amount of greenhouse gases emitted from human activity has drastically increased and is slowly warming the planet (1). The two most potent contributors to global warming are the greenhouse gases methane and carbon dioxide (CO_2), with methane having a global warming potential that is 25 times greater than carbon dioxide over a timescale of 100 years because it is more efficient at absorbing radiation than CO_2 (1, 2). Reducing these greenhouse gas emissions is imperative to mitigating the consequences of climate change. Methanogenic archaea are the primary source of biogenic methane produced from the metabolism of one-carbon compounds and/or CO_2 and hydrogen in a process called methanogenesis (3).

These methanogenic archaea, also referred to as methanogens, are anaerobic organisms that can only perform methanogenesis under anoxic conditions. Exposure of methanogens to oxygen leads to the generation of reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2) and superoxide that target and oxidize cellular components (4). These ROS can damage DNA, oxygen-sensitive proteins such as those involved in methanogenesis, and other molecules in the cell. Methanogens are transiently exposed to oxygen in their environment but can continue methanogenesis after oxidant exposure, which suggests they have mechanisms in place to protect oxygen-sensitive proteins from damage during oxidative stress (5). Understanding how methanogens combat oxidative stress is central to combating global warming by the

development of more oxygen-tolerant species in bioreactors for the production of methane as a renewable energy source.

A promising lead to elucidating how methanogens survive oxidative stress came from the differential expression of the *fpaA-rlp-rub* operon in *Methanothermobacter thermautotrophicus* through microarray analysis (6). The operon includes genes encoding a flavoprotein (*fpaA*), rubrerythrin-like protein (*rlp*) and a rubredoxin (*rub*) and is located downstream of *msvR*, a gene that is divergently transcribed from the operon (7). The *fpaA-rlp-rub* operon was found to be upregulated along with genes encoding for thioredoxin (Trx) and thioredoxin reductase (TrxR) after hydrogen peroxide (H_2O_2) exposure. This indicates that *M. thermautotrophicus* utilizes these enzymes to protect intracellular molecules from oxidation by ROS instead of more common antioxidant enzymes like superoxide dismutase or catalase (7).

The basal transcriptional machinery in *Archaea* is a simplistic version of the eukaryotic transcription system, although the gene regulatory mechanisms in *Archaea* are more similar to bacterial regulators (8-10). Archaeal transcriptional machinery consists of a transcription factor II B (TFB) homolog, a TATA-binding protein (TBP) and RNA polymerase (RNAP) that resembles RNAP II in eukaryotes (11, 12). The RNAP in Archaea have 11-13 subunits, depending on the species, and share a closer homology to eukaryotic RNAPII than to bacterial RNAP (13, 14). Archaeal transcription is initiated when TBP binds to the TATA box, a thymine and adenine-rich sequence located approximately 25 base pairs upstream of the transcriptional start site (15). TBP then recruits TFB to the promoter, where it binds to the purine-rich B-

recognition element (BRE) located immediately upstream of the TATA box (16). RNAP is then recruited to the promoter to complete the pre-initiation complex (12).

Although the archaeal basal transcriptional machinery is more similar to that in eukaryotes, the regulatory mechanisms and transcription regulators are more similar to bacterial regulators. In bacteria, a gene may be regulated by a single activator that binds upstream of the promoter region, or by a repressor that binds close to the promoter and directly blocks RNAP and transcription factors from binding to the promoter (17). Archaeal transcription regulators, which are predominantly repressors, function similarly to bacterial repressors and share the characteristic helix-turn-helix (HTH) domain that binds DNA. Archaeal repressors can either bind to the promoter to prevent recruitment of TFB and TBP to the pre-initiation complex or they can compete with RNAP for the transcriptional start site (10). Regulators that are involved in the oxidative stress response must be able to sense a redox shift within the cell. When ROS are present in the cell, a regulator senses the redox shift and either binds to its promoter region or releases itself from the DNA in order to activate or repress genes involved in the oxidative stress response. Redox sensing bacterial regulators such as OxyR, PerR and RexT have been studied more extensively than archaeal regulators and can provide insights into the redox-sensing mechanisms in archaeal regulators.

Numerous studies have shown that cysteine residues play a major sensory role in redox-sensing transcription regulators (18-20). Cysteine residues contain thiol groups that can react with oxygen-containing compounds to form disulfide bonds (21, 22). Formation of these disulfide bonds can cause a conformation change in the protein that alters its ability to bind DNA (23). OxyR is a tetrameric transcription regulator that was

originally characterized in *Escherichia coli* (24). H_2O_2 oxidizes one cysteine residue in OxyR and causes it to quickly form an intramolecular disulfide bond with a second cysteine (25). The resulting conformational change allows OxyR to interact with RNAP to activate or repress expression of genes in the OxyR regulon, which includes antioxidant gene products that combat ROS (26). The OxyR regulon also includes genes encoding glutathione and glutathione reductase 1, which enzymatically deactivate OxyR by reducing the disulfides back to free thiols (25). This autoregulation ensures that OxyR is only activated during oxidative stress.

RexT is a redox-sensing transcription regulator of the ArsR family that is involved in the oxidative stress response. It was first described in the cyanobacterium *Anabaena* sp. and was the first reported transcriptional regulator of the thioredoxin (*trx)* gene in cyanobacteria (27). Under normal growth conditions, RexT is bound to the promoter region of the *trxA2* gene. During oxidative stress, H_2O_2 inhibits the DNA binding activity of RexT by oxidizing two of the three cysteine residues to form a disulfide bond. RexT is released from the promoter and *trxA2* is derepressed. TrxA2 reduces RexT back to its DNA binding state to repress the *trxA2* gene once oxidant has been removed from the cell. Thus, similar to OxyR, RexT senses oxidant via redox-sensitive cysteines and induces expression of Trx to prevent damage in the cell (27).

PerR is a ferric uptake repressor homolog involved in the H_2O_2 stress response in *Bacillus subtilis* (28). Its regulon includes genes encoding catalase and alkylhydroperoxide reductase, both of which are involved in the detoxification of H_2O_2 (29). Although the typical mechanism for H_2O_2 sensing is through cysteine residues as described for OxyR and RexT, PerR senses H_2O_2 via metal-catalyzed oxidation of

histidine residues (29). PerR is a homodimer with two metal ions per subunit. Four cysteines coordinate zinc(II) and three histidines and two aspartates coordinate iron(II) (30). The incorporated iron(II) interacts with H_2O_2 to form iron(III), hydroxide ion and hydroxyl radical via the Fenton reaction ($H_2O_2 + Fe^{2+} \rightarrow Fe^{3+} + HO^- + {}^{\bullet}OH$). The hydroxyl radical then interacts with one of the histidines involved in coordinating the iron ion by incorporating oxygen into histidine to form a 2-oxo-histidine residue in each subunit (30). This induces a conformational change in PerR that releases it from DNA and derepresses its regulon for the detoxification of peroxides. PerR also represses the uptake of iron when H_2O_2 is present in the cell because iron can react with H_2O_2 to form a hydroxyl radical (25). It is unknown if 2-oxo-histidine can be reverted back to histidine, insinuating that PerR may be a sacrificial regulator during oxidative stress and cannot be regenerated following metal-catalyzed oxidation of histidine (30).

To date only three redox-sensing archaeal regulators are known: SurR, RosR and MsvR, the latter of which is the focus of this thesis. The first redox-sensing transcription regulator that was discovered in an archaeon was the sulfur response regulator SurR, which was found in the hyperthermophilic anaerobe *Pyrococcus furiosus* (31). This organism normally ferments carbohydrates to organic acids, CO_2 and hydrogen (H_2). When elemental sulfur (S^0) is present, metabolism in the cell switches from the production of H_2 to the production of hydrogen sulfide (H_2S) within minutes (32). SurR is a homodimer with each monomer containing N- and C-terminal winged HTH DNA binding domains and two cysteine residues in the motif CXXC (33). When S^0 is not present, SurR is in its reduced state and binds to multiple promoters in its regulon. It activates hydrogenases and other genes involved in H_2 production by

recruiting the basal transcriptional machinery to the promoter. SurR also represses transcription of numerous sulfur response genes by blocking the transcriptional machinery from being able to access the promoter. Upon introduction of S^0, the two cysteines in the CXXC motif in SurR become oxidized and form a disulfide bond. This causes a conformational change in the protein that releases it from the DNA, thus activating transcription of the sulfur response genes and causing the cell to shift from H_2 to H_2S production. As S^0 becomes depleted in the cell, SurR is converted back to its reduced state and H_2 production resumes (33).

The reactive oxygen species regulator RosR is unique to a small clade of halophilic archaea and was initially found in the extreme halophile *Halobacterium salinarum* (34). RosR has a central winged HTH domain flanked by effector domains at each terminus. It is a bifunctional regulator that plays an important role in a large regulatory network that enables *H. salinarum* to survive extreme oxidative stress by quickly responding to the presence of ROS and returning the cell to homeostasis. RosR directly regulates over 100 genes with putative functions in transcription, stress response, translation and other cellular functions. A structure for RosR has yet to be elucidated and no cysteines are present in the protein, so it is currently unknown how or if RosR senses oxidant directly and how the regulatory cascade in response to ROS is initiated (35).

The third redox-sensing archaeal transcription factor, MsvR, was initially identified in the *M. thermautotrophicus* genome where it is located upstream and divergently transcribed from the *fpaA-rlp-rub* operon (6). MsvR is a methanogen-specific homodimer with an ArsR family winged HTH DNA binding domain and a C-

terminal vinyl-4-reductase domain (36). An *msvR* homolog in *Methanosarcina acetivorans* (Ma *msvR*) is flanked by two genes encoding uncharacterized proteins and is currently only known to bind its own promoter (Figure 1B). *M. thermautotrophicus* MsvR (MthMsvR) contains five cysteines in the V4R domain, two of which are part of a CX_2CX_3H motif characteristic of some metal-binding proteins involved in redox-dependent transcription regulation (5). *M. acetivorans* MsvR (MaMsvR) contains ten cysteines, three of which are conserved in MthMsvR (Figure 2). It is postulated that some cysteine residues in each MsvR are involved in redox sensing and conformational change similar to the redox sensing mechanisms of OxyR. MsvR binds to its promoter in reduced conditions and after exposure to oxidant, it is believed that key cysteine residues are oxidized and form disulfide bond(s). This may induce a conformational change that releases MsvR from its promoter. In *M. thermautotrophicus*, release of MthMsvR from its promoter allows for expression of the *fpaA-rlp-rub* operon that is hypothesized to be involved in the oxidative stress response (37). This study aims to determine cysteine residues in MaMsvR involved in disulfide bond formation through site-directed mutagenesis of each cysteine and to identify genes in *M. acetivorans* that may be regulated by MaMsvR through RNA sequencing (RNA-seq).

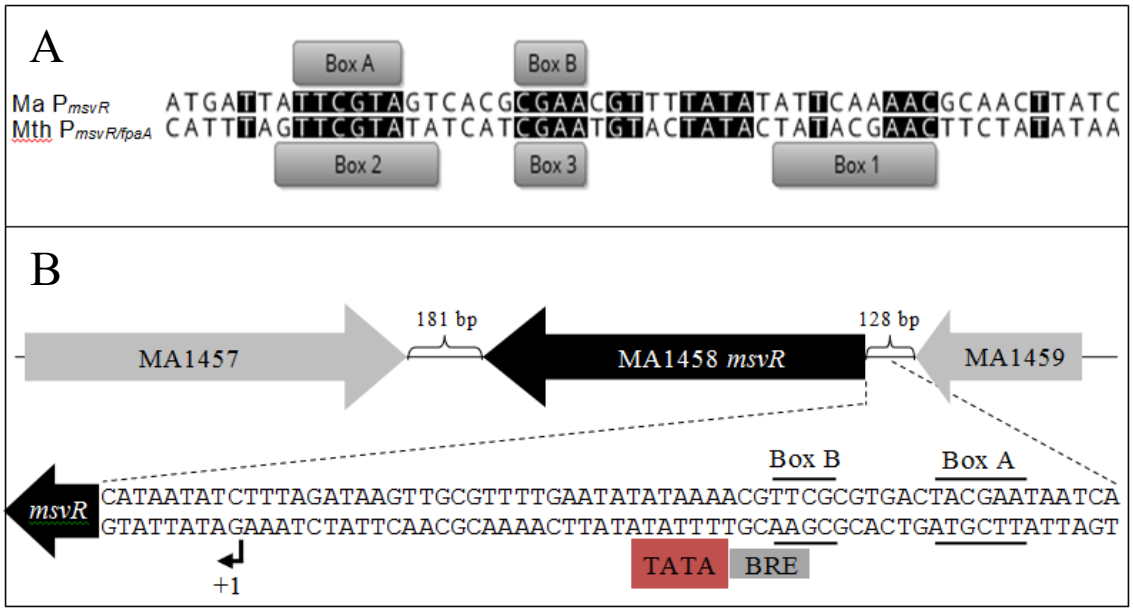


Figure 1. Alignment of MsvR binding boxes and genomic context for Ma *msvR*. This figure has been adapted from a previous study (5). **(A)** Alignment of binding boxes on Ma P_{msvR} to those on Mth $P_{msvR/fpaA}$. Gray boxes indicate MsvR binding boxes A and B on Ma P_{msvR} and binding boxes 1, 2 and 3 on Mth $P_{msvR/fpaA}$. Conserved nucleotides are shaded in black. **(B)** Genomic context for Ma *msvR*. Ma *msvR* is flanked by MA1457 and MA1459, with black brackets indicating the length of each intergenic region (181 bp and 128 bp). Arrows represent the direction of transcription for each gene. Dashed lines indicate the placement of the intergenic region just upstream of Ma *msvR* that is zoomed into below. MaMsvR binding boxes A and B are represented by solid black lines on each side the nucleotide sequence. The red and gray boxes represent the TATA box and BRE, respectively. The bent arrow and +1 indicate the transcriptional start site for Ma *msvR*, represented as a black arrow.

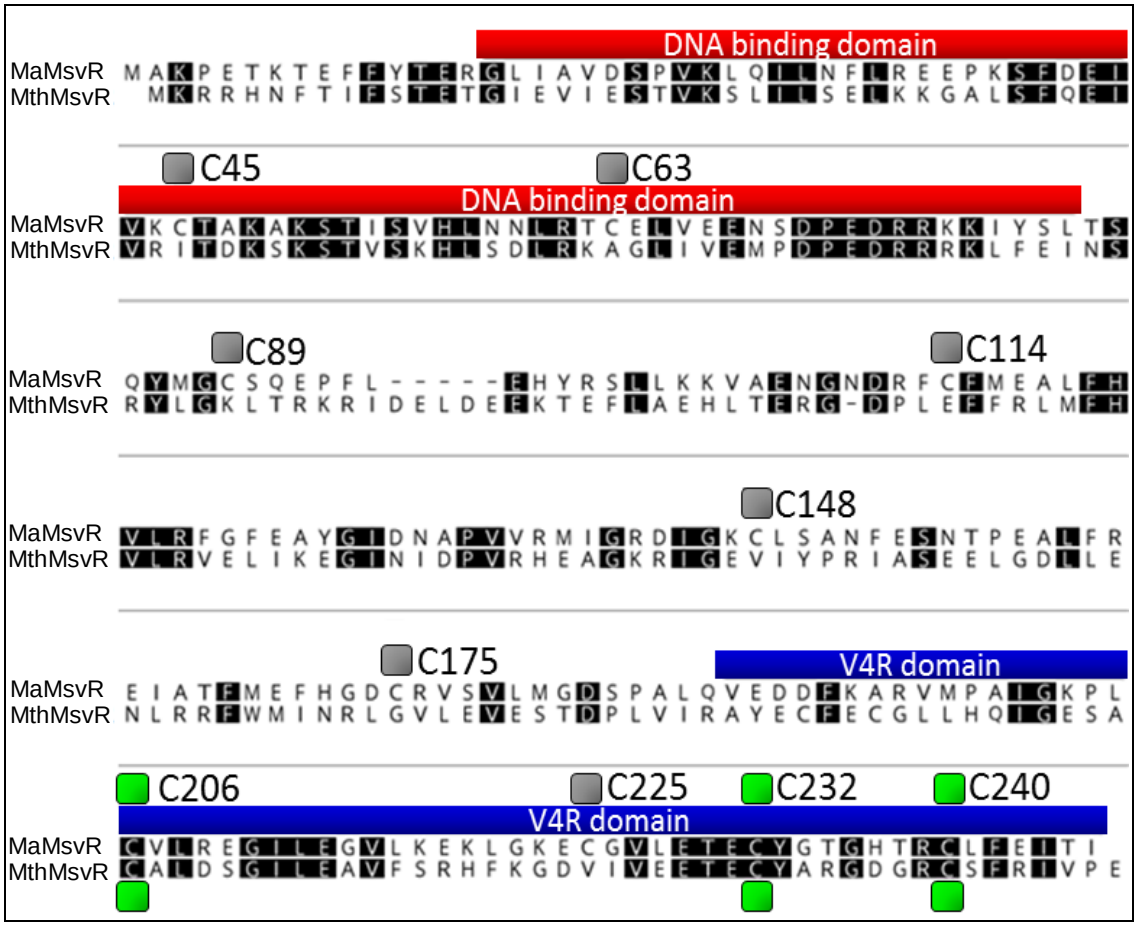


Figure 2. Amino acid alignment of MaMsvR and MthMsvR.

This figure has been adapted from a previous study (5). Geneious v8 was used to generate the alignment. The DNA binding domain and V4R domain are represented by red and blue boxes, respectively, over residues in each domain. Conserved cysteines are indicated with green boxes and non-conserved cysteines in MaMsvR are indicated by gray boxes, with the identifier for each cysteine displayed next to each box. Conserved residues are shaded in black.

Chapter 2: Substitutional Analysis of Cysteine Residues in MaMsvR, a Redox-Sensing Transcription Regulator in *Methanosarcina acetivorans*

Introduction

Like MthMsvR, MsvR homologs in *Methanobacteriales* and *Methanomicrobiales* are also divergently transcribed from an *fpaA-rlp-rub* operon. The intergenic regions between the *msvR* and *fpaA* homologs contain three regions of conserved inverted repeats designated as boxes 1-3 (37). Box 3, located in between boxes 1 and 2, overlaps with the transcription start site of *fpaA* (Figure 1A). MthMsvR represses P_{fpaA} by blocking access to the transcription start site and autoregulates its own expression by directly abrogating the TFB/TBP complex from binding to the P_{msvR} TATA box. MthMsvR binds to boxes 1-3 under both oxidizing and reducing conditions *in vitro* as well as to secondary binding sites that have the $TTCN_9GAA$ motif. Members of *Methanosarcinales* also have MsvR homologs. However, these *msvR* homologs are not divergently transcribed from *fpaA* and box 1 is not conserved in the promoter region (5). Ma *msvR* is flanked by two genes encoding uncharacterized proteins and is currently only known to bind to its own promoter. MaMsvR binds to boxes A and B which correspond to boxes 1 and 2 in *M. thermautotrophicus* (Figure 1A). All four binding boxes contain the partial inverted repeat $TTCGTAN_4TACGAA$. While MthMsvR binds P_{msvR} under both oxidizing and reducing conditions, MaMsvR only binds its promoter under reducing conditions, which implies an oxidized environment prevents MaMsvR from binding P_{msvR} (5).

The V4R domain in MthMsvR contains five cysteines, two of which are part of a CX_2CX_3H motif characteristic of some metal-binding proteins involved in redox-

sensing transcription regulators. MaMsvR is a homodimer containing ten cysteine residues including three cysteines in the V4R domain that are conserved in MthMsvR. Two cysteine residues are located in the DNA binding domain (C45 and C63), four are in the V4R domain (C206, C225, C232 and C240) and four are in the linker region (C89, C114, C145 and C178) (Figure 2). Similar to how OxyR functions, the cysteine residues in MaMsvR are believed to sense oxidant (e.g. H_2O_2) in the cell. It is hypothesized that oxidation of these residues induces disulfide bond formation between key cysteine residues and triggers a conformational change that releases MaMsvR from P_{msvR}.

After oxidant has been cleared, most redox-sensing transcription regulators are restored by disulfide reduction (36). In most organisms, a thioredoxin and/or glutathione system is present to reduce disulfides in the target protein to regenerate free thiols and reactivate the protein (37). Methanogens have thioredoxins, which are small proteins with a CXXC motif in the active site that catalyze oxidoreductase reactions through thiol-disulfide exchange (38). The two cysteines in thioredoxin form a disulfide bond and the free thiols in the target protein are regenerated. Thioredoxins must be enzymatically converted back to their reduced state by TrxR, which uses NADPH as an electron donor (39). *M. acetivorans* has seven putative Trx homologs (MaTrx1-7), one TrxR homolog (MaTrxR) and a complete NADPH-dependent thioredoxin system comprised of MaTrx7 and MaTrxR (40). MaTrx7 was demonstrated to specifically reduce disulfides in MaMsvR with reducing equivalents from MaTrxR and NADPH.

In a previous study, the four cysteines in the V4R domain of MaMsvR were individually mutated to alanine to test the impact each cysteine had on MaMsvR

binding to P_{msvR} (5). It was hypothesized that the three conserved cysteines (C206, C232 and C240) played an important role in redox sensing. C206 and C232 were determined to be important for MaMsvR redox sensing while C225 (the only non-conserved cysteine in the V4R domain of MaMsvR) was not. C240 appeared to be involved in redox sensing but the variant complex was not stable and did not provide conclusive results (5).

It is postulated that disulfide formation causes a conformational change in MaMsvR that releases it from its promoter and allows RNAP to bind P_{msvR} and proceed with transcription. Chapter 2 of this thesis focuses on which cysteines play a role in the conformational change between $MaMsvR^{red}$ and $MaMsvR^{ox}$. To accomplish this goal, alanine or serine was individually substituted for each cysteine. It was found that seven cysteine residues in MaMsvR are likely involved redox sensing after H_2O_2 exposure. If these are all involved in disulfide bonds, there is a possibility for seven intermolecular, three intramolecular and one intermolecular or some combination of intra/inter-molecular disulfide bonds to form.

Materials and Methods

DNA Purification and Quantitation

PCR products, gene fragments and digested DNA were purified using the Clean & Concentrator™ kit (Zymo Research). Purification of DNA for use in Electrophoretic Mobility Shift Assays (EMSA) was performed using the Wizard® SV Gel and PCR Clean-Up System (Promega). All plasmids were purified using the Zyppy™ Plasmid Miniprep Kit (Zymo Research). All DNA was quantified using the Qubit® dsDNA

Broad Range Assay Kit and the original version of the Qubit® fluorometer (Thermo Fisher Scientific). All purifications and quantifications were performed as per each manufacturer's instructions, with the exception of eluting DNA into 1 M Tris at pH 8.0 instead of H_2O.

Generation of MaMsvR Variants

Template:

E. coli DH5α strain number LK1341 harbors plasmid pLK314, which is comprised of full-length Ma *msvR* in the backbone vector pQE80LNS that encodes for an N-terminal *Strep*- tag II (Qiagen). A scrape of LK1341 was taken from the glycerol stock stored at -80°C and incubated for 16-18 hours at 37°C and shaken at 250 rpm in 25 mL of LB media (per liter of distilled deionized H_2O (ddH_2O): 10 g tryptone, 5 g yeast extract and 5 g NaCl) containing 100 µg mL^{-1} ampicillin. The plasmid pLK314 was purified and stored at -20°C until it was used as a template for generating cysteine to alanine/serine mutations in the Ma *msvR* construct. Wild-type (WT) strains are listed in Table 1 and all plasmids are listed in Table 2.

Primers:

Two pairs of primers were used for overlapping PCR: a pair of complementary primers that contained the point mutation and annealed to the desired mutation site on the template strands, and the primer pair LK414 and LK415 that flanked Ma *msvR* on the plasmid. All primers are listed in Table 3.

Overlapping PCR:

All overlapping PCR reactions used Phusion® High-Fidelity DNA polymerase (New England Biolabs®) to extend template strands. PCR reactions were run under the following thermal cycling conditions: 30 seconds at 98°C for initial denaturation followed by 30 cycles of 10 seconds at 98°C for denaturation, 30 seconds at 65 °C for primer annealing and 45 seconds at 72 °C for extension, and final extension at 72 °C for 5 minutes.

Two rounds of PCR were utilized to generate Ma *msvR* variants. The first round of PCR consisted of two separate PCR reactions. One reaction contained pLK314 along with LK414, the forward primer that annealed to the 5' end of the vector, and a second primer that contained the desired mutation and bound within Ma *msvR* at the site of mutagenesis. The second PCR reaction contained pLK314 along with LK415, which is the reverse primer that bound to the 3' end of the vector, and a primer containing the desired mutation that annealed within Ma *msvR*. PCR products were cleaned up and used as template for the second round of PCR.

Equal volumes of each reaction from the first round were combined for the second round of PCR. The complementarity of the internal primers allowed the template strands to anneal to each other and primers LK414 and LK415 allowed for extension of the template to produce amplicons of full-length Ma *msvR* that contained the desired mutation. The size of the product was confirmed by running an aliquot of each sample on a 1% agarose gel at 120 V and 350 mA for 30 minutes. The gel was imaged with short wave UV light on a GelDoc™ system (Bio-Rad) and image color was inverted for ease of viewing.

Cloning and Transformation:

Amplicons and pQE80LNS were digested using FastDigest® *BamHI* and *PstI* (Thermo Fisher Scientific) in amounts outlined by the manufacturer. The reaction mixtures were incubated at 37°C for 1-2 hours to allow for complete digestion at the restriction sites and the digested DNA was then cleaned up as previously described. The digested Ma *msvR* variant was ligated to digested pQE80LNS using T4 DNA ligase (New England Biolabs®). The reaction mixture was incubated at room temperature (RT) for 30 minutes and the remainder left after transformation was kept at RT overnight to allow for further ligation.

E. coli DH5α competent cells were prepared according to the Inoue method (41). Five µL of the ligation reaction was incubated with 100 µL of competent *E. coli* DH5α cells for 30 minutes on ice. The cells were heat shocked at 42°C for 90 seconds and then immediately transferred to ice for a 90 second incubation. Then 900 µL of LB was added to the cells and gently mixed by pipetting up and down. The tube of cells was shaken at 100 rpm inside an Erlenmeyer flask for 1 hour at 37°C. Cells were pelleted by centrifuging for 1 minute at 14,000 rpm. Then 900 µL of supernatant was removed and cells were resuspended in the remaining 100 µL of supernatant. Resuspended cells were plated on LB agar (per liter of ddH_2O: 10 g tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar) that contained 100 µg mL^{-1} ampicillin, and incubated for 16-18 hours at 37°C.

Colony PCR:

Colony PCR was utilized to screen for colonies of cells that contained the previously transformed plasmid. Individual colonies were randomly picked from the LB agar plate and resuspended in PCR master mix. Primers LK414 and LK415 and GoTaq® DNA polymerase (Promega) were added to the reaction mix. PCR products were run on a 1% agarose gel at 120 V and 350 mA for 30 minutes to determine which colonies contained the DNA insert. The gel was imaged with short wave UV light on a GelDoc™ system (Bio-Rad) and image color was inverted for ease of viewing.

Plasmid Sequencing:

Colonies that were shown to have the plasmid of interest were picked from the plate and suspended in 5 mL of LB media that contained 100 µg mL^{-1} of ampicillin. The cultures were incubated for 16-18 hours at 37°C and shaken at 250 rpm. Plasmids were purified and sent with primers LK414 and LK415 to the Oklahoma Medical Research Foundation DNA Sequencing facility for sequence confirmation. Sequences were compared to WT Ma *msvR* using Geneious v8 and one plasmid containing the Ma *msvR* insert with only the desired mutation was chosen to serve as the stock for that particular variant.

Overexpression of MaMsvR Variants

Transformation:

E. coli Rosetta™ (Novagen) competent cells were prepared according to the Inoue method (41). Two µL of pQE80LNS plasmid containing the sequence-confirmed Ma *msvR* variant was incubated with 50 µL of competent *E. coli* Rosetta™ cells for 30

minutes on ice. The cells were heat shocked at 42°C for 90 seconds and then immediately transferred to ice for a 90 second incubation. Then 950 µL of LB media was added to the cells and gently mixed by pipetting up and down. The tube of cells was shaken at 100 rpm inside an Erlenmeyer flask for 1 hour at 37°C. Cells were pelleted by centrifuging for 1 minute at 14,000 rpm. Then 900 µL of supernatant was removed and cells were resuspended in the remaining 100 µL of supernatant. Resuspended cells were plated on LB agar that contained 100 µg mL^{-1} of ampicillin and 100 µg mL^{-1} of chloramphenicol, and incubated for 16-18 hours at 37°C.

Auto-Induction:

Colonies were picked from the plate and suspended in 10 mL of LB media that contained 100 µg mL^{-1} of ampicillin and 100 µg mL^{-1} of chloramphenicol. The cultures were incubated for 16-18 hours at 37°C and shaken at 250 rpm. The following day, each 10 mL culture was pelleted down 2 mL at a time by centrifuging for 1 minute at 14,000 rpm. Supernatant was removed and discarded before pipetting 2 more mL into the tube for further centrifugation until 10 ml of culture was pelleted. A culture that served as the uninduced control was made by pipetting 70 µL of overnight culture into 7 mL of LB media that contained 100 µg mL^{-1} of ampicillin and 100 µg mL^{-1} of chloramphenicol.

Each 10 mL pellet was resuspended in 1 mL of ZY media (per liter: 10 grams tryptone, 5 grams yeast extract, 925 mL ddH_2O, 1 mL of 1 M $MgSO_4$, 1 mL of 1000x trace metals, 20 mL of 50x 5052, 50 mL of 20x NPS, 1 mL of 100 µg mL^{-1} ampicillin and 1 mL of 100 µg mL^{-1} chloramphenicol) and then poured into 1 L of ZY media (Table 5). The flasks containing the cultures for auto-induction and the culture of

uninduced cells were incubated for 5 hours at 37°C and 250 rpm and then for 16 hours at 22°C and 250 rpm.

After incubation, 1 mL of auto-induced culture and 1 mL of uninduced culture were pelleted via centrifugation at 14,000 rpm for 1 minute. The supernatant was removed and cells were resuspended in 1 mL of ddH_2O. Then 20 µL of each culture was transferred to a clean tube and 20 µL of Laemmli buffer (Bio-Rad) that contained 355 mM β-mercaptoethanol (per 1 mL of Laemmli buffer) was added to each culture. Cultures were boiled for 5 minutes and run on a 15% SDS-PAGE gel for 40-50 minutes at 200 V and 350 mA to confirm the auto-induction was successful. Auto-induced cultures were pelleted in 200 mL bottles at 8,900 rpm for 15 mins at 4°C. The supernatant was discarded and the pellet was scraped into a 50 mL conical tube, weighed, and then stored at -80°C.

Protein Purification

MaMsvR variants were purified under non-reducing conditions from *E. coli* cells using a *Strep*-Tactin Superflow Plus (Qiagen) column that bound the N-terminal *Strep*-tag II on the protein. Cell pellets were thawed on ice and resuspended in 5 mL of NP buffer (per liter: 6.9 g $NaH_2PO_4 \cdot H_2O$, 17.54 g NaCl, pH 8.0 with NaOH) per 1 g of pellet. Homogenized cells were then sonicated at 20% amplitude for 20 seconds on and 60 seconds off for a total of 5 minutes on time. Cellular debris was removed by centrifugation at 8,900 rpm for 15 minutes at 4°C. The resulting supernatant was used for protein purification. The 500 µL column was set up by pipetting 1 mL of *Strep*-Tactin Superflow Plus (50% slurry) (Qiagen) into a 20 mL polypropylene column. Each column was used up to 5 times before it was discarded. Cell lysate was run over the

column twice before the column was washed with 20 mL of NP buffer. Protein was eluted 6 times by pipetting 500 µL of NPD buffer (per liter: 6.9 g $NaH_2PO_4 \cdot H_2O$, 17.54 g NaCl, 0.54 g desthiobiotin, pH 8.0 with NaOH) over the column.

After purification, the column was stripped according to manufacturer's guidelines and stored at 4°C with NP buffer over the column. The six elutions, along with samples from the pellet, lysate, column flow through and wash were mixed with a 1:1 ratio of Laemmli buffer (Bio-Rad) that contained 355 mM β-mercaptoethanol (per 1 mL of Laemmli buffer) and boiled for 5 minutes. Samples were then run on a 15% SDS-PAGE gel for 40-50 minutes at 200 V and 350 mA to confirm protein was eluted and clean. Elutions that contained the highest concentrations of protein were combined in a DiaEasy™ Dialyzer MWCO 6-8 kDa (BioVision) for dialysis in PDB (per liter: 2.42 g Tris base, 14.91 g KCl, 0.95 g $MgCl_2$, 500 mL glycerol, brought to volume with H_2O) overnight at 4°C. Dialyzed protein was then quantified via Pierce™ Coomassie Plus (Bradford) Assay Kit (Thermo Fisher Scientific).

Size Exclusion Chromatography

MaMsvR variants C63A, C89A, C148, C175, C206A, 2CtoA and 3CtoA were run on size exclusion chromatography (SEC) at the Protein Production Core Facility at the University of Oklahoma using ÄKTA pure M1 with accompanying UNICORN software (GE Healthcare). Samples were run on a Superdex-200 Increase column in SEC buffer (150 mM NaCl, 20 mM tris hcl, pH 8.0) at 4°C. The column was standardized with the following markers: carbonic anhydrase from bovine erythrocytes (29 kDa), conalbumin from chicken egg white (75 kDa), ferritin from horse spleen (440 kDa) and blue dextran 2000 (2,000 kDa). The blue dextran marker gives the void

volume for the column. All standard are from the low molecular weight and high molecular weight gel filtration calibration kits (GE Healthcare).

Electrophoretic Mobility Shift Assay

Half of each protein stock was pretreated with 100x H_2O_2, 1 mM diamide (final conc.) or 1 mM paraquat (final conc.) and incubated at RT for 30 minutes prior to setting up the reactions. Pretreatment with an oxidizing agent was done to ensure complete oxidation of MaMsvR variants. Each EMSA master mix contained 6.5 µl ddH_2O, 2 µl of 10x Txn buffer (per 10 mL: 200 mM Tris pH 8, 100 mM $MgCl_2$, 1.2 M KCl), 3 µl PDB, 0.5 µl heparin (125 mg ml^{-1}), and 2 µl of a 500 nM stock of a 100-base pair (bp) fragment of P_{msvR}. Then 14 µl of master mix was pipetted into each reaction tube. For oxidized reactions, 4 µl of 10 µM oxidizing agent was added (H_2O_2, diamide or paraquat). For reduced reactions, 1 µl of 100 mM DTT and 3 µl of ddH_2O were added to each tube. Then 2 µl of MaMsvR (in PDB) was added to each reaction so that MaMsvR pretreated with a particular oxidant was used for EMSA reactions containing that same oxidant. MaMsvR concentrations are listed in figure legends. EMSA reactions were incubated at RT for 15 minutes while a 6-8% tris-borate (TB) gel pre-ran for 15 minutes at 200 V and 100 mA. After incubation, EMSA reactions were run on 6-8% TB gels for 35 minutes at 200 V and 100 mA. Gels were stained with SYBR Gold™ (Invitrogen) for 30 minutes and imaged with short wave UV light on a GelDoc™ system (Bio-Rad). Image coloration was inverted for ease of viewing.

Generation of MaMsvR Homology Model

The amino acid sequence of MaMsvR was uploaded to Protein Homology/analogY Recognition Engine v2.0 (Phyre2) to predict the tertiary structure of monomeric MaMsvR (42). Nine templates were utilized to model MaMsvR based on heuristics that maximized confidence, percentage identity and alignment coverage. The predicted structure included 99% of its residues modelled at 99% confidence (42). PyMOL was utilized to visualize the homology model and measure distances between cysteine residues (43). Further molecular graphics and surface topology analyses were performed with the UCSF Chimera package (44, 45).

Table 1. *E. coli* strains

Strain/ Reference	Plasmid Harbored	Cell Strain	Mutation(s)	Resistances
Invitrogen	None	DH5α	None	None
Novagen	None	Rosetta	None	Cam
LK1341	pLK314	DH5α	None	Amp
LK1359	pLK350	DH5α	C240S	Amp
LK1377	pLK350	Rosetta	C240S	Amp, Cam
LK1391	pLK360	DH5α	C232A	Amp
LK1392	pLK361	DH5α	C206A	Amp
LK1396	pLK360	Rosetta	C232A	Amp, Cam
LK1397	pLK361	Rosetta	C206A	Amp, Cam
LK1457	pLK393	Rosetta	C45A	Amp, Cam
LK1470	pLK401	DH5α	C63A	Amp
LK1487	pLK401	Rosetta	C63A	Amp, Cam
LK1488	pLK414	DH5α	C89A	Amp
LK1489	pLK415	DH5α	C114A	Amp
LK1490	pLK415	Rosetta	C114A	Amp, Cam
LK1491	pLK416	DH5α	C148A	Amp
LK1508	pLK414	Rosetta	C89A	Amp, Cam
LK1510	pLK416	Rosetta	C148A	Amp, Cam
LK1511	pLK426	DH5α	C175A	Amp
LK1512	pLK426	Rosetta	C175A	Amp, Cam
LK1515	pLK427	DH5α	C45A, C225A	Amp
LK1524	pLK427	Rosetta	C45A, C225A	Amp, Cam
LK1538	pLK440	DH5α	C45A, C114A, C225A	Kan
LK1540	pLK442	DH5α	C45A, C114A, C225A	Amp
LK1547	pLK442	Rosetta	C45A, C114A, C225A	Amp, Cam
LK1554	pLK452	DH5α	C45A, C114A, C225A, C240S	Amp
LK1555	pLK452	Rosetta	C45A, C114A, C225A, C240S	Amp, Cam

Table 2. Plasmids

Plasmid	**Backbone Vector**	**Function**	**Reference**
pQE80LNS	N/A	*Strep*-tag II labeling and expression vector	Qiagen
pLK314	pQE80LNS	MaMsvR^{WT} expression construct	This study
pLK350	pQE80LNS	MaMsvR^{C240S} expression construct	This study
pLK360	pQE80LNS	MaMsvR^{C232A} expression construct	This study
pLK361	pQE80LNS	MaMsvR^{C206A} expression construct	This study
pLK393	pQE80LNS	MaMsvR^{C45A} expression construct	This study
pLK401	pQE80LNS	MaMsvR^{C63A} expression construct	This study
pLK414	pQE80LNS	MaMsvR^{C89A} expression construct	This study
pLK415	pQE80LNS	MaMsvR^{C114A} expression construct	This study
pLK416	pQE80LNS	MaMsvR^{C148A} expression construct	This study
pLK426	pQE80LNS	MaMsvR^{C175A} expression construct	This study
pLK427	pQE80LNS	MaMsvR^{2CtoA} expression construct(contains substitutions C45A, C225A)	This study
pLK440	Zero blunt® TOPO®	MaMsvR^{3CtoA} cloning vector	Life technologies™
pLK442	pQE80LNS	MaMsvR^{3CtoA} expression construct (contains substitutions C45A, C114A, C225A)	This study
pLK452	pQE80LNS	MaMsvR^{4CtoN} expression construct (contains substitutions C45A, C114A, C225A, C240S)	This study

Table 3. Primers

Primer Number	Sequence (5' to 3')	Function
LK414	CCCGAAAAGTGCCACCTG	Forward primer for PQE80LNS vector
LK415	GTTCTGAGGTCATTACTGG	Reverse primer for PQE80LNS vector
LK588	TTCAGGGATCCATGGCAAAACCT GAGACCA	Ma *msvR* cloning with BamHI site
LK589	TTCAGCTGCAGTTATATTGTAATC TCAAAAAGACAG	Ma *msvR* cloning with PstI site
LK677	TTCAGCTGCAGTTATATTGTAATC TCAAAAAGACTGCG	LK589 extended for C240S change in Ma *msvR*
LK719	GGAAAACCTCTCGCTGTGCTTAG AGAAGGC	Forward primer for C206A variant in Ma *msvR*
LK720	GCCTTCTCTAAGCACAGCGAGAG GTTTTCC	Reverse primer for C206A variant in Ma *msvR*
LK721	GGTCCTGGAAACCGAGGCTTATG GGACCGG	Forward primer for C232A variant in Ma *msvR*
LK722	CCGGTCCCATAAGCCTCGGTTTC CAGGACC	Reverse primer for C232A variant in Ma *msvR*
LK723	GGGACCGGGCACACACGCGCTCT TTTTGAGATTAC	Forward primer for C240A variant in Ma *msvR*
LK724	GTAATCTCAAAAAGAGCGCGTGT GTGCCCGGTCCC	Reverse primer for C240A variant in Ma *msvR*
LK821	TGACGAAATTGTTAAAGCTACGG CAAAAGCCAAAT	Forward primer for C45A variant in Ma *msvR*
LK822	ATTTGGCTTTTGCCGTAGCTTTAA CAATTTCGTCA	Reverse primer for C45A variant in Ma *msvR*
LK823	CAATAATCTAAGAACAGCCGAGC TTGTTGAA	Forward primer for C63A variant in Ma *msvR*
LK824	TTCAACAAGCTCGGCTGTTCTTA GATTATTG	Reverse primer for C63A variant in Ma *msvR*
LK841	ACTTCCCAGTACATGGGCGCCTC TCAGGAGCCTTTT	Forward primer for C89A variant in Ma *msvR*
LK842	AAAAGGCTCCTGAGAGGCGCCC ATGTACTGGGAAGT	Reverse primer for C89A variant in Ma *msvR*
LK849	GGAAATGACAGGTTCGCCTTTAT GGAAGCCCTT	Forward primer for C114A variant in Ma *msvR*
LK850	AAGGGCTTCCATAAAGGCGAACC TGTCATTTCC	Reverse primer for C114A variant in Ma *msvR*
LK851	AGAGATATTGGGAAGGCTCTTTC GGCTAATTTT	Forward primer for C148A variant in Ma *msvR*
LK852	AAAATTAGCCGAAAGAGCCTTCC CAATATCTCT	Reverse primer for C148A variant in Ma *msvR*
LK853	GAATTTCACGGGGACGCCCGGGT CTCGGTCCTC	Forward primer for C175A variant in Ma *msvR*
LK854	GAGGACCGAGACCCGGGCGTCC CCGTGAAATTC	Reverse primer for C175A variant in Ma *msvR*

Table 4. MaMsvR variants with multiple substitutions

MaMsvR Variant Name	**Substitutions Present**
$MaMsvR^{2CtoA}$	C45A, C225A
$MaMsvR^{3CtoA}$	C45A, C114A, C225A
$MaMsvR^{4CtoN}$	C45A, C114A, C225A, C240S

Table 5. Recipes for solutions added to ZY media

1000x trace metals (per 100 mL):

Component	**Amount added**
$FeCl_3$	1.35 g
$CaCl_2$	0.294 g
$MnCl_2 \cdot 4\ H_2O$	0.198 g
$ZnSO_4 \cdot 7\ H_2O$	0.2875 g
$CoCl_2 \cdot 6\ H_2O$	0.0476 g
$CuCl_2 \cdot 2\ H_2O$	0.0341 g
$NiCl_2 \cdot 6\ H_2O$	0.0475 g
$Na_2MoO_4 \cdot 2\ H_2O$	0.0484 g
H_3BO_3	0.0124 g
12 M HCl	0.5 mL

Brought to volume with ddH_2O

50x 5052 (per 100 mL):

Component	**Amount added**
glycerol	25 g
ddH_2O	73 mL
glucose	2.5 g
α-lactose	10 g

20x NPS (per 100 mL):

Component	**Amount added**
$(NH_4)_2SO_4$	6.6 g
KH_2PO_4	13.6 g
Na_2HPO_4	28.61 g

Brought to volume with ddH_2O

Results

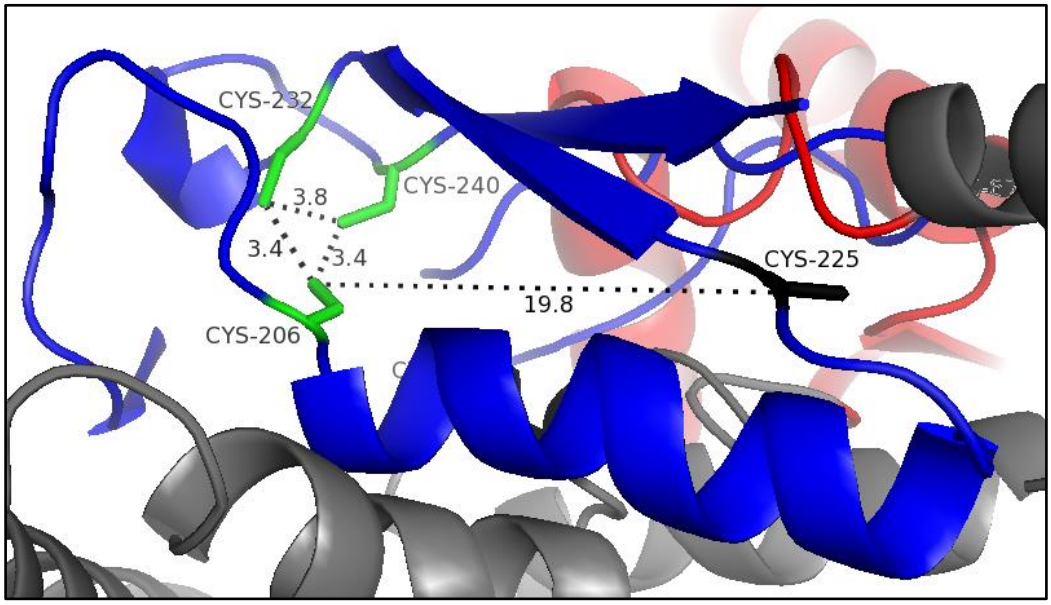
Rationale for MaMsvR Variant Analysis

A common redox-sensing mechanism in transcription regulators involves cysteine residues that become oxidized by ROS and often form intermolecular or intramolecular disulfide bonds (18-20). Formation of disulfide bonds causes a conformational change in the protein that affects its ability to bind its promoter. In the case of OxyR, one cysteine residue is oxidized by H_2O_2 and quickly interacts with the second cysteine to form a disulfide bond (25). This interaction induces a conformational change that allows OxyR to bind promoters in its regulon and activate or repress transcription of genes controlled by those promoters. However, in most cases the conformational change prevents the transcription regulator from binding its promoter and its regulon is derepressed, as is the case with PerR (29). Evidence suggests MaMsvR senses oxidant through a similar mechanism so the focus of this study was on elucidating the cysteines residues in MaMsvR that are involved in redox sensing (5).

MaMsvR is a homodimer containing ten cysteine residues in each monomer, providing dimeric MaMsvR with twenty cysteine residues that may be involved in the formation of disulfide bonds upon oxidation. The conventional distance cutoff for two cysteine residues to form a disulfide bonds is 2.3 Å (46, 47). However, that distance pertains mostly to permanent structural disulfide bonds (48). OxyR contains two cysteine residues (C199 and C208) that are separated by 17.3 Å in the reduced form of OxyR (49). The cysteines only interact after H_2O_2 oxidizes the thiol of C199 to form sulfenic acid (C199-SOH). C199-SOH then quickly interacts with C208 to form a disulfide bond that leads to a conformational change in the regulatory domain (50). The

formation of reversible disulfide bonds is influenced by the angle of a cysteine thiol, its accessibility to solvent and the intermediate formed upon thiol oxidation (48). When these variables were taken into account it was determined that cysteines separated by 6.2 Å or less have a high probability of forming a disulfide bond (48). This cutoff distance was used for analysis of cysteines in the MaMsvR homology model.

Although a crystal structure for MaMsvR has not been solved, the tertiary structure prediction program Phyre2 allowed for the amino acid sequence of MaMsvR to be compared to other similar structures from the Protein Data Bank. PyMol was utilized to visualize the homology model and look for cysteine residues that may be within the distance cutoff to form disulfide bonds. The three conserved cysteines in the V4R domain (C206, C232 and C240) were located 3.4-3.8 Å away from one another in the homology model. This indicates all three cysteines are within the cutoff distance to form disulfide bonds with one another, although disulfide partners cannot be determined from the model alone (Figure 3). The only nonconserved cysteine in the V4R domain (C225) is too far removed to be involved in an intramolecular disulfide bond with any of the cysteines in the V4R domain. The model presented here corroborates previous findings that C206, C232 and C240 may be involved in redox sensing and C225 is not. Interestingly, the structures used to model the V4R domain contain a bound zinc ion at the corresponding residues. To strengthen the argument for involvement of C240 in redox sensing and/or conformational change and to analyze the remaining cysteines in MaMsvR, alanine or serine was substituted for each cysteine (except for C225) to determine cysteines with possible involvement in redox sensing. These findings lead to double, triple and quadruple MaMsvR variants to validate preliminary results.



Distances between cysteines in V4R domain			
	C206	**C225**	**C232**
C225	19.8 Å		
C232	3.4 Å	21.5 Å	
C240	3.4 Å	20.5 Å	3.8 Å

Figure 3. Distances (Å) between cysteine residues in the V4R domain.

An intensive Phyre2 search was used to generate a full length homology model of MaMsvR using domain homologs in the Protein Data Bank. The DNA binding domain is colored in red and the V4R domain in blue. Conserved cysteines are colored in green and non-conserved cysteines in black. Distances (Å) between cysteines in the V4R domain are shown in the model and listed in the table.

Generation of Mutant Ma msvR Constructs

Overlapping PCR was utilized for site-directed mutagenesis of Ma *msvR*. Alanine was individually substituted for all cysteines in MaMsvR (except for C225) to remove the thiol group and abolish its ability to potentially form a disulfide bond upon oxidation. A second C240 variant was generated in which serine was substituted for cysteine ($MaMsvR^{C240S}$). The MaMsvR homology model placed C240 on the surface of the protein and the alanine substitution placed a hydrophobic residue on the solvent-exposed surface (Figure 4). This may have disrupted the structural integrity of MaMsvR and prevented it from functioning properly, leading to the inconclusive results previously discussed. $MaMsvR^{C240S}$ was created to avoid this issue while still eliminating the thiol group.

Expression and Purification of MaMsvR Variants in E. coli

Mutated Ma *msvR* amplicons from overlapping PCR were ligated into pQE80LNS, an expression vector containing a *lac* operator that regulates expression of the inserted gene. Resultant plasmids were transformed into the maintenance strain DH5α for plasmid storage. Plasmids were also transformed into the *E. coli* host strain Rosetta™ (Novagen), which has been engineered to enhance expression of proteins that contain codons rarely used in WT *E. coli*. MaMsvR was overexpressed via auto-induction, a method using lactose to induce expression of the gene regulated by the *lac* operator in the expression vector. Purified protein elutions were run on SDS-PAGE to confirm expression and purification were successful (Figure 5). SEC was used to remove protein contaminants from some of the purified MaMsvR variants (Figure 6).

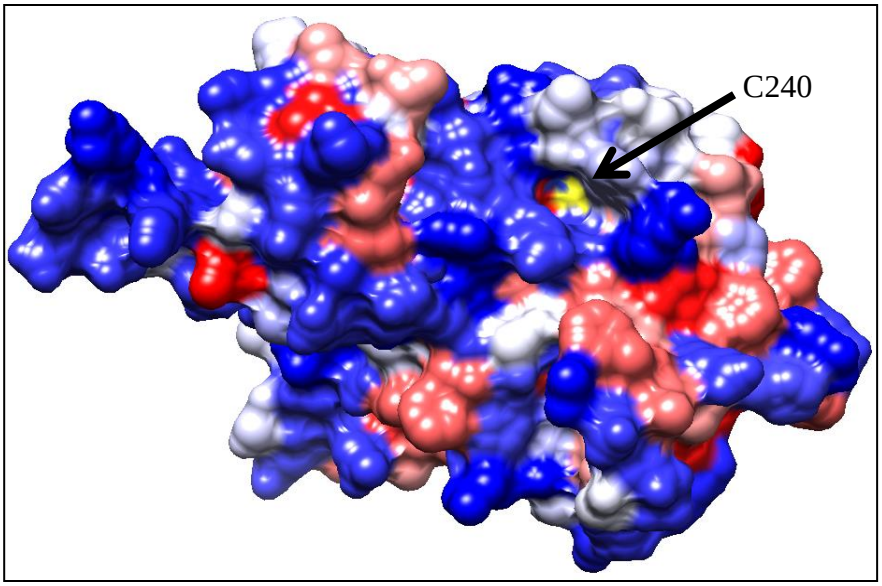


Figure 4. Molecular surface representation of MaMsvR homology model.

An intensive Phyre2 search was used to generate a full length homology model of MaMsvR using domain homologs in Protein Data Bank. Chimera was used to color hydrophobic surface areas in red and hydrophilic surfaces in blue. C240 is labeled and represented in yellow.

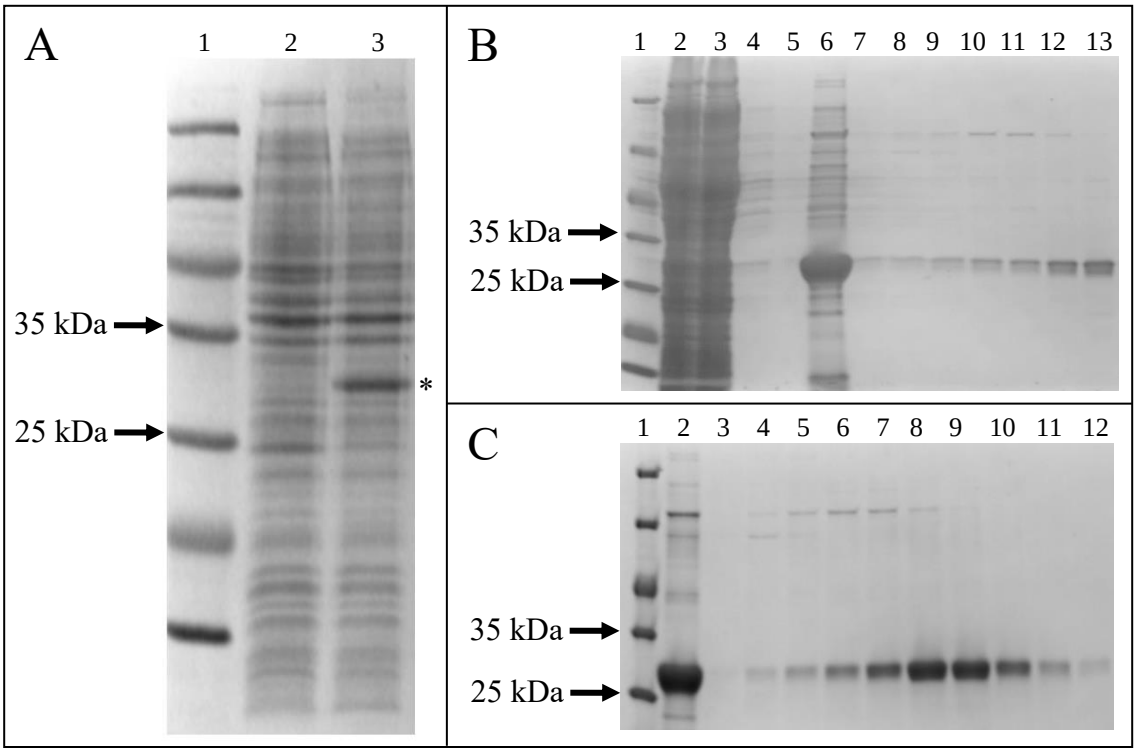


Figure 5. Representative MaMsvR autoinduction, purification and SEC gel images. *Strep*-tag II MaMsvR is 29.2 kDa. **(A)** Representative image of autoinduction trials of MaMsvR variants. Lane 1 shows the Thermo Fisher Scientific Prestained Protein Molecular Weight Marker. Lane 2 shows uninduced cell lysate. Lane 3 shows induced cell lysate. The band indicating MaMsvR overexpression is labeled with an asterisk. **(B)** Representative image of MaMsvR purification. Lane 1 shows the Thermo Fisher Scientific Prestained Protein Molecular Weight Marker. Cell pellet (Lane 2), lysate loaded onto column (Lane 3), column flow through (Lane 4), pooled sample of column washes (Lane 5), all elutions that were pooled for SEC (Lane 6) and individual elutions (Lanes 7-13) are shown. **(C)** Representative gel image of SEC fractions (Lanes 3-12) and fractions that were pooled for analysis (Lane 2). Lane 1 shows the Thermo Fisher Scientific Prestained Protein Molecular Weight Marker.

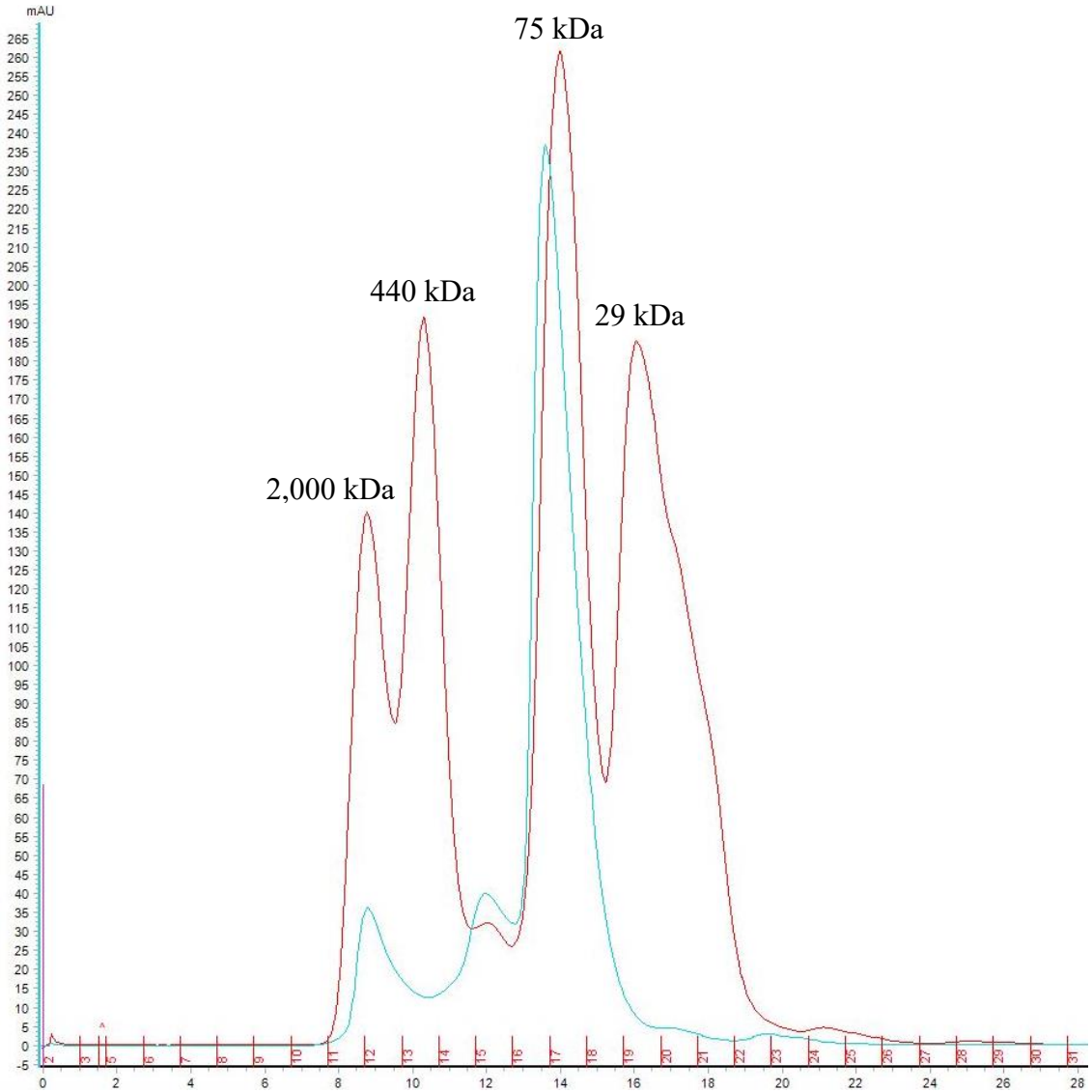


Figure 6. Representative chromatogram of MaMsvR variant after SEC.

The standard is colored in red and the sample in light blue. Fraction numbers are indicated in red on the x-axis and volume in mL is indicated in black on the x-axis. The y-axis represents absorbance at 280 nm.

Analysis of P_{msvr} Binding by Single MaMsvR Variants Oxidized with H_2O_2

The protein-DNA interactions of MaMsvR and its promoter (P_{msvR}) were tested through EMSA. The EMSA technique is based on the fact that protein-DNA complexes migrate more slowly through non-denaturing polyacrylamide than free DNA. A functional DNA-binding protein is incubated with target DNA and run on a polyacrylamide gel to determine if the protein can bind the target DNA in the buffer conditions provided. If the protein binds the DNA fragment, the movement of the complex through the gel is retarded and a band that is higher in the gel than the band produced from unbound DNA is produced. The band representing protein-DNA complexes is called a shift.

A lane for free DNA was always run with the DNA-protein binding reactions to generate a baseline for unbound DNA to which the shifts could be compared. Purified protein was treated with 100x H_2O_2 and incubated at RT for 30 minutes to ensure complete oxidation of the protein. Oxidized MaMsvR variant stocks were utilized for EMSA analysis to ensure the protein remained functional after oxidation. The concentrations of MaMsvR variants were titrated over P_{msvR} in reduced and oxidized conditions to determine which variants altered MaMsvR binding of P_{msvR}.

MaMsvR variants that behaved like WT MaMsvR ($MaMsvR^{WT}$) during EMSA indicated the substitution of alanine/serine for cysteine did not affect MaMsvR binding to P_{msvR}. MaMsvR variants that bound P_{msvR} in both oxidized and reduced conditions suggested the substituted cysteine may be a key player in redox sensing and/or conformational change because a disulfide bond either did not form or, if a disulfide bond did form, it did not impact MaMsvR binding of P_{msvR}. It was hypothesized that

cysteines in the V4R domain were involved in redox sensing because three of the four cysteines are conserved in *M. thermautotrophicus* and V4R domains are sometimes involved in redox sensing (51, 52). Analysis started with variants MaMsvR^{C206A}, MaMsvR^{C232A} and MaMsvR^{C240S} in the V4R domain. MaMsvR^{C225A} was not analyzed due to previous findings that it is not involved in redox sensing. Previously published results showed that C206 and C232 impacted MaMsvR binding to P_{msvR} and C240 may impact binding (5). However, the MaMsvR^{C240A} variant complex appeared to be unstable so the results were inconclusive.

Analysis of Variants in V4R Domain: MaMsvR^{C206A}, MaMsvR^{C232A} and MaMsvR^{C240S}

MaMsvR^{C206A} was previously shown to bind P_{msvR} in both oxidized and reduced conditions, implicating it as a key cysteine in disulfide bond formation that induced a conformational change in MaMsvR (5). At 20-fold and 40-fold excess MaMsvR^{C206A} over P_{msvR}, MaMsvR^{C206A} was shown again to bind P_{msvR} when oxidized with H_2O_2 and reduced with DTT (Figure 7). Variants MaMsvR^{C232A} and MaMsvR^{C240S} also bound P_{msvR}, at 20-fold excess protein over P_{msvR}, under oxidized and reduced conditions (Figure 8). Clear results for all three conserved cysteines in the V4R domain were obtained and because all three variants bound P_{msvR} when oxidized with H_2O_2 and reduced with DTT, C206, C232 and C240 appear to play a role in redox sensing and/or conformational change. The three residues are also within the cutoff distance to form disulfide bonds with one another, although disulfide bond partners cannot be determined.

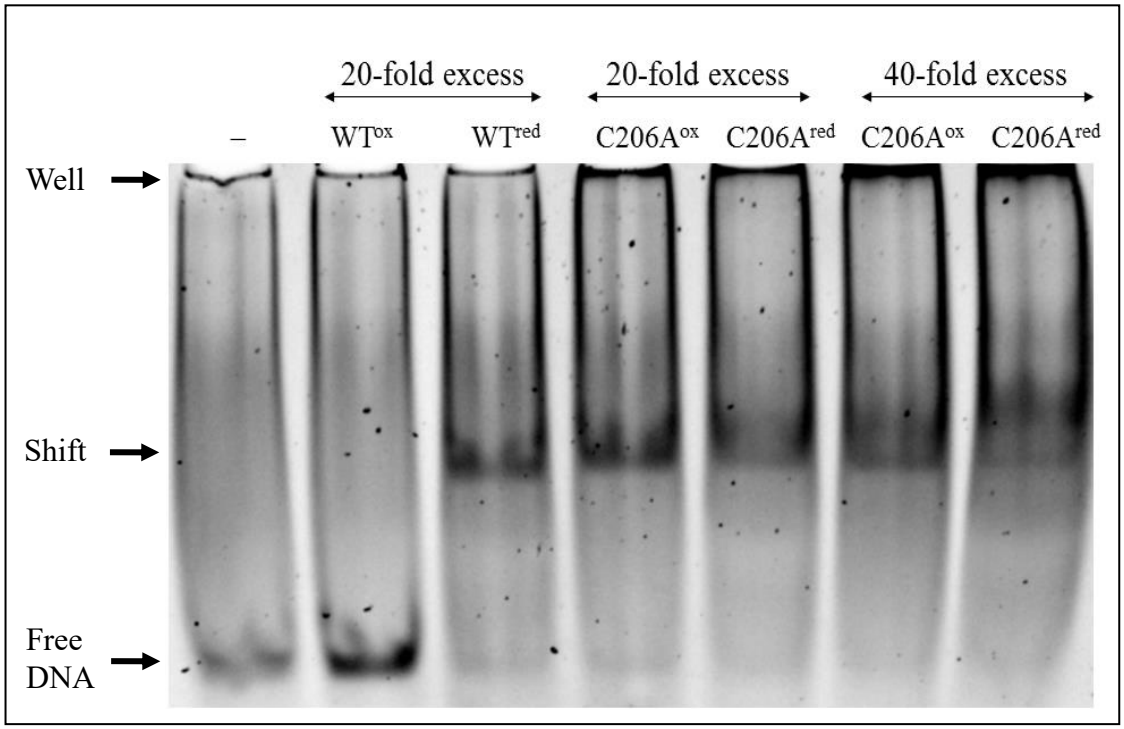


Figure 7. EMSA of MaMsvR^{C206A} with P_{msvR}.

EMSA to test binding of MaMsvR^{C206A} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

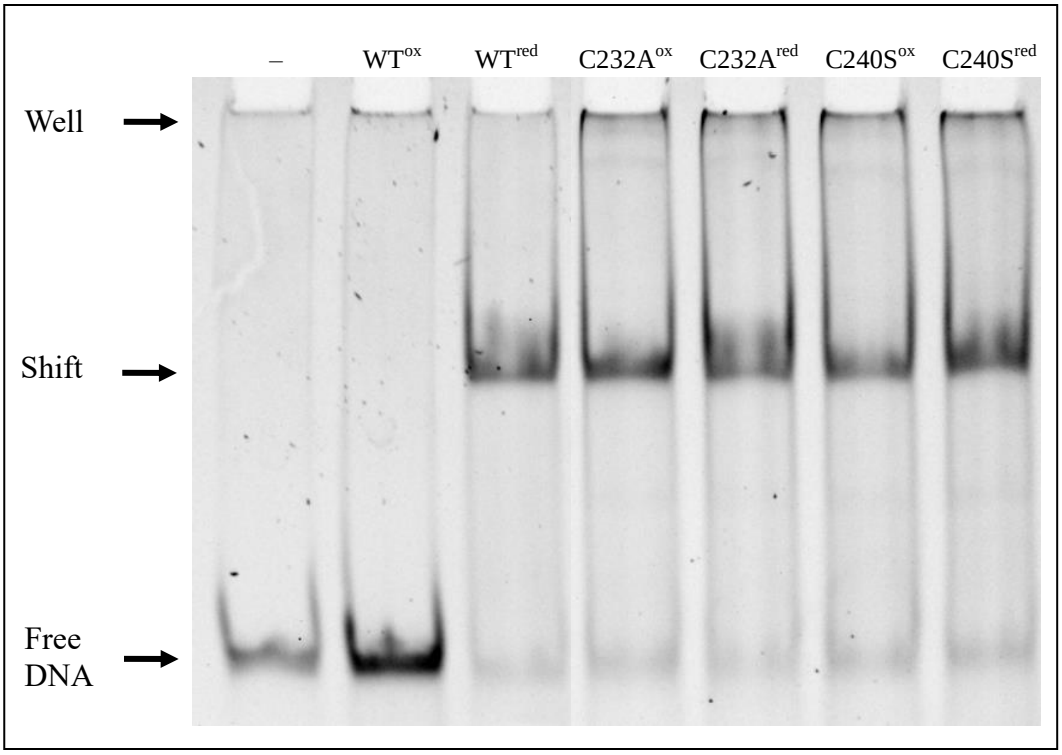


Figure 8. EMSA of MaMsvR^{C232A} and MaMsvR^{C240S} with P_{msvR}.

EMSA to test binding of MaMsvR^{C232A} and MaMsvR^{C240S} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR (20-fold excess) complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

Analysis of Variants in DNA Binding Domain: MaMsvR^{C45A} and MaMsvR^{C63A}

Based on the MaMsvR homology model generated using Phyre2, both C45 and C63 are not located within the 6.5 Å distance cutoff to form intramolecular disulfide bonds with any other cysteine residue (Figure 9). C175 is the nearest cysteine residue to C45 and they are separated by 16.5 Å. The nearest cysteine to C63 with putative involvement in redox sensing and/or conformational change is C89, and they are separated by 18.2 Å. Binding of MaMsvR^{C45A} and MaMsvR^{C63A} to P_{msvR} (at 20-fold excess protein over P_{msvR}) under oxidized and reduced conditions was tested through EMSA. MaMsvR^{C45A} behaved like MaMsvR^{WT} and only produced a shift when reduced with DTT (Figure 10). This demonstrates that C45 is not a likely participant in redox sensing or disulfide bond formation. MaMsvR^{C63A} produced shifts under both oxidized and reduced conditions, indicating that C63 may be involved in redox sensing and/or conformational change.

Taken together, the EMSA results for MaMsvR^{C45A} and the isolation of C45 from other cysteines residues based on the MaMsvR homology model suggest C45 is not involved in redox sensing or disulfide bond formation. Although C63 is not located within the 6.5 Å cutoff to form intramolecular disulfide bonds with any other cysteine residues, the EMSA results suggest it may be involved in redox sensing and/or conformational change. C63 is located at the c-terminal end of the DNA binding domain and is separated from C63' by a distance of approximately 36 Å (not shown). These cysteines could possibly be involved in an intermolecular disulfide bond. However, a significant conformational change would need to occur in order for the disulfide bond between the two C63 residues to form.

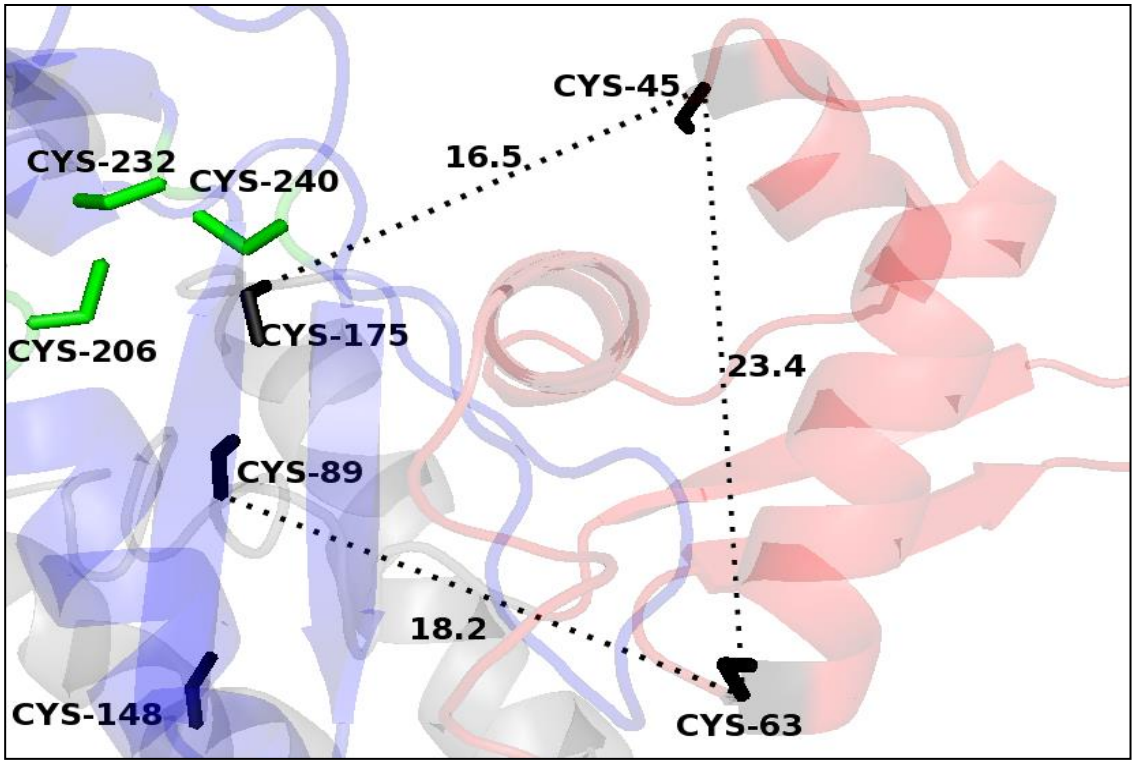


Figure 9. Location of C45 and C63 in the DNA binding domain of MaMsvR with distances (Å) to neighboring cysteines.

An intensive Phyre2 search was used to generate a full length homology model of MaMsvR using domain homologs in the Protein Data Bank. The DNA binding domain is colored in red and the V4R domain in blue. Conserved cysteines are colored in green and non-conserved cysteines in black. Distances (Å) from C45 or C63 to the nearest cysteine residues and between C45 and C63 are labeled.

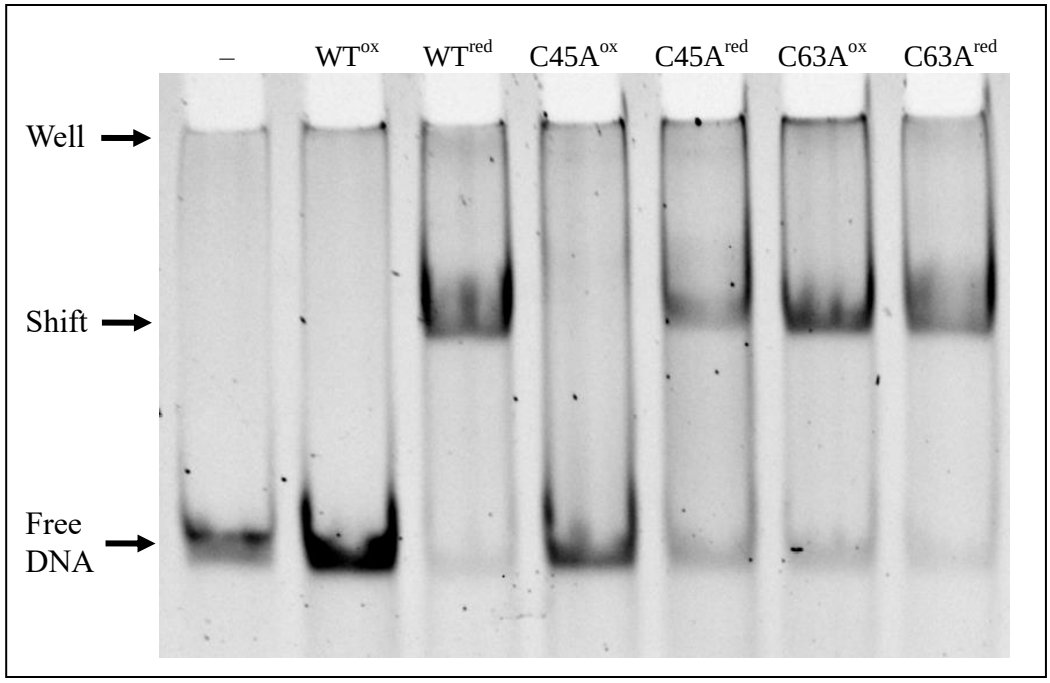


Figure 10. EMSA of MaMsvR^{C45A} and MaMsvR^{C63A} with P_{msvR}.

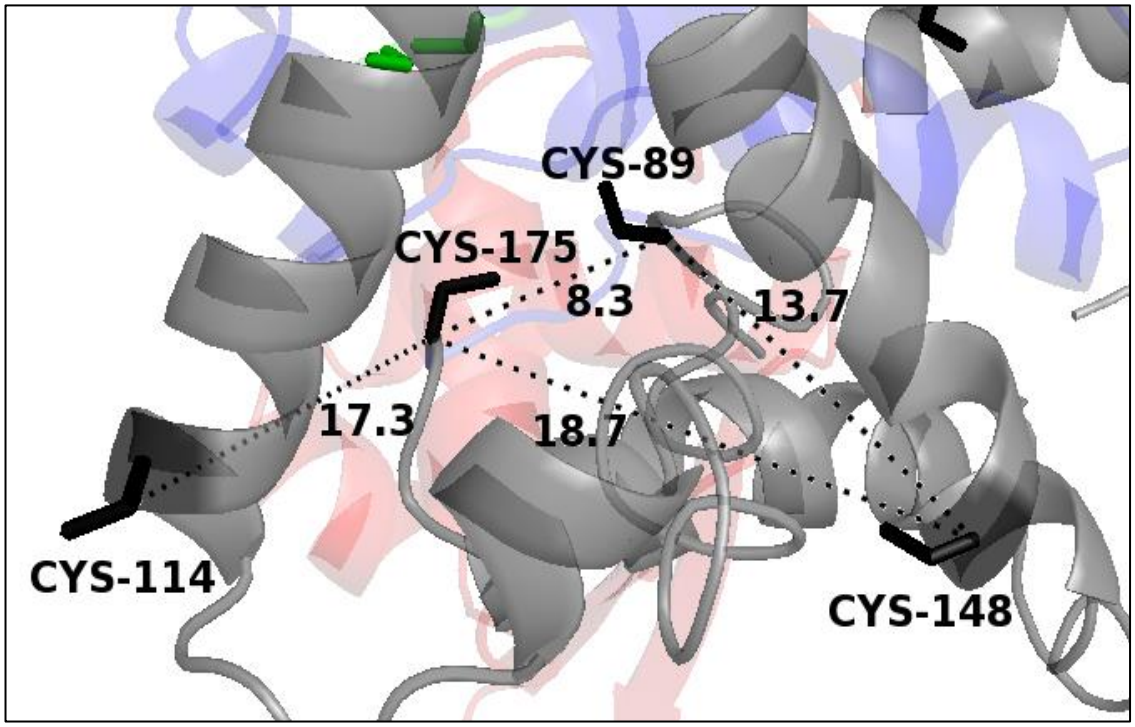
EMSA to test binding of MaMsvR^{C45A} and MaMsvR^{C63A} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR (20-fold excess over P_{msvR}) complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

Analysis of Variants in the Linker Region: MaMsvR^{C89A}, MaMsvR^{C114A}, MaMsvR^{C148A} and MaMsvR^{C175A}

Based on the MaMsvR homology model, C114 and C148 are located on the surface of MaMsvR and C89 and C175 are located near the three cysteines in the V4R domain that have been implicated in redox sensing (Figure 11). The nearest cysteine to C114 is C175 and they are separated by a distance of 17.3 Å. C89 is the nearest cysteine to both C175 and C148; C89 is located 8.3 Å away from C175 and 13.7 Å away from C148. EMSA results indicated that MaMsvR^{C89A} bound P$_{msvR}$ at 10, 20 and 40-fold excess over P$_{msvR}$ under oxidized conditions, although the shifts are faint (Figure 12). MaMsvR^{C89A} was the only variant that consistently produced faint shifts from binding P$_{msvR}$ under oxidized conditions and this may the result of weak interactions between MaMsvR^{C89A} and P$_{msvR}$. Regardless of the faint shifting pattern, C89 may be involved in redox sensing and/or conformational change.

At 20-fold and 40-fold excess MaMsvR^{C114A} over P$_{msvR}$, MaMsvR^{C114A} behaved like MaMsvR^{WT} by only binding P$_{msvR}$ when reduced with DTT (Figure 13). Both MaMsvR^{C148A} and MaMsvR^{C175A} bound P$_{msvR}$ under oxidizing and reducing conditions at 20-fold excess protein over P$_{msvR}$ (Figure 14). However, the shifted bands for both variants are smeared and MaMsvR^{C175A} produced a supershifted band in both oxidized and reduced conditions. The supershifted band from MaMsvR^{C175A} may be due to binding stoichiometry that differs from all other variants. The shifted band for MaMsvR^{WT} is also smeared, which may indicate smearing produced by the variant complexes was not caused by protein degradation or complex disassociation.

MaMsvR^{C114A} is the only variant in the linker region that did not bind P_{msvR} under oxidized conditions and is likely not involved in disulfide bond formation. Variants MaMsvR^{C89A}, MaMsvR^{C148A} and MaMsvR^{C175A} bound P_{msvR} under both oxidized and reduced conditions, implicating C89, C148 and C175 in disulfide bond formation. However, disulfide bond partners cannot be determined based on the results presented.



Distances between cysteines in the linker region and neighboring cysteines						
	C206	**C232**	**C240**	**C89**	**C114**	**C148**
C89	10.2 Å	14.2 Å	12.1 Å			
C114	21.2 Å	22.8 Å	22.5 Å	23.7 Å		
C148	20.8 Å	24.2 Å	22.4 Å	13.7 Å	26.4 Å	
C175	7.7 Å	12.6 Å	9.0 Å	8.3 Å	17.3 Å	18.7 Å

Figure 11. Location of cysteines in the MaMsvR linker region with distances (Å) to neighboring cysteines.

An intensive Phyre2 search was used to generate a full length homology model of MaMsvR using domain homologs in the Protein Data Bank. The DNA binding domain is colored in red, the V4R domain in blue and the linker region in gray. Conserved cysteines are colored in green and non-conserved cysteines in black. Distances (Å) between cysteines in the linker region to neighboring cysteine residues are labeled and shown in the table.

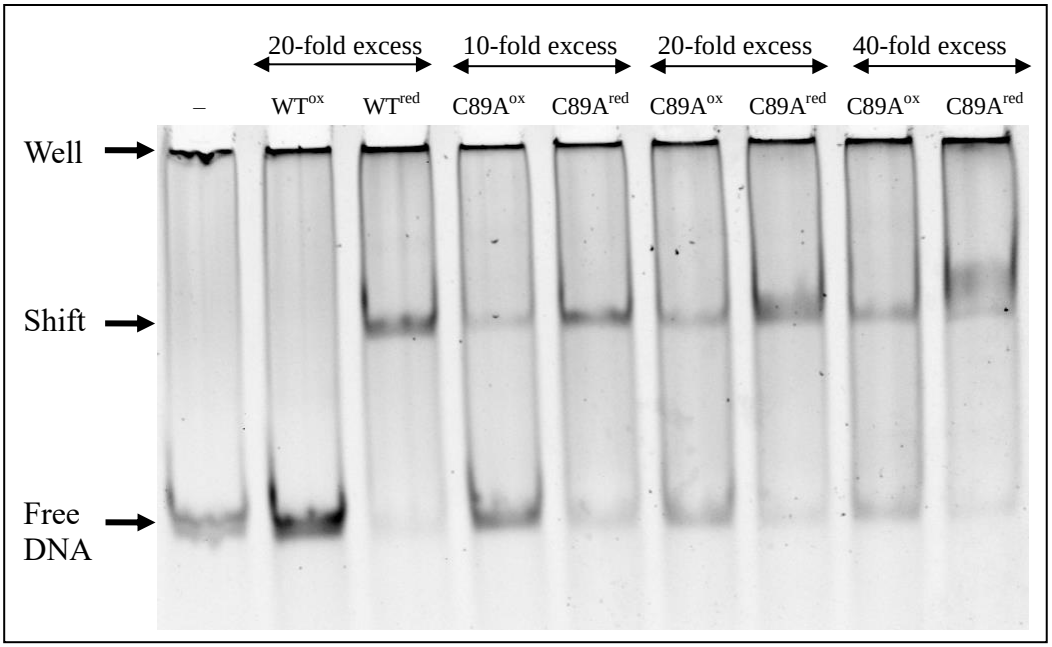


Figure 12. EMSA of MaMsvR^{C89A} titrated over P_{msvR}.

EMSA to test binding of MaMsvR^{C89A} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents MaMsvR^{WT} or MaMsvR^{C89A} complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

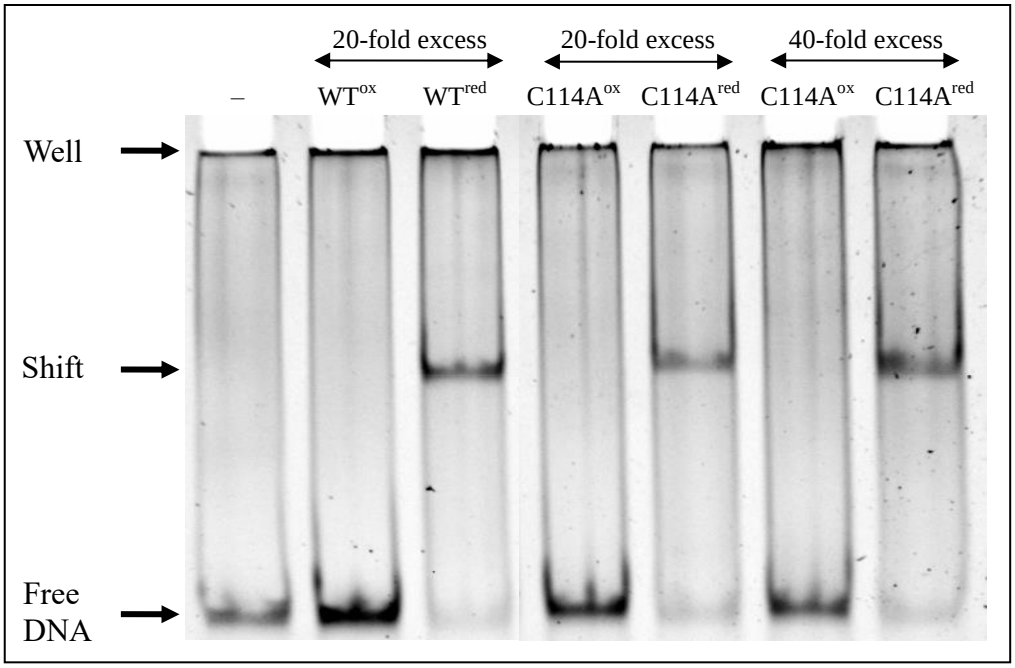


Figure 13. EMSA of MaMsvR^{C114A} titrated over P$_{msvR}$.

EMSA to test binding of MaMsvR^{C114A} to P$_{msvR}$ (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents MaMsvR^{WT} or MaMsvR^{C114A} complexed with P$_{msvR}$. The control lane (–) contains no protein. Unbound P$_{msvR}$ is labeled as free DNA. Gel wells are indicated.

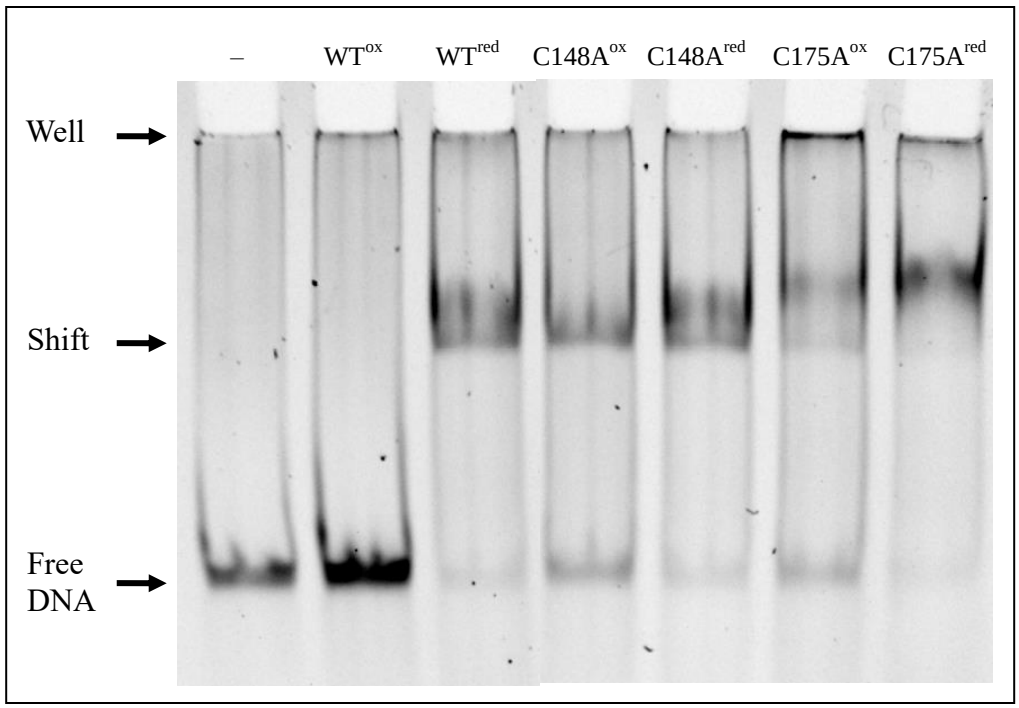


Figure 14. EMSA of MaMsvR^{C148A} and MaMsvR^{C175A} with P_{msvR}.

EMSA to test binding of MaMsvR^{C148A} and MaMsvR^{C175A} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR (at 20-fold excess over P_{msvR}) complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

Overall Comparisons of Single MaMsvR Variants Oxidized with H_2O_2

Variants MaMsvR^{C63A}, MaMsvR^{C89A}, MaMsvR^{C148A}, MaMsvR^{C175A}, MaMsvR^{C206A}, MaMsvR^{C232A} and MaMsvR^{C240S} produced shifts under oxidized and reduced conditions. The findings indicate seven cysteines, C63, C89, C148, C175, C206, C232 and C240, are possible players in redox sensing and/or conformational change. Disulfide bond partners cannot be elucidated from the results presented and further research is necessary to determine which cysteines interact with one another. The three remaining single variants, MaMsvR^{C45A}, MaMsvR^{C114A} and MaMsvR^{C225A}, did not produce a shifted band with P_{msvR} in oxidized conditions and are likely not involved in redox sensing or conformational change. To confirm these findings a double and triple variant was generated for analysis through EMSA.

Analysis of Promoter Binding by MaMsvR Variants with Multiple Substitutions

MaMsvR variants with multiple substitutions were generated to confirm findings that C45, C114 and C225 are not involved in redox sensing and to allow for removal of cysteines not relevant to function to facilitate downstream experiments. MaMsvR^{2CtoA} included the first two cysteines (C45 and C114) not involved in disulfide bond formation and MaMsvR^{3CtoA} contained C45, C114 and C225 (Table 4). A third variant consisted of MaMsvR^{3CtoA} plus a C240S variant and is referred to as MaMsvR^{4CtoN} here, where N denotes an A/S mutation. MaMsvR^{4CtoN} was generated to test if removing one cysteine involved in disulfide bond formation would cause the variant to bind P_{msvR} under both oxidized and reduced conditions.

EMSA results indicated the double and triple mutants both behaved like MaMsvR^{WT} at 20-fold excess protein over P_{msvR} (Figure 15). The findings confirm that C45, C114 and C225 are likely not involved in disulfide bond formation and those residues can be removed to facilitate downstream experiments. Variant MaMsvR^{4CtoN} bound under both conditions (Figure 16). These results imply that a substitution for any of the cysteines involved in disulfide bond formation will alter MaMsvR binding to P_{msvR}. The wells of the gel for MaMsvR^{4CtoN} are darkened from protein that has aggregated in the wells. This implies the variant is not stable which may be due to the number of mutations altering stability of the protein structure.

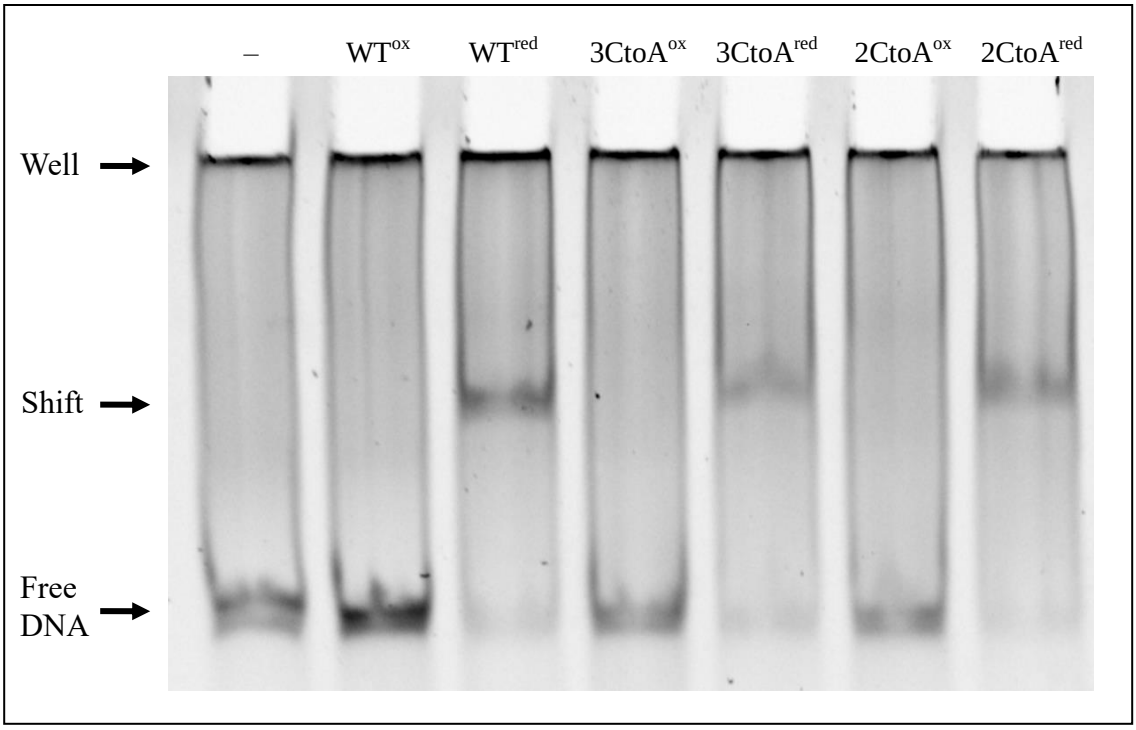


Figure 15. EMSA of MaMsvR^{2CtoA} and MaMsvR^{3CtoA} with P_{msvR}.

EMSA to test binding of MaMsvR^{2CtoA} and MaMsvR^{3CtoA} to P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). All protein samples were at 20-fold excess over P_{msvR}. Each shift represents MaMsvR^{WT} or variant MaMsvR complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

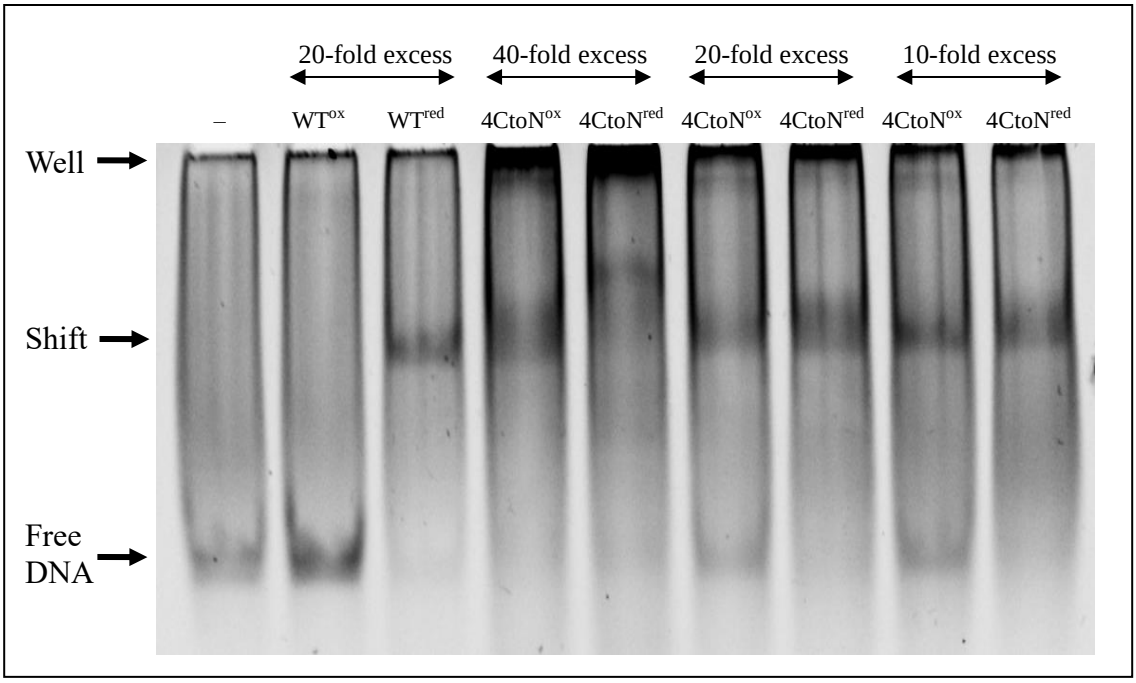


Figure 16. EMSA of MaMsvR^{4CtoN} with P_{msvR}.

EMSA to test binding of MaMsvR^{4CtoN} titrated over P_{msvR} (50 nM) when oxidized with H_2O_2 (2 µM) or reduced with DTT (5 mM). Each shift represents MaMsvR^{WT} or MaMsvR^{4CtoN} complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

Analysis of P_{msvr} Binding by MaMsvR Variants Oxidized with Paraquat and Diamide

Oxidation of cysteine thiols depends on the specificity of oxidizing agents, solvent accessibility, thiol pKa, and a multitude of other factors (48). To get a better idea of which cysteines in MaMsvR are susceptible to oxidation, single MaMsvR variants were oxidized with diamide and paraquat and compared to EMSA results from H_2O_2-induced cysteine oxidation. Diamide is a thiol-specific oxidizing agent that promotes disulfide bond formation (53). Paraquat induces oxidative stress by producing the superoxide anion ($^{\bullet}O_2^-$) that reacts with free thiols to produce disulfide bonds (54). Variants MaMsvR^{C63A}, MaMsvR^{C89A}, MaMsvR^{C148A} and MaMsvR^{C175A} were purified and run through SEC prior to EMSA analysis. All EMSAs of MaMsvR variants oxidized with diamide or paraquat were only performed once so results from these data are preliminary.

MaMsvR^{WT} (20-fold excess over P_{msvR}) did not bind P_{msvR} when oxidized with diamide, H_2O_2 or paraquat (Figures 17-20). This indicates all three oxidizing agents successfully oxidized MaMsvR^{WT} under the conditions tested. At 20-fold excess MaMsvR^{C63A} over P_{msvR}, variant MaMsvR^{C63A} bound P_{msvR} when oxidized with H_2O_2 and paraquat, but did not bind P_{msvR} when oxidized with diamide (Figure 17). Variant MaMsvR^{C89A} partially bound P_{msvR} when oxidized with H_2O_2, fully bound with paraquat and did not bind when oxidized with diamide (Figure 18). Variant MaMsvR^{C148A} bound P_{msvR} when oxidized with H_2O_2 but did not bind when oxidized with diamide or paraquat (Figure 19).

Combined, the results suggest that diamide is a powerful oxidizing agent since MaMsvR^{C63A}, MaMsvR^{C89A} and MaMsvR^{C148A} all bound P_{msvR} either fully or partially

when oxidized with H_2O_2 but none of them bound P_{msvR} when oxidized with diamide. Paraquat was not a strong agent for oxidizing MaMsvR variants because MaMsvR^{C63A} and MaMsvR^{C89A} still bound P_{msvR} after treatment with paraquat (2 µM). Paraquat was not used as an oxidizing agent for analyses of other MaMsvR variants. The results suggest each cysteine in MaMsvR is sensitive to different oxidants and the activation of MaMsvR may depend on which agent is causing oxidative stress. However, it is important to note that the analyses were preliminary and have not yet been repeated to confirm findings.

Variants MaMsvR^{C175A}, MaMsvR^{C114A} and MaMsvR^{C206A} were oxidized with diamide (2 µM) for analysis by EMSA (Figure 20). Similar to other MaMsvR variants, MaMsvR^{C175A} and MaMsvR^{C114A} did not bind P_{msvR} when oxidized with diamide. MaMsvR^{C206A} was the only variant that bound P_{msvR} when oxidized with diamide. Two shifts were produced for MaMsvR^{C206A} binding P_{msvR} under both oxidized and reduced conditions and all shifts with MaMsvR^{C206A} represent partial binding to P_{msvR}. Another lab member purified the MaMsvR^{C206A} variant used for this analysis and the variant produced the same two shifts when oxidized with H_2O_2 (not shown), indicating that something was wrong with this preparation. One explanation is partial degradation of the protein resulted in complexes of varying molecular weight or aggregated protein. The results support the hypothesis that C206 is a key player in redox sensing and/or conformational change because all other variants that previously bound P_{msvR} when oxidized with H_2O_2 were more sensitive than MaMsvR^{C206A} to diamide oxidation.

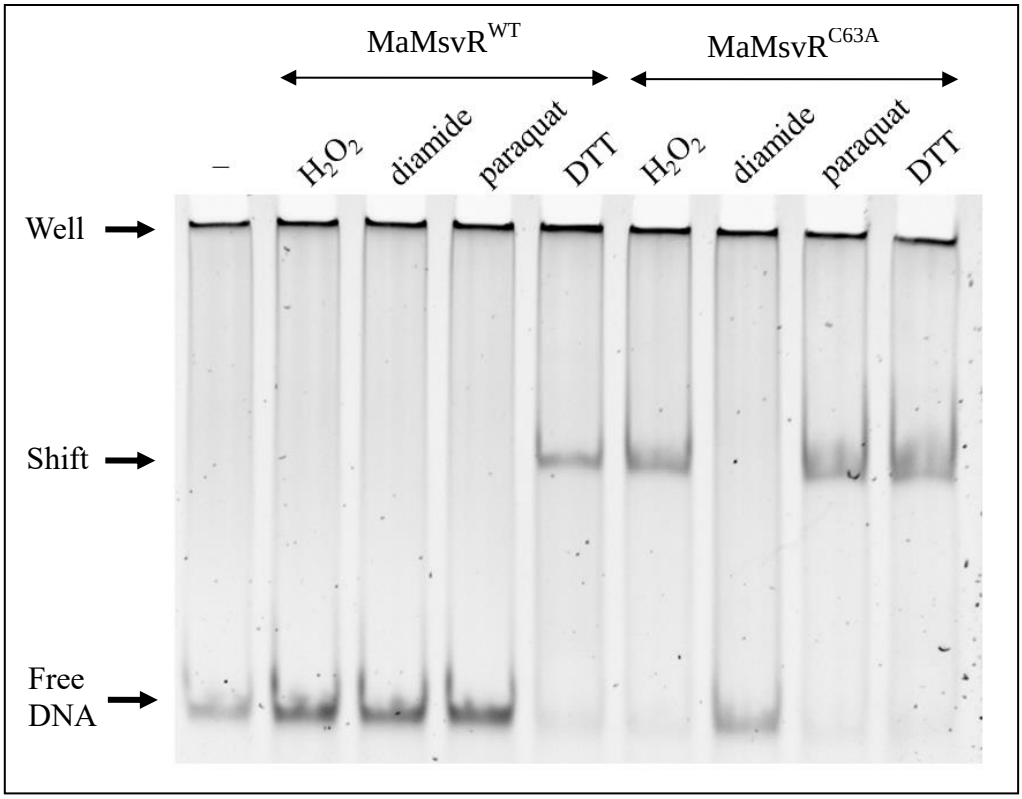


Figure 17. EMSA of MaMsvR^{C63A} with P_{msvR} after exposure to different oxidants. EMSA to test binding of MaMsvR^{C63A} (20-fold excess) to P_{msvR} (50 nM) when oxidized with H_2O_2, diamide or paraquat (all at 2 µM concentration) or reduced with DTT (5 mM). Each shift represents MaMsvR^{WT} or MaMsvR^{C63A} complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA Gel wells are indicated.

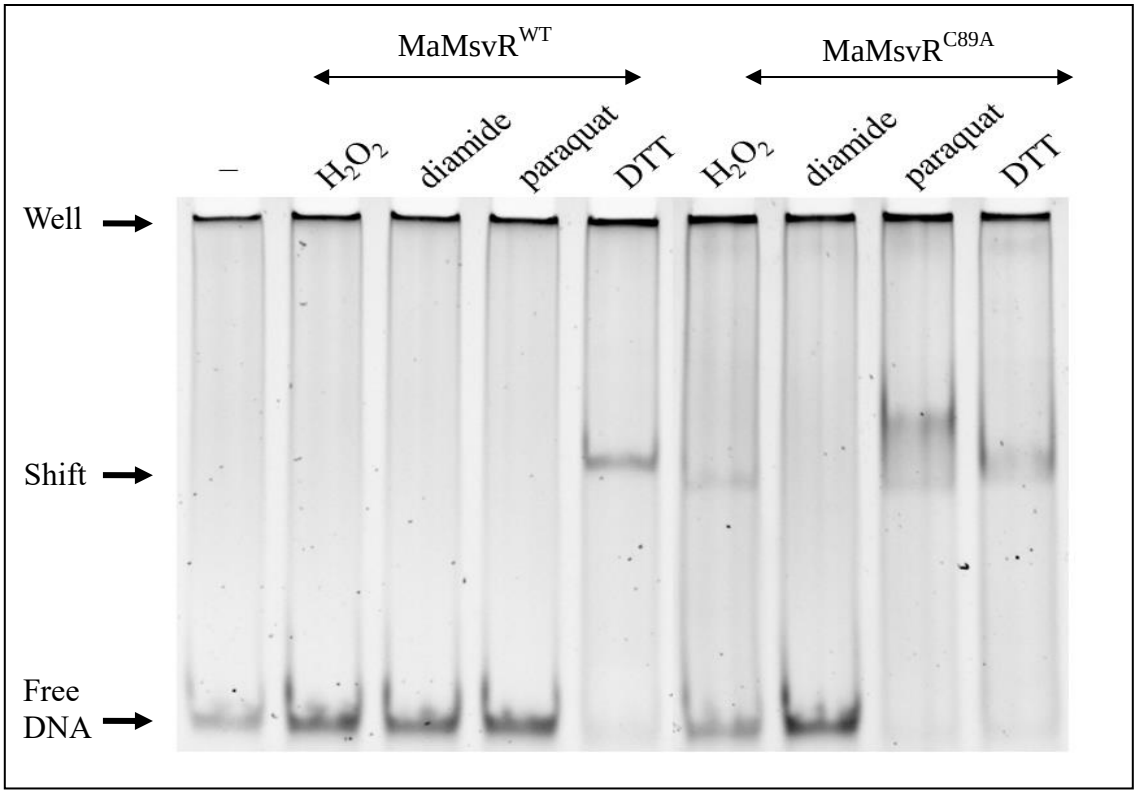


Figure 18. EMSA of MaMsvR^{C89A} with P_{msvR} after exposure to different oxidants. EMSA to test binding of MaMsvR^{C89A} (20-fold excess) to P_{msvR} (50 nM) when oxidized with H_2O_2, diamide or paraquat (all at 2 µM concentration) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

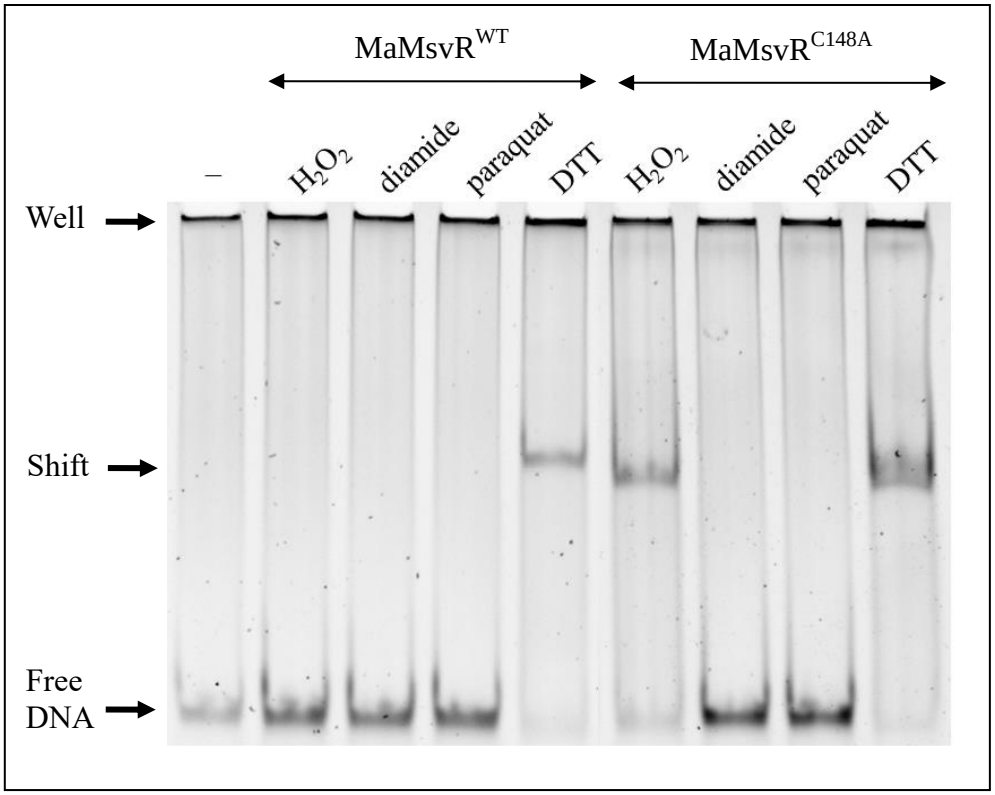


Figure 19. EMSA of MaMsvR^{C148A} with P_{msvR} after exposure to different oxidants. EMSA to test binding of MaMsvR^{C148A} (20-fold excess) to P_{msvR} (50 nM) when oxidized with H_2O_2, diamide or paraquat (all at 2 µM concentration) or reduced with DTT (5 mM). Each shift represents WT or variant MaMsvR complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

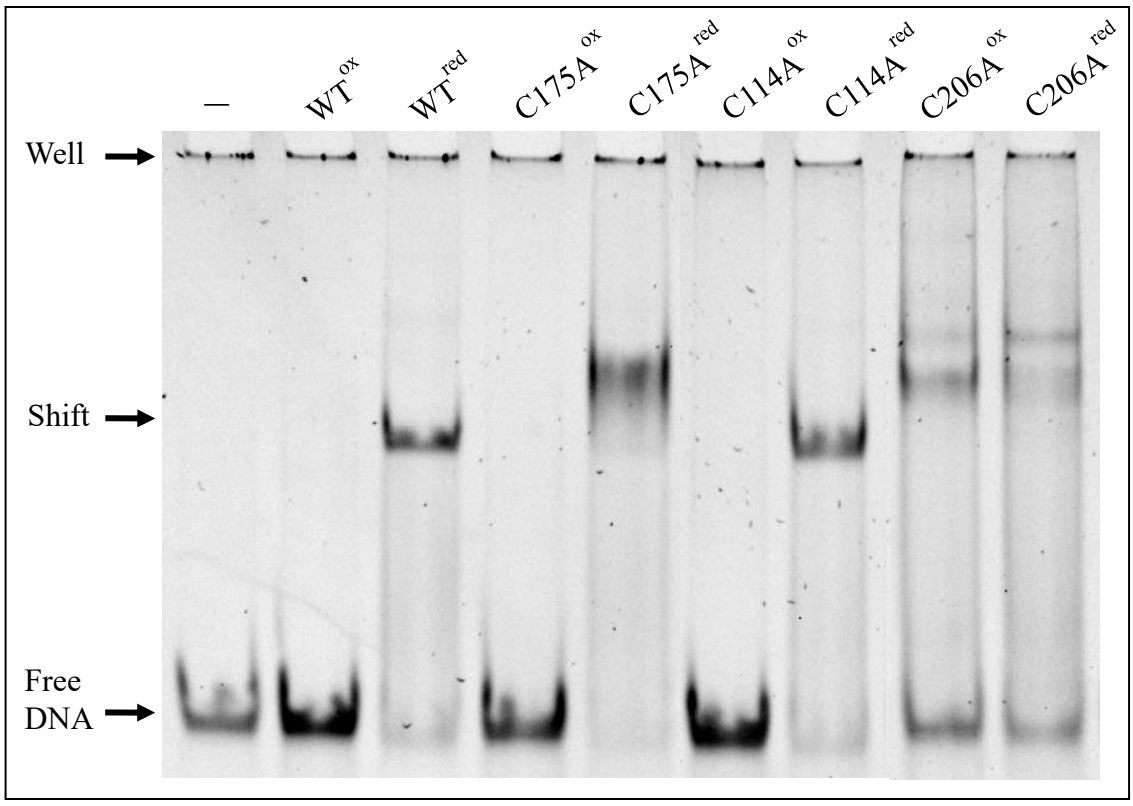


Figure 20. EMSA of MaMsvR variants with P_{msvR} after exposure to diamide. EMSA to test binding of $MaMsvR^{C175A}$, $MaMsvR^{C114A}$ and $MaMsvR^{C206A}$ (all 20-fold excess over P_{msvR}) to P_{msvR} (50 nM) when oxidized with diamide (2 μM) or reduced with DTT (5 mM). Each shift represents $MaMsvR^{WT}$ or variant MaMsvR complexed with P_{msvR}. The control lane (–) contains no protein. Unbound P_{msvR} is labeled as free DNA. Gel wells are indicated.

Conclusion

The EMSA results for variants MaMsvR^{C63A}, MaMsvR^{C89A}, MaMsvR^{C148A}, MaMsvR^{C175A}, MaMsvR^{C206A}, MaMsvR^{C232A} and MaMsvR^{C240S} produced shifts when oxidized with H_2O_2, indicating that the substitutions in these variants prevented the conformational change to release MaMsvR from P_{msvR}. This implicates seven cysteines, C63, C89, C148, C175, C206, C232 and C240, as key residues for redox sensing and/or conformational change in MaMsvR. Variants MaMsvR^{C45A}, MaMsvR^{C114A} and MaMsvR^{C225A} did not produce shifted bands when oxidized with H_2O_2, indicating C45, C114 and C225 are not involved in redox sensing or conformational change.

The lack of binding under oxidized conditions to P_{msvR} by the triple variant containing MaMsvR^{C45A}, MaMsvR^{C114A} and MaMsvR^{C225A} further suggests that C45, C114 and C225 are not involved in MaMsvR binding P_{msvR}. Binding to P_{msvR} under oxidized conditions was restored in the quadruple variant containing MaMsvR^{C45A}, MaMsvR^{C114A}, MaMsvR^{C225A} and MaMsvR^{C240S} and further implicates C240 as a key residue in MaMsvR binding P_{msvR}. The results from oxidizing MaMsvR variants with H_2O_2, diamide and paraquat suggest that diamide is the strongest of the three oxidizing agents for the oxidation of MaMsvR. The only variant that still bound P_{msvR} when oxidized with diamide was MaMsvR^{C206A}. This further implicates C206 as a key residue in redox sensing and/or conformational change in MaMsvR.

Of the seven cysteine residues putatively involved in MaMsvR redox sensing and/or conformational change, only three residues, C206, C232 and C240, are within the distance cutoff to form disulfide bonds. These three cysteines are located in the V4R domain of MaMsvR and are conserved in MthMsvR. The V4R domain has been

previously implicated in redox sensing and it can be postulated that C206, C232 and/or C240 may be involved in redox sensing. The formation of at least one intermolecular bond is probable because seven cysteines are implicated in disulfide bond formation. Further studies are needed to determine the number of intra- and intermolecular disulfide bonds, identify disulfide partners and elucidate the redox-sensing mechanism of MaMsvR.

Chapter 3: Preliminary Analysis of the MaMsvR Regulon by RNA-seq

Introduction

Methanosarcinae are a diverse group of methanogens that are unparalleled in terms of metabolism, physiology and environmental versatility. The *M. acetivorans* genome was sequenced in 2002 and the species' versatility was demonstrated in its 5.7 million bp genome, which is the largest genome known for an archaeon (55). Genes encoding function in two methanogenesis pathways were found, allowing *M. acetivorans* to grow on a variety of substrates. The two methanogenesis pathways are methylotrophic and acetoclastic, which produce methane from methylated compounds and acetate, respectively.

In the methylotrophic methanogenesis pathway, the methyl group from methylated compounds (CH_3–X) is transferred to a corrinoid protein by a methyltransferase (MT1) that is specific for a particular substrate (56). A second substrate-specific methyltransferase (MT2) then demethylates the corrinoid protein to transfer the methyl group to coenzyme M (CoM) to produce the intermediate methyl-CoM (CH_3–CoM) that is central to the pathway. In the acetoclastic methanogenesis pathway, acetate is activated to acetyl-coenzyme A (acetyl-CoA) via two reactions catalyzed by acetate kinase (Ack) and phosphotransacetylase (Pta) (57). The acetyl-CoA decarbonylase/synthase (ACDS) complex then transfers the methyl group from acetyl-CoA to tetrahydrosarcinapterin (H_4SPt) to generate the intermediate methyl-H_4SPt and releases CoA and CO_2. *M. acetivorans* has two copies of the six-gene operon encoding for the ACDS complex and it has been postulated that this unusual duplication

may aid in differential expression of genes encoding acetoclastic methanogenesis functions (55).

The genome also indicated that *M. acetivorans* has a wide range of responses to external stimuli via a two-component signal transduction system (55). Classical two-component regulatory systems consist of a histidine kinase and a response regulator that function in signal transduction (58, 59). Signal transduction cascades may also include the second messenger 3', 5'-cyclic adenosine monophosphate (cAMP), which modulates activity of cAMP-dependent protein kinases and is one of the initiating components in these signal transduction cascades (60). The level of intracellular cAMP is controlled by cAMP phosphodiesterases (Cpd), which are controlled by extracellular signals. Histidine kinases sense specific environmental stimuli and autophosphorylate a key histidine residue in the kinase domain (61). The phosphate is then transferred to an aspartate residue on the receiver domain of the cognate response regulator. The activated response regulator then regulates transcription of genes necessary to respond to the environmental stimulus that was initially sensed by histidine kinase. The two-component signal transduction system in *M. acetivorans* contains 50 genes encoding histidine kinases and 18 response regulators (55).

Although many of pathways in *M. acetivorans* have been elucidated, little is known about the transcription regulators that modulate them. RNA-seq has been a vital tool in the elucidation of regulons of some transcription regulators in *M. acetivorans* to provide insights into the regulation of certain pathways. RNA-seq is a technique that uses deep-sequencing technologies for transcriptome profiling (62). RNA is extracted from cells and converted to a library of complementary DNA (cDNA) fragments. The

fragments are sequenced to obtain short sequences from both ends that are between 30-400 bp in length (62). Reads are then mapped to a reference genome and mRNA abundance is quantified to determine levels of gene expression. RNA-seq of a WT strain and a strain containing a gene deletion has been used to determine the role the gene product plays within the cell.

MreA (*Methanosarcina r*egulator of *e*nergy-converting metabolism) was characterized through RNA-seq as a global regulator of energy-converting metabolism in *M. acetivorans* (63). MreA was shown to be involved in the activation of genes encoding acetoclastic methanogenesis functions and repression of genes encoding methylotrophic methanogenesis functions when cells were grown on acetate. MreA was also shown to either directly or indirectly regulate five regulators in the methanol-specific regulator (Msr) family. Msr regulators modulate expression of genes encoding methyltransferases in the methylotrophic methanogenesis pathway (64). The findings implicate MreA as a regulator that facilitates the switch between growth substrates in *M. acetivorans* (63).

MaMsvR was postulated to be involved in the oxidative stress response in *M. acetivorans* due to its redox sensitivity *in vivo* and *in vitro* and the involvement of MsvR homologs in oxidative stress in hydrogenotrophic methanogens (37). RNA-seq was utilized for preliminary analysis of the MaMsvR regulon to gain insights into pathways it may regulate and determine its possible involvement in redox stress.

Materials and Methods

Culturing WT and Ma ΔmsvR Strains

The RNA-seq analysis compared transcript levels from a *M. acetivorans* strain containing an *msvR* deletion (Ma Δ*msvR*) to a WT *M. acetivorans* strain (Table 6). The WT strain was donated from the Metcalf lab at the University of Illinois and the Ma Δ*msvR* strain from the Lessner lab at the University of Arkansas. Anaerobic culture tubes were filled with 10 mL of high salt (HS) media and had a headspace containing 80% nitrogen and 20% CO_2 (Tables 7 and 8). Then 0.1 mL of 2% Na_2S stock solution and 0.125 mL of 100% methanol were added to each tube prior to inoculating the media with cells. Prior to inoculation, 0.1 mL of a puromycin stock solution (250 μg mL^{-1}) was added to tubes intended for inoculation with the Ma Δ*msvR* strain. A syringe was used to transfer 0.1 mL of each strain to a culture tube. Strains were grown in triplicate. Cultures were incubated at 37 °C and optical density (OD) was measured and recorded daily. Cultures were transferred weekly in the anaerobic chamber until they reached the desired OD.

RNA Extraction

When cell cultures reached an OD of approximately 0.5, cells were pelleted by centrifugation at 6,000 rpm for 2 minutes. Supernatant was removed and cells were resuspended in 1 mL of TRI Reagent® Solution (Ambion®) to lyse cells. Cells were stored at -80 °C overnight to aid in cell lysis. Once all samples were ready for extraction, 1 mL of 100% ethanol was added to each sample. All six of the samples were then split into three samples to serve as technical replicates for the RNA extraction

process. RNA was extracted using the Quick-RNA™ MiniPrep kit (Zymo Research) according to manufacturer's protocol. After RNA elution, the three technical replicates were pooled together so each biological replicate contained 180 µl of eluted RNA.

RNA Quantification and Preliminary Verification of RNA Quality

RNA was quantified using the NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). The quality of total RNA was initially checked by running samples on a 1% agarose gel at 120 V and 350 mA for 30 minutes. The gel was imaged with short wave UV light on a GelDoc™ system (Bio-Rad) and image color was inverted for ease of viewing.

RNA Quality Verification on Bioanalyzer

To ensure the extracted RNA was not degraded, RNA samples were analyzed using the RNA 6000 Nano Kit (Agilent) according to manufacturer's instructions. Samples were run on the 2100 Bioanalyzer System (Agilent).

cDNA Library Preparation and RNA-seq

All six RNA samples were sent to Tufts University Core Facility Genomics Core for preparation of cDNA libraries and RNA-seq. The Ribo-Zero rRNA Removal Kit (Illumina) was used to deplete rRNA from the samples. Then TruSeq Stranded mRNA Library Prep kit (Illumina) was used to prepare a strand-specific and directional library from mRNA. Samples were sequenced on the HiSeq 2500 system (Illumina) on high output mode to generate single reads of 50 base pairs.

RNA-seq Analysis

The reads for each sample were mapped to the *M. acetivorans* genome using the Bowtie 2 plugin in Geneious v8 (65). Genome assemblies that the reads were mapped to were then imported into SeqMonk v0.34.1 (Babraham Bioinformatics) for RNA-seq analysis. Data was quantitated using the RNA-seq quantitation pipeline. The R-based filter DESeq2 pipeline was used to determine differentially expressed genes based on raw counts for each gene (66). The false discovery rate (FDR) was calculated for each transcript using the DESeq2 pipeline. Differentially expressed genes with a fold-change (FC) less than 2.0 were excluded from the results unless they were deemed relevant to this study. Genes were separated into two groups depending on whether they were upregulated or downregulated in the Ma Δ*msvR* strain compared to the WT strain.

Table 6. *M. acetivorans* strains used for RNA-seq

Strain	**Genotype**	**Reference**
WWM73	WT	67
DJL36	Δ*msvR*	This study

Table 7. Recipe for anaerobic HS media

Anaerobic HS Media (per 1 L):

Component	**Amount added**
NaCl	23.4 g
$NaHCO_3$	3.8 g
KCl	1.0 g
$MgCl_2 \cdot 6\ H_2O$	11.0 g
$CaCl_2 \cdot 2\ H_2O$	0.3 g
Trace elements	10 mL
5x vitamin solution	2.0 mL
0.2% resazurin	0.5 mL

After heating at 55-60 °C with N_2/CO_2 bubbling for 30 minutes, add:

NH_4Cl	1 g
Cysteine	0.5 g
1 M KH_2PO_4	5 mL

Table 8. Recipes for solutions added to HS media

Trace elements (per 1 L):

Component	**Amount added**
nitrolotriacetic acid	1.5 g
$MgSO_4 \cdot 7\ H_2O$	3 g
$MnSO_4 \cdot H_2O$	0.5 g
NaCl	1 g
Ingredients were dissolved in ddH2O and the pH of the solution was brought up to ~7.0	
Post pH adjustment:	
$FeSO_4 \cdot 7\ H_2O$	0.1 g
$CoSO_4$ or $CoCl_2$	0.1 g
$CaCl_2$	0.1 g
$ZnCl_2$	0.08 g
$CuSO_4 \cdot 5\ H_2O$	0.1 g
$AlK(SO_4)_2$	0.1 g
H_3BO_3	0.1 g
Na molybdate	0.025 g
$NiCl_2 \cdot 6\ H_2O$	0.024 g

5x vitamin solution (per 1 L):

Component	**Amount added**
biotin	0.01 g
folic acid	0.01 g
pyridoxine-HCl	0.05 g
riboflavin	0.025 g
thiamine	0.025 g
nicotinic acid	0.025 g
pantothenic acid	0.025 g
vitamin B12	0.025 g
p-aminobenzoic acid	0.025 g
thioctic (lipoic) acid	0.025 g

Results

Rationale for RNA-seq Analysis

Elucidating the MaMsvR regulon is imperative to understanding the biological role of MaMsvR in *M. acetivorans*. The mutational analysis of individual cysteines in MaMsvR strengthened the hypothesis that MaMsvR senses oxidant *in vitro* via redox-sensing cysteine residues but it is currently not known how the redox state of MaMsvR affects gene expression or which genes are regulated by MaMsvR. RNA-seq was chosen for preliminary analysis of the MaMsvR regulon because it offers accurate quantitation of a wide range of expression levels, there is little to no background noise and it has higher specificity and sensitivity than microarrays which allows for detection of transcripts in low abundance (62).

RNA was extracted from WT and Δ*msvR M. acetivorans* strains grown on methanol for RNA-seq analysis (Table 5). The quality of total RNA was initially checked by running aliquots of RNA elutions on a non-denaturing agarose gel (Figure 21). The sharp and clear 16S and 23S rRNA bands indicated the RNA was intact, and the absence of higher molecular weight bands suggested the absence of DNA contamination. Although a DNA ladder was run alongside the RNA samples, the placement of RNA bands with respect to the DNA ladder was consistent with the expected rRNA sizes specified by Thermo Fisher Scientific. RNA samples were then run on the 2100 Bioanalyzer System (Agilent) for more in-depth analysis of RNA integrity. The results (not shown) corroborated previous findings that all RNA samples were intact.

Preliminary RNA-seq analysis was completed using the R-based package DESeq2 (Bioconductor) available through SeqMonk v0.34.1 (Babraham Bioinformatics). DESeq2 estimates variance-mean dependence in raw count data to test for differential expression based on a negative binomial distribution model (68). SeqMonk v0.34.1 has limited functionality for in-depth and exploratory analyses but was chosen as the tool for initial data preparation and quality control (69). Future RNA-seq analysis will utilize other bioinformatics tools that will provide a deeper analysis of differentially expressed genes between WT and Ma Δ*msvR* strains.

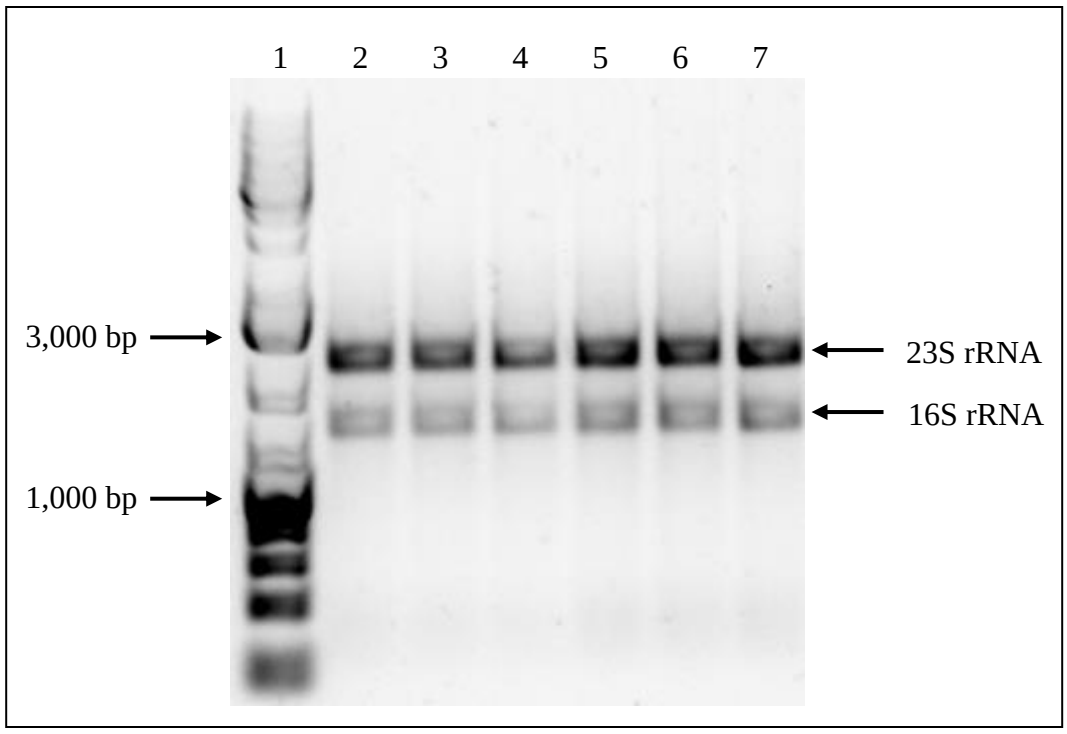


Figure 21. RNA integrity checked by agarose gel electrophoresis.

Lane 1 shows the GeneRuler™ 1 kilobase (kb) DNA ladder (Thermo Fisher Scientific).

Lanes 2-4 show 16S and 23S rRNA bands for the three Ma Δ*msvR* biological replicates.

Lanes 5-7 show rRNA bands for the three WT *M. acetivorans* replicates.

General RNA-seq Analysis of Differentially Expressed Genes

From the DESeq2 analysis, 432 genes ($p<0.05$, Wald test) were found to be differentially expressed in the Ma Δ*msvR* versus WT *M. acetivorans* strains. Of those 432 genes, 252 genes showed elevated expression and 180 genes showed depressed expression in the Ma Δ*msvR* strain compared to WT (Tables 10 and 11 in the appendix). A fold-change cutoff based on biological relevance was set at 2.0, which lead to 103 genes that passed the cutoff. Of those 103 genes, 61 genes were upregulated and 42 genes were downregulated (Ma Δ*msvR*/WT). However, all FC cutoffs are arbitrary and some genes that were below the cutoff but relevant to the study based on previous findings were analyzed further. The 103 differentially expressed genes with a FC≥2.0 were grouped based on arCOG functional categories (Table 9) (70). Notably, genes with putative involvement in the MaMsvR regulon function in methylotrophic and acetoclastic methanogenesis, transcription regulation and signal transduction.

Table 9. arCOG functional categories of differentially expressed genes (Ma Δ*msvR*/WT) with FC≥2.0.

arCOG Functional Category		**Upregulated Gene Count**	**Downregulated Gene Count**
C	Energy production and conversion	21	4
E	Amino acid transport and metabolism	8	1
G	Carbohydrate transport and metabolism	1	0
H	Coenzyme transport and metabolism	0	2
I	Lipid transport and metabolism	1	0
P	Inorganic ion transport and metabolism	1	5
K	Transcription	1	1
T	Signal transduction mechanisms	2	2
M	Cell wall/membrane/envelope biogenesis	1	1
R	General function prediction only	7	10
S	Function unknown	18	16

Preliminary Analysis of Genes Involved in Methanogenesis

Ten genes involved in methylotrophic methanogenesis were differentially expressed and 15 genes involved in acetoclastic methanogenesis showed elevated expression in the Ma Δ*msvR* strain compared to WT. For methylotrophic methanogenesis genes, only one gene was differentially expressed for each substrate-specific methyltransferase and corrinoid protein (Figure 22). MT1 homologs specific for methanol (MA4392), methylamine (MA2972), dimethylamine (MA2425) and trimethylamine (MA0528) metabolism were differentially expressed in the Ma Δ*msvR* strain compared to WT. This trend was also seen in genes encoding for corrinoid proteins, where corrinoid proteins specific for methanol (MA4391), methylamine (MA2971) and dimethylamine (MA2424) were differentially expressed (Ma Δ*msvR*/WT). One gene encoding a corrinoid protein specific for an unknown substrate (MA4164) was upregulated 4.1-fold in the Ma Δ*msvR* strain as well. A methanol-specific MT2 homolog was upregulated 1.6-fold in the Ma Δ*msvR* strain and two MT2 homologs specific for unknown substrates (MA2163 and MA4165) were differentially expressed. Two genes encoding dimethylsulfide-specific proteins with fused MT2 and corrinoid domains, *mtsD* (MA0859) and *mtsF* (MA4384), were upregulated 8.9-fold and 5.2-fold in the Ma Δ*msvR* strain, respectively.

The indirect or direct regulation of one isoform for each substrate-specific MT1, MT2 and corrinoid proteins involved in methylotrophic methanogenesis by MaMsvR suggests that MaMsvR may modulate expression of these genes based on which isozyme functions best under the growth conditions provided. A previous study found that *mtsF* is upregulated under a variety of redox stress conditions, so it is plausible that

MaMsvR represses *mtsF* in a redox-dependent manner (71). It cannot be determined from the data provided if MaMsvR regulates methylotrophic methanogenesis genes based on substrate availability, redox conditions, and/or other cellular factors.

Almost all of the genes involved in acetoclastic methanogenesis were upregulated in the Ma Δ*msvR* strain (Figure 23). All 12 genes comprising the duplicate ACDS operon along with a third, lone copy of CO dehydrogenase/acetyl-CoA synthase (*cdhA*) were upregulated by FCs greater than 2.0 in the Ma Δ*msvR* strain. The genes encoding Ack (MA3606) and Pta (MA3607) also showed elevated expression in the Ma Δ*msvR* strain compared to WT with FCs of 2.0 and 2.3, respectively. In a previous study, *M. acetivorans* cells grown on acetate were shown to have greater expression of genes involved in oxidative stress response compared to cells grown on methanol, indicating that an oxidative stress response is induced in *M. acetivorans* cells grown on acetate (72). This finding may explain the elevated expression levels of genes involved in acetoclastic methanogenesis in the Ma Δ*msvR* strain versus the WT strain. MaMsvR may be involved in repressing, either directly or indirectly, genes encoding acetoclastic methanogenesis function to prevent the cell from metabolizing acetate and thus protect it from undergoing oxidative stress. It can be postulated that adapted growth of the Ma Δ*msvR* strain to acetate would be significantly increased because the acetoclastic pathway is already activated.

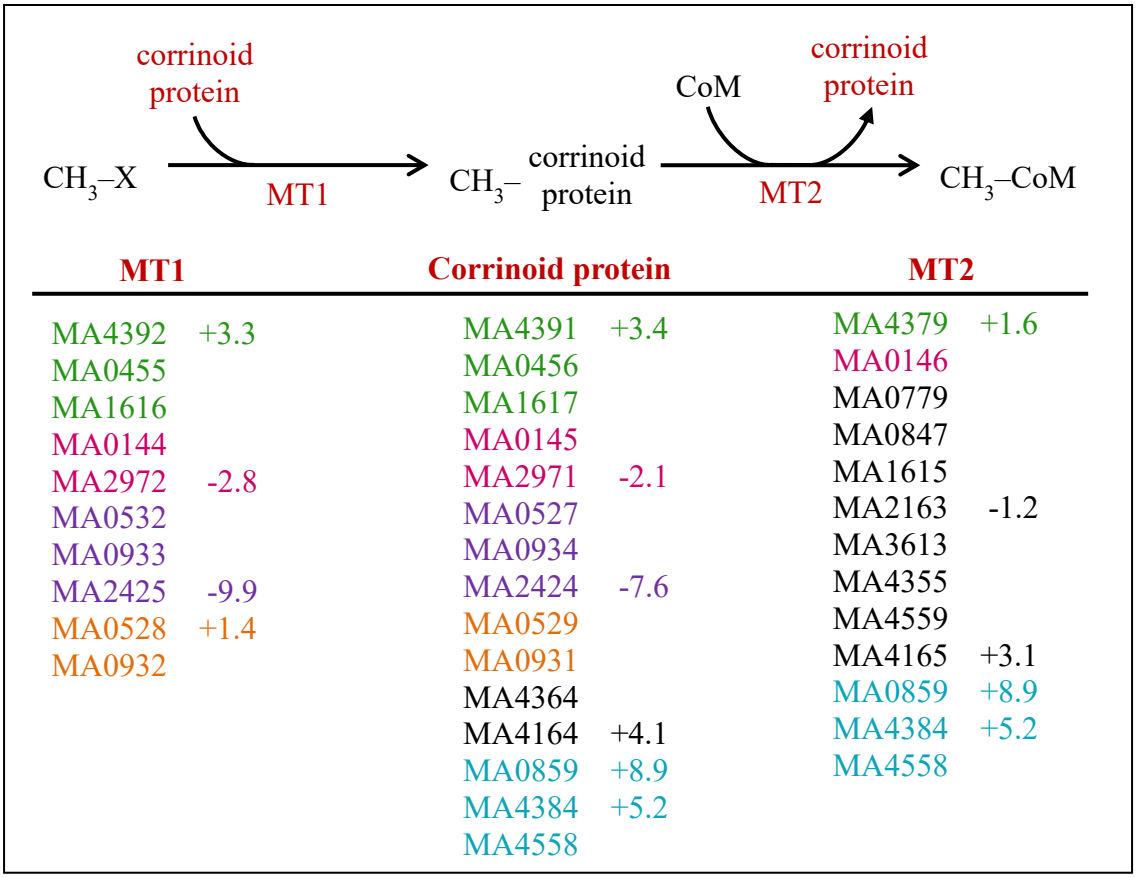


Figure 22. Differential expression of redundant genes encoding methylotrophic methanogenesis functions.

This figure has been adapted from a previous study (55). The methylotrophic methanogenesis pathway for the conversion of methylated compounds to the intermediate methyl-CoM. The gene number for substrate-specific methyltransferases and corrinoid proteins involved in methanol (green), methylamine (magenta), dimethylamine (purple) and trimethylamine (orange) metabolism are listed. Three genes (cyan) encode for dimethylsulfide-specific proteins with fused MT2 and corrinoid domains. The substrates for genes listed in black have not been identified yet. The expression FC (Ma Δ*msvR*/WT) for differentially expressed genes is listed to the right of the gene number.

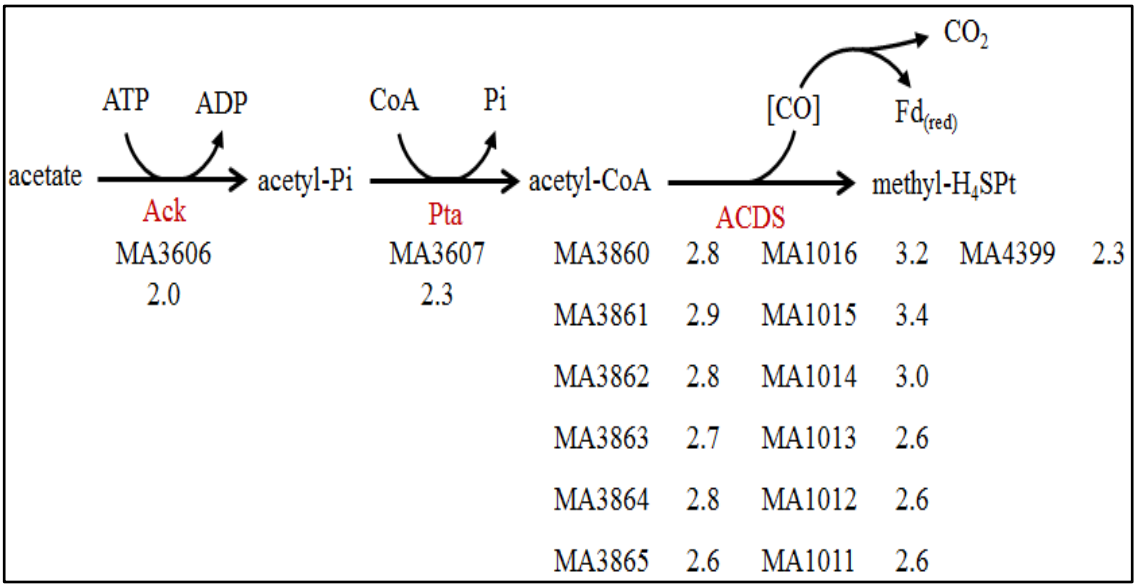


Figure 23. Upregulated genes in the Ma Δ*msvr* strain that encode acetoclastic methanogenesis functions.

This figure has been adapted from a previous study to only include data from this study (55). The acetoclastic methanogenesis pathway for the conversion of acetate to the intermediate methyl-H_4SPt. The gene numbers for Ack and Pta are listed with the expression FC (Ma Δ*msvr*/WT) beneath each gene number. Genes encoding two nearly identical copies of the ACDS complex are listed with the expression FC to the right of the gene number. A third, lone copy of *cdhA* (MA4399) is listed to the right of the other *cdhA* homologs.

Preliminary Analysis of Differentially Expressed Signal Transduction Genes

M. acetivorans contains 50 genes encoding histidine kinases and 18 response regulators (55). Two response regulators (RR), MA1366 and MA4671, were downregulated (Ma Δ*msvR*/WT) 1.8-fold and 1.7-fold, respectively. Although the FCs do not pass the cutoff of 2.0, it is probable that Ma *msvR* expression may be linked to direct or indirect activation of these two response regulators. A total of 11 histidine kinase genes were differentially regulated but only 3 passed the FC cutoff. The histidine kinase encoded by MA1957 was downregulated -3.5-fold and histidine kinase genes MA0620 and MA0203 were upregulated 2.9-fold and 2.0-fold, respectively. The remaining 8 genes encoding histidine kinases had FCs less than 2.0 and included 6 upregulated genes (MA0619, MA2553, MA1274, MA0777, MA1463 and MA2890) and 2 downregulated genes (MA1645 and MA1645).

One gene (MA1686) encoding cAMP phosphodiesterase (*cpdA*) was downregulated -3.5-fold (Ma Δ*msvR*/WT). This further implicates the potential involvement of MaMsvR in regulatory signal transduction cascades that affect gene expression. The higher number of upregulated histidine kinases suggest MaMsvR may be directly or indirectly involved in repressing these signal transduction kinases when grown on methanol. *M. acetivorans* cells grown in the presence of oxygen or other stressors may show increased expression of these signal transduction genes. It can be imagined that MaMsvR may be involved in a signal transduction regulatory cascade in which MaMsvR regulates expression of histidine kinases and response regulators that then regulate expression of genes involved in stress response(s). Two response regulators have been included in the gene regulatory network (Figure 24) with other

transcription regulators putatively involved in the MaMsvR regulon to depict a possible hierarchy of regulators in *M. acetivorans*.

Preliminary Analysis of Differentially Expressed Transcription Regulators

Two genes encoding putative transcription regulatory proteins were differentially expressed (FC≥2.0) in *M. acetivorans*. An ArsR family transcription regulator (MA0504) was expressed 2.1-fold greater in the Ma Δ*msvR* strain versus WT and SmtB (MA4344), an ArsR family efflux system transcriptional regulator, had a FC of -3.1. The findings indicate that the deletion of Ma *msvR* modulates gene expression, either directly or indirectly, by affecting the expression of other transcription regulators. This suggests MaMsvR may be involved in a regulatory hierarchy, although the specific functions of ArsR and SmtB in *M. acetivorans* and their interactions with MaMsvR have not been elucidated yet.

Five other transcription regulators were found to be differentially expressed but did not pass the FC cutoff of 2.0. Three regulators (MA3302, MA2914 and MA1424) showed upregulated expression and two regulators (MA1101 and MA2782) were downregulated.in the Ma Δ*msvR* strain versus WT. A global regulator involved in energy-converting metabolism, *mreA* (MA3302), had 1.5-fold greater expression in the Ma Δ*msvR* strain versus WT (63). MreA was found to modulate expression of genes encoding methanogenesis functions by activating acetoclastic methanogenesis genes and repressing methylotrophic methanogenesis genes when cells were grown on acetate, although it is unknown if this regulation is direct or indirect (63). The data suggests that MaMsvR represses expression of *mreA* either directly or indirectly. This supports the findings here that MaMsvR may be involved in modulating acetoclastic methanogenesis

when grown on methanol by repressing genes functioning in the pathway as well as repressing expression of at least one other transcription regulatory protein that regulates methanogenesis.

From the *mreA* RNA-seq analysis it was determined that Ma *msvR* had an expression ratio (Δ*mreA*/WT) of 1.7 when strains were grown on acetate (63). Taking this into account, along with the 1.5-fold greater expression (Ma Δ*msvR*/WT) of *mreA* when cells were grown on methanol, implies a mutual repression of both *mreA* and Ma *msvR* under different growth conditions. Both regulators modulate expression of many of the same genes and appear to be involved in the regulation of energy-converting metabolism as well as other regulators. A gene regulatory network has been constructed to depict the connection between MaMsvR and MreA and the transcription regulators differentially expressed in both data sets as well as genes involved in signal transduction and methanogenesis pathways (Figure 24). However, it is important to note that cells were grown on acetate for RNA-seq analysis of *mreA* and on methanol for Ma *msvR* RNA-seq analysis. It cannot be determined if regulation by either regulator is direct or indirect, so the mechanisms linking the two regulators cannot be elucidated yet. It is possible that the upregulation of genes encoding acetoclastic methanogenesis function in the Ma Δ*msvR* strain was due to *mreA* being upregulated in that strain as well. Further studies are needed to determine genes directly regulated by MaMsvR.

The two other predicted transcription regulators with elevated expression in Ma Δ*msvR* compared to WT, MA2914 and MA1424, both belong to the HTH-3 family of transcription regulators and had expression FCs of 1.35 and 1.26, respectively. These regulators have not been characterized yet but their upregulated expression in the Ma

Δ*msvR* strain suggests that the repertoire of gene regulation by MaMsvR may be expanded by its modulation of other transcription factors and response regulators.

Two regulators (MA1101 and MA2782) were downregulated in the Ma Δ*msvR* strain compared to WT. An uncharacterized regulator (MA1101) was downregulated 1.3-fold and a TetR family regulator (MA2782) was downregulated 1.7-fold. The TetR family of regulators is involved in transcription control of genes encoding stress response functions such as multidrug efflux pumps, osmotic stress response and removal and degradation of toxins (73). The deletion of Ma *msvR* from *M. acetivorans* causing decreased expression of MA2782 suggests that MaMsvR may be involved in a regulatory hierarchy of genes encoding stress response functions.

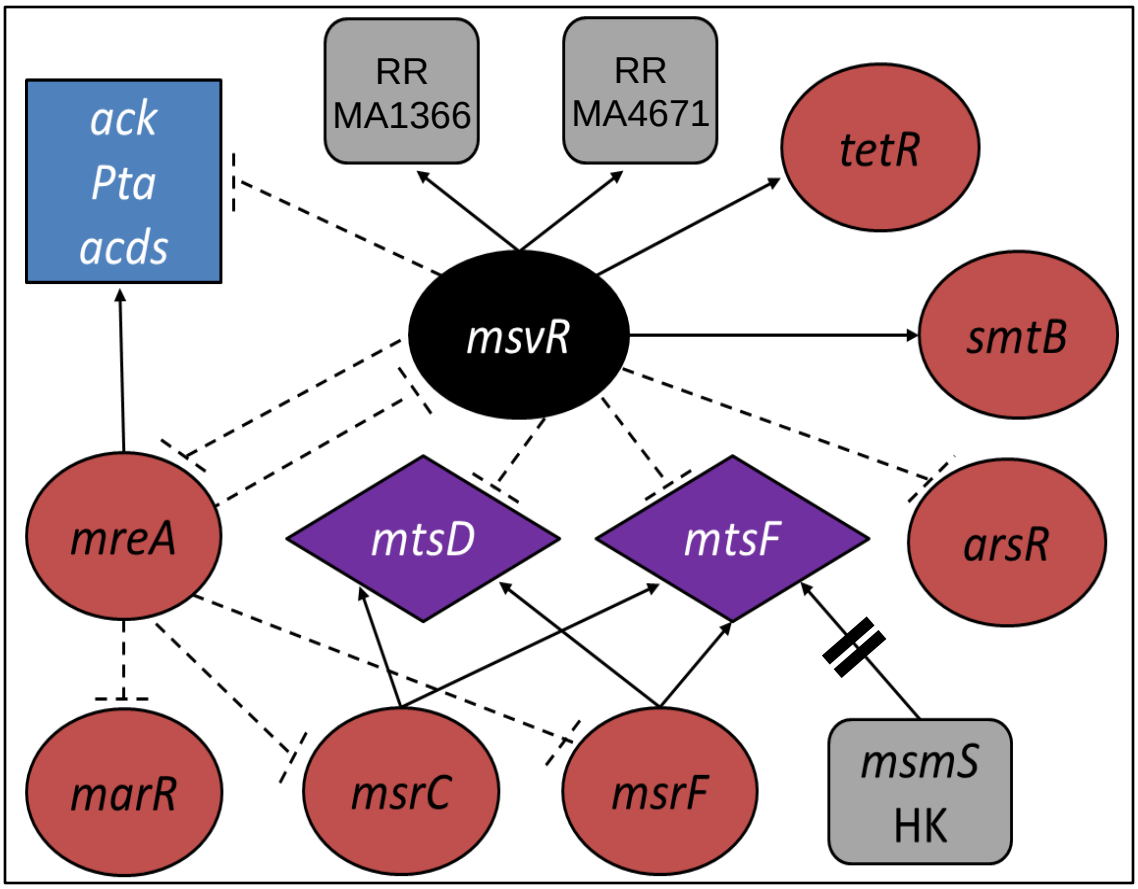


Figure 24. Predicted MaMsvR gene regulatory network.

The black oval represents Ma *msvR*. Transcription regulators putatively involved in MaMsvR and/or MreA regulons are depicted as dark red ovals. Genes encoding for two response regulators and one histidine kinase are depicted as gray rectangles. The blue rectangle represents genes involved in acetoclastic methanogenesis. Two genes encoding for methyltransferases involved in methylotrophic methanogenesis are depicted as purple diamonds. Dashed lines with blunt ends represent repression and solid arrows indicate activation or derepression by a transcription regulator. The double line indicates indirect regulation. All other regulatory connections may be either direct or indirect.

Conclusion

The preliminary RNA-seq results suggest MaMsvR directly or indirectly regulates genes encoding function in methanogenesis, signal transduction, transcription regulation and various other cellular functions. MaMsvR was initially believed to be involved in the oxidative stress response in *M. acetivorans* but the RNA-seq data implicates MaMsvR as a redox-sensitive regulator of central metabolism due to its direct or indirect involvement in: (i) repression of acetoclastic methanogenesis genes when cells were grown on methanol, (ii) activation and repression of genes encoding methylotrophic methanogenesis function and (iii) regulation of MreA, a regulator of both acetoclastic and methylotrophic methanogenesis pathways. In this way, MaMsvR may function similarly to SurR, a redox-sensitive regulator in *P. furiosus* that senses S^0 in the environment and regulates the switch from the production of H_2 to the production of H_2S (33). The findings indicate MaMsvR may be involved in the switch between growth substrates and/or the redox-dependent regulation of methanogenesis genes.

Further experiments are needed to better assess the MaMsvR regulon. RNA-seq experiments of *M. acetivorans* cells grown on different substrates and under different redox conditions will aid in the elucidation of the MaMsvR regulon and the determination of its involvement in redox stress.

References

1. IPCC. 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC. Cambridge, UK, Cambridge University Press.

2. **Ragsdale SW**. 2016. Targeting methanogenesis with a nitrooxypropanol bullet. PNAS. **113**(22):6100-6101.

3. **Moss AR, Jouany JP, Newbold J.** 2000. Methane production by ruminants: its contribution to global warming. Ann Zootech. **49**(3):231-253.

4. **Imlay JA**. 2002. How oxygen damages microbes: oxygen tolerance and obligate anaerobiosis. Adv Microb Physiol. **46**:111–153.

5. **Isom CE, Turner JL, Lessner DJ, Karr EA.** 2013. Redox-sensitive DNA binding by homodimeric *Methanosarcina acetivorans* MsvR is modulated by cysteine residues. BMC Microbiol. **13**:163.

6. **Kato S, Kosaka T, Watanabe K.** 2008. Comparative transcriptome analysis of responses of *Methanothermobacter thermautotrophicus* to different environmental stimuli. Environ Microbiol. **10**:893-905.

7. **Seedorf H, Dreisbach A, Hedderich R, Shima S, Thauer RK.** 2004. $F_{420}H_2$ oxidase (FprA) from *Methanobrevibacter arboriphilus*, a coenzyme F_{420}-dependent enzyme involved in O_2 detoxification. Arch Microbiol. **182**:126-137.

8. **Bell SD, Jackson SP.** 1998. Transcription and translation in Archaea: a mosaic of eukaryal and bacterial features. Trends Microbiol. **6**:222-228.

9. **Bell SD.** 2005. Archaeal transcriptional regulation – variation on a bacterial theme? Trends Microbiol. **13**:262-265.

10. **Karr EA.** 2014. Transcription Regulation in the Third Domain. Adv Appl Microbiol. **89**:101-133.

11. **Soppa J**. 1999. Transcription initiation in Archaea: facts, factors and future aspects. Mol Microbiol. **31**:1295–1305.

12. **Bell SD, Magill CP, and Jackson SP**. 2001. Basal and regulated transcription in Archaea. Biochem Soc Trans. **2**:392–395.

13. **Zillig W, Stetter KO, Tobien M.** 1978. DNA-dependent RNA polymerase from *Halobacterium halobium*. Eur J Biochem. **91**:193–199.

14. **Huet J, Schnabel R, Sentenac A, Zillig W.** 1983. Archaebacteria and eukaryotes possess DNA-dependent RNA polymerases of a common type. Embo J. **2**:1291–1294.

15. **Gohl H, Gröndahl B, Thomm M.** 1995. Promoter recognition in archaea is mediated by transcription factors: identification of transcription factor aTFB from *Methanococcus thermolithotrophicus* as archaeal TATA-binging protein. Nucl Acids Res. **23**:3837-3841.

16. **Hausner W, Wettach J, Hethke C, Thomm M**. 1996. Two transcription factors related with the eucaryl transcription factors TATA-Binding Protein and Transcription Factor IIB direct promoter recognition by an archaeal RNA polymerase. J Biol Chem. **270**:30144-30148.

17. **Browning DF, Busby SJW**. 2004. The regulation of bacterial transcription initiation. Nature Rev Microbiol. **2**:57-65.

18. **Abate C, Patel L, Rauscher III FJ, Curran T.** 1990. Redox regulation of Fos and Jun DNA-binding activity *in vitro*. Science. **249**:1157-1161.

19. **Sen CK.** 1998. Redox signaling and the emerging therapeutic potential of thiol antioxidants. Biochem Pharmacol. **55**:1747-1758.

20. **Marshall HE, Merchant K, Stamler JS.** 2000. Nitrosation and oxidation in the regulation of gene expression. FASEB J. **14**:1889-1900.

21. **Gilbert HF.** 1990. Molecular and cellular aspects of thiol-disulfide exchange. Adv Enzymol Relat Area Mol Biol. **63**:69-172.

22. **Pineda-Molina E, Klatt P, Vazquez J, Marina A, Garcia de Lacoba M, Perez-Sala D, Lamas S.** 2001. Glutathionylation of the p50 subunit of NF-κB: a mechanism for redox-induced inhibition of DNA binding. Biochemistry. **40**:14134-14142.

23. **Kim SO, Merchant K, Nudelman R, Beyer WF Jr, Keng T, DeAngelo J, Hausladen A, Stamler JS.** 2002. OxyR: a molecular code for redox-related signaling. Cell. **109**:383–396.

24. **Christman MF, Storz G, Ames BN.** 1989. OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. Proc Natl Acad Sci USA. **86**(10):3484-3488.

25. **Zheng M, Åslund F, Storz G.** 1998. Activation of the OxyR transcription factor by reversible disulfide bond formation. Science. **279**:1718–1721.

26. **Michán C, Manchado M, Dorado G, Pueyo**. 1999. In vivo transcription of the *Escherichia coli oxyR* regulon as a function of growth phase and in response to oxidative stress. J Bacteriol. **181**(9):2759-2764.

27. **Ehira S, Ohmori M.** 2012. The Redox-sensing Transcriptional Regulator RexT Controls Expression of Thioredoxin A2 in the Cyanobacterium *Anabaena* sp. Strain PCC 7120. J Biol Chem. **287**(48):40433-40440.

28. **Bsat N, Herbig A, Casillas-Martinez L, Setlow P, Helmann JD.** 1998. *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. Mol Microbiol. **29**(1):189-198.

29. **Lee JW, Helmann JD.** 2006. The PerR transcription factor sense H_2O_2 by metal-catalyzed histidine oxidation. Nature. **440**:363-367.

30. **Jacquamet L, Traoré DAK, Ferrer J, Proux O, Testemale D, Hazemann J, Nazarenko E, El Ghazouani A, Caux-Thang C, Duarte V, Latour J**. 2009. Structural characterization of the active form of PerR: insights into the metal-induced activation of PerR and Fur proteins for DNA binding. Mol Microbiol. **73**(1):20-31.

31. **Lipscomb GL, Keese AM, Cowart DM, Schut GJ, Thomm M, Adams MW, Scott RA.** 2009. SurR: a transcriptional activator and repressor controlling hydrogen and elemental sulphur metabolism in *Pyrococcus furiosus*. Mol Microbiol. **71**:332–349.

32. **Fiala G, Stetter KO**. 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch Microbiol. **145**:56–61.

33. **Yang H, Lipscomb GL, Keese AM, Schut GJ, Thomm M, Adams MW, Wang BC, Scott RA.** 2011. SurR regulates hydrogen production in *Pyrococcus furiosus* by a sulfur-dependent redox switch. Mol Microbiol. **77**:1111-1122.

34. **Sharma K, Gillum N, Boyd JL, Schmid AK.** 2012. The RosR transcription factor is required for gene expression dynamics in response to extreme oxidative stress in a hypersaline-adapted archaeon. BMC Genomics. **13**:351.

35. **Tonner PD, Pittman AMC, Gulli JG, Sharma K, Schmid AK.** 2015. A regulatory hierarchy controls the dynamic transcriptional response to extreme oxidative stress in Archaea. PLoS Genet. **11**(1):e1004912.

36. **Sheehan R, McCarver AC, Isom CE, Karr EA, Lessner DJ**. 2015. The *Methanosarcina acetivorans* thioredoxin system activates DNA binding of the redox-sensitive transcriptional regulator MsvR. J Ind Microbiol Biotechnol. **42**:965-969.

37. **Karr EA.** 2010. The methanogen-specific transcription factor MsvR regulates the *fpaA-rlp-rub* oxidative stress operon adjacent to *msvR* in *Methanothermobacter thermautotrophicus*. J Bacteriol. **192**:5914-5922.

38. **Quan S, Schneider I, Pan J, Von Hacht A, Bardwell JC**. 2007. The CXXC motif is more than a redox rheostat. J Biol Chem. **282**(39):28823-28833.

39. **Arnér ES, Holmgren A**. 2000. Physiological functions of thioredoxin and thioredoxin reductase. Eur J Biochem. **267**(20):6102-6109.

40. **McCarver AC, Lessner DJ.** 2014. Molecular characterization of the thioredoxin system from *Methanosarcina acetivorans*. FEBS J. **281**(20):4598–4611.

41. **Sambrook J, Russell DW.** 2001. Molecular Cloning: A laboratory manual. Third ed. Cold Spring Harbor Laboratory Press, New York. pp1.112-1.115.

42. **Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE**. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. **10**:845-858.

43. The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

44. **Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE**. 2004. UCSF Chimera – a visualization system for exploratory research and analysis. J Comput Chem. **25**(13):1605-1612.

45. **Sanner MF, Olson AJ, Spehner JC**. 1996. Reduced surface: an efficient way to compute molecular surfaces. Biopolymers. **38**(3):305-320.

46. **Rubinstein R, Fiser A.** 2008. Predicting disulfide bond connectivity in proteins by correlated mutations analysis. Bioinformatics. **24**(4):498-504.

47. **Petsko GA, Ringe D**. 2004. Protein Structure and Function. Sunderland, MA: New Science Press Ltd.

48. **Sanchez R, Riddle M, Woo J, Momand J.** 2008. Prediction of reversibly oxidized protein cysteine thiols using protein structure properties. Protein Sci. **17**(3):473-481.

49. **Choi HJ, Kim SJ, Mukhopadhyay P, Cho S, Woo JR, Storz G, Ryu SE.** 2001. Structural basis of the redox switch in the OxyR transcription factor. Cell. **105**:103-113.

50. **Lee C, Lee SM, Mukhopadhyay P, Kim SJ, Lee SC, Ahn WS, Yu MH, Storz G, Ryu SE.** 2004. Redox regulation of OxyR requires specific disulfide bond formation involving a rapid kinetic reaction path. Nat Struct Mol Biol. **11**:1179-1185.

51. **Martìnez-Antonio A, Collado-Vides J.** 2003. Identifying global regulators in transcriptional regulatory networks in bacteria. Curr Opin Microbiol. **6**:482–489.

52. **Storz G, Tartaglia LA, and Ames BN.** 1990. Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. Science. **248**:189–194.

53. **Obin M, Shang F, Gong X, Handelman G, Blumberg J, Taylor A.** 1998. Redox regulation of ubiquitin-conjugating enzymes: mechanistic insights using the thiol-specific oxidant diamide. FASEB J. **12**(7):561-569.

54. **Liu D, Yang J, Li L, McAdoo DJ**. 1995. Paraquat - A superoxide generator - Kills neurons in the rat spinal cord. Free Radic Biol Med. **18**(5):861-867.

55. **Galagan JE, Nusbaum C, Roy A, Endrizzi MG, Macdonald P, FitzHugh W, Calvo S, Engels R, Smirnov S, Atnoor D, Brown A, Allen N, Naylor J, Strange-Thomann N, DeArellano K, Johnson R, Linton L, McEwan P, McKernan K, Talamas J, Tirrell A, Ye W, Zimmer A, Barber RD, Cann I, Graham DE, Grahame DA, Guss AM, Hedderich R, Ingram-Smith C, Kuettner HC, Krzycki JA, Leigh JA, Li W, Liu J, Mukhopadhyay B, Reeve JN, Smith K, Springer TA, Umayam LA, White O, White RH, Conway de Macario E, Ferry JG, Jarrell KF, Jing H, Macario AJL, Paulsen I, Pritchett M, Sowers KR, Swanson RV, Zinder SH, Lander E, Metcalf WM, Birren B.** 2002. The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. Genome Res. **12**(4):532-42.

56. **Ferry JG.** 1999. Enzymology of one-carbon metabolism in methanogenic pathways. FEMS Microbiol Rev. **23**(1):13-38.

57. **Ferry JG**. 2011. Acetate kinase and phosphotransacetylase. Methods Enzymol. **494**:219-231.

58. **Ronson CW, Nixon BT, Ausubel FM**. 1987. Conserved domains in bacterial regulatory proteins that respond to environmental stimuli. Cell. **49**:579-581.

59. **Gao R, Stock AM**. 2009. Biological insights from structures of two-component proteins. Annual Rev Microbiol. **63**:133-154.

60. **Thomason PA, Traynor D, Cavet G, Chang WT, Harwood AJ, Kay RR**. 1998. An intersection of the cAMP/PKA and two-component signal transduction systems in *Dictyostelium*. EMBO J. **17**:2838-2845.

61. **Hoch JA**. 2000. Two-component and phosphorelay signal transduction. Curr Opin Microbiol. **3**(2):165-170.

62. **Wang Z, Gerstein M, Snyder M**. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet. **10**(1):57-63.

63. **Reichlen MJ, Vepachedu VR, Murakami KS, Ferry JG**. 2012. MreA functions in the global regulation of methanogenic pathways in *Methanosarcina acetivorans*. mBio. **3**(4):e00189-12.

64. **Bose A, Metcalf WW**. 2008. Distinct regulators control the expression of methanol methyltransferase isozymes in Methanosarcina acetivorans C2A. Mol Microbiol. **67**(3):649-661.

65. **Langmead B, Salzberg S**. 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods. **9**:357-359.

66. **Anders S, Huber W**. 2010. Differential expression analysis for sequence count data. Genome Biol. **11**:R106.

67. **Guss AM, Rother M, Zhang JK, Kulkarni G, Metcalf WW**. 2008. New methods for tightly regulated gene expression and highly efficient chromosomal integration of cloned genes for Methanosarcina species. Archaea. **2**:193–203.

68. **Love MI, Huber W, Anders S**. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. **15**:550.

69. **Younesy H, Möller T, Lorincz MC, Karimi MM, Jones SJM**. 2015. VisRseq: R-based visual framework for analysis of sequencing data. BMC Bioinformatics. **16**(Suppl 11):S2.

70. **Makarova KS, Sorokin AV, Novichkov PS, Wolf YI, Koonin EV**. 2007. Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. Biol Direct. **2**:33.

71. **Molitor B, Stassen M, Modi A, El-Mashtoly SF, Laurich C, Lubitz W, Dawson JH, Rother M, Frankenberg-Dinkel N.** 2013. A heme-based redox sensor in the methanogenic archaeon *Methanosarcina acetivorans*. J Biol Chem. **288**(25):18458-18472.

72. **Li L, Li Q, Rohlin L, Kim U, Salmon K, Rejtar T, Gunsalus RP, Karger BL, Ferry JG**. 2007. Quantitative proteomic and microarray analysis of the archaeon *Methanosarcina acetivorans* grown with acetate versus methanol. J Proteome Res. **6**:759-771.

73. **Ramos JL, Martínez-Bueno M,Molina-Henares AJ, Terán W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R**. 2005. The TetR family of transcription repressors. Microbiol Mol Biol Rev. **69**(2):326-356.

Appendix: RNA-seq Data Tables

Table 10. Upregulated transcripts in Ma Δ*msvR*

Locus ID	Description	arCOG	FDR	Expression Difference (Ma Δ*msvR*-WT)	Fold Change
MA0859	Methanogenic corrinoid protein with fused MT2 domain (*mtbC1*)	C	0	3.15	8.87
MA4563	hydantoinase (*hyuA*)	E	0	2.61	6.09
MA1124	hypothetical protein	S	7.09E-22	2.59	6.03
MA4384	Methanogenic corrinoid protein with fused MT2 domain (*mtbC1*)	C	3.91E-17	2.39	5.24
MA0848	hypothetical protein	S	3.19E-15	2.20	4.60
MA3319	hypothetical protein	S	1.05E-14	2.15	4.44
MA3318	hypothetical protein	S	3.56E-25	2.11	4.33
MA4164	Methanogenic corrinoid protein (*mtbC1*)	C	7.95E-19	2.05	4.15
MA4391	Methanogenic corrinoid protein (*mtaC2*)	C	0.0028	1.80	3.48
MA1015	ACDS subunit epsilon (*cdhB*)	C	8.57E-21	1.75	3.36
MA4392	methyltransferase isozyme 1 (*mtaB2*)	C	0.0018	1.74	3.33
MA1016	ACDS subunit alpha (*cdhA*)	C	1.80E-35	1.71	3.28
MA4008	Uncharacterized transmembrane protein	S	4.40E-17	1.71	3.28
MA1316	Na+/proline symporter, proline permease (*putP*)	E	1.81E-07	1.67	3.18
MA4165	methylcobalamin:CoM methyltransferase isozyme A (*cmtA*)	C	3.91E-16	1.62	3.08
MA1014	ACDS subunit gamma (*cdhC*)	C	1.71E-18	1.61	3.05
MA3753	Transmembrane oligosaccharyl transferase	M	0	1.58	3.00
MA0505	hypothetical protein	S	1.48E-11	1.56	2.95
MA3861	ACDS subunit epsilon (*cdhE*)	C	5.73E-07	1.56	2.95
MA0620	histidine kinase & HK-like ATPase	T	1.24E-10	1.51	2.86
MA3862	ACDS subunit gamma (*cdhC*)	C	3.77E-09	1.51	2.85
MA2986	Anthranilate/para-aminobenzoate synthase component II	E	8.76E-04	1.49	2.82
MA3864	ACDS subunit delta (*cdhD*)	C	3.70E-23	1.49	2.81
MA4166	sugar transport family protein	G	7.42E-32	1.49	2.81
MA1012	ACDS subunit delta (*cdhD*)	C	2.32E-21	1.48	2.79
MA3860	ACDS subunit alpha (*cdhA*)	C	2.27E-08	1.48	2.79
MA2990	tryptophan synthase subunit alpha (*trpA*)	E	0.0085	1.44	2.71
MA0858	Hydantoinase (*hyuA*)	E	1.47E-23	1.43	2.70
MA4394	Zinc transporter	R	0.0390	1.43	2.69
MA3863	CO dehydrogenase accessory	C	1.47E-16	1.42	2.68

	protein				
MA2987	tryptophan synthase subunit epsilon	E	0.0017	1.42	2.68
MA1013	CO dehydrogenase nickel-insertion accessory protein	C	4.94E-18	1.41	2.66
MA4385	hypothetical protein	S	6.28E-04	1.41	2.66
MA2991	Tryptophan synthase beta superfamily	E	0.0054	1.40	2.64
MA2139	hypothetical protein	S	5.34E-18	1.40	2.64
MA4393	Uncharacterized transmembrane protein	S	5.61E-04	1.39	2.63
MA3865	ACDS subunit epsilon (*cdhE*)	C	5.46E-26	1.39	2.63
MA4324	Peptidase family M3B Oligopeptidase F (*pepF*)	S	4.11E-17	1.39	2.62
MA2992	Indole-3-glycerol phosphate synthase (IGPS)	E	0.0110	1.38	2.59
MA1011	ACDS subunit gamma (*cdhC*)	C	9.69E-28	1.35	2.55
MA4256	hypothetical protein	S	0.0091	1.33	2.52
MA1123	hypothetical protein	S	4.05E-38	1.32	2.49
MA2783	hypothetical protein	S	2.91E-11	1.30	2.46
MA4257	hypothetical protein	S	0.0017	1.24	2.36
MA3607	Phosphate acetyltransferase involved in acetoclastic methanogenesis	C	3.56E-11	1.23	2.35
MA1102	Formylglycine-generating enzyme required for sulfatase activity	R	1.61E-35	1.23	2.34
MA4017	Nicotianamine synthase protein	R	7.04E-12	1.21	2.31
MA4399	ACDS subunit alpha (*cdhA*)	C	8.92E-09	1.20	2.30
MA1418	ABC transport system for cobalt	P	8.79E-25	1.17	2.26
MA4634	hypothetical protein	S	5.56E-04	1.15	2.22
MA2959	hypothetical protein	S	1.97E-22	1.15	2.22
MA1417	hypothetical protein	S	1.48E-10	1.14	2.21
MA3668	hypothetical protein	S	3.36E-09	1.12	2.18
MA2003	sodium-dependent transporter	R	6.46E-20	1.11	2.16
MA0345	hypothetical protein	S	0.0031	1.10	2.14
MA4365	ABC-2 family transporter protein,	R	2.92E-06	1.09	2.12
MA0504	ArsR family transcriptional regulator	K	1.24E-09	1.08	2.12
MA2244	coenzyme A ligase	I	2.54E-06	1.06	2.08
MA0203	signal-transducing histidine kinase	T	9.31E-29	1.03	2.04
MA3606	Acetate kinase A invovled in acetoclastic methanogenesis	C	2.68E-05	1.02	2.03
MA1690	CrtC superfamily	R	1.08E-05	1.02	2.03
MA3198	Tryptophan synthase subunit beta (*trpB*)	E	6.67E-17	0.99	1.99
MA2690	phosphoenolpyruvate carboxylase	C	0.0204	0.98	1.97
MA2272	ABC transporter, permease protein	R	1.21E-10	0.98	1.97

MA1028	long-chain fatty-acid-CoA ligase	I	9.69E-28	0.97	1.96
MA3666	hypothetical protein	S	3.99E-11	0.95	1.93
MA1027	hypothetical protein	S	9.85E-15	0.93	1.90
MA0161	ferritin	R	0.0059	0.93	1.90
MA1161	macrolide-efflux protein	V	4.11E-17	0.92	1.90
MA1030	NAD(+) synthase (glutamine-hydrolyzing) (*nadE*)	H	1.21E-08	0.92	1.90
MA1553	hypothetical protein	S	1.78E-14	0.92	1.90
MA3667	hypothetical protein	S	6.31E-09	0.90	1.87
MA3665	hypothetical protein	S	3.02E-07	0.88	1.84
MA0162	universal stress protein	R	0.0092	0.87	1.83
MA2750	hypothetical protein	S	1.17E-08	0.85	1.81
MA0431	ferredoxin	C	7.40E-07	0.84	1.78
MA1550	hypothetical protein	S	3.24E-05	0.83	1.78
MA0533	hypothetical protein	S	0.0091	0.81	1.75
MA2324	hypothetical protein	S	0.0077	0.80	1.75
MA2945	hypothetical protein	S	9.50E-09	0.80	1.74
MA3121	hypothetical protein	S	3.83E-15	0.79	1.73
MA0163	hypothetical protein	S	1.23E-04	0.79	1.73
MA2453	hypothetical protein	S	2.54E-06	0.79	1.73
MA4475	hypothetical protein	S	5.18E-05	0.77	1.70
MA2273	hypothetical protein	S	6.09E-16	0.76	1.69
MA0557	hypothetical protein	S	7.35E-12	0.75	1.68
MA2152	ubiquinone/menaquinone biosynthesis methyltransferase (*ubiE*)	H	0.0217	0.74	1.67
MA2021	hypothetical protein	S	3.90E-07	0.74	1.67
MA1463	sensory transduction histidine kinase	T	2.64E-04	0.74	1.67
MA1953	hypothetical protein	S	2.53E-06	0.73	1.66
MA0945	hypothetical protein	S	2.85E-05	0.73	1.66
MA0556	hypothetical protein	S	1.13E-09	0.72	1.65
MA0092	hypothetical protein	S	0.0271	0.71	1.64
MA2964	hypothetical protein	S	0.0385	0.71	1.64
MA0107	gluconate 5-dehydrogenase	R	2.08E-18	0.70	1.63
MA4379	Methyltransferase isozyme 2 (*mtaA*)	C	0.0016	0.70	1.62
MA2275	ABC transporter, ATP-binding protein	R	7.95E-19	0.70	1.62
MA4478	cell surface protein (*slg*)	M	0.0094	0.69	1.62
MA0425	hypothetical protein	S	2.54E-07	0.69	1.61
MA1545	hypothetical protein	S	0.0137	0.67	1.59
MA4562	hydantoinase (*hyuA*)	E	0.0271	0.67	1.59
MA0608	pyruvate phosphate dikinase (*ppdK*)	G	2.93E-05	0.66	1.58
MA1551	hypothetical protein	S	7.57E-04	0.65	1.57

MA1542	hypothetical protein	S	0.0021	0.64	1.56
MA0857	hsp60-5	O	0.0062	0.64	1.56
MA1725	phenylacetate-CoA ligase	Q	8.24E-13	0.63	1.55
MA2314	hypothetical protein	S	0.0031	0.63	1.55
MA0016	*cheY2*	T	3.56E-08	0.60	1.52
MA2812	hypothetical protein	S	0.0489	0.60	1.51
MA3302	*Methanosarcina* regulator of energy-converting metabolism (*mreA*)	K	3.84E-04	0.59	1.51
MA0707	caffeoyl-CoA O-methyltransferase	E	3.67E-04	0.59	1.51
MA0275	tetrahydromethanopterin S-methyltransferase, subunit D (*mtrD*)	H	1.48E-10	0.59	1.50
MA3443	hypothetical protein	S	0.0055	0.59	1.50
MA1500	$Fe_{420}H_2$ dehydrogenase subunit (*fpoI*)	C	0.0432	0.58	1.50
MA1726	indolepyruvate oxidoreductase, subunit beta (*iorB*)	C	1.52E-14	0.58	1.49
MA4036	ABC transporter, ATP-binding protein	R	3.35E-07	0.58	1.49
MA0860	hypothetical protein	S	1.21E-05	0.57	1.49
MA1244	oligopeptide ABC transporter, oligopeptide-binding protein	R	0.0015	0.57	1.48
MA2911	2-oxoisovalerate ferredoxin oxidoreductase, gamma subunit	C	6.72E-04	0.57	1.48
MA0199	hypothetical protein	S	6.39E-04	0.56	1.48
MA0674	pyruvate carboxylase subunit B	G	0.0021	0.56	1.47
MA0273	Tetrahydromethanopterin S-methyltransferase, subunit beta	H	1.93E-08	0.56	1.47
MA1022	indolepyruvate ferredoxin oxidoreductase, subunit alpha	C	6.99E-05	0.55	1.47
MA1660	hypothetical protein	S	1.22E-06	0.54	1.45
MA4386	hsp60-4	O	0.0020	0.53	1.45
MA4028	hypothetical protein	S	9.51E-04	0.53	1.44
MA2406	transposase	X	6.38E-04	0.53	1.44
MA1536	ABC transporter, ATP-binding protein	R	1.78E-05	0.53	1.44
MA1373	Fructose-1-phosphate kinase	G	9.85E-10	0.53	1.44
MA4349	Branched-chain amino acid aminotransferase/4-amino-4-deoxychorismate lyase	E	7.73E-05	0.53	1.44
MA3122	surface antigen gene	M	0.0027	0.53	1.44
MA4570	multiple resistance/pH regulation related protein C (Na+/H+ antiporter) (*mrpC*)	R	3.93E-04	0.53	1.44
MA1508	cell surface protein	M	0.003083604	0.52	1.44
MA4430	Coenzyme F_{420}-dependent 5, 10-methenyltetrahydromethanopterin dehydrogenase	C	4.51E-04	0.52	1.43

MA4035	hypothetical protein	S	0.0013	0.52	1.43
MA1498	$F_{420}H_2$ dehydrogenase subunit (*fpoD*)	C	0.0054	0.51	1.43
MA0693	hypothetical protein	S	0.0345	0.51	1.43
MA1653	integral membrane protein	S	0.0134	0.51	1.43
MA0675	pyruvate carboxylase subunit A (*pycA)*	G	0.0031	0.51	1.43
MA0528	methyltransferase isozyme 1	C	0.0440	0.51	1.43
MA4156	F0F1-type ATP synthase, epsilon subunit	C	0.0390	0.51	1.42
MA3988	transposase	X	0.0012	0.51	1.42
MA2429	hypothetical protein	S	1.48E-06	0.51	1.42
MA0269	Tetrahydromethanopterin S-methyltransferase, subunit H (*mtrH*)	H	4.13E-06	0.51	1.42
MA2909	2-oxoisovalerate ferredoxin oxidoreductase, alpha subunit (*vorA*)	C	7.38E-04	0.50	1.42
MA0980	hypothetical protein	S	0.0265	0.50	1.42
MA1248	*ubiE*/COQ5 methyltransferase	H	3.05E-04	0.50	1.41
MA4578	hypothetical protein	S	2.03E-04	0.50	1.41
MA1497	$F_{420}H_2$ dehydrogenase subunit (*fpoC*)	C	0.0061	0.49	1.41
MA3891	hypothetical protein	S	0.0219	0.49	1.41
MA0271	Tetrahydromethanopterin S-methyltransferase subunit	H	8.12E-07	0.49	1.40
MA1400	hypothetical protein	S	1.44E-08	0.49	1.40
MA1877	hypothetical protein	S	3.72E-04	0.49	1.40
MA2890	sensory transduction histidine kinase	T	2.83E-12	0.49	1.40
MA1495	$F_{420}H_2$ dehydrogenase subunit	C	0.0190	0.49	1.40
MA1698	hypothetical protein	S	0.0425	0.48	1.40
MA0560	hypothetical protein	S	1.69E-06	0.48	1.39
MA4159	F0F1-type ATP synthase, subunit alpha (*atpB*)	C	0.0245	0.48	1.39
MA3369	hypothetical protein	S	2.93E-05	0.47	1.39
MA2106	hypothetical protein	S	0.0156	0.47	1.39
MA4611	hypothetical protein	S	0.0356	0.47	1.39
MA1940	hypothetical protein	S	0.0053	0.47	1.38
MA3293	hypothetical protein	S	0.0031	0.46	1.38
MA2630	alcohol dehydrogenase (*adh*)	R	0.0123	0.46	1.38
MA2238	hypothetical protein	S	1.24E-08	0.46	1.38
MA2912	AMP-binding protein	R	3.93E-04	0.46	1.38
MA4160	F0F1-type ATP synthase, beta subunit	C	0.0331	0.46	1.37
MA2913	hypothetical protein	S	0.0086	0.45	1.37
MA1544	hypothetical protein	S	0.0326	0.45	1.37

MA1249	oligopeptide ABC transporter, oligopeptide-binding protein	R	0.0053	0.45	1.37
MA4663	multidrug resistance efflux pump	V	4.61E-04	0.45	1.37
MA3731	nicotinamide-nucleotide adenylyltransferase	H	0.0074	0.45	1.36
MA2656	hypothetical protein	S	3.93E-05	0.44	1.36
MA3728	hypothetical protein	S	0.0178	0.44	1.36
MA4566	multiple resistance/pH regulation related protein G (Na+/H+ antiporter) (*mrpG*)	R	0.0126	0.44	1.36
MA0270	Tetrahydromethanopterin S-methyltransferase, subunit G	H	2.52E-04	0.44	1.35
MA3424	hypothetical protein	S	9.21E-07	0.44	1.35
MA2914	Transcriptional regulator, Hth-3 family	K	2.12E-04	0.44	1.35
MA2488	ABC transporter, ATP-binding protein	R	0.0260	0.44	1.35
MA0201	Isocitrate/isopropylmalate dehydrogenase (*leuB*)	C	3.68E-05	0.44	1.35
MA4152	F0F1-type ATP synthase subunit (*atpH*)	C	0.0158	0.43	1.35
MA2478	peptide ABC transporter, ATP-binding protein	R	0.0184	0.43	1.35
MA0336	cell surface protein	M	0.0036	0.43	1.35
MA4477	hypothetical protein	S	0.0075	0.43	1.35
MA1590	surface antigen gene	M	0.0208	0.43	1.35
MA2497	Fumarate hydratase (*fumA*)	C	9.95E-04	0.43	1.34
MA3006	Formaldehyde-activating enzyme in methanogenesis	C	9.14E-04	0.42	1.34
MA4476	hypothetical protein	S	0.0027	0.42	1.34
MA4380	hypothetical protein	S	2.93E-05	0.42	1.33
MA0831	transposase	X	0.0286	0.41	1.33
MA2536	Carbon anhydrase (*cam*)	P	0.0012	0.41	1.33
MA1483	Trk-type K+ transport system (*trkH*)	P	1.85E-06	0.41	1.33
MA0619	sensory transduction histidine kinase	T	6.38E-04	0.40	1.32
MA1423	hypothetical protein	S	0.0291	0.40	1.32
MA1374	hypothetical protein	S	3.30E-05	0.39	1.31
MA2156	periplasmic binding protein	R	4.17E-04	0.39	1.31
MA1069	hypothetical protein	S	0.0183	0.39	1.31
MA1919	adenylosuccinate synthase	F	6.12E-04	0.39	1.31
MA1742	hypothetical protein	S	0.0017	0.39	1.31
MA1611	Leucyl-tRNA synthetase (*leuS*)	J	0.0203	0.39	1.31
MA2015	hypothetical protein	S	0.0017	0.39	1.31
MA0021	high-affinity gluconate transporter (gluconate permease)	G	0.0214	0.38	1.30
MA3239	hypothetical protein	S	0.0286	0.38	1.30

MA3496	pyruvoyl-dependent arginine decarboxylase	E	0.0089	0.38	1.30
MA3174	hypothetical protein	S	0.0273	0.38	1.30
MA0593	hypothetical protein	S	0.0390	0.37	1.29
MA1276	N-ethylammeline chlorohydrolase	R	0.0405	0.36	1.28
MA1543	hypothetical protein	S	0.0091	0.36	1.28
MA1055	hypothetical protein	S	0.0333	0.35	1.27
MA3437	branched chain amino acid transport protein (*azlC*)	E	3.73E-06	0.35	1.27
MA1466	4-thiouridine synthase (*thiI*)	J	0.0177	0.35	1.27
MA1475	hypothetical protein	S	0.0390	0.34	1.27
MA3999	hypothetical protein	S	0.0315	0.34	1.27
MA2454	hypothetical protein	S	3.55E-04	0.34	1.27
MA1424	transcriptional regulator, Hth-3 family	K	1.43E-04	0.34	1.26
MA2582	hypothetical protein	S	0.0332	0.34	1.26
MA4237	Heterodisulfide reductase, subunit B	C	0.0354	0.33	1.26
MA2740	hypothetical protein	S	0.0495	0.32	1.25
MA2153	periplasmic binding protein	R	0.0035	0.32	1.25
MA2742	nitroreductase	C	0.0135	0.32	1.25
MA4154	Corrinoid protein involved in methanogenesis (*atpC*)	C	0.0299	0.32	1.25
MA3791	Acetolactate synthase subunit (*ilvH*)	E	2.39E-04	0.31	1.24
MA1264	DNA-directed RNA polymerase subunit B (*rpoB1*)	K	0.0137	0.31	1.24
MA1537	hypothetical protein	S	0.0133	0.31	1.24
MA3176	hypothetical protein	S	0.0120	0.30	1.24
MA3790	Ketol-acid reductoisomerase (*ilvC*)	E	2.85E-04	0.30	1.23
MA2115	hypothetical protein	S	0.0164	0.30	1.23
MA2553	sensory transduction histidine kinase	T	0.0160	0.30	1.23
MA3175	hypothetical protein	S	0.0390	0.29	1.23
MA0561	hypothetical protein	S	0.0340	0.29	1.22
MA2431	Isoleucyl-tRNA synthetase (*ileS*)	J	0.0239	0.29	1.22
MA1274	sensory transduction histidine kinase	T	1.53E-04	0.29	1.22
MA1031	hypothetical protein	S	0.0069	0.29	1.22
MA0777	signal-transducing histidine kinase	T	0.0016	0.29	1.22
MA4240	hypothetical protein	S	0.0489	0.28	1.22
MA1103	integral membrane protein	R	7.85E-04	0.28	1.22
MA1841	hypothetical protein	S	0.0049	0.28	1.21
MA1061	polysaccharide biosynthesis protein	M	0.0484	0.27	1.21
MA0003	sodium:proline symporter (*putP*)	E	0.0275	0.27	1.20
MA1263	DNA-directed RNA polymerase subunit D (*rpoA1*)	K	0.0141	0.26	1.20

MA1095	Preprotein translocase subunit (*secY*)	U	0.0037	0.25	1.19
MA0558	hypothetical protein	S	0.0498	0.24	1.18
MA0750	hypothetical protein	S	0.0340	0.23	1.17
MA1096	Adenylate kinase (*adk*)	F	0.0073	0.23	1.17
MA1813	hypothetical protein	S	0.0113	0.23	1.17
MA2520	Orotate phosphoribosyltransferase (*pyrE*)	F	0.0484	0.22	1.17
MA0925	aromatic amino acid transferase (*tyrB*)	E	0.0132	0.22	1.16
MA0916	erythrocyte band 7 integral membrane protein	R	0.0135	0.20	1.15
MA1175	mannosyltransferase	M	0.0349	0.20	1.15
MA0915	nodulation protein (*nfeD*)	O	0.0378	0.20	1.15

Table 11. Downregulated transcripts in Ma Δ*msvR*

Locus ID	Description	arCOG	FDR	Expression Difference (Ma Δ*msvR*-WT)	Fold Change
MA1458	MaMsvR	K	0	-12.57	-6070.85
MA2425	Methyltransferase isozyme 1 (*mtbB*)	C	5.21E-05	-3.31	-9.94
MA3895	Nitrogenase subunit (*nifH*)	P	0.0042	-3.16	-8.92
MA2424	Corrinoid protein involved in methanogenesis (*mtbC*)	C	1.47E-12	-2.94	-7.65
MA3471	hypothetical protein	S	0.0190	-2.66	-6.32
MA2288	hypothetical protein	S	1.97E-10	-2.42	-5.34
MA4207	ammonium transporter	P	5.63E-06	-2.18	-4.52
MA1707	hypothetical protein	S	4.38E-13	-2.12	-4.33
MA4199	succinoglycan biosynthesis regulator (*exsB*)	M	0	-1.91	-3.77
MA0776	hypothetical protein with TaqI-like C-terminal specificity domain	R	3.61E-41	-1.85	-3.60
MA4208	P-II family nitrogen regulatory protein (*glnB*)	E	2.39E-04	-1.85	-3.60
MA1957	sensory transduction histidine kinase	T	8.62E-05	-1.82	-3.53
MA1686	3',5'-cyclic AMP phosphodiesterase (*cpdA*)	T	5.34E-06	-1.81	-3.50
MA4535	hypothetical protein with methyltransferase domain	R	6.76E-16	-1.72	-3.28
MA4198	hypothetical protein	S	4.25E-20	-1.66	-3.16
MA4344	efflux system transcriptional regulator, ArsR family (*smtB*)	K	2.71E-04	-1.65	-3.13
MA4340	biotin synthesis BioY protein	H	2.58E-17	-1.56	-2.94
MA4197	hypothetical protein	S	3.74E-11	-1.54	-2.91
MA3936	Uncharacterized inner membrane protein YnfA	R	0.0286	-1.54	-2.90
MA1624	hypothetical protein	S	8.81E-04	-1.53	-2.88
MA4532	iron compounds ABC transporter, ATP-binding protein	P	1.95E-18	-1.52	-2.86
MA4536	ABC transporter, solute-binding protein	R	7.35E-16	-1.50	-2.83
MA2972	Methyltransferase isozyme 1 (*mtmB2*)	C	0.0037	-1.48	-2.79
MA2801	hypothetical protein	S	2.82E-08	-1.44	-2.72
MA4196	6-pyruvoyltetrahydropterin synthase	H	1.56E-10	-1.41	-2.66
MA3051	haloacid dehalogenase-like hydrolase	R	1.04E-04	-1.35	-2.54
MA4346	hypothetical protein with PspC domain	R	0.0031	-1.30	-2.46
MA2355	hypothetical protein	S	0.0024	-1.24	-2.36
MA1214	P-II family nitrogen regulatory protein (*nifI1*)	P	3.83E-04	-1.23	-2.35
MA4531	hypothetical protein with S-layer	S	1.10E-05	-1.22	-2.32

	family duplication domain				
MA1239	molybdenum-pterin binding protein	R	8.24E-13	-1.19	-2.29
MA1626	hypothetical protein	S	0.0022	-1.14	-2.21
MA2781	hypothetical protein with methyltransferase domain	R	0.0326	-1.12	-2.17
MA2971	Corrinoid protein (*mtmC*)	C	0.0248	-1.10	-2.15
MA1234	iron(III) ABC transporter, permease protein	P	6.39E-04	-1.07	-2.10
MA2267	hypothetical protein	S	0.0017	-1.05	-2.07
MA3240	hypothetical protein	S	1.03E-17	-1.04	-2.06
MA3044	SAM-dependent methyltransferase	R	0.0315	-1.04	-2.06
MA0546	hypothetical protein	S	6.49E-07	-1.04	-2.06
MA2802	hypothetical protein	S	1.75E-04	-1.04	-2.05
MA2803	hypothetical protein	S	1.47E-05	-1.02	-2.03
MA2974	hypothetical protein	S	2.00E-04	-1.01	-2.02
MA1200	iron ABC transporter, solute-binding protein	R	0.0097	-0.97	-1.96
MA4446	hypothetical protein	S	4.49E-04	-0.94	-1.91
MA3234	hypothetical protein	S	0.03964	-0.93	-1.90
MA4372	hydrogenase-3, subunit E (HycE homolog)	C	0.0188	-0.91	-1.87
MA3105	hypothetical protein	S	0.001748 324	-0.89	-1.85
MA4216	Glutamine synthetase (*glnA*)	E	4.39E-13	-0.88	-1.84
MA1366	response regulator receiver	T	5.77E-04	-0.86	-1.81
MA0261	Thiamine biosynthesis protein (*thiC*)	H	1.62E-23	-0.82	-1.77
MA4671	response regulator receiver	T	0.0020	-0.80	-1.74
MA4373	hydrogenase-3, subunit G (HycG homolog)	C	0.0343	-0.80	-1.74
MA3046	oligopeptide ABC transporter, oligopeptide-binding protein (*oppD*)	R	5.26E-04	-0.80	-1.74
MA4341	Nickel and Cobalt transporter (*cbiO*)	P	3.38E-18	-0.79	-1.73
MA2804	hypothetical protein	S	9.53E-05	-0.79	-1.73
MA1212	P-II family nitrogen regulatory protein (*nifI1*)	P	0.0049	-0.77	-1.71
MA0547	winged HTH domain-containing protein/riboflavin kinase	H	1.83E-04	-0.76	-1.70
MA1271	hypothetical protein	S	0.0019	-0.75	-1.68
MA1211	P-II family nitrogen regulatory protein (*nifI2*)	P	2.44E-04	-0.74	-1.67
MA2782	TetR family transcriptional regulator	K	0.0182	-0.73	-1.65
MA4067	hypothetical protein	S	1.87E-06	-0.72	-1.65
MA4425	hypothetical protein	S	0.0016	-0.72	-1.64
MA2287	hypothetical protein	S	0.0086	-0.70	-1.63

MA1233	iron(III) ABC transporter, ATP-binding protein	R	0.0406	-0.70	-1.63
MA4442	hypothetical protein	S	0.0208	-0.67	-1.59
MA0544	Protein-L-isoaspartate (*pcm-1*)	O	6.86E-04	-0.67	-1.59
MA1748	cobalt transport proteins (*cbiQ*)	P	1.06E-05	-0.66	-1.58
MA4058	hypothetical protein	S	0.0190	-0.63	-1.54
MA4343	cobalt transport family protein	P	1.25E-09	-0.62	-1.54
MA1238	Molybdopterin oxidoreductase component (*mop*)	R	1.21E-10	-0.60	-1.52
MA2657	hypothetical protein	S	0.0183	-0.59	-1.50
MA3943	hypothetical protein	S	0.0438	-0.58	-1.50
MA3664	cell surface protein	R	0.0073	-0.57	-1.49
MA1205	Nitrogenase subunit NifH (ATPase)	P	0.0078	-0.56	-1.48
MA0653	hypothetical protein	S	0.0141	-0.56	-1.47
MA3459	hypothetical protein	S	0.0073	-0.55	-1.46
MA4316	hypothetical protein	S	0.0488	-0.54	-1.46
MA0643	H/ACA RNA-protein complex component Nop10p	J	6.54E-09	-0.54	-1.46
MA4465	hypothetical protein	S	0.0272	-0.53	-1.45
MA1740	hypothetical protein	S	5.60E-04	-0.53	-1.44
MA1210	Nitrogenase iron-iron protein alpha chain (*anfD*)	C	0.0039	-0.52	-1.44
MA2188	transposase	X	0.0073	-0.52	-1.44
MA3632	Cadmium resistance protein (*cadA*)	P	0.0327	-0.52	-1.43
MA0303	hypothetical protein	S	0.0129	-0.52	-1.43
MA3852	transposase	X	2.29E-04	-0.52	-1.43
MA1290	transposase	X	0.0076	-0.51	-1.42
MA2748	transposase	X	1.57E-04	-0.50	-1.41
MA3529	transposase	X	2.10E-04	-0.50	-1.41
MA2815	GCN5-related N-acetyltransferase	R	9.51E-04	-0.49	-1.40
MA2204	hypothetical protein	S	1.23E-04	-0.48	-1.39
MA3166	hypothetical protein	S	2.71E-04	-0.48	-1.39
MA3550	hypothetical protein	S	0.0025	-0.48	-1.39
MA0513	maltose O-acetyltransferase (*maa*)	G	9.35E-04	-0.47	-1.39
MA2774	iron ABC transporter, permease	R	0.0425	-0.47	-1.38
MA2203	cellulosomal protein	S	3.72E-07	-0.47	-1.38
MA0644	Translation initiation factor 2, alpha subunit (*eIF-2A*)	J	1.74E-13	-0.46	-1.38
MA4214	benzodiazepine receptor TspO	R	0.0264	-0.46	-1.38
MA0042	Peptide chain release factor (*erf1*)	J	0.0023	-0.46	-1.37
MA0424	hypothetical protein	S	0.0017	-0.46	-1.37
MA0812	hypothetical protein	S	0.0305	-0.45	-1.37
MA0519	hypothetical protein	S	0.0477	-0.45	-1.37
MA0646	Ribosomal protein L44E (*rpl44e*)	J	2.02E-14	-0.45	-1.37

MA0265	hypothetical protein	S	9.81E-04	-0.45	-1.37
MA1977	Molybdopterin-guanine dinucleotide biosynthesis protein A (*mobA*)	H	0.0211	-0.45	-1.37
MA1209	nitrogenase, subunit delta (*anfG*)	C	0.0378	-0.45	-1.36
MA1732	hypothetical protein	S	0.0082	-0.44	-1.36
MA2280	ABC-type molybdate transport system, periplasmic component (*modA*)	P	0.0326	-0.44	-1.35
MA0703	hypothetical protein	S	0.0059	-0.43	-1.35
MA0642	3-isopropylmalate dehydratase	E	2.49E-04	-0.43	-1.35
MA0470	hypothetical protein	S	0.0135	-0.43	-1.35
MA0605	metallo-beta-lactamase	R	0.0272	-0.43	-1.34
MA1101	transcriptional regulator	K	0.0317	-0.43	-1.34
MA4440	polysaccharide deacetylase	G	0.0451	-0.42	-1.34
MA0543	Protein-L-isoaspartate (*pcm-2*)	O	0.0135	-0.42	-1.34
MA3291	hypothetical protein	S	0.0078	-0.42	-1.34
MA0302	oligopeptide ABC transporter, oligopeptide-binding protein (*oppA*)	R	0.0125	-0.42	-1.34
MA0606	bifunctional short chain isoprenyl diphosphate synthase (*idsA*)	R	0.0477	-0.42	-1.34
MA0763	hypothetical protein	S	0.0104	-0.41	-1.33
MA0462	hypothetical protein	S	0.0297	-0.41	-1.33
MA3904	*ubiE*/COQ5 methyltransferase	R	0.0258	-0.41	-1.33
MA4592	3-dehydroquinate synthase	E	0.0193	-0.40	-1.32
MA4591	Deoxyribose-phosphate aldolase (*deoC*)	F	0.0351	-0.40	-1.32
MA3315	hypothetical protein	S	0.0340	-0.39	-1.31
MA0189	SNF2 family helicase	K	0.0011	-0.39	-1.31
MA0051	hypothetical protein	S	0.0208	-0.38	-1.30
MA0152	hypothetical protein	S	0.0214	-0.37	-1.30
MA2679	transposase	X	0.0167	-0.37	-1.29
MA4609	DNA-(apurinic or apyrimidinic site) lyase	L	0.0182	-0.37	-1.29
MA3906	oligopeptide ABC transporter, oligopeptide-binding protein (*oppB*)	R	0.0186	-0.36	-1.29
MA2816	GCN5-related N-acetyltransferase	R	0.0191	-0.36	-1.28
MA4195	Fe-S oxidoreductase family (*nifB*)	C	0.0337	-0.36	-1.28
MA1961	cell surface glycoprotein (S-layer protein)	M	0.0031	-0.35	-1.28
MA1363	phosphoglycolate phosphatase	R	2.29E-04	-0.35	-1.28
MA4060	hypothetical protein	S	0.0137	-0.35	-1.28
MA0008	Pheromone shutdown protein (*traB*)	S	0.0073	-0.34	-1.27
MA4424	hypothetical protein	S	0.0132	-0.34	-1.27
MA0757	phosphoesterase	R	0.0031	-0.34	-1.27
MA3021	sodium/calcium exchanger protein	R	2.82E-04	-0.34	-1.27

MA0747	hypothetical protein	S	0.0072	-0.34	-1.26
MA1915	D-aminopeptidase (*dppA*)	E	0.0494	-0.34	-1.26
MA0909	ATPase	R	0.0378	-0.34	-1.26
MA1741	ribosomal biogenesis protein	J	0.0020	-0.34	-1.26
MA0645	Ribosomal protein S27E (*rps27e*)	J	8.45E-04	-0.34	-1.26
MA2609	hypothetical protein	S	0.0390	-0.34	-1.26
MA0452	glycosyltransferase	M	0.0018	-0.33	-1.26
MA0803	hypothetical protein	S	0.0015	-0.33	-1.25
MA4615	Isopropylmalate/homocitrate/citramalate synthase (*leuA*)	E	0.0019	-0.33	-1.25
MA3378	hypothetical protein	S	0.0042	-0.32	-1.25
MA0866	hypothetical protein	S	0.0166	-0.32	-1.25
MA3338	hypothetical protein	S	0.0259	-0.32	-1.25
MA1949	glutamate decarboxylase	E	0.0273	-0.31	-1.24
MA2628	glycogen debranching enzyme (GDB1)	G	0.0075	-0.31	-1.24
MA1310	hypothetical protein	S	7.38E-04	-0.30	-1.23
MA1645	sensory transduction histidine kinase	T	5.57E-04	-0.29	-1.23
MA1026	hypothetical protein	S	0.0276	-0.29	-1.23
MA0030	hypothetical protein	S	0.0349	-0.29	-1.22
MA0684	Flavodoxin reductase (ferredoxin-NADPH reductase) family 1 (*fprB*)	C	0.0469	-0.29	-1.22
MA2259	hypothetical protein	S	0.0408	-0.29	-1.22
MA1425	transposase	X	0.0349	-0.28	-1.22
MA2163	Methyltransferase isozyme 2 involved in methanogenesis (*ctmM*)	C	0.0089	-0.28	-1.22
MA0120	hypothetical protein	S	8.58E-04	-0.28	-1.21
MA0210	hypothetical protein	S	0.0477	-0.27	-1.21
MA4317	phycocyanin alpha phycocyanobilin lyase related protein	R	0.0188	-0.27	-1.20
MA3853	Cell division ATPase (*ftsA*)	D	0.0471	-0.27	-1.20
MA4111	Ribosomal protein L24 (*rplX)*	J	0.0045	-0.26	-1.20
MA1352	hypothetical protein	S	0.0331	-0.25	-1.19
MA4507	Superfamily II DNA helicase (RecQ)	L	0.0117	-0.25	-1.19
MA4020	ABC transporter, ATP-binding protein	R	0.0027	-0.25	-1.19
MA4573	hypothetical protein	S	0.0059	-0.25	-1.19
MA3975	type IV secretion system protein	R	0.0018	-0.25	-1.19
MA0720	hypothetical protein	S	0.0193	-0.24	-1.18
MA4112	Translation initiation factor 6 (*eIF6*)	J	0.0135	-0.24	-1.18
MA3729	Phosphoenolpyruvate synthase (*ppsA*)	G	0.0289	-0.23	-1.18
MA0081	Polyphosphate kinase (*ppk*)	P	0.0270	-0.23	-1.17
MA4475	Dihydrodipicolinate synthase/N-	E	0.0020	-0.22	-1.17

	acetylneuraminate lyase (*dapA*)				
MA4376	response regulator receiver	T	0.0080	-0.21	-1.16
MA2968	hypothetical protein	S	0.0180	-0.20	-1.15
MA4114	Ribosomal protein L39E (*rpl39e*)	J	0.0400	-0.19	-1.14
MA4110	prefoldin, subunit alpha (*pfdA*)	O	0.0474	-0.18	-1.14