GENETIC MANIPULATIONS AND GENE REGULATION OF S1 PEPTIDASE (*ctp*A) EXPRESSION IN *Mycoplasma mycoides* SUBSP. *capri*

By

AYMAN B. ALLAM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

ACKNOWLEDGMENTS

I thank my committee members Drs. Mary Brown, Leticia Reyes, Ann Progulske-Fox, John Dame and Thomas Rowe for their time and advice. I am grateful for Dr. Mary Brown for making a home for me in her laboratory, her mentorship, and for creating a supportive environment for my professional development. I would like to thank Dr. Leticia Reyes for her friendship and for her help addressing and treating some of the shortcomings.

I also thank members of our lab; Jan Stevenson, Lori Wendland, and Fiona Maunsell for their friendship and support. I would like to acknowledge the Interdisciplinary Core of Biotechnology Research (ICBR) - Proteomics Core Facility for their help with 2D gels, iTRAQ™ and MS data.

The majority of the work was funded by USDA CSREES grant 99-35204 and subcontract 6415-100607 from the University of South Florida Center for Biological Defense (contract W911SR-06-C-0023).

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

GENETIC MANIPULATIONS AND GENE REGULATION OF S1 PEPTIDASE *(CTP*A) EXPRESSION IN *MYCOPLASMA MYCOIDES* SUBSP. *CAPRI*

By

Ayman Bialy Allam

August 2010

Chair: Mary B. Brown
Co chair: Leticia Reyes
Major: Veterinary Medical Sciences

Mycoplasmas are exquisitely simple with the smallest genomes of free-living organisms, yet the range of infections and host species are among the most diverse in the microbial world. A major limiting factor in unraveling the virulence factors of these microbes is the paucity of tools for genetic manipulation. As a target gene for disruption, we chose the *Mycoplasma mycoides* subsp. *capri* (*M. mycoides capri*) S1 peptidase (*ctpA*) that confers a proteolytic phenotype that can be readily determined by growth on agar supplemented with casein. We developed a suicide plasmid employing *RecA* from *Escherichia coli* under the direction of a mycoplasmal promoter and with *tetM* as a selection marker. The desired targeted double-crossover event was obtained in 24% of transformants obtained using pExp1-*ctpA*::*tetM-recA*ec, which represents a substantive 140-fold improvement in the percentage of transformants with the desired gene disruption as well as demonstrates targeted double crossover events in a mycoplasma species other than *Mycoplasma genitalium*.

The expression profile of mutant versus the wild type was determined using two proteomic technologies: 2D-gel and iTRAQ™. Results from the proteomic data showed

significant difference in the expression profile of mutant in comparison to the wild type. Genes involved in glycolysis were downregulated while genes participating in glycerol metabolism were upregulated. Also many of the ribosomal proteins were upregulated. Taken together the data suggest that the disrupted gene may influence or modulate the expression of stress response genes, especially those expressed during oxidative stress.

CHAPTER 1
INTRODUCTION

Overview

Mycoplasmal infections present one of the most intriguing paradoxes in infectious diseases. On one hand, these microbes are exquisitely simple with the smallest genomes of free-living organisms (Fadiel et al. 2007, Krause et al. 2004, Rocha et al. 2002). Conversely, the range of infections and host species are among the most diverse in the microbial world. Mycoplasmas are members of the family Mollicutes and are prokaryotes that lack a cell wall. This essential difference between mycoplasmas and prokaryotes, including cell wall-deficient forms of other bacteria, is their inherent inability to make any cell wall components due to their lack of the genes for cell wall biosynthesis. In contrast, other cell wall-deficient bacteria have the genes, and thus the potential, for making a cell wall (Mattman 2001).

Phylogenetically, 16S rRNA sequences indicate that mycoplasmas are eubacteria emerging from or related to Gram positive bacteria (Woese et al. 1980). They are thought to have originated from Gram positive bacteria through degenerative or reductive evolution (Razin 1992, Neimark 1986, Maniloff 1983). Specifically, mycoplasmal genomes are thought to have evolved by attrition from the genomes of species in the Lactobacillus group (Maniloff 2002, Maniloff 1992). Mycoplasmas are currently considered to be the smallest and simplest self-replicating organisms (Razin et al. 1985). The full genome sequence is available for a large number of mycoplasma species and a dedicated site is available for comparative analyses (see cbi.labri.fr/eng/molligen.htm). The small genome size of these bacteria is thought to define the lower limit of naturally existing independent life. The analysis of the minimal

gene complement of some mycoplasmas is one of the sources of synthetic biology, a new discipline of biology (Gibson et al. 2008, Glass et al. 2006).

Another characteristic that sets mycoplasmas apart from many other bacteria is the low CG ratio of mycoplasma genome, which can go as low as 23% (Jones et al. 1963). Through their evolution and likely as a direct result of the low GC content, mycoplasmas deviated from the conventional genetic code of using UGA as a universal stop codon. Instead the UGA is a code for tryptophan (Blanchard 1990), a feature shared by mitochondria and ciliates that also use UGA to code for tryptophan (Fox 1987). Therefore, mycoplasmas have only one release factor, RF-1, which recognizes both UAA and UAG stop codons (Inagaki et al. 1996, Inagaki et al. 1993). Release factor RF -2, which recognizes both UAA and UGA stop codons and is present in most eubacteria, is absent in mycoplasmas.

Mycoplasmas also have undergone a rapid rate of evolution. It has been estimated that mycoplasmal phylogenetic tree branches have accumulated an average of about 50% more base changes than has the Lactobacillus group (Woese et al. 1985). In the course of their evolution, mycoplasmas appear to have lost many of the genes and, in turn, enzymatic functions for synthesis of macromolecule precursors and pathways including biochemical pathways for cell wall biosynthesis, amino acids biosynthesis, nucleotides, and fatty acids biosynthesis. Instead, mycoplasmas possess a number of hydrolytic and degrading enzymes, including both proteases and nucleases (Maniloff 1992). It is likely that many of the precursors for macromolecules are derived from the host, and mycoplasmas do maintain a large complement of transporter genes. Mycoplasmas have retained a variety of genes for intermediary metabolism and energy

production. Those include pathways for the degradation of organic substrates to generate small molecules for biosynthetic reactions as well as genes for energy production via substrate-level phosphorylation (Pollack 2002, Maniloff 1992).

The number of genes for transcription and translation apparatus in mycoplasmas is comparable to those present in other bacteria. These genes comprise more 20% of the total ORFs annotated in mycoplasma genomes (Glass et al. 2000, Himmelreich et al. 1997, Himmelreich et al. 1996, Fraser et al. 1995). Mycoplasmal RNA polymerase resembles those RNA polymerases in other bacteria, and they seem to recognize promoter sequences similar to those found in other bacteria and recognized by their σ^{70} type in *E. coli* or σ^{A} vegetative type in Gram positive bacteria. Those promoters contain a -35 region (TTGACA) and -10 Pribnow Box (TATAAT) with about 17 bp spacer between them (Razin et al. 1998, Muto et al. 1987). In fact, mycoplasma promoters have been recognized by *E. coli* RNA polymerase both *in vivo* and *in vitro* as well (Taschke et al. 1988). Analysis of many promoters upstream of mycoplasma genes showed that the -10 region seems to be conserved while there are variations in the sequence of the -35 region (Halbedel et al. 2007, Weiner et al. 2000, Waldo et al. 1999). Ribosomal Binding Sites (RBS) or Shine-Dalgarno (SD) sequences have been identified in the 5' upstream region of most mycoplasma genes (Bove 1993, Muto et al. 1987). Termination of transcription in mycoplasmas is Rho-independent. The gene for Rho terminator has not been detected in mycoplasma genomes while the stem loop structure followed by a string of T(U)s has been found downstream of many ORFs. Genes for termination modulators NusA and NusG have been found in mycoplasmas (Fraser et al. 1995, Glass et al. 2000, Himmelreich et al. 1996, Muto et al. 1987). Similar

to other bacteria, mycoplasmas have many tRNA and rRNA genes that are clustered in operons (Yamao et al. 1988, Muto 1987, Samuelsson et al. 1985, Yamao et al. 1985). However, unlike most bacteria, mycoplasmas have little or no redundancy for rRNAs and tRNAs (Glass et al. 1992, Andachi et al. 1989). Mycoplasmas have all three initiation factors IF1, IF2, and IF3; they also have all four elongation factors EF-Tu, EF-Ts, EF-G, and EF-P (Glass et al. 2000, Himmelreich et al. 1996, Bork et al. 1995, Fraser et al. 1995).

Genetic Regulation in Mycoplasmas

The current view of gene regulation in mycoplasmas is that, with few notable exceptions (including variable surface proteins), mycoplasmas do not regulate the expression of their genes. Mycoplasmas are thought to lack refined on-off switching mechanisms and global regulation for transcriptional adaptation to environmental changes. Their genes are generally thought to be constitutively expressed rather than regulated (Muto and Ushida 2002), albeit coordinated with the growth rate of the cell through stringent control (Cashel et al. 1996, Gourse et al. 1996). Furthermore, most known or common bacterial regulatory mechanisms are absent in mycoplasmas. For example, unlike *Escherichia coli*, *Bacillus subtilis* and many other bacteria, mycoplasmas have only one sigma factor even though they contain many stress response genes (Muto et al. 1987, Razin et al. 1998). The presence of these stress response genes in mycoplasmas suggests that either these genes are not under regulatory control or that the regulation might occur by a yet unknown mechanism not involving the alternate Sigma factors.

Very few repressor-like proteins have been reported in mycoplasma genomes (Bork et al. 1995, Fraser et al. 1995, Glass et al. 2000, Himmelreich et al. 1996, Glass

et al. 2006). The limited number of these proteins implies that classic regulatory mechanism for turning genes on/off may be limited or lacking in mycoplasmas. Also absent from the genomes of mycoplasmas are genes for quorum sensing and two-component systems (Simmons et al. 2007). Taken together, these findings imply that in addition to limited individual gene regulation mechanisms, mycoplasmas appear to lack commonly identified global regulatory mechanisms for influencing gene expression. No global regulator has been reported to date, and a recently emerging view is that mycoplasmas implement post-translational regulation as the primary mechanism for controlling biological activities in response to changing environmental conditions (Schmidl et al. 2010). Mycoplasmas are therefore thought to use stochastic mechanisms similar to those controlling phase and size variations to genetically regulate processes as complex as biofilm formation (Simmons et al. 2007).

The notable exceptions regarding gene regulation in mycoplasma are the stochastic mechanisms responsible for phase and size variations (reviewed in Citti et al. 2005). Mycoplasmas exhibit an extremely high rate ($>10^{-3}$) of antigenic variation of many surface proteins (Rosengarten et al. 2000, Rosengarten et al. 1990). The mechanisms for phase, size, and antigenic variations are well-studied and include slipped-strand mispairing, site-specific recombination, and gene conversion. The specific method of regulation is dependent upon the individual mycoplasmal species, and not all mechanisms apply to each species.

Some genes contain short homo (poly A) or heteropolymeric (poly AT) tracts that undergo frequent and reversible changes in their nucleotide number. These mutational changes are thought to be the result of slipped-strand mispairing (SSM) during DNA

replication or processes that require DNA synthesis. The consequence of these changes is to switch on or off the corresponding gene expression. As described below, SSM can affect the gene expression at the level of transcription or translation, depending on the location of these poly A/poly AT tracts. Genes that are transcriptionally affected by SMM mechanism contain their poly tracts in the promoter region. For example, the *vlp* gene family of *M. hyorhinis* contains a poly A tract in the promoter between the -35 and -10 regions. Expression of the gene is dependent on the length of poly A tract, with a specific nt tract length (N=17 A) required for gene expression. Reversible deletion or insertion of a single nucleotide abolishes the transcription of the gene (Citti et al. 1995). A similar mechanism is thought to control the phase variation of Vmm of *M. mycoides* and MAA2 of *M. arthritidis* (Persson et al. 2002).

In the case of translational control by SSM mechanism, the homo or hetero poly tract is located in the coding sequence of the gene. A reversible addition or deletion of a single nucleotide by SSM results in a reversible frameshift mutation, which in turn produces an in-frame stop codon leading to a premature termination of the translation process. The final outcome is a truncated gene product. This mechanism has been reported for Vaa adhesion surface protein of *M. hominis* (Zhang et al. 1997) and for P78 in *M. fermentans* (Theiss et al. 1997).

Site-specific recombinases and integrases can control the expression of some surface proteins through a “cut and paste” type mechanism. In this mechanism, four DNA strands are broken, exchanged, and then religated at specific locations. The relative positions of the two targeted recombination sites dictate the outcome of this

type of recombination. Inverted recombination sites result in DNA inversion; in contrast, directly repeated recombination sites lead to excision. This mechanism has been documented in the *vsa*, *vsp*, and *vpma* gene families of *M. pulmonis*, *M. bovis*, and *M. agalactiae* respectively (Bhugra et al. 1995, Lysnyansky et al. 2001, Glew et al. 2002).

To date, the only example of gene conversion has been described in *M. synoviae*. The *M. synoviae* genome contains a single copy of the full length *vlhA* gene, but it also possesses a large family of pseudogenes that encodes different extents of the expressed *vlhA*. Variants of the *vlhA* gene are generated by unidirectional recombination between the expressed gene and the pseudogenes. The recombined pseudogene sequence was shown to duplicate, while the region replaced in the expressed gene was lost (Noormohammadi et al. 2000).These conversions events appear to be mediated through highly conserved regions at the boundaries of each pseudogene. This suggests that the recombination events are carried out by a yet unidentified site-specific recombinase.

Regulation of a few heat shock proteins has been reported (Chang et al. 2008, Musatovova et al. 2006). CIRCE (*C*ontrolling *I*nverted *R*epeat of *C*haperon *E*xpression) is a cis-DNA regulatory element that consists of a 9 bp inverted repeat that is separated by a 9 bp spacer. HrcA (*H*eat *R*egulation at *C*IRCE) is a repressor protein that binds to CIRCE sequence and negatively regulates the downstream gene by preventing its transcription (Zuber et al. 1994). Although mycoplasmas do have heat shock proteins, CIRCE elements in the upstream of their promoters have been identified only in *M. genitalium*, *M. pneumoniae*, and *M. hypopneumoniae* (Himmelreich R. et al. 1996). Interestingly, Musatovova et al. showed that some of the typical prokaryotic heat shock

genes (for example, *groEL*, *groES*) did not respond to heat shock. Transcriptome studies in mycoplasmas that lacked HrcA-CIRCE upstream of the promoter demonstrated that some heat shock proteins were responsive to temperature shock, implying that the HrcA-CIRCE mechanism is not universally implemented by mycoplasmas to regulate response to heat shock.

However, with the exception of gene regulation in variable surface proteins and limited heat shock response genes, a conundrum remains. On one hand, a large number of mycoplasmal genomes have been sequenced, yet because of the absence of regulatory proteins there is still a dearth of knowledge regarding the mechanisms that control gene expression in these microbes. Despite this absence of clearly identified global regulators, recent *in vivo* and *in vitro* transcriptome and proteomic studies reported global responses to environmental or host stimuli (Oneal et al. 2008, Madsen et al. 2008, Schafer et al. 2007, Cecchini et al. 2007, Pinto et al. 2007, Madsen et al. 2006a, Madsen et al. 2006b). Importantly, even with a very limited number of transcriptional factors, genome wide analyses of some of the smallest and simplest mycoplasmas revealed an unexpected level of complexity and versatility in their in metabolic responses to environmental conditions. Their adaptation seems to be similar to that of more complex bacteria, providing hints that other, unknown regulatory mechanisms might exist (Yus et al. 2009). Furthermore, proteome complexity could not be directly inferred from the composition and organization of their minimal genomes or even their extensive genome wide transcriptional analysis (Kuhner et al. 2009). In addition, transcriptome analysis also showed surprisingly unanticipated diversity and heterogeneity in mycoplasma transcription profiles including the presence of many

operons, the production of alternative transcripts in response to environmental perturbations, and the high frequency of antisense RNA (Guell et al. 2009). Therefore, it is becoming increasingly clear that mycoplasmas are able to respond to environmental cues and regulate gene expression, but the underlying mechanisms for their global and individual gene regulation are not understood.

Tools for Genetic Manipulation in Mycoplasmas

A major challenge to understanding the fundamental biology of *Mollicutes* at the molecular level and at the host:pathogen interface is our limited capacity to genetically manipulate mycoplasmas ((Renaudin 2002, Halbedel et al. 2007). The lack of genetic tools to specifically manipulate and disrupt genes in mycoplasmas or to carry out complementation studies has been a major impediment in understanding their genetics and elucidating regulatory mechanisms ((Renaudin 2002, Halbedel et al. 2007). At present, four basic approaches have been reported for genetic manipulation of mycoplasmas: random transposon mutagenesis, targeted double-crossover homologous recombination using suicide plasmids, targeted recombination using *oriC* plasmids, and creation of new mycoplasma strains from genomes that have been cloned and engineered in yeast.

The most widely used approach to genetic manipulation of the *Mollicutes* has been the use of transposon-based mutagenesis (Chopra-Dewasthaly et al. 2005, French et al. 2008, Glass et al. 2006, Halbedel et al. 2007, Hutchison et al. 1999, and Voelker, et al. 1998). Transposons have been used successfully in a global approach of gene inactivation to define the minimally essential genes for sustaining life and also to generate mutants of interest (Glass et al. 2006, Hutchison et al. 1999). Recently, Janis et al. 2008 incorporated the *γδ* resolvase gene into an *OriC* plasmid and used this

construct to transform a randomly generated mutant of *M. mycoides* subsp *mycoides*, resulting in a mutant which lacked the majority of the transposed sequences introduced by the first random insertional mutagenesis step. All of these approaches are limited in that random insertion of transposons does not allow the specific targeting of a gene of interest and requires screening of large numbers of transformants to identify an insertion in a specific gene. In contrast to random insertional mutagenesis, targeted gene disruption has been far less successful in mycoplasmas.

Two different strategies have been used for targeted gene disruption in Mollicutes: the use of replicating plasmids and the use of non-replicating (suicide) plasmids. Replicating plasmids based on *oriC,* a chromosomal region that harbors the *dnaA* gene and adjacent DnaA boxes have been developed for many *Mollicutes* species (Chopra-Dewasthaly et al. 2005a, Chopra-Dewasthaly et al. 2005b, Chopra-Dewasthaly et al. 2008, Cordova et al. 2002, Cox, 2007, Duret et al. 2005, Duret et al. 2003, Halbedel and Stulke 2007, Janis et al. 2005, Jarhede et al. 1995, Lartigue et al. 2003) although for some mycoplasmas efforts to create replicating plasmids have been unsuccessful. These plasmids were able to replicate within their respective hosts (Renaudin et al. 2005).These *oriC* plasmids have been used for heterologous gene expression as well as for targeted gene disruption by single-crossover recombination (Chopra-Dewasthaly et al. 2005b, Chopra-Dewasthaly et al. 2008, Duret et al. 2005, Duret et al. 2003, Janis et al. 2005, Lartigue et al. 2003). The underlying concept behind the use of homologous OriC-based plasmid as a means for targeted gene integration is that replicative plasmids increase the time over which homologous

recombination can take place compared to non-replicative “suicide” plasmids (Duret et al. 1999, Janis et al. 2005).

Homologous OriC plasmids have been developed for some mycoplasmas including the mycoides cluster (Lartigue et al. 2003), *M. pulmonis* (Cordova et al. 2002), *M. agalactiae* (Chopra-Dewasthaly et al. 2005), and *S. citri* (Ye et al. 1994). These replicating plasmids use minimal sequences of the respective chromosomal OriC in order to limit their integration at the homologous chromosomal OriC region and favor their chance to integrate at the site of interest (Janis et al. 2005, Renaudin et al. 1995). Although this approach has met with limited success, several confounding issues regardless of the type of *oriC* used in the plasmid construct do exist.

Even when integration of the OriC plasmid at the gene of interest occurred, consistently the frequency was still very low and occurred through a single cross-over event (Chopra-Dewasthaly et al. 2005a, Lartigue et al. 2003, Cordova et al. 2002). In other cases, homologous OriC plasmids replicated freely and did not integrate at either the desired gene or chromosomal origin of replication, despite the presence of long sequences of perfect homology (Janis et al. 2005, Cordova et al. 2002). Eventually, this led to loss of the gene copy from the integrative plasmid, presumably through deletion. Another complication was that the disruption of the gene of interest occurred, but there was another intact copy of the gene present on the freely replicating plasmid; this made it difficult to recognize the mutant phenotype and also confounded the interpretation and the analysis of the experiments (Janis et al. 2005). The instability of the mutants was another concern and ensued as a consequence of the incompatibility of the two OriC

regions carried on the same chromosome, resulting in the resolution and then the loss of the integrated material (Cordova et al. 2002, Ogasawara et al. 1991).

Classical double-crossover homologous recombination using a suicide plasmid is potentially a powerful technique to delete mycoplasma genes. However, it has been reported previously only in *Mycoplasma genitalium* (Burgos et al. 2008, Dhandayuthapani et al. 1999; Dhandayuthapani et al. 2001), and *M. pneumoniae* (Krishnakumar et al. 2010). To date, targeted gene disruption in *M. genitalium* stands out as the only successful attempt to inactivate genes through a classical double-crossover homologous recombination using a suicide plasmid (Dhandayuthapani et al. 1999, Dhandayuthapani et al. 2001). The majority of other reported attempts to create double-crossover mutations in target genes, including the *lppA* gene in *M. capricolum* (Janis et al. 2005, Muto et al. 2002), the hemolysin A(*hlyA*) gene in *M. pulmonis* (Cordova et al. 2002), the motility gene *scm1* in *Spiroplasma citri* (Duret et al. 1999), failed to produce any transformants at all, not even a single cross-over integration despite extensive sequence homology with the target gene. A few groups did produce transformants in *M. gallisepticum* and *Acholeplasma laidlawii*. Two studies were performed in *M. gallisepticum.* The first was aimed at proving the ability the of *M. gallisepticum* to carry out homologous recombination. The suicide plasmids, carrying randomly cloned restriction fragments with uninterrupted perfect homology up to 5 Kb in size, were able to integrate into the chromosome (Cao et al. 1994). The second study inactivated the surface protein P47 through a single-crossover plasmid integration (Markham et al. 2003). Disruption of the *A. laidlawii recA* gene was accomplished through a single crossover recombination event, and the plasmid integration was found

to depend on the presence of a fully functional RecA. The transformation frequency, however, was 10 fold less than that of the replicating plasmid or that of suicide plasmid containing Tn916 (Dybvig et al. 1992).

Recently, a new method for manipulating *Mollicutes* genomes was reported where the genome of *Mycoplasma mycoides* subsp. *capri* was cloned as a YAC. Once in a yeast cell, the mycoplasmal genome could readily be manipulated using the significant power of yeast genetic tools. Then the altered yeast artificial chromosome (YAC) was transplanted back into a recipient *Mycoplasma capricolum* cell to produce a new strain of *M. mycoides* subsp. *capri* (Lartigue et al. 2009). This approach is applicable for large scale-genome alterations, but may be expensive and technically challenging for investigators interested in specific gene targets.

The Mycoides Cluster

The *Mycoplasma mycoides* "cluster" of organisms is among the most virulent of the mycoplasmas. They are of major world-wide economic significance in diseases of cattle and goats (Rodreguez et al. 1995). There are six groups recognized within the cluster, and these mycoplasmas are closely related as evidenced by serological cross reactions and a 50-60% DNA homology (Abou-Groun et al. 1994, Cottew 1987). Two members of the cluster, *M. mycoides* subsp. *mycoides* Small Colony type and *M. capricolum* subspecies *capripneumoniae* (formerly F38 group), are listed as class B agents of special concern by USDA and APHIS (Federal Register 67, No. 155, 9 CFR 121.2b) and cause contagious bovine and caprine pleuropneumonia, respectively. Based on recent DNA analysis separate subspecies designation has been given to *M. capricolum* subsp. *capricolum* and (formerly *M. mycoides* subsp. *mycoides* Large

Colony type) and distinguishes them from the causative agents of contagious pleuropneumonia (Manso-Silván et al. 2009).

Mycoplasma mycoides subsp. *capri* GM12 mainly causes pneumonia in sheep and goats. It is also associated with mastitis and arthritis of sheep and goats as well as septicemia in goats (Frey 2002). One phenotypic difference between the USDA class B agents causing contagious pleuropneumonia and the two closely related subspecies is the ability to degrade casein, an activity that is quite uncommon in mycoplasmas. We have identified the gene encoding the proteolytic activity in both *M. mycoides* subsp. *capri* GM12 type ATCC 35297 (DaMassa et al. 1983) and *M. capricolum* ATCC 27343 (Leach et al. 1993). The predicted gene (ZP_02512724) product is a membrane protease, has a tail specific protease (TSPc) domain, and confers a proteolytic phenotype that can be readily determined by growth on agar supplemented with casein. Because of its homology with the carboxyl tail specific processing domain and other similar proteins in other bacterial species, we have referred to this gene as ctpA.
Our long term goal is to understand mechanisms of gene regulation in mycoplasmas. Critical to this goal is the ability to target specific genes for mutagenesis, preferably by double cross-over events. Because of the lack of selection and phenotypic markers in most mycoplasmas, we chose to target a gene that conferred a readily identified phenotype (casein hydrolysis).

Goals of Study and Summary of Findings

The specific goals of this study are to develop a new approach for targeted mutagenesis using a suicide plasmid, to characterize the mutant with the gene disruption, and to determine if this gene could be used as a potential model for gene regulation in mycoplasmas.

We developed a suicide plasmid employing *RecA* from *Escherichia coli* under the direction of a mycoplasmal promoter and with *tetM* as a selection marker. The ctpA gene of *M. mycoides* subs. *capri* was targeted for disruption. We consistently obtained a marked improvement in the percentage of transformants with the desired gene disruption, all of which occurred by targeted double crossover event. To our knowledge, this is the only mycoplasma species other than *Mycoplasma genitalium* in which a targeted double crossover mutation has been obtained. We also demonstrated that the disruption of this gene had a significant impact on the proteomic profile of the microbe, most notably with respect to proteins associated with oxidative stress, protein repair, and a shift form glycolysis to glycerol metabolism. Finally, we determined that the mutation altered the manner in which *M. mycoides* subs. *capri* responds to heat shock.

CHAPTER 2
TARGETED HOMOLOGOUS RECOMBINATION IN *MYCOPLASMA MYCOIDES* SUBSP. *CAPRI* IS ENHANCED BY INCLUSION OF HETEROLOGOUS *RECA*

Introduction

Mycoplasmal infections present one of the most intriguing paradoxes in infectious diseases. On one hand, these microbes are exquisitely simple with the smallest genomes of free-living organisms (Fadiel et al. 2007, Krause and Balish 2004, Rocha and Blanchard 2002). Conversely, the range of infections and host species are among the most diverse in the microbial world (Fadiel et al. 2007). Perhaps the greatest challenge to understanding the fundamental biology of *Mollicutes* at the molecular level and at the host:pathogen interface is our limited capacity to genetically manipulate mycoplasmas (Halbedel and Stulke 2007, Pilo et al. 2007). At present four basic approaches have been reported: random transposon mutagenesis, targeted double-crossover homologous recombination using suicide plasmids, targeted recombination using *oriC* plasmids, and creation of new mycoplasma strains from genomes that have been cloned and engineered in yeast.

The most widely used approach to genetic manipulation of the *Mollicutes* has been the use of transposon-based mutagenesis (Chopra-Dewasthaly et al. 2005b, French et al. 2008, Glass et al. 2006, Halbedel and Stulke 2007, Hutchison et al. 1999, Voelker and Dybvig 1998). Transposons also have been used successfully in a global approach of gene inactivation to define the minimally essential genes for sustaining life and also to generate mutants of interest (Glass et al. 2006, Hutchison et al. 1999). Recently, Janis et al. 2008 (Janis et al. 2008) incorporated the *γδ* resolvase gene into an *OriC* plasmid and used this construct to transform a randomly generated mutant of *M. mycoides* subsp *mycoides*, resulting in a mutant which lacked the majority of the

transposed sequences introduced by the first random insertional mutagenesis step. All of these approaches are limited in that random insertion of transposons does not allow the specific targeting of a gene of interest and requires screening of large numbers of transformants to identify an insertion in a specific gene.

In contrast to random insertional mutagenesis, targeted gene disruption has been far less successful in mycoplasmas. Replicating plasmids based on *oriC,* a chromosomal region that harbors the *dnaA* gene and adjacent DnaA boxes, have been developed for *Mollicutes* species (Chopra-Dewasthaly et al. 2005a, Chopra-Dewasthaly et al. 2005b, Chopra-Dewasthaly et al. 2008; Cordova et al. 2002, Cox 2007, Duret et al. 2005, Duret et al. 2003, Halbedel and Stulke 2007, Janis et al. 2005, Jarhede et al. 1995, Lartigue et al. 2003) and have been used for heterologous gene expression as well as for targeted gene disruption by single-crossover recombination (Chopra-Dewasthaly et al. 2005b, Chopra-Dewasthaly et al. 2008, Duret et al. 2005, Duret et al. 2003, Janis et al. 2005, Lartigue et al. 2003). Unfortunately, this approach still has relatively limited success and also has several confounding issues, regardless of the type of *oriC* used in the plasmid construct. Although classical double-crossover homologous recombination using a suicide plasmid is potentially a powerful technique, recombination by double-crossover has been reported only for *M. genitalium*, and even then it occurs at a very low frequency (Burgos et al. 2008, Dhandayuthapani et al. 1999, Dhandayuthapani et al. 2001). Recently a new method for manipulating *Mollicutes* genomes was reported where the genome of *Mycoplasma mycoides* subspecies *capri* was cloned as a YAC. Once in a yeast cell, the mycoplasmal genome could readily be manipulated using the significant power of yeast

genetic tools. The altered YAC was then transplanted back into a recipient *Mycoplasma capricolum* cell to produce a new strain of *M. mycoides* subspecies *capri* (Lartigue et al. 2009).

We report here the development of a novel suicide plasmid employing *RecA* from *E. coli* under the direction of a mycoplasmal promoter and with *tetM* as a selection marker that resulted in consistent recovery of targeted stable double-crossover mutants, with a 140 fold increase over random insertional mutagenesis with respect to the number of desired gene disruptions relative to the number of total transformants generated.

Materials and Methods

Mycoplasma Strains and Cultivation

Mycoplasma mycoides subsp. *capri* GM12 type ATCC 35297 (*M. mycoides capri* (DaMassa *et al.*, 1983), formerly known as *M. mycoides* subsp. *mycoides* Large Colony type (Manso-Silvan et al. 2009); and *Mycoplasma capricolum* ATCC 27343 (Leach et al. 1993) were grown at 37 °C in SP4 medium (Tully *et al.*, 1979); SP4 casein agar was supplemented with a final concentration of 1% skim milk (DIFCO, Detroit, MI). For growth of mutants, tetracycline (final concentration, 5µg/ml) was added to the media.

PCR Primers and Conditions.

The primers used for PCR reactions are given in Table 2-1. The reactions used either 1µg of genomic DNA or 100 ng of plasmid DNA as a template for the PCR reactions. For a given reaction, 30 pmol of primer (Genosys Sigma-Aldrich Co) was used. To facilitate cloning, restriction enzyme recognition sites were created in selected primers (Table 2-1). The PCR reactions were performed using Applied Biosystems Perkin Elmer GeneAmp 2400 PCR System (see supplementary material for details).

Cloned Pfu DNA Polymerase (Stratagene) was used with primer pair 1; Taq DNA polymerase (Invitrogen) was used with primer pairs 2, 3 and 5; TaqPlus® Long PCR System (Stratagene) was used for primer pair 4. All reactions were held at 95^{o}C for 3 min and subjected to 25 cycles of template denaturation at 95°C for 1 min, except *tetM* amplifications which were allowed to denature for 3 min. Primer annealing occurred for 1 min at 42°C (primer pairs 1 and 2), 50°C (primer pair 3) or at 40°C (primer 4) or for 30 sec at 44°C (primer pair 5). Polymerization at 72°C occurred for 2 min (primer pairs 3 and 5), 3 min (primer pairs 1 and 2), or 10 min (primer pair 4). Polymerization was then followed by a final extension for 7 min at 72°C for all reactions.

Construction of Plasmids

An overview of the constructs used is shown in Figure 2-1. To construct pExp1-*ctpA,* the *tetM* gene, obtained from plasmid pIVT-1 and flanked by 828 bp of the 5' end of the *ctpA* gene and 456 bp of the *ctpA* 3' end, was cloned into pExp1-*ctpA* to produce pExp1-*ctpA*::*tetM* (Figure 2-1). Specifically pExp1-*ctpA* was digested with BclI and PstI, the two fragments were separated by gel electrophoresis, and the larger fragment was gel purified. Plasmid pIVT-1 was digested with BamHI and PstI, the fragment containing the *tetM* gene was gel purified and ligated to the pExp1-*ctpA* backbone to produce pExp1-*ctpA*::*tetM* (Figure 2-2, A and B). Genomic DNA was prepared from *M. mycoides capri* using the Epicentre DNA extraction kit according to the instructions provided. The *ctpA* gene was amplified using primer pair 1 (Table 2-1), gel purified, and cloned into Topo ENTER vector (Invitrogen). The plasmid pENTER*ctpA* was propagated in *E. coli* Top10 strain and transferred to the cloning vector EXP1-DEST using the LR Clonase kit (Invitrogen). The presence and direction of the cloned *ctpA* gene was confirmed by restriction enzyme digestion at each step (Figure 2-2). The size of the flanking

sequences was chosen to provide sufficient homologous sequences to facilitate recombination in the target gene (Dhandayuthapani *et al.*, 1999; Duret *et al.*, 2005). The direction as well as the insertion was verified by digestion with KpnI (Figure 2-2B).

To construct pEXP-*ctpA::tetM-recAec*, the *E. coli recA* gene was cloned under the upstream sequences and promoter region of *ctpA*. The upstream sequences and promoter region (USE) as well as the first 12 nt of *ctpA* coding sequence were amplified using primer pair 2 (Table 2-1), gel purified , and cloned in the Topo vector pXL-PCR (pXL-USE*ctpA*). PCR primer pair #2 (Table 2-1) amplifies the upstream sequences, the promoter region, and the *ctpA* coding sequence. Construction of pExp1-*ctpA::tetM-recAec* is shown in Figure 2-2, C and D. The PCR amplified fragment (USE*ctpA*) was gel purified, cloned in the Topo vector pXL-PCR, and the insertion of USE*ctpA* as well as the correct direction was confirmed by digestion with BamHI. The coding sequence of the *E. coli recA* gene was amplified from genomic DNA using primer pair 3 (Table 2-1). The PCR product containing the coding sequence of the *recA* gene was gel purified and cloned in Topo vector pXL-PCR, producing the plasmid pXL-*recA*. Plasmid pXL*recA* was digested with SpeI, and the smaller fragment containing the *recA* gene was gel purified. Plasmid pXL-USE*ctpA* was digested with SpeI, and the backbone (larger fragment) was gel purified. The backbone fragment of pXL-USE*ctpA* and the *E. coli recA* were ligated together, and *E. coli* strain Stbl2 was transformed. The resulting plasmid pXL-USE*recA* contained the *E. coli recA* gene inserted in frame behind the first four amino acids of *M. mycoides capri ctpA* gene. The expression of *recAec* was regulated by the promoter and the upstream sequences of *M. mycoides capri ctpA*.

The capability of the promoter and the upstream sequences of *M. mycoides capri ctpA* to drive expression of an *E. coli* gene was confirmed by placing *lacZ* under the direction of *M. mycoides capri ctpA* USE and $ctpA_p$ (Figure 2-3). Using primer pair #5 (Table 2-1) and identical methodology as for *recA*, *E. coli lacZ* was cloned into the same construct and position as *recA* to produce plasmid pXL-USE*lacZ*. After PCR amplification of *lacZ* with an engineered SpeI site at the 5' end, the PCR product was cloned into pCR-XL Topo vector to produce pXL-(SpeI)*lacZ*. pXL-(SpeI)*lacZ* was digested with SpeI to produce the (SpeI)*lacZ* fragment. The fragment was gel purified and ligated to the same plasmid backbone (pXL-USE$ctpA_p$) used to create pXL-USE$ctpA_p$*recA*ec. This placed the *lacZ* expression under the regulatory control of the USE and promoter of the *ctpA* gene. The construct was digested with NdeI and XbaI to verify the correct orientation of the cloned fragment. The plasmid was introduced into *E. coli* Top10 F′ genotype {F´{lacIq Tn*10* (TetR)} *mcr*A ∆(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ*∆M15 ∆*lac*X74 *recA*1 *ara*D139 ∆(*ara-leu*)7697 *gal*U *gal*K *rps*L *end*A1 *nup*G}(Invitrogen). Transformants were plated on LB agar coated with X-gal; a blue color confirmed the capacity of the *ctpA* promoter and the USE to drive the expression of the downstream gene. The same construct, but without USE and $ctpA_p$, was used as a control and was ß-galactosidase negative.

PEG8000-Mediated Chemical Transformation.

An overnight culture of *M. mycoides capri* was diluted 1/1000 in 40 ml of fresh SP4 broth. The culture was grown until the mid-log phase (~7-8 hrs) and then placed on ice for 2 min. *M. mycoides capri* was pelleted by centrifugation at 12,000g for 30 min at 4 ºC, the supernatant was decanted, and the cells were washed in 10 mM Tris buffer, pH 6.5. Cells were centrifuged at 12,000g for 30 min at 4 ºC, resuspended in 1 ml of 0.1 M $CaCl_2$, and incubated on ice for 1 hr. Yeast t-RNA (10 µg, Sigma-Aldrich), plasmid DNA (30 µg), and 9 ml of 60% (w/v) PEG8000 in 10 mM Tris pH 6.5 (final concentration, 54% PEG8000) were added, and the cells were incubated for 2 min at room temperature. Twenty-five ml of 10 mM Tris buffer, pH 6.5, was added, and the cells were centrifuged at 12,000g for 30 min at 4ºC. The supernatant was decanted, and the cells were suspended in 2 ml of warm (37ºC) SP4 broth supplemented with 2 mM $MgCl_2$ (Lavery et al. 1992, Hoffman et al. 2000, Goryshin et al. 1998, Goryshin et al. 2000) and incubated at 37ºC for 2 hrs. Following the 2 hr recovery time, the cells were plated on SP4 agar plates containing 5 µg/ml tetracycline and 1% casein. Colony growth was observed after 48 hr. All colonies were picked, expanded in SP4 broth, and genomic DNA obtained. PCR was performed using primer pair 4 to detect the disruption of *ctpA* gene.

Random and Targeted Transformation

For random mutagenesis, *M. mycoides capri* was transformed with Tn4001T, a gift from K. Dybvig, (plasmid pIVT-1(Dybvig et al. 2000) by PEG8000 (Sigma-Aldrich)-mediated chemical transformation as previously described (Dybvig et al. 2000, Lartigue et al. 2009),. For targeted mutagenesis, *M. mycoides capri* GM12 was transformed with pExp1-*ctpA::tetM* or pExp1-*ctpA*::*tetM-recA*ec using PEG8000-mediated chemical

transformation. *M. capricolum* was transformed by electroporation. Disruption of *ctpA* was confirmed by loss of phenotype on casein agar and by PCR of the flanking regions of the *ctpA* coding sequence. The location, precise site of insertion, and direction of insertion of *tetM* into the coding sequence of *ctpA* were determined by sequencing.

The random mutagenesis of *M. capricolum* was performed at the J. Craig Venter Institute in collaboration with Dr. John Glass, and selected mutants were provided to us. Ten ml cultures of log phase *M. capricolum* cells were pelleted and washed twice with electroporation buffer (EB) comprised of 8 mM HEPES + 272 mM sucrose at pH 7.4. Cells were then scraped into 25 ml of 4°C EB, and the suspension triterated to break up any cell clumps, after which the cells were pelleted by centrifuging for 10 min at 4^{o}C at 4575g. The supernatant was decanted, and the centrifuge tubes were inverted a few minutes to insure there was no residual supernatant, which could contain enough salts from the medium to compromise electroporation. Cells were resuspended in a total volume of 200-300 ml EB. On ice, 100 ml cells were mixed with 30 mg pIVT-1 plasmid DNA, transferred to a 2 mm chilled electroporation cuvette (BioRad, Hercules, CA), and electroporated using 2500 V, 25 mF, and 100 Ω. After electroporation cells were resuspended in 1 ml of 37°C SP4 medium and the cells were allowed to recover for 2 hours at 37^{o}C with 5% CO_2. Aliquots of 200 ml of cells were spread onto SP4 agar plates containing 2 mg/l tetracycline hydrochloride (VWR, Bridgeport, NJ). The plates were incubated for 3-4 weeks at 37^{o}C with 5% CO_2 until colonies were visible.

Confirmation of Disruption of *ctpA* and Location of Insertion Sites

For *M. mycoides capri*, genomic DNA was sequenced using primer pair # 6 (Table 2-1) to verify *tetM* in these mutants. The precise site of insertion was further pinpointed by sequencing the junction between the *tetM* gene and the flanking sequence of the *ctpA* gene using primer 6 and primer pair # 4 (Table 2-1). The location and direction of insertion of *tetM* into the coding sequence of *ctpA* was confirmed by sequencing. This was done for both *M. mycoides capri* ctpA::tetM mutant created by random mutagenesis using Tn4001T (Table 2-2) and a representative mutant created by homologous recombination using pExp1-*ctpA::tetM-recAec* (Tables 2-3 and 2-4) . In the *M. capricolum* ctpA::tetM mutant generated by random mutagenesis, Tn4001T inserted 1269 bp from the 5′ end of the 1,968 bp *ctpA* gene (MCAP_0240, YP_424227, J. Glass, personal communication). *tetM* insertion sites in *M. capricolum* were located by DNA sequencing from genomic templates containing ~0.5 µg of genomic DNA. A 30 base oligonucleotide GTACTCAATGAATTAGGTGGAAGACCGAGG (Integrated DNA Technologies, Coralville, IA) primer that binds in the *tetM* gene103 base pairs from one of the transposon/genome junctions was used, and, the insertion site was located on the *M. capricolum* genome using BLAST.

Southern and Northern Blots

For Southern blots, genomic DNA samples were prepared (DNeasy Tissue kit, Qiagen) from *M. mycoides capri* (wt), *M. mycoides capri ctpA::tetM* created via homologous recombination using pExp1-*ctpA*::*tetM-recA*ec and from *M. mycoides capri ctpA*::Tn4001T created by random mutagenesis. Following electrophoresis on a 0.8% agarose gel and transfer to a positively charged 1.2µm nylon membrane (Roche), membranes were hybridized with labeled probes (DIG High Prime DNA labeling and

Detection Starter Kit I, Roche) to either a 1,157 bp *tetM* fragment (Table 2-5) obtained from plasmid pMP05 (Cordova et al. 2002), courtesy of A. Blanchard and J. Renaudin, or a 1,018 bp fragment (Table 2-6) with homology to the plasmid backbones used in preparation of the constructs**.**

For Northern blots, total RNA was extracted from early stationary phase *M. mycoides capri* or *M. mycoides capri ctpA::tetM* using RiboPure-Bacteria Kit (Ambion) according to the manufacturer's instructions. The NorthernMax kit (Ambion) protocols were followed for gel separation, transfer of RNA to the membrane, and hybridization steps. A 1,446 bp *ctpA* DIG-labeled probe (Table 2-7) was generated by random primed DNA labeling using DIG High Prime DNA labeling and Detection Starter Kit I (Roche) according to the manufacturer's instructions.

Results

Inclusion of *recA* from *E. coli* in disruption plasmid II (pExp1-*ctpA*::*tetM-recA*ec) greatly enhanced the recovery of *ctpA* mutants (Table 2-8). In an earlier study in our laboratory, random mutagenesis of *M. mycoides capri* using the conjugative transposon Tn916 (Dybvig and Cassell 1987, Dybvig and Alderete 1988, Whitley and Finch 1989) yielded 1,776 transformants, but only two (0.11%) had an insertion in the gene of interest. When the smaller plasmid, Tn4001T (Dybvig et al. 2000) was used in this current study, the efficiency of obtaining the desired mutant was increased to 1/674 (0.15%) for *M. mycoides capri* and 1/384 (0.26%) for *M. capricolum*. In three independent experiments using pExp1-*ctpA*::*tetM* to create a targeted mutation, only one *ctpA::tetM* mutant of 152 tetracycline-resistant clones (0.66% efficiency) was obtained as a result of homologous recombination. Although this approach required fewer transformants be screened (1/152 for an overall efficiency of 0.65%) than for

random mutagenesis, only one of three experiments yielded a ctpA::tetM mutant and this lack of consistency was a concern. When *recA* was added to the construct, a dramatic and consistent increase in the efficiency of obtaining a ctpA::tetM mutant was observed. Using pExp1-*ctpA::tetM-recAec,* we obtained ctpA::tetM mutants in each of three independent transformation experiments, with 21.4 % to 26.7% of transformants having a disruption in the targeted gene. Although the overall number of transformants obtained with pExp1-*ctpA::tetM-recAec* was low (<20 per experiment), the high recovery rate of transformants with the target gene disrupted far outweighed the minor difficulty of having to use at least a log higher number of wild type *M. mycoides capri* for transformation with pExp1-*ctpA::tetM-recAec* than for transformation with Tn4001T.

Disruption of *ctpA* resulted in loss of the proteolytic phenotype (Figures 2-4 and 2-5). Insertion of *tetM* into the *ctpA* gene via targeted mutagenesis with pExp1-*ctpA::tetM-recAec* (Figure 2-4, A and C) resulted in the loss of proteolytic activity on casein agar (Figure 2-4, B, and Figure 2-5). All *M. mycoides capri ctpA::tetM* mutants had identical inserts of the expected size (Figure 2-5, C) as well as loss of the proteolytic phenotype. Disruption by Tn4001T random insertional mutagenesis of both the *M. mycoides capri ctpA gene* and the *ctpA* homologous gene in *Mycoplasma capricolum* (YP_424227) resulted in loss of the proteolytic phenotype (see Figure 2-5, panels B, C and E). The presence of *tetM* in all mutants, regardless of method of transformation, was confirmed by PCR (Figures 2-4, A, C and 2-5, lanes 4, 5 and 7).

The presence and location of the *tetM* gene in *ctpA::tetM* mutants was confirmed by Southern blot (Figure 2-6) and sequencing (Tables 2-3 and 2-4), respectively**.** The acquisition and copy number of the *tetM* gene in *ctpA::tetM* mutants and the presence

of the *tetM* gene inserted by Tn4001T-mediated random mutagenesis was verified by Southern blot (Figure. 2-6, A). Based on the size of the PCR fragment obtained using primers specific for the flanking regions of the *ctpA* gene (Table 1, primer pair 4) and sequencing (Tables 2-3 and 2-4) provided in supplementary material), a double crossover homologous recombination was determined to have occurred when *M. mycoides capri* was transformed with pExp1-*ctpA::tetM-recAec*. The plasmid backbone was absent (Figure 2-6 B, lane 3) in the ctpA::tetM mutant generated by pExp1-*ctpA::tetM-recAec*, while the ctpA::tetM mutant generated by Tn4001T random insertional mutagenesis retained the plasmid backbone (Figure 2-6 B, Lane 4).

CtpA appears to be monocistronic. The absence of transcription of *ctpA* in the mutants was confirmed by Northern blot (Figure 2-7). The length of the transcript in *M. mycoides capri* was the same size as the mature ORF, supporting the monocistronic transcription of *CtpA*. Additionally, disruption of either the upstream or downstream gene had no impact of the proteolytic phenotype.

Discussion

The likelihood of obtaining an insertion in the gene of interest by random insertional mutagenesis varies depending on the mycoplasma species and the specific gene of interest. Even if the desired insertion occurs, screening of large numbers of transformants is required. As an example, less than 0.2% of *M. mycoides capri* or *M. capricolum* transformants obtained using Tn4001T contained insertions in the *ctpA* gene. Further, random insertional mutagenesis does contain an inherent bias for preferred insertion sites (Boesen et al. 2004, Glass et al. 2006).

Cloning and engineering mycoplasma genomes in *Saccharomyces cerevisiae* offers all the genetic tools available to yeast biologists for seamless manipulation of

mycoplasma genomes. Thus multiple modifications are possible. The main drawback of this method is that it is technically very demanding and laborious, and costly. Additionally, the yeast cloning process takes two to three weeks to complete (Lartigue et al. 2009).

Targeted gene disruption in *Mollicutes* using replicating plasmids based on *oriC* has been reported for a number of mycoplasma species, including members of the Mycoides, Pneumoniae, and Hominis subgroups (Chopra-Dewasthaly et al. 2005b, Chopra-Dewasthaly et al. 2008, Duret et al. 2005, Duret et al. 2003, Janis et al. 2005, Lartigue et al. 2003, Lee et al. 2008). However, mutagenesis by single-crossover homologous recombination using oriC plasmids has met with limited success. Even in *M. gallisepticum* and the Mycoides group species in which the OriC plasmids work best, replicating plasmids require many passages before the plasmid integrates into the chromosome (Cordova et al. 2002, Lee et al. 2008) and when obtained, the mutants are not stable. Replicative plasmids based on *OriC* actually were less efficient than suicide plasmids in generating targeted mutants (Dybvig and Woodard 1992, Kannan and Baseman 2006) and also less effective than our results using both pExp1-*ctpA::tetM* and pExp-1 *ctpA::tetM:recAec*.

In mycoplasmas, as in other bacteria, *recA* is the only recombination gene universally present, although it is not part of the essential gene set (French et al. 2008, Glass et al. 2006). Studies have confirmed the critical role of RecA in recombination events in *Mollicutes* (Ogasawara et al. 1991, Rocha et al. 2005). In *Acholeplasma laidlawii,* the *recA* gene was disrupted through a single crossover recombination event (Dybvig and Woodard 1992) and plasmid integration was found to depend on the

presence of a fully functional RecA. Therefore, it seems that homologous recombination in mycoplasmas, when it occurs, is likely RecA-dependent (Dybvig and Woodard 1992). Thus, we reasoned we might be able to augment homologous recombination by expression of a heterologous source of RecA with a high GC content to minimize hybridization with the indigenous *RecA*.

In our study, inclusion of the *E. coli RecA* in our construct resulted in a 140-fold increase in recovery of the desired mutant over other techniques currently used for genetic manipulation of mycoplasmas in general and members of the *Mycoides* cluster specifically. Although the overall number of transformants obtained with pExp1-*ctpA::tetM-recA*ec was relatively low (<20 transformants) and the number of wild type *M. mycoides capri* needed for transformation was high, >20% of the transformants in each experimental repetition had the insertion in the targeted gene. Importantly, all ctpA::tetM mutants that were obtained using pExp1-*ctpA::tetM-recA*ec occurred by double crossover homologous recombination and are therefore far more likely to be stable. Unlike the mutant obtained by random insertional mutagenesis using Tn4001T, the mutants obtained with pExp1-*ctpA::tetM-recA*ec did not contain remnants of the plasmid backbone, which may also enhance stability. To the best of our knowledge, a successful classical double-crossover homologous recombination using a suicide plasmid has been reported previously only in *M. genitalium* (Burgos *et al.*, 2008; Kannan and Baseman 2006). Importantly, this simple yet elegant approach provides a significant advance in our ability to manipulate the *M. mycoides capri* genome. Like *OriC*, this technique may have applicability in other mycoplasmal species. This technique provides a new and promising approach that provides an additional genetic

tool to use in unraveling the pathogenic mechanisms by which these microbes induce disease.

Table 2-1. Primer pairs used for PCR.

Primers	5' →3' sequence
1 For	*ctpA* Coding sequence For: CACCATGAAACTAGTTAAAAAAATAG
1 Rev	*ctpA* Coding sequence Rev CTTACTAAGGTTAATTTTTGTTATTTTAA
2 For	[a]USE*ctpA*For : GGATCCCATTTTTATACAAATGTTTTTTTCTTTGG
2 Rev	*ctpA* Coding sequence Rev: CTTACTAAGGTTAATTTTTGTTATTTTAA
3 For	[b]*E.coli RecA* For: ACTAGTTATGGCTATCGACGAAAACAAACAG
3 Rev	[c]*E.coli RecA* Rev: CTCTAGAGATGCGACCCTTGTGTATCAAAC
4 For	*CtpA* Flanking For(thi): GGATCCGTTTATAAAAATGGAGAATTTGCACAA
4 Rev	*ctpA* Flanking Rev: CTATTCACTATAGTATTAGGAAAAAAA
5 For	*tetM* BB14 : CGTATATATGCAAGACG
5 Rev	*tetM* BB15 : TTATCAACGGTTTATCAGG
6	Tn4001tSeq : GTACTCAATGAATTAGGTGGAAGACCGAGG

Restriction enzyme sites that were created for cloning purposes are underlined. [a]GGATCC is an introduced BamHI site. [b]ACTAGT is an introduced SpeI site. [c]TCTAGA is an introduced XbaI site.

Table 2-2. Sequence for *M. mycoides capri ctpA::tetM* mutant created by random mutagenesis using Tn4001t.

Nucleotide	Sequence
1	CCGAGGCACT GCATAACATC TTCCGCAGTA CCGCCCGATT CCACCTGTAT
51	AATCGCAAGA AGTATGTTGG GACTTTTACA CAATTATACG GACTTTATCA
101	TTATGTTTTA ACTTCACCAT TTTCATTCTC AGCTGGAAAC ATTTTCCCTC
151	AACTAGTTAA AGATAATAAT GTTGCAAAAG TAATTGGATT TAAAACTGCA
201	GGTGGAGCTT CAGCGATTAG TCAAGCAATT CTACCAACTG GAGATATTAT
251	TCAATTAAGT AGTAATAATG TTTTAACTAA TAAATCTCAT CAAAGTTTAG
301	AATATGGTGT TAATCCAGAT ATTACACTTG GATTTGATCC ATTCAAACAA
351	ACTGAAAAAT TCTTTGATTC AGCTTATATT CAACAAGCTA TTAATAAAGA
401	TACAAATACA TTAAATTCAA TTCCAGCTAC TCATTCTAGT GTTGTTGAAC
451	CAAATTATGT ACATAAACTT GTAGAACAAC CTCAACCACT ACAATTAAGT
501	AGAAAAACTG ATGAAACTGA AATAAAAAAT CTTAATAATT TATTTTCTAG
551	TATAAAAGAA ACCGAAAGAA AAGATGCATA TTTTGTACTT GGAGCACTTG
601	GTGTTGTTAT TAGTTTAGCA ATCTCATTTG TAATTATTAA AAAAATATTA
651	AAATAACAAA AATTAACCTT AGTAAGAAAT ACTAAGGTTT TTTTATTTTT
701	TTTAAAAAAA TAGGTCACTT TTAAAATTTT TTTCCTAATA CTATAGTGAA
751	TAGAAAAAAG AAAGGAATTA ATTATGAAAA AAGCTTTAAA AGTATTTTCG
801	ATTTATTCAT TAGACCCAGT TGA

The dotted underlined sequence is the Tn4001 sequence. The solid underlined sequence indicates the resumption of the *ctpA* coding sequence at nt1400 *ctp.* The normal text indicates the start of sequences downstream from *ctpA*.

Table 2-3. Sequence for *M. mycoides capri ctpA::tetM* mutant created by homologous recombination using disruption plasmid II (*pExp1-ctpA::tetM-recAec)*.

Nucleotide	Sequence
1	ACCGAAAAAA ACCTTTTTTA TACAAATGTT TTTTTCTTTG GATTTTTTTA
51	TTTAAAAAAC GTCTAAAAAA CCAGTATAGA TCTTGTATAA GCAAAAAAAT
101	GGACTAATAT TTAAATAAAG CAAAAAGGAG AAAAAGATTA TGAAACTAGT
151	TAAAAAAATA GGCTTTTTAA GCTTAAGTGC AATTAGCATA TTAGGACCAC
201	TAGCTATGAT TAACAATCTA ACTACTGATA ATAATCTTTT AATAACCAAA
251	AGGTTTTTAA GTAGTTCAAA CTCTAATGTT GGGTTAAAAT CATATGATTA
301	TATAAACTTA ATTAACAACA AATATATACC TACAAAAATT AATTTACACG
351	ATCATAATGG GATTGCTTAT ATTGGAGTTA AAGAATTCTT AAAGTCTCTA
401	GACGGACTTA TTAGTTTTTC TAAAATAAAA GTAAGACCAT ATCAAAACGC
451	AAATTTTTAT AAAGAAAAAG AAATTAGTTA TAATTACAAA AACAATAAAG
501	TTGTTTTAAA CTCAATTAGT AAATATTCAA ATAATAATAA AACTACTAAC
551	TATCAACTAG AAATTGATAG TAAAAATAAA ACTATTACAG TATCTGATAA
601	TGACTTTTTT ACAGATATTT TCACTTTTTA TAGACGTGGT GAAGAAGATT
651	TAAATATTGA CTTTTTAAAT ACTGAAATTG TAAATAAAAA TAAACATATA
701	GTATTTGATT TAAACAAGTA TGGAATCGAA ATTTTAAATG ATCAAAATGA
751	CTTGTATTTA CCATTAGTAC TAATTAATCA ACTATTTTTA AATCAATCAA
801	ATGTGCAATT GTATTTTAAT GGACAATCTG TTAATTTATT TGCATACAGT
851	AAAACACTTG GAAAAGTTGA ATTATTAAAG CAATTAAAAC ACTCATATTT
901	AATAATCAGA ATCATATACC AGCGGTTTAA AAGATTTTCA ATATAAATAT
951	TTAGGATTTT ATT**TGATCC**C CGACCTCCAA CAAACCGCCA TTTGGAAAGT
1001	AATATACAAT ATTTTAAACA GCGTAAATAG CAACTACCAT TATACGGTTT
1051	TTTTAATTGG CGTTTAGTAA G

The sequence shows the insertion site of *tetM* into the coding sequence. The solid underlining indicates the upstream sequences of *ctpA* that are not found in the disruption construct. The nnormal text indicates sequences of *ctpA* that are present in the disruption construct. The double underlining indicates *tetM* sequences present in the disruption construct. The bold, underlined sequence indicates the insertion site for *tetM*.

Table 2-4. Internal sequence for *M. mycoides capri ctpA::tetM* mutant created by homologous recombination using disruption plasmid II (*pExp1-ctpA::tetM-recAec)*.

Nucleotide	Sequence
1	AATATTCCCG AGGATGCATA ACATCTTCCG CAGTACCGCC CGATTCCCCT
51	GTATAATCGC AAGAAGTATG TTGGGACTTT TACACAATTA TACGGACTTT
101	ATCCTTTCTG ATGTATTAGA AGTAACAGGT ATTGTTTGGA ATGCAACTTT
151	AACTTTTGTG GCTGTTATTC TTATTTCATT AATATTAGAT GAAATTGGTT
201	TTTTTGAATG GTCTGCGATA CATATGGTCA AGGCTTCAAA CGGTAATGGC
251	TTAAAAATGT TTGTTTTTAT TATGTTACTT GGGGCAATTG TAGCAGCATT
301	TTTCGCAAAT GATGGTGCAG CTTTAATCTT AACGCCTATT GTATTAGCGA
351	TGGTAAGGAA TCTAGGATTT AATCAAAAAG TGATTTTCCC CTTTATTATT
401	GCCAGTGGTT TTATTGCTGA TACTACATCA CTTCCCTTAA TTGTAAGTAA
451	CTTAGTTAAT ATCGTTTCTG CAGGTGGAGC TTCAGCGATT AGTCAAGCAA
501	TTCTACCAAC TGGAGATATT ATTCAATTAA GTAGTAATAA TGTTTTAACT
551	AATAAATCTC ATCAAAGTTT AGAATATGGT GTTAATCCAG ATATTACACT
601	TGGATTTGAT CCATTCAAAC AAACTGAAAA ATTCTTTGAT TCAGCTTATA
651	TTCAACAAGC TATTAATAAA GATACAAATA CATTAAATTC AATTCCAGCT
701	ACTCATTCTA GTGTTGTTGA ACCAAATTAT GTACATAAAC TTGTAGAACA
751	ACCTCAACCA CTACAATTAA GTAGAAAAAC TGATGAAACT GAAATAAAAA
801	ATCTTAATAA TTTATTTTCT AGTATAAAAG AAACCGAAAG AAAAGATGCA
851	TATTTTGTAC TTGGAGCACT TGGTGTTGTT ATTAGTTTAG CAATCTCATT
901	TGTATTATTA AAAAAAATAT TAAAATACAA AAATTAACCT TAGTGAAGAA
951	ATACTAAGTT TTATTTTTTT TAAAAATAGG TCACTTTAA

The sequence shows the end of the *tetM* gene and the resumption of the *ctpA* coding sequence. The double underlining indicates tetM sequences. The single underlined CTG CAG denotes the end of the *tetM* sequence and the resumption of the downstream *ctpA* coding sequence.

Table 2-5. Sequence of tetM probe.

CATATGGATTTCTTAGCAGAAGTATATCGTTCAT
TATCAGTTTTAGATGGGGCAATTCTACTGATTTCTGCAAAAGATGGCGTA
CAAGCACAAACTCGTATATTATTTCATGCACTTAGGAAAATGGGGATTCC
CACAATCTTTTTTATCAATAAGATTGACCAAAATGGAATTGATTTATCAA
CGGTTTATCAGGATATTAAAGAGAAACTTTCTGCCGAAATTGTAATCAAA
CAGAAGGTAGAACTGTATCCTAATGTGTGTGTGACGAACTTTACCGAATC
TGAACAATGGGATACGGTAATAGAGGGAAACGATGACCTTTTAGAGAAAT
ATATGTCCGGTAAATCATTAGAAGCATTGGAACTCGAACAAGAGGAAAGC
ATAAGATTTCAGAATTGTTCTCTGTTCCCTCTTTATCATGGAAGTGCAAA
AAGTAATATAGGGATTGATAACCTTATAGAAGTTATTACTAATAAATTTT
ATTCATCAACACATCGAGGTCCGTCTGAACTTTGCGGAAATGTTTTCAAA
ATTGAATATACAAAAAAAAGACAACGTCTTGCATATATACGCCTTTATAG
TGGAGTACTACATTTACGAGATTCGGTTAGAGTATCAGAAAAGGAAAAAA
TAAAAGTTACAGAAATGTATACTTCAATAAATGGTGAATTATGTAAGATT
GATAGAGCTTATTCTGGAGAAATTGTTATTTTGCAAAATGAGTTTTTGAA
GTTAAATAGTGTTCTTGGAGATACAAAACTATTGCCACAGAGAAAAAAGA
TTGAAAATCCGCACCCTCTACTACAAACAACTGTTGAACCGAGTAAACCT
GAACAGAGAGAAATGTTGCTTGATGCCCTTTTGGAAATCTCAGATAGTGA
TCCGCTTCTACGATATTACGTGGATTCTACGACACATGAAATTATACTTT
CTTTCTTAGGGAAAGTACAAATGGAAGTGATTAGTGCACTGTTGCAAGAA
AAGTATCATGTGGAGATAGAAATAACAGAGCCTACAGTCATTTATATGGA
GAGACCGTTAAAAAATGCAGAATATACCATTCACATCGAAGTGCCGCCAA
ATCCTTTCTGGGCTTCCATTGGTCTATCTGTATCACCGCTTCCGTTGGGA
AGTGGAATGCAGTATGAGAGCTC

The NdeI (CATATG) and SacI (GAGCTC) restriction sites are underlined.

Table 2-6. Sequence of the plasmid backbone probe.

TCGGAAAAAG
AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTG
CAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTAC
GGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATC
AAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAG
TATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTC
AGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTAC
GATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTC
ACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGG
TCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAG
AAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTG
AGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGC
GCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACT
CTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCG

Table 2-7. Sequence of the *ctpA* probe used for the Northern blot.

ACTAGTTAAAAAAAT
AGGCTTTTTAAGCTTAAGTGCAATTAGCATATTAGGACCACTAGCTATGA
TTAACAATCTAACTACTGATAATAATCTTTTAATAACCAAAAGGTTTTTA
AGTAGTTCAAACTCTAATGTTGGGTTAAAATCATATGATTATATAAACTT
AATTAACAACAAATATATACCTACAAAAATTAATTTACACGATCATAATG
GGATTGCTTATATTGGAGTTAAAGAATTCTTAAAGTCTCTAGACGGACTT
ATTAGTTTTTCTAAAATAAAAGTAAGACCATATCAAAACGCAAATTTTTA
TAAAGAAAAAGAAATTAGTTATAATTACAAAAACAATAAAGTTGTTTTAA
ACTCAATTAGTAAATATTCAAATAATAATAAAACTACTAACTATCAACTA
GAAATTGATAGTAAAAATAAAACTATTACAGTATCTGATAATGACTTTTT
TACAGATATTTTCACTTTTTATAGACGTGGTGAAGAAGATTTAAATATTG
ACTTTTTAAATACTGAAATTGTAAATAAAAATAAACATATAGTATTTGAT
TTAAACAAGTATGGAATCGAAATTTTAAATGATCAAAATGACTTGTATTT
ACCATTAGTACTAATTAATCAACTATTTTTAAATCAATCAAATGTGCAAT
TGTATTTTAATGGACAATCTGTTAATTTATTTGCATACAGTAAAACACTT
GGAAAAGTTGAATTATTAAAGCAATTAAAACACTCATATTTAAATAATCA
GAATCATATACCAGCAGGTTTAAAAGATTTTCAATATAAATATTTAGGAT
TTTTATTTGATCATTTTTATGGTATTAAATTAGATAAAAATGCTTCATAT
AAAGATTTATTTAAAAAATATGAAAAATACATTAAAGCAGATAATACTAC
TCACTACTTAACAAGTAGATATTTAATTGAACAATTAGATGATTTACATT
CATCATATTTATTAACAGGATATTATAATAAAGATTTAGAAACAATTAAT
AAAGCTGTTTTAAAAACAACAACACCTAGATCTGATAGATTTAAAGATAT
TGCAAGAAGATTAAGCGCATATTATGATAAAGAGTTAAACTATAAAAATG
TTTATACTCCAGATAGAAAAACAAGTGTTATTTCATTTAAAAACTTTGAA
GCTAATTCAGCTTTTAAAATCGAAGAAAGCTTAAAACAAGCTCAAAGAGA
TGGTATTAAAAATATTGTTTTAGATGTAAGCTTTAATAGAGGTGGTTATT
TAGGAACTGCTTTTGAAATCATGGGATTTTTAACAGATAAACCATTTAAA
TCTTATTCATATAATCCTTTAACAAAAGAACAACAAGTTGAAACTATTAA
ATCAAGATTTAAAAAATATGATTTTAACTATTATGTTTTAACTTCACCAT
TTTCATTCTCAGCTGGAAACATTTTCCCTCAACTAGT

The Spe I sites are underlined.

Table 2-8. Generation of *ctpA* mutants by random insertion mutagenesis and homologous recombination.

Species[a]	Method[b]	Number *ctpA* mutants/ Number transformants (%)
Mmc	RIM with Tn916	2/1776 (0.11)
Mmc	RIM with Tn4001T	1/674 (0.15)
Mcap	RIM with Tn4001T	1/384 (0.26)
Mmc	Experiment 1: HR with Pexp1-*ctpA*::*tetM*	1/59 (1.69)
Mmc	Experiment 2: HR with Pexp1-*ctpA*::*tetM*	0/42 (0)
Mmc	Experiment 3: HR with Pexp1-*ctpA*::*tetM*	0/51 (0)
Mmc	Experiment 1: HR with Pexp1-*ctpA*::*tetM-recAec*	3/14 (21.4)
Mmc	Experiment 2: HR with Pexp1-*ctpA*::*tetM-recAec*	4/15 (26.7)
Mmc	Experiment 3: HR with Pexp1-*ctpA*::*tetM-recAec*	4/17 (23.5)

[a] *Mmc* indicates *Mycoplasma mycoides* subsp. *capri* GM12. *Mcap* indicates *Mycoplasma capricolum.* [b] RIM indicates random insertion mutagenesis. HR indicates homologous recombination.

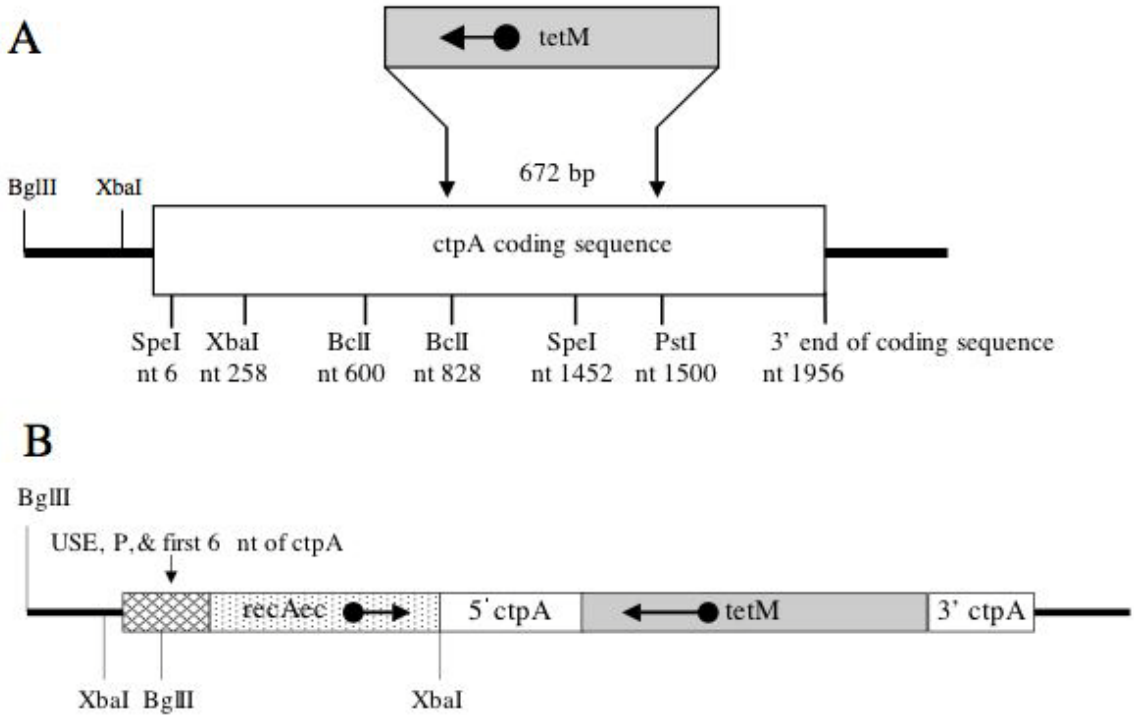

Figure 2-1. Constructs used for targeted mutagenesis of *Mycoplasma mycoides* subsp. *capri*. A) Diagrammatic information on preparation of pExp1-*ctpA::tetM*. The 672 kb removed from *ctpA* and subsequent insertion of *tetM* is shown between BclI site at nt 828 and PstI site at nt 1500 of the coding sequence. B) Diagrammatic information on preparation of pExp1-*ctpA::tetM-recAec*. *RecAec* under the direction of the upstream elements and promoter region of *ctpA* was cloned following nt12 of the *ctpA* coding sequence and ligated into the XbaI site of *pExp1-ctpA::tetM*.

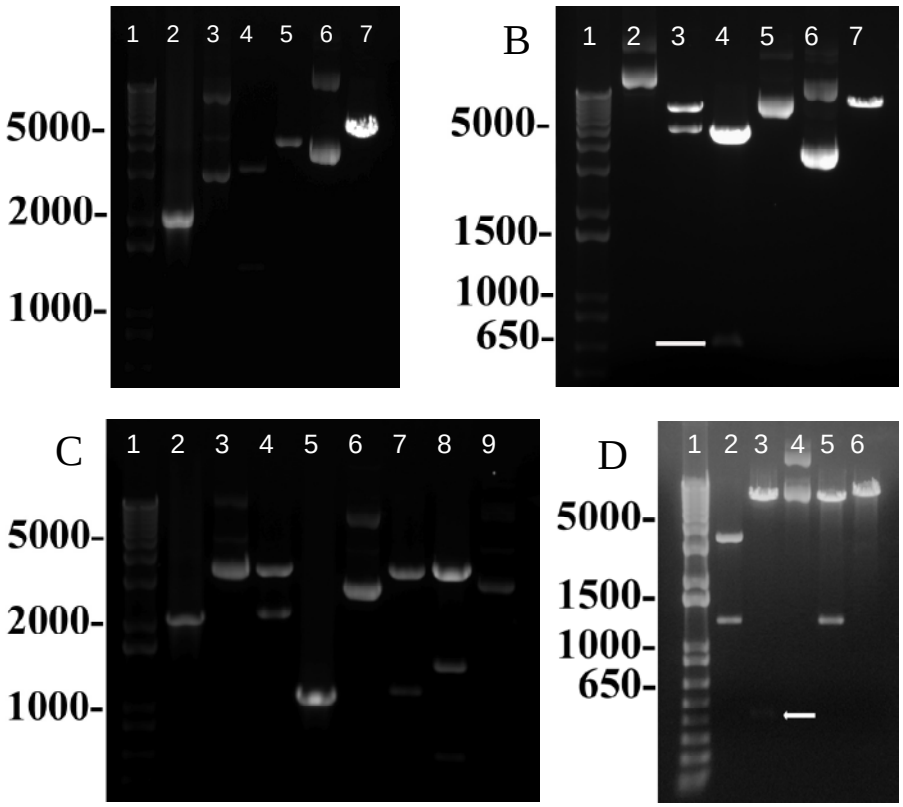

Figure 2-2. Construction of *pExp1-ctpA::tetM* and pExp1-*ctpA::tetM-recAec.* A) Gel demonstrating the key features of preparation of pExp1-*ctpA*. Lane1: 1KbPlus DNA Marker (Invitrogen); Lane 2: PCR of *ctpA* coding sequence; Lane 3: Undigested pENTER*ctpA*; Lane 4: pENTER*ctpA*+SpeI; Lane 5: pENTER*ctpA*+NotI and XbaI; Lane 6: Undigested pExp1-*ctpA*; Lane 7: pExp1-*ctpA*+NdeI. B) Gel demonstrating the key features of preparation of pExp1-*ctpA::tetM*. Lane 1: 1 Kb Plus DNA Marker (Invitrogen); Lane 2: Undigested pIVT-1; Lane 3: pIVT-1+BamHI and PstI; Lane 4: pExp1-*ctpA*+BclI and PstI, showing 672 bp removed from *ctpA* gene (arrow); Lane 5: Undigested pExp1-*ctpA::tetM*; Lane 6: pExp1-*ctpA*+KpnI; Lane 7: pExp1-*ctpA::tetM*+KpnI, showing linearization of pExp1-*ctpA::tetM* due to introduction of KpnI site found in *tetM*. C) Gel demonstrating the key features of preparation of construction of pXL-USE*ErecAec*. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: PCR of USE*ctpA* from *M. mycoides capri* wild type; Lane 3: Undigested pXL-URE*ctpA*; Lane 4: pXL-URE*ctpA*+BamHI; Lane 5: PCR of *recA* gene from *E. coli*; Lane 6: Undigested pXL-*recAec*; Lane 7: pXL-*recAec*+SpeI; Lane 8: pXL-USE*ctpA*+SpeI; Lane 9: Undigested pXL-USE*recAec*. D) Gel demonstrating the key features of preparation of pExp1-*ctpA::tetM-recAec*. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: pXL-USE*recAec*+XbaI; Lane 3: pExp1-*ctpA::tetM*+XbaI; Lane 4: Undigested pExp1-*ctpA::tetM-recAec*; Lane 5: pExp1-*ctpA::tetM-recAec*+XbaI; Lane 6: pExp1-*ctpA::tetM-recAec*+BglII. The arrow denotes the 342 bp fragment that was deleted from pExp1-*ctpA::tetM*.

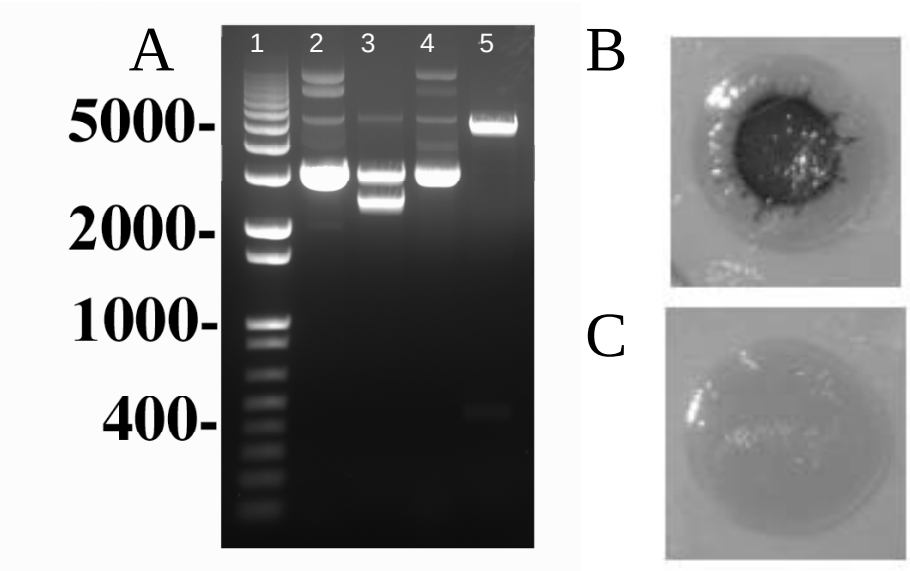

Figure 2-3. Construction of pXL-USEcptA$_p$lacZ. A) Gel demonstrating the key features in the preparation of *pXL-USEcptA$_p$lacZ*. Lane 1: 1 kbPlus DNA marker (Invitrogen); Lane 2: pXL-(SpeI)*lacZ*; Lane 3: pXL-(SpeI)*lacZ* + SpeI; Lane 4: pXL-USE*cptA$_p$lacZ*; Lane 5 : pXL-USE*cptA$_p$lacZ* + NdeI and XbaI. B) *E. coli* transformed with pXL-USE*cptA$_p$lacZ* showing capability of *M. mycoides capri ctpA* USE and promoter to drive expression of *LacZ*. C) *E. coli* transformed with *pXL-lacZ*, showing no *LacZ* expression in the absence of the *M. mycoides capri ctpA* USE and promoter.

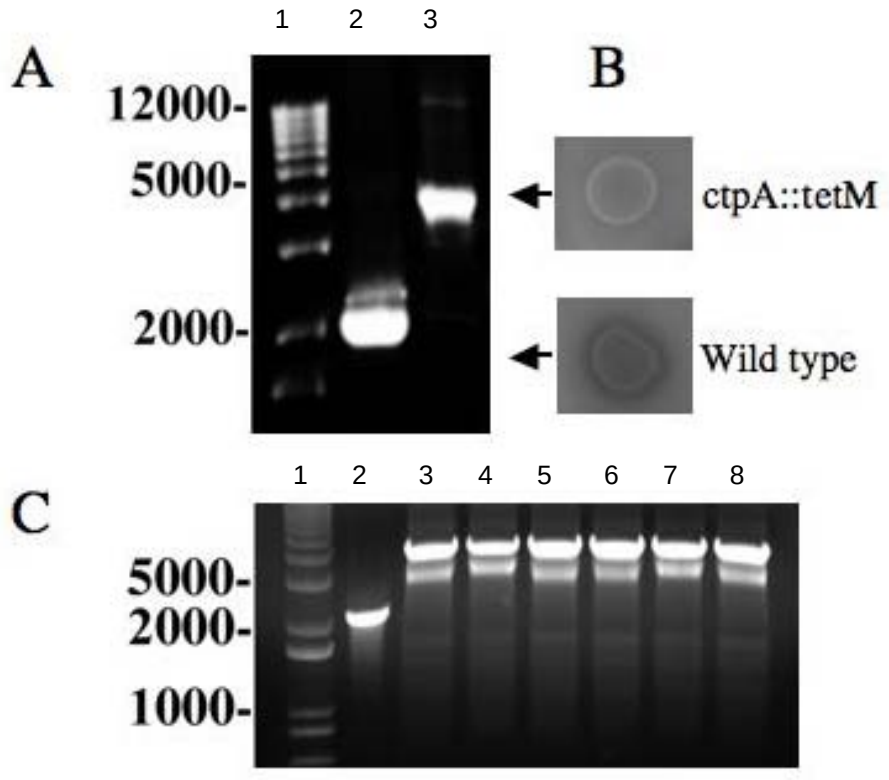

Figure 2-4. Disruption of the *ctpA* gene through homologous recombination results in loss of proteolytic phenotype. A) PCR amplification of genomic DNA from *M. mycoides capri* wild type and *M. mycoides capri ctpA::tetM* mutant obtained by homologous recombination using pExp1-*ctpA::tetM-recAec*. The PCR primers were specific for the genes flanking the *ctpA* gene. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: *M. mycoides capri* wild type; Lane 3: *M. mycoides capri ctpA::tetM* showing the 4 kb insertion in the *ctpA* gene. B) Growth of *M. mycoides capri ctpA::tetM* (top) and *M. mycoides capri* wild type (bottom) on casein agar. Note absence of proteolytic zone around *M. mycoides capri ctpA::tetM*. C) PCR amplification of genomic DNA from *M. mycoides* capri wild type and six *ctpA::tetM* mutants obtained through homologous recombination using either pExp1-*ctpA::tetM* or pExp1-*ctpA::tetM-recAec*. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: *M. mycoides capri* wild type; Lane 3: *M. mycoides* capri *ctpA::tetM* mutant obtained with pExp1-*ctpA::tetM*; Lanes 4-8: Representative *M. mycoides capri ctpA::tetM* mutants obtained with pExp1-*ctpA::tetM-recAec*. All inserts were of the expected size and occurred at the same location ± 1 nt in the ctpA coding sequence.

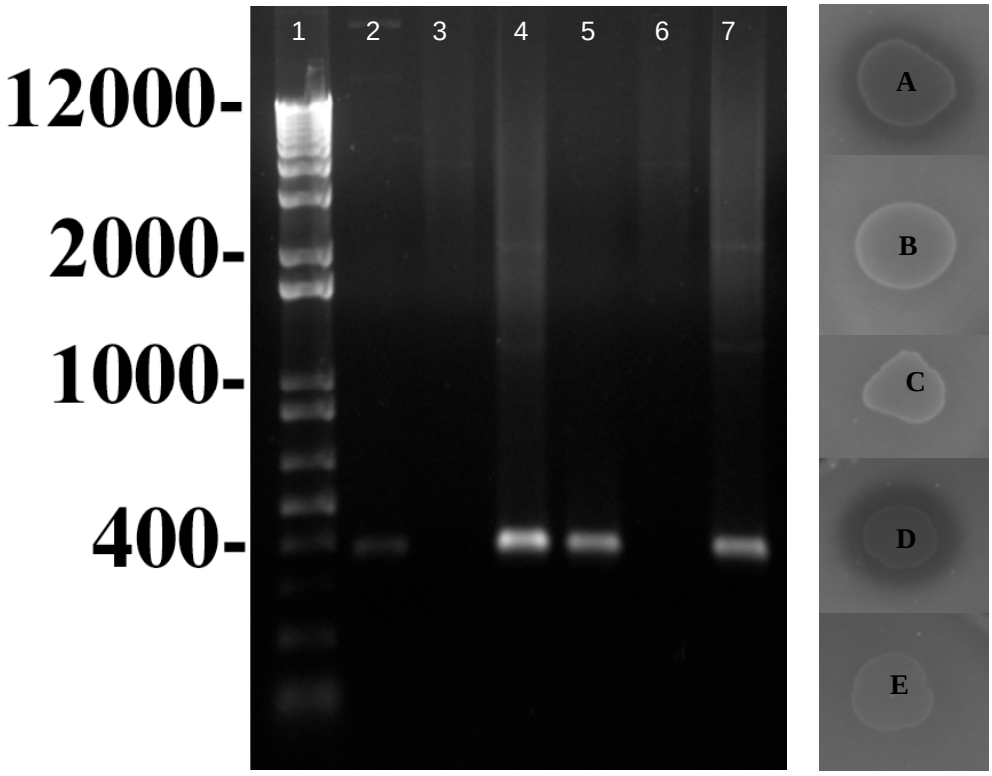

Figure 2-5. Amplification of *tetM* gene and corresponding phenotype. The *tetM* gene was amplified from genomic DNA of wild type and mutant strains to confirm insertion (gel on left): Lane1: 1KbPlus Marker (Invitrogen); Lane 2: pIVT-1; Lane 3: *M. mycoides capri* wild type genomic DNA; Lane 4: *M. mycoides capri ctpA::tetM* mutant generated by homologous recombination; Lane 5: *M. mycoides capri* :: Tn4001t(*ctpA*) generated by random insertional mutagenesis; Lane 6: *Mycoplasma capricolum* wild type genomic DNA; Lane 7: *M.capricolum::*Tn4001t*(ctpA)* generated by random insertional mutagenesis. The phenotype was confirmed by growth on casein agar (right). Shown are *M. mycoides capri* wild type (Lane 3, Panel A), *M. mycoides capri ctpA::tetM* mutant (Lane 4, Panel B), *M. mycoides capri ctpA::*Tn4001t (Lane 5 and Panel C), *M. capricolum* wild type (Lane 6, Panel D), and *M. capricolum ctpA::*Tn4001t (Lane 7, Panel E).

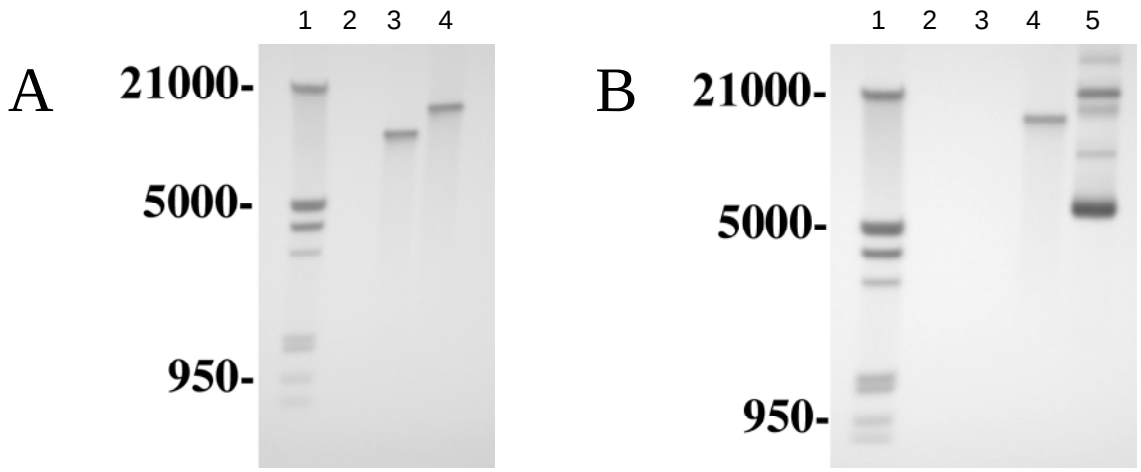

Figure 2.6. Presence of *tetM* and plasmid vector backbone in *M. mycoides capri* GM 12 wild type and *M. mycoides capri ctpA::tetM* mutant. SpeI was used for restriction digest because Tn4001T lacks this site. A) Southern blot was developed using a probe specific for *tetM*. Both the *M. mycoides capri ctpA::tetM* mutants obtained by HR and RIM had a single copy of *tetM*. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: SpeI restriction digest of genomic DNA from *M. mycoides capri* wild type; Lane 3: SpeI restriction digest of genomic DNA from *M. mycoides capri ctpA::tetM* mutant obtained by a double cross-over homologous recombination (HR) event using *pExp1-ctpA::tetM-recAec*; Lane 4: SpeI restriction digest of genomic DNA from *M. mycoides capri ctpA::tetM* mutant obtained by random insertional mutagenesis (RIM) using Tn4001T. B) Southern blot was developed using a probe specific for sequences present in the plasmid backbones of both *pExp1-ctpA::tetM-recAec* and Tn4001T. Lane 1: 1KbPlus Marker (Invitrogen); Lane 2: SpeI restriction digest of genomic DNA from *M. mycoides capri* wild type; Lane 3: SpeI restriction digest of genomic DNA from *M. mycoides capri ctpA::tetM* mutant obtained by a double cross-over HR event using *pExp1-ctpA::tetM-recAec*; Lane 4: SpeI restriction digest of genomic DNA from *M. mycoides capri ctpA::tetM* mutant obtained by RIM using pIVT-1 Tn4001T; Lane 5: Tn4001T transposon, which lacks a SpeI site, as a positive control. Note that the plasmid backbone was not retained in *M. mycoides capri ctpA::tetM* obtained by HR (Lane 3) but was present in the *M. mycoides capri ctpA::tetM* mutant obtained by RIM (Lane 4).

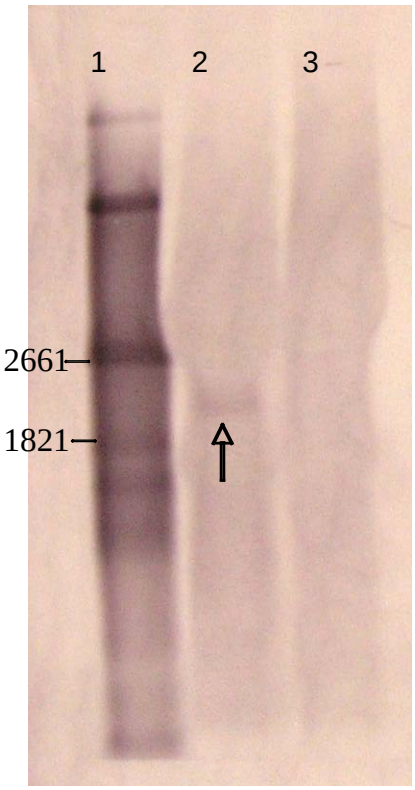

Figure 2-7. Northern blot confirming the absence of a *ctpA* transcript in *M. mycoides capri ctpA::tetM* mutant. Lane 1: DIG–labeled RNA MW marker I (Roche); Lane 2: Total RNA extracted from *M. mycoides capri* wild type; Lane 3: Total RNA extracted from *M. mycoides capri ctpA::tetM* mutant. Probe: Spe I gel purified fragment of *ctpA* gene.

CHAPTER 3
TARGETED DELETION OF S41 PEPTIDASE IN MYCOPLASMA MYCOIDES SUBSP. CAPRI RESULTS IN STRESS RESPONSE, AND PERTURBED GLYCOLYSIS/GLUCONEOGESIS

Introduction

Carboxyl-terminal proteases belong to a group of proteases that were initially identified by genetic complementation analysis of specific photosynthetic mutants of the cyanobacterium Synechocystis sp. strain PCC 6803. CtpA protein was shown to be responsible for cleavage of the D1 precursor (pD1) polypeptide of photosystem II (PSII) (Nixon et al. 1992). This C-terminal processing is essential for the subsequent water oxidation and the generation of oxygen molecules in oxygenic photosynthetic organisms hence it is important for correct functioning of the PSII complex (Diner et al. 1988). Subsequently they have been identified in the chloroplasts of algae and higher plants.

TSP proteins have also been identified in many bacterial pathogens including *Borrelia, Chlamydia, Shigella, Vibrio,* and *Yersinia*. Although their overall physiological functions in bacteria are not well understood; they have been reported to play important roles in the modification of some bacterial proteins with consequences to their virulence or their physiology (Baumler et al 1994). For example the inactivation of TSP of *Borrelia burgdorferi*, the etiologic agent of Lyme borreliosis, affected the synthesis and the processing of the outer membrane protein P13 and the inactivation of CtpA protein had a pleiotropic effect on the protein expression profile in *Borrelia* (Ostberg et al. 2004). in Chlamydiae, two TSP proteins CT441 and CT858 have been found to target and interfere with host proteins that are involved in the host immune response against microbial infection. CT858 degrades the regulatory factor X5(RFX5) and Upstream Stimulation Factor 1(USF-1), these transcription factors are required for the expression

of the major histocompatibilty complex (Zhong et al. 2001). CT441 has been reported to cleave p65 protein which is an important regulator for the NF-κB pathway of the inflammatory response (Lad et al. 2007).

They are considered to be unique among serine-type proteases due to their resistance to the conventional protease inhibitors and their characteristic catalytic center (Ekici et al. 2008, Liao et al. 2000, Paetzel et al. 1997). They selectively target and cleave the non-polar C-termini of many precursor proteins (Silber et al. 1992) either during the course of their maturation (Gollin et al. 1992, Hatchikian et al. 1999. Islam et al. 1993, Menon et al. 1993, Rossmann et al. 1994) or during their transportation into other organelles or export to the periplasm (Diner et al. 1988, Nagasawa et al. 1989). They may also target and degrade damaged or aberrant proteins (Keiler et al 1996). In Gram positive bacteria, C-terminal processing is a required step in the transpeptidation mechanism to anchor some surface proteins to the cell envelope (Mazmanian et al. 2001).

The gene ctpA (ZP_02512724) has been identified through random transposon mutagenesis and then specifically disrupted through targeted mutagenesis as explained in chapter 2. *In silico* analysis of the DNA sequence and the predicted protein product indicated that it belongs to Carboxyl-terminal processing proteases with the presence of a tail specific protease domain (TSPc) located between amino acid residues 340 and 544 (Figure 3-1). This domain has been shown to recognize and cleave specific hydrophobic residues of the substrate (Beebe et al. 2000). It also contains a signal peptide (aa 1-24) that overlaps with the transmembrane domainTMD (aa 7-29) (Figure 3-1 and 3-2); these domains indicate that the protein is targeted and anchored to the

cell membrane. The protein with its domain composition resembles another *E. coli* periplasmic TSP-containing protease known as Prc (processing involving C-terminal cleavage). Prc is a 76-kDa periplasmic protein involved in processing of penicillin binding protein 3 (PBP3) an integral protein in the inner membrane; Inactivation of *Prc* in *E. coli* results in a mutant that exhibits sensitivity to high temperature and leakage of periplasmic proteins under osmotic stress (Hara et al. 1991). Other characterized CtpA proteases also contain N-terminal signal sequences to transport it across the thylakoid membrane in chloroplast (Karnauchov et al. 1997, Mitchell et al. 1997).

Proteomic Studies in Mycoplasmas

Mycoplasmas are very attractive and suitable microorganisms for the proteomic studies and this fact was recognized very early. Two of the smallest mycoplasmas *M. genitalium* and *M. pneumoniae* were among the first organisms for which proteomic profiles were determined (Jaffe et al 2004, Wasinger et al 2000, Ueberle et al 2002). In addition, there are advantages of performing proteomic analysis in mycoplasmas. Those include the availability of several fully and partially sequenced mycoplasma genomes; this knowledge of the genomic sequences and potential protein coding sequences is almost an absolute prerequisite for proteomic technologies.

The small size of the mycoplasma genomes and their relatively few proteins they encode (less than 1000) should make a high experimental coverage possible, while the paucity of their transcriptional factors and regulators (Razin et al 1998) should permit the detection of nearly all of the gene products made by the organism in a relatively small number of environmental stimuli or culture conditions.

As part of this study, differential proteome profiling was carried out followed by pattern discovery analysis; the objective of these proteomic experiments and analyses

was to gain a perspective as to what the role of *ctpA* may be and the effect of its disruption on the overall protein expression profile of *Mycoplasma mycoides* subsp. *capri*. Determining some of the proteins impacted by the *ctpA* disruption would shed light on the function of this gene and the type of pathways and reactions wherein it may participate or perform.

Materials and Methods

Mycoplasma Strains and Their Cultivation

Mycoplasma mycoides subsp. *capri* GM12 type ATCC 35297 (*M. mycoides capri*) (Manso-Silván et al. 2009) and *Mycoplasma mycoides* subsp. *capri* GM12 *ctpA::tetM* were used for this study. *M. mycoides* subsp. *capri* GM12 *ctpA::tetM* was generated by double cross-over homologous recombination as described in Chapter 2.

For all experiments both mycoplasmas strains were cultivated in parallel at 37 °C in the same batch of SP4 medium. The SP4 medium that was used to culture *M. mycoides ctpA::tetM* was supplemented with 10 µg/ml of tetracycline. Microbial growth was monitored by optical density readings that were performed at 640 nm and all cultures were harvested at late log phase (OD_{640} = 0.08). All microbial cultures contained 10^{12} CFU per ml of media at the time of harvest, which was confirmed by culture. Cultivation of mycoplasmas was performed 3 times and each culture was considered a biological replicate.

Preparation of Protein Extracts for Proteomics

Bacterial suspensions were divided into 2 aliquots so that one aliquot was used for 2 dimensional electrophoresis and the second was used for amine specific peptide based labeling (iTRAQ™) followed by tandem mass spectrometry. Bacterial suspensions were pelleted by centrifugation at 8,000 x g, for 30 minutes at 4^0 C, and pellets were washed with wash buffer solution (Calbiochem ProteoExtract® kit, San Diego, CA).

Protein from pellets that were to be analyzed by 2 - dimensional electrophoresis were extracted with Trizol (Invitrogen Corp., Carlsbad, CA) according to the manufacturer's protocol. Pelleted protein extracts were allowed to air dry and were stored at -20ºC. Protein from samples that were to be analyzed by iTRAQ™ analysis were extracted with ProteoExtract® Complete Mammalian Proteome Extraction Kit (Calbiochem, San Diego, CA).

Proteome Profiling by 2-Dimensional Differential Gel Electrophoresis (2D-DIGE)

Protein pellets were dissolved in buffer containing 8 M urea, 2 M thiourea, 4% CHAPS, and 10 mM Tris, pH 8.5. Total protein from each biological replicate was first determined by Lowry assay. The protein concentration of each sample was adjusted so that both protein samples, *M. mycoides ctpA::tetM* and wild type *M. mycoides* contained the same amount of protein before labeling with CyDye™ Fluor minimal dyes (GE Healthcare, Piscataway, NJ). *M. mycoides ctpA::tetM* was labeled with Cy5 and wild type *M. mycoides* (control) was labeled with Cy3 fluorescent dyes. The Cy2 labeled mixture that was composed of an equal amount of *M. mycoides ctpA::tetM* and wild type *M. mycoides* was used as an internal standard. Fifty micrograms of each sample was mixed and loaded onto a 24 cm, pH 3 -11 IPG strip (GE Healthcare, Piscataway, NJ).

Electrofocusing of proteins was performed overnight at 1000 volts with a temperature controlled IPGphor Isoelectrofocusing unit (GE Healthcare, Piscataway, NJ). The separated IPG strip was then loaded into a pre-cast 8 to 16% SDS-polyacrylamide gel (Jule, Inc. Milford. CT). Three 2- dimensional electrophoresis experiments were performed, and each experiment contained a different biological replicate.

Differential Image Analysis of Protein Gels

Gel images were obtained with Typhoon 9600 Variable Mode Imager (GE Healthcare) which is designed to optimally detect each protein bound to different CyDye™ Fluor. Cy3 was detected with a 532 nm excitation laser and a 580 BP 30 emission filter. Cy5 was detected with a 633 nm excitation laser and a 670 BP 30 emission filter. Cy2 was detected with a 488 nm excitation laser and a 520 BP 40 emission filter. The digital image information from each gel was acquired and analyzed with DeCyder 2D version 7.0 software (GE Healthcare, Piscataway, NJ). Specifically, the Cy2 internal standard was used to co-detect, match and normalize protein spots in all 3 gels. Protein ratios for each gel spot were generated by dividing the total area of *M. mycoides ctpA::tetM* spot by the total area of the corresponding wild type *M. mycoides* spot. For statistical analysis, protein ratios from all 3 biological replicates were analyzed by Student's t test and 1 way ANOVA. A composite image of all 3 biological replicates was generated with the Biological Variation Analysis module from DeCyder 2D version 7.0 software (GE Healthcare, Piscataway, NJ).

Protein Spot Excision

Only protein ratios that were 2 fold or greater in difference were considered for further analysis. In order to visually detect protein spots of interest, all 3 gels were counter stained with DeepPurple (GE Healthcare) fluorescent stain. Protein spot coordinates that were obtained from DeCyder 2D version 7.0 software (GE Healthcare, Piscataway, NJ) were used to locate the specific protein spot of interest in each gel. An automated spot picker, ProPic Workstation (Digilab Genomic Solutions Inc., Ann Arbor, MI) selected protein targets. The same protein spot from each biological replicate was pooled for processing and identification by tandem mass spectrometry.

Protein destaining and enzymatic digestion was performed as previously described (Stone et al. 1997). Briefly, gel spots were destained and washed in 50% acetonitrile in 25 mM ammonium bicarbonate buffer and dehydrated in a speedvac centrifuge for 15 minutes. Protein was reduced and alkylated by incubating with 45 mM dithiothreitol and incubation for 30 minutes at room temperature followed by incubation with 100 mM iodoacetamide for 30 min in darkness. Gel pieces were then washed in 50% acetonitrile in 50 mM ammonium bicarbonate and dehydrated in the speedvac for 15 minutes prior to digestion with trypsin enzyme cocktail (12.5 ng/µl of trypsin in 50 mM ammonium bicarbonate pH 8.4, 5 mM CaCl2). Enzyme cocktail was added to the sample at a ratio of 1:20, enzyme to protein. Digested protein samples were separated and identified using LC-MS/MS as described for the iTRAQ™ protein fractions.

Quantitative Proteomics Using Peptide Labeling and 2D-LC-MS/MS

The total protein concentration of all protein samples was determined with the Non-Interfering Protein Assay™ Kit (Calbiochem, San Diego, CA). In order to minimize variability, the protein extracts from all 3 biological replicates of wild type *M. mycoides*

subsp. *capri* (control) were combined and the total protein concentration of the pooled sample was adjusted to match the total protein concentration of each *M. mycoides ctpA::tetM* biological replicate.

Bacterial protein extracts were processed and labeled with an amine specific peptide-based labeling system, iTRAQ™, according to the manufacturer's instructions (Applied Biosystems, Foster City, CA). Briefly, a 60 µg aliquot of each sample was dissolved in 20 µL of dissolution buffer (0.5M triethylammonium bicarbonate) and reduced with reducing agent (50 mM tris-2-carboxyethyl phosphine) at 60°C for 1 hour. After reduction, cysteines were blocked with 200 mM methyl methanethiosulfonate for 10 minutes at room temperature. Ten microliters of a trypsin solution (Promega Corporation, Madison, WI) was added to each sample and incubated overnight at 37°C. After digestion, the pooled wild type control sample was labeled with 114 reagent, and biological replicates 1, 2, and 3 of *M. mycoides ctpA::tetM* was labeled with 115, 116 or 117 reagent respectively. All labeled samples were combined, and desalted by using a macrospin column Vyadac Silica C18 (The Nest Group Inc, Southboro, MA) prior to strong cation exchange (SCX) procedure.

SCX fractionation of desalted iTRAQ™ labeled peptides was performed with a polysulfoethyl A column (100 x 2.1 mm, 5 µm, 300 Å). Peptides resuspended in buffer A (75% 0.01 M ammonium formate, 25% ACN) were eluted during a linear gradient of 0-20% buffer B (75% 0.5 M ammonium formate, 25% ACN) and detected at an absorbance of 280 nm. Eluted fractions were further separated by capillary reverse phase HPLC using an LC Packing C18 Pep Map column (DIONEX, Sunnyvale, CA). Mass spectrometric analysis of column elute was performed inline with a hybrid

quadrupole-TOF mass spectrometer QSTAR (Applied Biosystems Inc). The focusing potential and ion spray voltage was set to 275 V and 2600 V, respectively. The information-dependent acquisition mode of operation was employed in which a survey scan from *m/z* 400–1200 was acquired followed by collision induced dissociation of the three most intense ions.

Tandem mass spectra were extracted by Analyst (v 1.1. ; Applied Biosystem Inc). The National Center for Biotechnology Information bacterial protein database (concatenation of the forward and random sequences) was used for protein identification. Searches were performed using MS/MS data interpretation algorithms from Protein Pilot™ (Paragon™ algorithm, v 3.0, Applied Biosystem Inc) and Mascot (v 2.2, Matrix Science, London, UK). The Paragon algorithm from Protein Pilot™ was set up to search iTRAQ™ 4-plex samples as variable modifications with methyl methanethiosulfonate as a fixed modification. The Protein Pilot™ algorithm was selected to search automatically for biological modifications such as homocysteines. Additional information on this algorithm can be found in (Shilov et al. 2007). The confidence level for protein identification was set up to 1.3 (95%) which is the default for the detected protein threshold in a Paragon™ method. The differential expression ratios for protein quantitation were obtained from Protein Pilot™ which calculates protein ratios using only ratios from the spectra that are distinct to each protein, excluding the shared peptides of protein isoforms. Peptides with low spectral counts were also excluded from the calculation of averages by setting the intensity threshold for the sum of the signal-to-noise ratio for all the peak pairs at greater than 9. All the quantitative ratios were then corrected for bias automatically by Protein Pilot™ when

processing the data to create the Pro Group™ algorithm results. The bias factor calculated for the iTRAQ™ ratios were 0.3650 for 115/114; 0.4357 for 116/114; and 0.4286 for 117/114. Each protein that was quantified was identified by a minimum of three spectra that had an error factor (EF) less than 2. The EF is a measure of the variation between the different iTRAQ™ ratios (the greater the variation, the greater the uncertainty) and represents the 95% uncertainty range for a reported ratio. The P-value is calculated based on the 95% confidence interval. Only protein ratios with a P value ≤ 0.01 were considered significant.

Proteins were grouped according to global biologic functions as assigned in the Molligen 2.0 database (Barré et al. 2004). Enrichment analysis using Fisher exact test with false discovery rate correction for multiple comparisons was used to identify any biological function categories that were significantly over or underrepresented in the *M. mycoides ctpA::tetM* peptidase mutant.

Growth Curves for *M. mycoides* subsp. *capri* Wild Type and *ctpA* Mutant at Different Temperatures

Both *M. mycoides* subsp. *capri* wild type and *ctpA* mutant were grown to log phase (OD640=0.06) in SP4 medium, these cultures were then used to inoculate fresh media for constructing growth curves. For each strain, 1 ml of the inoculum culture was added to 1 liter of fresh SP4 medium. This 1L culture medium was then divided into two 500 ml cultures; one of them was incubated at 37^{o}C and the other at 42^{o}C. The cultures were monitored for the bacterial growth by measuring OD640 and CFU count.

Results

The 2 - dimensional gel images are presented in Figure 3-3. Only 10 proteins were identified with 99% or greater accuracy. They are listed in Table 3-1. Most proteins

were detected were found to be significantly decreased in *M. mycoides ctpA::tetM*. These were: preprotein translocase SecA, adenylsuccinate succinate synthase, phosphoglycerate kinase, and hypothetical proteins MSC 0133 and 0539. The only proteins found to be significantly increased in the mutant were the transcription anti-termination protein NusG and seryl-tRNA synthetase.

Two hundred twenty-one proteins were identified with the iTRAQ™ system. Of these, 61 proteins exhibited a significant change as a result of the S41 gene deletion (see Table 3-2). Proteins that are involved in several biological pathways were affected. However, the only biological function category to exhibit a significant perturbation as a result of the S41 gene deletion was the glycolysis pathway ($P \leq 0.04$). See Figure 3-4. These proteins included the glucose-specific IIA component of the phosphotransferase system, the glycolysis core module, and pyruvate metabolism. Although enrichment analysis did not recognize a significant perturbation in genetic information process, it was interesting to note that at least 18 of the 41 ribosomal genes that were detected were significantly increased in the mutant.

At 37^{o}C the growth rates of the wild type and ctpA were comparable; however at higher temperature 42^{o}C the ctpA mutant had significantly slower growth than the wild type and showed significant loss viability much sooner than the wild type (Figure 3-5).

Discussion

The disruption of *ctpA* gene has a pleiotropic effect on the expression profile of the mutant in comparison to the wild type strain. A similar effect has been reported for disruption of the *ctpA* gene in *Borrelia burgdorferi* (Ostberg et al 2004). Proteins involved in the gylcolysis/gluconeogenesis pathway have been impacted the most. Their expression was downregulated in the mutant in comparison to those of the wild type. Because of the interconnections between the glycolysis pathway and other central metabolic pathways (Deutscher et al. 2002); proteins and enzymes participating in nucleotides, amino acids, and lipids metabolism have also been affected by the disruption of *ctpA* gene.

Dihydroxyacetone kinase showed a two-fold increase in its expression in the mutant. This enzyme catalyses the phosphorylation of dihydroxyacetone (Dha) into the glycolytic intermediate dihydroxyacetone phosphate (DhaP) (Garcia-Alles 2004). In bacteria Dha is produced via oxidation of glycerol (Forage et al. 1982) and possibly by aldol cleavage of fructose 6-phosphate (Schurmann et al. 2001). Therefore, disruption of the *ctpA* gene in *Mycoplasma mycoides* subsp. *capri* might signify a shift in carbon source and energy metabolism toward glycerol. This shift toward glycerol would have implications for the cell not only in terms of energy and growth but also in coping with the byproducts and consequences of its metabolic activity, most prominently the increase in the level of H_2O_2 and oxidative stress.

In addition to the ubiquitous and constitutively expressed glycerol facilitator, GlpF, (Vilei et al. 2000) *Mycoplasma mycoides* subsp. *capri*, like the more virulent *Mycoplasma mycoides* subsp. *mycoides* Small Colony, has the genes for a glycerol import protein, an ABC transporter (*gtsABC*) and the gene for l-α-glycerophosphate

oxidase (*glpO*), which catalyzes the oxidation of glycerol-3-phosphate (G3P) to yield dihydroxyacetone phosphate (DhaP) with the concomitant release of H_2O_2. GlpO was identified as the membrane protein that plays a central role in cytotoxicity of *M. mycoides* subsp. *mycoides* SC strains towards embryonic calf nasal epithelial (ECaNEp) cells (Pilo et al. 2005). Glycerol metabolism and the subsequent release of H_2O_2 were important in the cytotoxicity of *M. pneumoniae* (Hames et al 2009).

The production of H_2O_2, a reactive oxygen species (ROS), has been proposed as the causative agent for the host cellular and tissue damage, (Tryon et al. 1992), and an inducer of the inflammatory process (Bischof et al. 2008). ROS causes cell damage through its reactions with major components of the cell including lipids, proteins, and nucleic acids. ROS causes peroxidation of unsaturated fatty acids which in turn leads to the release of more radicals and fatty aldehydes opening the way for more modifications of cellular structures. ROS also causes chemical modification of protein backbones, chain fragmentation and the oxidation of amino acid side chains. ROS are potentially mutagenic because of their reactions with DNA at purine bases that lead to strand breaking.

Lactate dehydrogenase was downregulated in the mutant while the level of acetate kinase expression did not change in the mutant versus the wild type. This imbalance between the two reactions in the mutant may have repercussions on the redox potential and consequently the level of oxidative stress within the cell.

Mycoplasmas can utilize glucose, fructose and glycerol as a carbon and energy sources. They are catabolized to pyruvate via glycolysis. Pyruvate then can be converted to acetyl-CoA by the pyruvate dehydrogenase complex and finally to acetate

by the enzymes phosphotransacetylase and acetate kinase. Alternatively, pyruvate can be reduced to lactate by lactate dehydrogenase. Both reactions have their specific advantage and downside. While the conversion of pyruvate to acetate yields two additional molecules of ATP and two more molecules NADH per molecule glucose, the regeneration of NAD^+ from NADH is not possible.

The recycling of NADH is important for balancing the redox potential and is expected to be a crucial point for the adjustment of Mycoplasmas metabolism, since they do not possess an electron respiratory chain that can be used for this purpose (Pollack et al. 1981). Re-oxidation of NADH into NAD^+ can only occur either in the course of the reduction of pyruvate to lactate by lactate dehydrogenase; this reaction does not generate H_2O_2 (Halbedel et al 2007b) or through the oxidation of NADH by the enzyme NADH oxidase which in this process converts molecular oxygen to hydrogen peroxide. Because of the decrease in the lactate dehydrogenase expression the mutant is expected to either suffer the imbalance in the redox potential or to rely more on the NADH oxidase activity for the regeneration of NAD^+.

Glycerol metabolism offers another source for generating hydrogen peroxide; it does so via the oxidation of glycerol-3-phosphate by the enzyme glycerol-3-phosphate oxidase which also uses molecular oxygen as the electron acceptor (Pilo et al. 2005). Due to the increase in NADH oxidase – mediated regeneration of NAD^+ and the oxidation of glycerol, the mutant would generate more H_2O_2 hence increase the level of oxidative stress compared to the wild type. Therefore the presence of an intact ctpA gene in the wild type might help strike the balance between the production of ATP

(conversion of pyruvate to acetate) and the regeneration of NAD^+ without generating ROS (conversion of pyruvate to lactate) or disruption of the Redox potential.

An alternative interpretation of the proteomic data could be predicated on the premise that glycerol synthesis rather than its catabolism occurs in the mutant as a result of downregulation of the glycolysis pathway and disruption of the redox potential. The rational for this view is that, first, SP4 medium is supplemented with glucose not glycerol, and the proteomic data showed that glycerol kinase expression did not increase in the mutant in comparison to the wild type. This is consistent with the finding that glycerol kinase activity could be modulated to prevent excessive uptake of glycerol in response to the presence of another more preferred carbon source such as glucose (Holms 1996). The glycerol kinase level should have been higher in the mutant as a result of an increase in the glycerol metabolism. Secondly, in *Mycoplasma pneumoniae*, proteomic studies showed that the presence of glycerol in the medium has a repressive effect on the expression of acetate kinase, while glucose increased its expression. This is consistent with the fact that in the presence of glucose, pyruvate conversion into acetate produces an additional ATP molecule (Miles 1992). The effect of carbon source on lactate dehydrogenase was quite the opposite, with glucose having the repressive effect while glycerol increased its expression (Halbedel et al. 2007b). Since the proteomic data shows a decrease in the lactate dehydrogenase expression but no decrease in the acetate kinase level, it is unlikely that glycerol uptake and catabolism is increased in the mutant.

Considering the fact that dihydroxyacetone (Dha) can be produced from aldol cleavage of fructose 6-phosphate (Schurmann et al. 2001), an alternative scenario

might be that glycerol can be produced intracellularly, where Dha is converted into DhaP via the action of dihydroxyacetone kinase, and then can be reduced to glycerol-3-phosphate (through action of glycerol -3-phosphate dehydrogenase) (Cocks et al. 1985). Glycerol-3-phosphate can then be used in the synthesis of lipids, lipoproteins, and nucleotides.

One very important and pertinent corollary to the shift toward glycerol synthesis in the mutant is the conversion of excess NADH to yield NAD^+ , an essential step to redress the balance in the redox potential. Unlike glycerol catabolism, in the case of glycerol synthesis the regeneration of NAD^+ does not lead to the release of H_2O_2. In addition, glycerol is involved in the recycling of inorganic phosphate consumed in the glycolysis (Ansell et al. 1997).

This metabolic shunt may also have a compensatory effect to the downregulation of the enzymes participating in nucleotide biosynthesis. Finally, since many of the ribosomal proteins in the mutant were upregulated, and so was the signal recognition particle protein, glycerol biosynthesis may indicate change in the composition or an increase in the biogenesis of the cell membrane of the mutant.

Disruption of the *ctpA* gene resulted in increased susceptibility toward high temperature in the mutant. This is similar to the effect of knocking out Prc, another TSP-containing protease in protein in *E. coli*. Disruption of *prc* resulted in *E. coli* mutants that were sensitive to high temperature and osmotic pressure (Hara et al. 1991). It is not clear if there is a connection between the impact of CtpA on metabolism and its role in heat tolerance.

Disruption of *ctpA* has a pleiotropic impact on the expression profile of *M. mycoides* subsp. *capri*; notably its effect on the glycolysis/gluconeogenesis pathway. The CtpA knock-out mutant has a two fold increase in the expression of dihydroxyacetone kinase, which suggests that CtpA may be involved in maintaining the balance between utilization of glucose and glycerol as a carbon and energy source. This balance is important for managing oxidative and minimizing the production of potentially harmful H_2O_2.

Table 3-1. Proteins that significantly differed ($P < 0.006$) in *M. mycoides* subsp. *capri ctpA::tetM* as determined by 2D-DIGE.

Spot	pI/MW	Ratio	Accession	Function	Protein name	Prot. ID prob
1468	4.64/32252	7.48	gi|42561478	Genetic information process	NusG	99.9%
1471	4.8/32239	12.03	gi|42561478	Genetic information process	NusG	100%
1472	4.98/32222	4.46	gi|42561478	Genetic information process	NusG	100%
206	5.11/105679	-6.26	gi|42560648	Genetic information process	SecA	100%
1844	3.35/18579	-6.32	gi|42560687	Metabolism l	Hypothetical protein MSC_0133	99%
602	5.41/70969	7.74	gi|42560626	Metabolism	Seryl-tRNA synthetase	100%
962	6.54/48816	-2.26	gi|42561370	Metabolism	Adenylosuccinate synthase	100%
1001	5.93/50641	-2.08	gi|42561204	Metabolism	Phosphoglycerate kinase	100%
123	3.9/129473	-8.89	gi|42561072	Unclassified	Hypothetical protein MSC_0539	100%
517	5.76/75928	4.74	gi|108795342	Antibiotic selection gene	TetM	100%

P values were obtained by Students t test (N = 3). Numbers correspond to spot identifications on Figure 3-3. Isoelectric point (pI) and molecular weight (MW) in Daltons are actual experimental values obtained from the 2D gel. Ratios were obtained by dividing the spot area of the *M. mycoides* subspecies *capri* mutant sample (N = 3) by the spot area of wild type *M. mycoides* subspecies *capri.*

Table 3-2. Proteins in Mycoplasma *mycoides* subsp. *capri* that are present in the glycolysis/gluconeogenesis pathway.

EC number	Gene	Product	Expression in *ctpA* mutant	*Mmcap* locus	*Mcap* locus
5.3.1.9	pgi	glucose-6-phosphate isomerase	Decreased	MMCAP1_0445	MCAP_0465
2.7.1.11	pfkA	6-phosphofructokinase	Decreased	MMCAP1_0220	MCAP_0220
4.1.2.13	fba	fructose-1,6-bisphosphate	Decreased	MMCAP1_0131	MCAP_0136
5.3.1.1	tpiA	triose-phosphate isomerase	Decreased	MMCAP1_0727	MCAP_0750
1.2.1.12	gap	glyceraldehyde-3-phosphate dehydrogenase, type I	Decreased	MMCAP1_0607	MCAP_0632
2.7.2.3	pgk	phosphoglycerate kinase	Decreased	MMCAP1_0606	MCAP_0631
5.4.2.1	gpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	Decreased	MMCAP1_0729	MCAP_0752
4.2.1.11	eno	phosphopyruvate hydratase	Decreased	MMCAP1_0213	MCAP_0213
2.7.1.40	pyk	pyruvate kinase	Decreased	MMCAP1_0221	MCAP_0221
1.1.1.27	ldh	L-lactate dehydrogenase (probable)	Decreased	MMCAP1_0475	MCAP_0439
2.3.1.12	pdhC	dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex	Decreased	MMCAP1_0227	MCAP_0227
1.2.4.1	pdhA	pyruvate dehydrogenase (acetyl-transferring) E1 component, alpha subunit	Decreased	MMCAP1_0225	MCAP_0225
1.2.4.1	pdhB	pyruvate dehydrogenase E1 component subunit beta (S complex, 36 kDa subunit	Decreased	MMCAP1_0226	MCAP_0226
1.8.1.4	lpdA	dihydrolipoyl dehydrogenase	Decreased	MMCAP1_0228	MCAP_0228

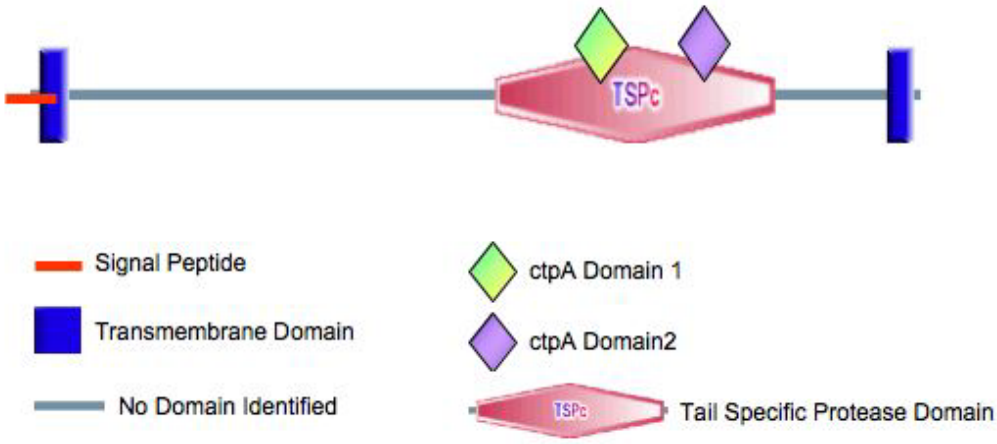

Figure 3-1. Diagrammatic representation of domains identified in the S4 peptidase (CtpA) protein of *M. mycoides* subsp. *capri*. Domains were identified using the SMART algorithm search (Letunic et al. 2009), and the figure is adapted from the SMART output.

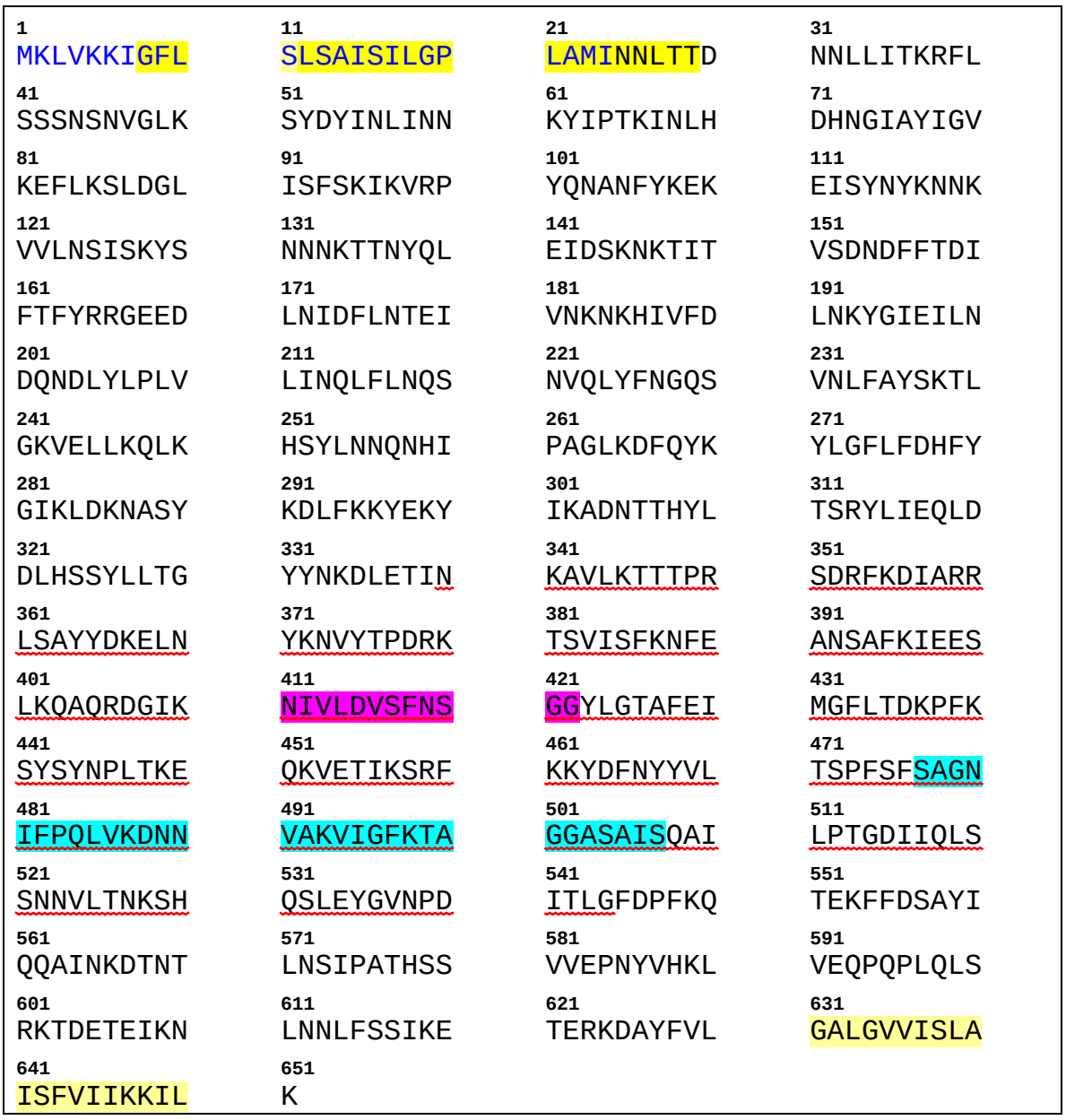

Figure 3-2. Complete amino acid sequence of the S41 peptidase (CtpA) of *M. mycoides* subsp. *capri* MCAP 0241 protein showing identified domains. The protein is 651 amino acids in length, with a predicted molecular weight of 75.09 KD and pI of 9.25. The signal peptide domain (aa 1-24) is shown in blue font. The two transmembrane domains (aa 7-29 and 631-650) are highlighted in yellow. Overlaps between the signal peptide and first transmembrane region are denoted by blue font with yellow highlight. The tail specific protease (TSPc) domain from aa 340 to 544 is denoted by red underline. The two ctpA domains are within the TSPc domain: ctpA domain 1 (pink highlight) extends from aa 411-422; ctpA domain 2 (turquoise highlight) extends from aa 477-507.

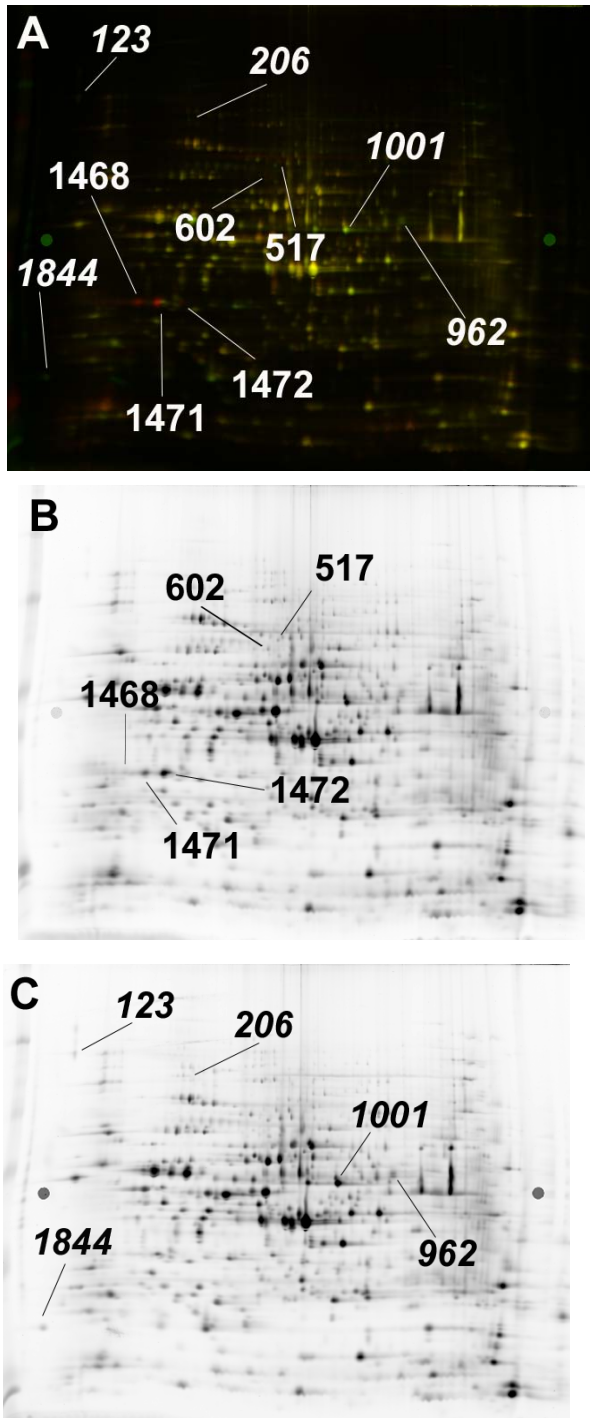

Figure 3-3. Differential 2- dimensional electrophoresis of *M. mycoides* subsp *capri* GM12 and *M. mycoides ctpA::tetM*. Panel A shows the distribution of proteins with Cy3 (red) and Cy5 (green) labeling. Panel B shows the distribution of protein spots that were significantly increased in *M. mycoides ctpA::tetM*. Panel C shows the distribution of protein spots that were significantly increased in *M. mycoides* subsp. *capri*.

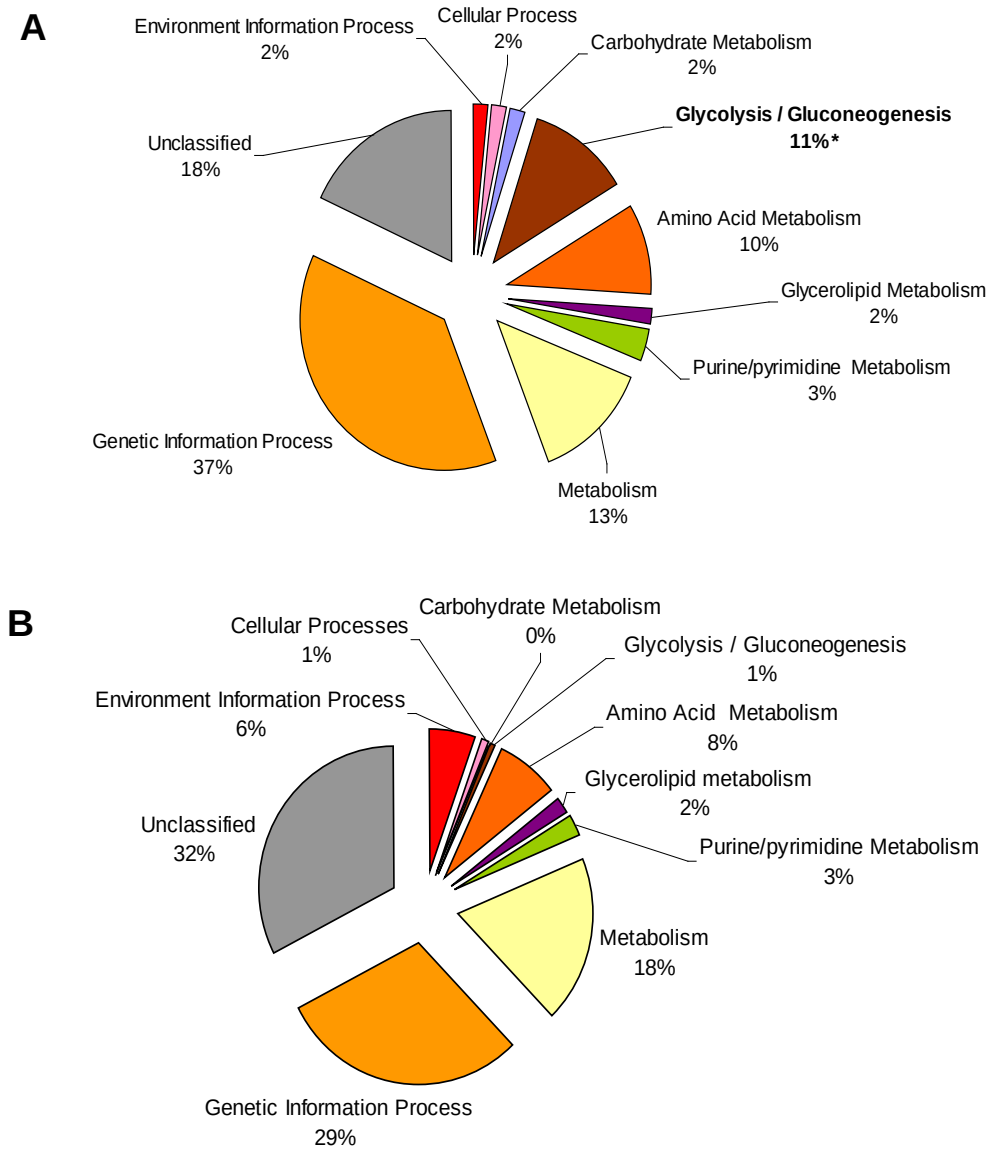

Figure 3-4. Distribution of biological function categories of *Mycoplasma mycoides* subsp. *capri* GM12 proteins that were significantly altered (A) or unchanged (B) by genetic deletion of S41 peptidase gene. Values represent the percent of proteins assigned to each biological function group that were significantly altered by deletion of ctpA gene (A), and the percent of proteins that were not affected by the gene deletion (B). Gene ontology designations were obtained from the NCBI protein database or Molligen 2.0 database. *Biological function categories significantly different ($P \leq 0.04$).

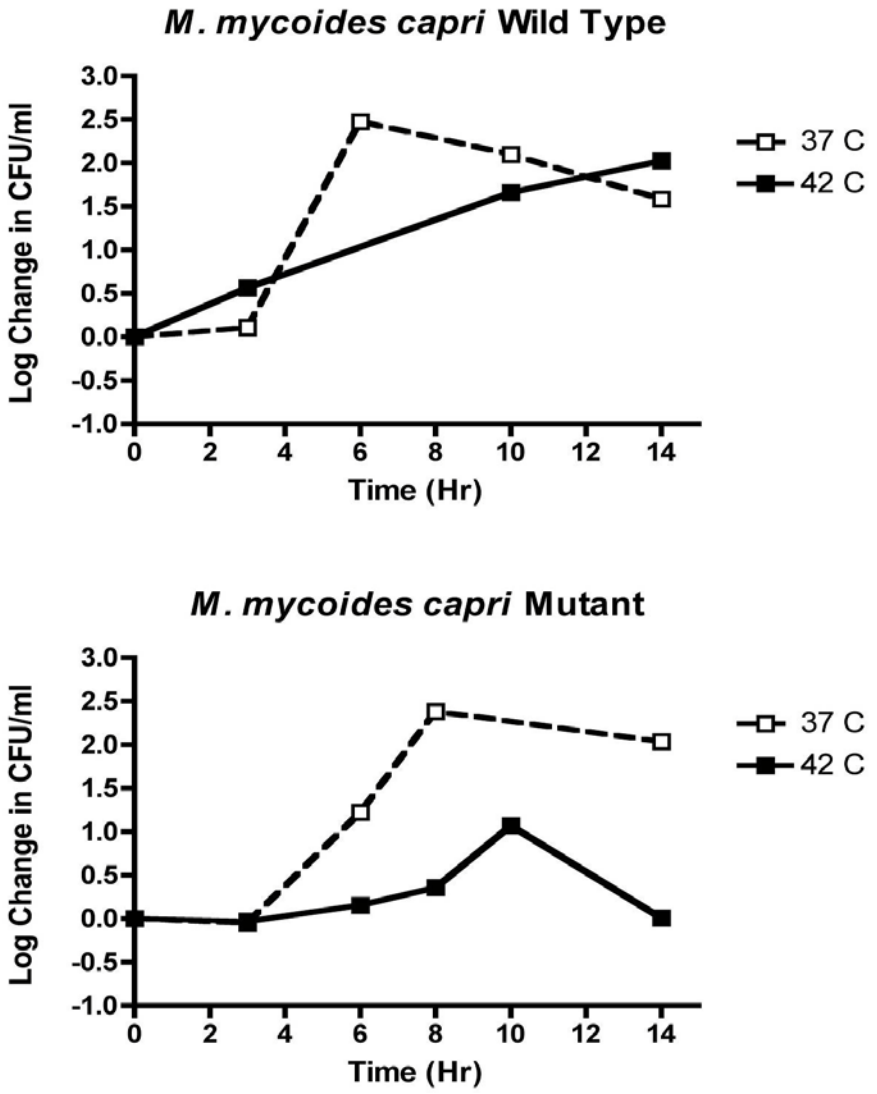

Figure 3-5. Growth curves for *M. mycoides* subsp. *capri* wt and the *ctpA* mutant at 37°C and 42°C. The mutant grew more slowly at 42°C in comparison to the wild type. Also the mutant growth declined rapidly at 42°C.

CHAPTER 4
TRANSCRIPTION OF S41 PEPTIDASE *(CTPA)* GENE

Introduction

The current view of gene regulation in mycoplasmas is that, with few notable exceptions (including variable surface proteins), mycoplasmas do not regulate the expression of their genes (refer to Chapter 1 for an in depth discussion). Most known or common bacterial regulatory mechanisms are absent in mycoplasmas, and therefore these microorganisms are thought to lack refined on-off switching mechanisms and global regulation for transcriptional adaptation to environmental changes. Conceptually, mycoplasmas lack most of the conventional regulatory components and factors that other bacteria have. The absence of the conventional regulatory components and factors makes it difficult to understand how mycoplasmas regulate their gene expression in response to environmental and host signals or to adapt to different habitats and host niches. For instance, unlike other bacteria, mycoplasmas do not possess the putative repressor and activator proteins (Muto et al. 1987, Razin et al. 1998) nor do they have two-component systems (Simmons et al. 2007). Transcription factors and proteins constitute only about 0.5% of the protein coding sequences in mycoplasma genomes (SchmidI et al. 2010), while other bacteria reserve as much as 10% of their coding capacity to transcriptional regulation (Greenberg 2000). Moreover, although mycoplasmas have many genes involved in the stress response, they have only one sigma factor, that makes it unclear how they regulate these genes in response to stress conditions (Dandekar et al. 2000).

Despite the absence of clearly identified global regulators, recent *in vivo* and *in vitro* transcriptome and proteomic studies reported global responses to environmental or

host stimuli (Oneal et al. 2008, Madsen et al. 2008, Schafer et al. 2007, Cecchini et al. 2007, Pinto et al. 2007, Madsen et al. 2006a, Madsen et al. 2006b). Importantly, even with a very limited number of transcriptional factors, genome wide analyses of some of the smallest and simplest mycoplasmas revealed an unexpected level of complexity and versatility in their in metabolic responses to environmental conditions. Their adaptation seems to be similar to that of more complex bacteria, providing hints that other, unknown regulatory mechanisms might exist (Yus et al. 2009). Furthermore, proteome complexity could not be directly inferred from the composition and organization of their minimal genomes or even their extensive genome wide transcriptional analysis (Kuhner et al. 2009). In addition, transcriptome analysis also showed surprisingly unanticipated diversity and heterogeneity in mycoplasma transcription profiles including the presence of many operons, the production of alternative transcripts in response to environmental perturbations, and the high frequency of antisense RNA (Gardner and Minion, 2010, Guell et al. 2009). Therefore, it is becoming increasingly clear that mycoplasmas are able to respond to environmental cues and regulate gene expression, but the underlying mechanisms for individual and global gene regulation are not understood.

In part because of the dearth of classical transcriptional factors, most investigators have concluded that mycoplasmas regulate their genes mainly at the post-translational level (Schmidl et al. 2010) and lack a refined transcriptional regulation; instead mycoplasmas are believed to employ stochastic processes similar to those responsible for generating antigenic variations (Simmons et al. 2007). However, mycoplasmas have recently been shown to respond to different stress factors at the transcriptional as well as post-transcriptional levels (Oneal et al. 2008, Madsen et al.

2008, Schafer et al. 2007, Cecchini et al. 2007, Pinto et al. 2007, Madsen et al. 2006a, Madsen et al. 2006b).

Mycoplasmas also transcriptionally control some of their gene expression according to the type of carbon source available in the medium (Helbedel et al. 2007). Most recently, expression of the *Mycoplasma pneumoniae* community-acquired respiratory distress syndrome (CARDS) toxin was shown to be regulated at the transcriptional level in response to the growth phase or the attachment to the host cell (Kannan et al. 2010). We also noted that the expression of the *ctpA* gene in *M. mycoides* subsp. *capri* appeared to correlate with the growth phases, with the highest expression occurring at the stationary phase when pH of media drops to 5.5. Proteomic studies described in Chapter 3 suggested that CtpA may be a part of the stress response with emphasis on oxidative stress and pH. In addition, CtpA knock-out mutant exhibits sensitivity to heat shock. Since CtpA seems to be involved in stress response to pH and heat shock, we hypothesized that *ctpA* expression would be coordinated with increasing the acidity of the medium and shifting to higher temperature.

The objectives of the current studies were (1) to determine if the *ctpA* gene expression is regulated in response to stress conditions, (2) to determine if *ctpA* gene expression is regulated at the transcriptional level, and (3) to develop an alternative model system using the more efficient and genetically defined *E. coli* to study the regulation of *ctpA* expression.

Materials and Methods

Mycoplasma Strain and Cultivation

Mycoplasma mycoides subsp. capri GM12 type ATCC 35297 (DaMassa et al. 1983), formerly known as *M. mycoides* subsp. *mycoides* Large Colony type (Manso-

Silvan et al. 2009) was grown at 37 °C in SP4 medium (Tully et al. 1979); SP4 casein agar was supplemented with a final concentration of 1% skim milk (DIFCO, Detroit, MI). For growth of mutants, tetracycline (final concentration, 5µg/ml) was added to the media.

PEG8000-Mediated Chemical Transformation.

For random mutagenesis, *M. mycoides capri* was transformed with Tn4001T (plasmid pIVT-1), a gift from K. Dybvig, (Dybvig et al. 2000), by PEG8000 (Sigma-Aldrich)-mediated chemical transformation as previously described (Dybvig et al. 2000, Lartigue et al. 2009). For targeted mutagenesis (additional details are provided in Chapter 2), *M. mycoides capri* GM12 was transformed with pExp1-*ctpA*::*tetM-recA*ec using PEG8000-mediated chemical transformation.

An overnight culture of *M. mycoides capri* was diluted 1/1000 in 40 ml of fresh SP4 broth. The culture was grown until the mid-log phase (~7-8 hrs) and then placed on ice for 2 min. *M. mycoides capri* was pelleted by centrifugation at 12,000g for 30 min at 4 ºC, the supernatant was decanted, and the cells were washed in 10 mM Tris buffer, pH 6.5. Cells were centrifuged at 12,000g for 30 min at 4 ºC, resuspended in 1 ml of 0.1 M $CaCl_2$, and incubated on ice for 1 hr. Yeast t-RNA (10 µg, Sigma-Aldrich Co), plasmid DNA (pIVT, random insertional mutagenesis or pExp1-*ctpA*::*tetM-recA*ec, targeted homologous recombination; 30 µg), and 9 ml of 60% (w/v) PEG8000 in 10 mM Tris pH 6.5 (final concentration, 54% PEG8000) were added, and cells were incubated for 2 min at room temperature. Twenty-five ml of 10 mM Tris buffer, pH 6.5, was added, and the cells were centrifuged at 12,000g for 30 min at 4ºC. The supernatant was decanted, and the cells were suspended in 2 ml of warm (37ºC) SP4 broth supplemented with 2 mM $MgCl_2$ (Lavery et al. 1992, Hoffman et al. 2000, Goryshin et al. 1998, Goryshin et al.

2000) and incubated at 37ºC for 2 hrs. Following the 2 hr recovery time, the cells were plated on SP4 agar plates containing 5 µg/ml tetracycline and 1% casein. Colony growth was observed after 48 hr. All colonies were picked and expanded in SP4 broth.

The clones were then grown to the same early log phase (OD640 = 0.03) and plated on SP4 agar plate supplemented with 1% casein. The diameters of the clear zones around the different clones were measured, and those showing altered or no proteolytic activity were further expanded. Genomic DNA was extracted from the expanded colonies using the DNeasy Blood and Tissue Kit (catalogue # 69504) from Qiagen (Valencia, CA). The point of insertion and the disrupted genes were identified by sequencing from the 3' end of the Tn4001T transposon past the point of insertion using the sequencing primer Tn4001tSeq. DNA sequencing was carried out at TIGR and the University of Florida DNA sequencing Core Facility.

Transcription of *ctpA* in *Mycoplasma mycoides* subsp. *capri* Wild Type and in Mutants with Altered Proteolytic Activity

Mycoplasma mycoides subsp. *capri* wild type and 9 mutants created by random insertional mutagenesis with Tn4001T as well as an earlier mutant created by random insertional mutagenesis with Tn916 (Rosentel, 2003) were grown from overnight cultures to mid-log phase (OD640 = 0.06). The cells were collected by centrifugation as described above and total RNA was extracted from each culture as described below. For each mutant, three replicates were assayed for transcription of ctpA; one replicate was used for the 50S ribosomal protein L13 gene that served as the internal reference gene.

Effects of pH and Heat Shock on Transcritpton of *ctpA.*

Mycoplasma mycoides subsp. *capri* wild type was grown to log phase (OD_{640} = 0.04), and the cells were harvested by centrifugation at 10,000 rpm for 10 minutes at 4^{o}C. To determine the effects of pH on *ctpA* expression, cells were resuspended in fresh SP4 medium that had been adjusted to pH 5.0, 7.0, or 9.0. To assess the effects of heat shock, temperature shock was done at 42 ^{o}C, with 37 ^{o}C as a reference. The cells were subjected to growth under the different pH and temperature points for 1 hour, harvested by centrifugation as described before, and total RNA extracted as described below. For each experiment, five independent replicates were assayed for ctpA transcription; ribosome L013 served as the internal reference gene.

Extraction of RNA and Quantitative Real Time-PCR for *ctpA* Transcription

Total RNA was extracted using the RiboPure-Bacteria kit (catalogue # 1925) from Ambion (Austin, Texas) according to the instruction manual. Primers for the qRT-PCR (Table 4-1) were ordered from Genosys Sigma-Aldrich Co. Quantitative RT-PCR was performed using QuantiTect SYBR Green RT-PCR Kit (catalogue # 204243) from Qiagen (Valencia, CA). In a 50 µl PCR reaction, 500 ng of total RNA and 30 pM of each primer were used. The remaining steps and cycling conditions were performed according to the instruction manual. Briefly, a reverse transcription step was carried out for 20 minutes at 50^{o}C, followed by a PCR initial activation step for 15 minutes at 95^{o}C. All reactions were subjected to 55 cycles of template denaturation at 94°C for 15 sec. Primer annealing occurred for 30 sec. at 50°C. Extension at 72°C occurred for 30 sec. Data collection was performed during the extension step.

PCR cycles were performed in an iCycler iQ Real-Time PCR Detection System (catalogue # 170-8740) from BioRad (Hercules, CA). Three genes (50S ribosomal

protein L13, 50S ribosomal protein L09, and transcriptional elongation factor) were selected as internal control candidates under the experimental conditions. These genes were chosen based on their steady expression in the proteomic data of *M. mycoides capri* wild type and the *ctpA:tetM* mutant (Chapter 3). Ribosome L013 was chosen and served as the internal reference gene for all data normalization.

Construction of Transcriptional *lacZ* Fusions of *ctpA* Gene

The promoterless *lacZ* gene was amplified by PCR using pET200/D/lacZ as a template and cloned in frame behind the upstream nucleotide sequence and promoter region of ctpA and the first three coding sequences of *ctpA* to create the plasmid construct pXL-USE*ctpA-lacZ* (see Chapter 2 for details). pXL-USE$_{ctpA}$-*lacZ*, which contained the full upstream nucleotide sequence and promoter region, was used as a template to create five constructs with deletions in the upstream nucleotides sequence and promoter regions. The constructs were created using PCR and forward primers specifically designed for five different regions of the USE*ctpA* in the pXL- USE$_{ctpA}$-*lacZ* (Figure 4.3). The reverse primer was common and recognized the terminator of *lacZ* (Table 4.1). In order to facilitate the identification of the positive clone as well as to confirm the correct orientation of the cloned fragment, the forward primers contained an engineered BamHI site at the 5′ end of the coding sequence being amplified. In addition, the *lacZ* coding sequence was also amplified and cloned in frame with the native *E. coli lacZ* promoter and RBS, creating a construct wherein the *lacZ* reporter gene was driven by its own promoter P_{lac}, (Table 4.1).

For a 50 µl PCR reaction, the reaction mixture contained 0.5 µl of Taq Polymerase (5U/µl), 1.5 µl of $MgCl_2$ (50 mM), 5.0 µl of 10X buffer, 1.5 µl of dNTPs (40mM), 20 pMol

of forward and reverse primers, and 50 ng of the template plasmid pXL-USEctpA-LacZ. PCR was performed on a BioRad Icycler (catalogue # 170-8720, Hercules, CA). All reactions were held at 95^{o}C for 3 min and subjected to 25 cycles of template denaturation at 95°C for 30 sec. Primer annealing occurred for 40 sec at 50°C. Polymerization at 68°C occurred for 4 min. Polymerization was then followed by a final extension for 7 min at 68°C.

The PCR fragments were cloned by ligation into pXL-PCR-Topo vectors (catalogue # k4750-10, Invitrogen, Carlsbad, CA) according to the instruction manual. The ligation reactions were used to chemically transform *E. coli* Stbl2 strain (catalogue # 10268-019, Invitrogen, Carlsbad, CA, which does not contain *LacZ.* The transformed clones were selected on LB agar plates supplemented with 50 µg/ml kanamycin. Clones were then picked and expanded. Plasmids were purified using a QIAprep Spin Miniprep kit (catalogue # 27104) from Qiagen (Valencia, CA). The resulting plasmids were digested with BamHI and NcoI (New England Biolabs, Ipswich, MA) to confirm the presence of the cloned fragment. If the cloned fragment was present, then plasmids were digested with XbaI and NdeI (New Englands Biolabs, Ipswich, MA) to confirm the right orientation of the cloned fragment. All the cloned fragments contained *lacZ* cloned in frame after the first three amino acids of the coding sequence of *ctpA*, but the fusion constructs differed in the nucleotide sequence upstream of the coding sequence (Figure 4.3)

The six *E. coli* Stbl2 clones that contained the transcriptional fusion of the reporter *lacZ* gene driven by different truncated versions of the nucleotides upstream of the coding sequence of ctpA were screened for LacZ activity using the bromo-chloro-

indolyl-galactopyranoside (X-gal) Blue/White test. Clones were plated on LB agar plate containing 20 mg/ml X-gal and screened for the ability to produce blue-pigmented colonies.

Determination of Activity of LacZ Fusion Constructs

The activity of the LacZ expressed from the fusion constructs were determined using the "High Sensitivity β–Galactosidae Assay Kit™" (catalogue # 200710 Statagene, La Jolla, CA). In all the experiments, *E. coli* Stble2 was the host strain and pXL-PCR Topo was the plasmid vector. Activity was assessed in four different clones: (1) RBS-*lacZ* that served as a negative control; (2) P_{lac}-*lacZ* as an internal control wherein *lacZ* expression is under the control of its own native promoter and therefore is not regulated by either pH or temperature (Wilson, et al. 2007), (3) experimental clone USE-*lacZ* in which LacZ expression is driven by the entire upstream nucleotide sequence plus the promoter of ctpA and finally (4) experimental clone Pro_{ctpA}-*lacZ* wherein the upstream nucleotides sequence were deleted so that *lacZ* is driven only by the full promoter region of *ctpA*.

Expression of *lacZ* Reporter Gene in Response to pH and Heat Shocks

Each *E. coli* clone was grown in 50 ml LB medium (Luria-Bertani medium) at 37^{o}C. When the culture reached log phase (OD600 = 0.5), the cells were equally divided into five tubes (10 /per tube) and pelleted at 4000 rpm for 10 minutes. The supernatants are discarded. To determine pH shock effects, pellets from tubes 1-3 were resuspended in10 ml each of LB medium adjusted to different pH values (5.0, 7.0, and 9.0). Each of these cultures was immediately divided into five 2-ml cultures, resulting in 5 replicas for each pH point. The pH shock was performed at 37 ^{o}C for 10 minutes (Coll et al. 1994). For heat shock studies, cells from the remaining two tubes (4 and 5) were resuspended

in in10 ml each of LB medium, pH 7.0, immediately divided into five 2-ml cultures per treatment. Heat shock (N=5) was carried out at 42^{o}C for 30 minutes; controls were incubated at 37 ^{o}C (N=5).

After the pH and heat shocks were carried out, 500 µl of each culture was taken to measure the total protein concentration. The total protein concentration was measured using the Non-Interfering Protein Assay Kit (catalogue # 488250, Calbiochem, San Diego, CA). An additional 100 µl of each culture was pelleted and resuspended in the β–galactosidase assay lysis buffer. All reagents were provided in the kit, and the β–galactosidae assay was performed according to the instructions provided with the kit.

Results

Identification of Mutants with Altered Proteolytic Phenotypes

A total of 674 mutants were obtained by random insertional mutagenesis. Nine mutants of *M. mycoides* subsp. *capri*::Tn4001T were selected on a SP4-casein agar plate because of their altered proteolytic activities as indicated by the diameter of the clear zones relative to that of the wild type. One additional mutant (J8) was available from a previous study using Tn916 (Rosentel, 2003). All mutants were sequenced. The DNA sequence data were searched using BLAST, and the homologues of the disrupted genes were identified (Table 4.2). The full genome of *M. mycoides capri* GM12 can be accessed via either GenBank or the dedicated Molligen database (http://cbi.labri.fr/outils/molligen/).

Mutant #152 was the only mutant that showed a complete loss of proteolytic activity. Subsequent DNA sequencing and BLAST analysis indicated that this mutant had Tn4001T inserted in the coding sequence of the *ctpA* gene, MMCAP_0241. Mutant *M. mycoides* subsp. *capri*::Tn916 (J8) had Tn916 inserted in the noncoding sequence

upstream of the *ctpA* gene. The Tn916 insertion in mutant J8 was 23 bp upstream of a presumptive - 35 and - 10 transcription promoter sequence (TTGTAT and TAATAT, respectively) and 67 bp upstream of a probable ribosome binding (Shine-Dalgarno) sequence (GGAG). Mutant *M. mycoides* subsp. *capri*::Tn916 (J8) showed significant reduction in proteolytic activity.

The distribution of the disrupted genes was throughout the *M. mycoides* subsp. *capri* genome (Table 4-3). Mutant #10 (MMCAP_0157, *mgtE*) was the only mutant that showed a slightly increased protease phenotype as indicated by a greater zone size than the wild type (Table 4-2). The remaining 7 mutants had a relatively reduced proteolytic activity (Table 4-2).

Three mutants (mutants # 2, 153, and 154) had disrupted genes coding for hypothetical proteins (Table 4-2). The remaining disrupted genes had known functions. Mutant # 7 had a disruption in MMCAP_0038, *ftsH*. FtsH is an ATP- and Zn^{++}-dependent metalloproteinase that belongs to AAA proteases (ATPases associated with a variety of cellular activities) and is also considered a Charonin (proteins with protease-associated chaperone activity) (Van Melderen et al. 1996). Mutant # 9 (MMCAP_0169) had a disrupted gene that codes for *opp*, an ATP-dependent oligopeptide ABC transporter that plays a role in capturing and transporting peptides, providing a source for amino acids for growth (Monnet 2003). Mutant # 38 (MMCAP_0792) had a disruption in *atp*, the gene coding for ATP synthase. *Nox*, the gene coding for NADH oxidase, was disrupted in Mutant # 155 (MMCAP_0223). NADH oxidase is a flavin-containing enzyme that is involved in the transfer of electrons to oxygen as the terminal electron acceptor.

With the exception of mutant J8, all disruptions occurred within the coding sequence of the gene (Table 4-3).

Transcription of *ctpA* in Mutants with Altered Proteolytic Phenotypes.

The level of *ctpA* transcription in the mutants was investigated using Real Time RT-PCR. The gene for 50S ribosomal protein L13 was chosen as an internal control while the level of ctpA transcript in the mutants was normalized to that of the wild type. Despite the differences in the identities and functions of the disrupted genes, all the mutants except the *mgtE,* magnesium transport knock-out mutant exhibited a reduction in transcription of ctpA (Figure 4.1). In the *mgtE* mutant, transcription of *ctpA* was not significantly different from that of the wild type (Figure 4.1).

Effect of pH and Temperature on the Transcription of *ctpA*: Real Time RT-PCR.

The gene for transcription elongation factor was chosen as an internal control in the pH shock experiments while the gene for 50S ribosomal protein L09 was selected as an internal control in heat shock experiments. Quantitative RT-PCR showed an increase in the transcription of *ctpA* gene in response to the shift from neutral pH 7.0 to acidic pH 5.0 (Figure 4.5 A). In contrast, transcription declined in response to alkaline shift (pH 7.0 to pH 9.0). Real Time RT-PCR also showed an increase in the transcription of ctpA in response to heat shock at 42^{o}C (Figure 4.5 B). Thus, it appears likely that transcription of *ctpA* is regulated in response to both pH and temperature.

Expression of Transcriptional Fusion of *lacZ* Reporter Gene in *E. coli*

LacZ expression was observed with all constructs except when the promoter region was deleted (Figure 4.4). The *lacZ* reporter was transcriptionally driven by both complete and different truncations of the upstream nucleotide sequences and the

promoter region. The presence of the RBS alone was not sufficient to drive the transcription of *lacZ* gene.

Effect of pH and Temperature on the Transcription of *lacZ* reporter Gene in *E. coli*

In the presence of complete upstream nucleotides sequence, β-galactosidase activity increased in response to acidic pH 5.0 and decreased with alkaline pH 9.0 (Figure 4.6 Upper panel). In addition, a detectable increase in β-galactosidase activities was measured in response to heat shock to 42^{o}C (Figure 4.6 Lower panel). This pattern was similar to transcription of *ctpA* in response to pH and heat shock in *M. mycoides* subsp. *capri* as indicated from the qRT-PCR data.

Deletion of upstream sequence element did not abrogate the expression of the reporter *lacZ* gene; indicating that the presence of the promoter alone is sufficient for expressing the downstream gene. This coincides with results from the X-gal Blue/White test (Figure 4.4). However, when the upstream nucleotide sequence was removed, *lacZ* expression was constitutive across all the experimental conditions.

Discussion

Mycoplasmas are thought to lack refined on-off switching mechanisms and global regulation for transcriptional adaptation to environmental changes. Their genes are generally thought to be constitutively expressed rather than regulated (Muto and Ushida 2002), albeit coordinated with the growth rate of the cell through stringent control (Cashel et al. 1996, Gourse et al. 1996). Furthermore, most known or common bacterial regulatory mechanisms are absent in mycoplasmas. For example, unlike *Escherichia coli*, *Bacillus subtilis* and many other bacteria, mycoplasmas have only one sigma factor even though they contain many stress response genes (Muto et al. 1987, Razin et al. 1998). In addition, very few repressor-like proteins have been reported in mycoplasma

genomes (Bork et al. 1995, Fraser et al. 1995, Glass et al. 2000, Himmelreich et al. 1996, Glass et al. 2006). Despite the absence of clearly identified global regulators, recent *in vivo* and *in vitro* transcriptome and proteomic studies reported global responses to environmental or host stimuli (Oneal et al. 2008, Madsen et al. 2008, Schafer et al. 2007, Cecchini et al. 2007, Pinto et al. 2007, Madsen et al. 2006a, Madsen et al. 2006b). Results from our study confirm that the transcription of *ctpA* is likely modified in response to pH and temperature, and also is impacted by products of other genes as well.

Mycoplasma mycoides subsp. *capri* mutants (N=10) with altered proteolytic phenotypes were generated by random mutagenesis using either Tn4001T or Tn916. Although only a direct disruption of the *ctpA* gene caused a complete loss of the proteolytic phenotype, other gene disruptions affected the protease phenotype quantitatively as demonstrated by RT-PCR. The disrupted genes were diverse with respect to identity, function, and genomic distribution. Therefore, we expected that the effects on altered proteolytic activity would occur at different levels, i.e. transcription translation, post-translational modification, indirect effects, etc. However, qRT-PCR data showed that, with one exception (*M. mycoides* subsp. *capri mgtE::Tn4001t)*, all mutants had a reduced level of *ctpA* transcription in comparison to the wild type. These results, while somewhat surprising, may underscore the importance of transcription regulation as a key component for controlling *ctpA* expression. A possible explanation for these unexpected results may be that the proteins affecting *ctpA* expression are involved in fundamental processes such as energy metabolism, maintenance, stress response, and homeostasis. As such, their disruption would have a profound impact on

the cell as a whole. The consequence of their disruption may not necessarily be exclusive or specific to the *ctpA* gene, but rather a part of their pleiotropic effect on the expression of several genes.

Three genes encoded conserved hypothetical proteins of unknown function but with homologs in other members of the mycoides cluster. Of the genes with known function, *ftsH* encodes an ATP- and Zn^{++}-dependent metalloproteinase which belongs to the AAA protease family (ATPases associated with a variety of cellular activities) and may also be a putative Charonin, or protein with protease-associated chaperone activity (Van Melderen et al. 1996). The FtsH protein also is involved in the degradation of cytosolic regulatory proteins and proteins with short nonpolar carboxyl termini. Thus, FtsH may be part of a quality control system to avoid potentially harmful accumulation of free subunits of membrane-embedded protein complexes (Schumann 1993).

The *opp* gene (mutant 9) codes for an oligopeptide ABC transporter, an ATP-dependent transporter that plays a role in capturing and transporting peptides as a nutritional source for amino acids (Monnet 2003). It has also been implicated in sensing the environment and signaling processes through transporting dipeptides (Abouhamad et al. 1991). This is particularly intriguing since *ctpA* transcription is quantitatively responsive to at least two environmental signals (pH and temperature). Mutant # 38 has a disrupted *atp* synthase. In mycoplasmas, ATP- synthase acts mainly as an ATPase, hydrolyzing ATP in order to export H^+, thereby generating and maintaining the proton motive potential (Cirillo 1993, Shirvan et al 1993). Finally, the gene product NADH Oxidase (mutant 155) plays an important role in energy metabolism and ATP generation (Pollack et al. 1997). This flavin-containing enzyme is involved in the transfer of

electrons to oxygen as the terminal electron acceptor. Additionally, NADH Oxidase has a particularly important metabolic role in the cytochrome-less mycoplasmas because it regenerates NAD^+ from NADH (Pollack 2002).

The transcription of the *ctpA* gene is modulated not only by cellular metabolism and energy status, but also is responsive to environmental stimuli and stress conditions. Both pH and temperature shifts impacted the transcription of *ctpA*. qRT-PCR showed an increase in *ctpA* transcription in response to acidic pH (5.0) and heat shock at 42^oC; while shifting to alkaline pH (9.0) reduced the transcription level. Mutant #J8 has an insertion of Tn916 that occurred 23 bp upstream of a presumptive - 35 and - 10 transcription promoter sequence (TTGTAT and TAATAT, respectively) and 67 bp upstream of a probable ribosome binding (Shine-Dalgarno) sequence (GGAG). Therefore, a unique feature of this mutant is that it has a disruption in the non-coding upstream nucleotide sequence while the promoter and coding sequence of *ctpA* gene are left intact. Importantly, as shown by qRT-PCR, this disruption adversely affected the transcription of *ctpA* gene. This finding led us to question the possible role that this noncoding nucleotide sequence may play in modulating *ctpA* transcription

To address the hypothesis that the upstream nucleotides sequence and promoter region may be instrumental in regulating transcription of *ctpA*, mutation analysis was performed. Different truncated versions of these regions were created as in frame transcriptional fusions of *lacZ* reporter gene, and expression of *LacZ* in these constructs was used to investigate the consequences of deletions in the upstream region. These were made and cloned in plasmid constructs (pXL-USE_{ctpA} – P_{ctpA} *lacZ*) in *E. coli* clones (see material and methods).

Initially, I introduced these different fusions into *M. mycoides* subsp. *capri*, using Tn4001T carried on plasmid pIVT-1 as a cloning vector. Unfortunately, although the expression of *lacZ* from these constructs in *M. mycoides* subsp. *capri* was visible enough to discern a pattern of expression correlated with the extent of truncations in USE and promoter regions, the reaction was too weak to document using the β-galactosidase assay. Deletion of the entire upstream nucleotide sequence did not affect expression of *lacZ*, but removing the promoter abolished all traces of *lacZ* expression. However, this expression system in *M. mycoides* subsp. *capri* was unsuitable to use for accurate and quantitative analysis. The weakness of *lacZ* expression can be attributed to many factors. The *lacZ* gene is a 3 Kb fragment cloned from *E. coli* with relatively higher GC content than the mycoplasma genomes which may adversely impact the translation efficiency in mycoplasmas. The gene was likely introduced as a single or low copy number in the genome; that in turn may limit the expression of LacZ. It is also possible that either cloned sequence or Tn4001T when used as cloning vector may have instability issues (Gawron-Burke et al. 1984). Transcriptional attenuation by the flanking sequence on the cloned gene of interest may also be a factor (Su et al. 1992). Finally, for convenience, the cloning of *lacZ* took place at the expense of the signal and leader peptide, which prevented the secretion of LacZ into the medium. This, in addition to the low copy number, may have exacerbated the sensitivity issue and made it more difficult to be more perceptually visible and quantifiable. Apart from the weakness of reporter gene expression, another confounding factor is the concomitant random insertion mutagenesis of transposons which would make it difficult to draw meaningful

conclusions about the quantitative response of gene expression to the mutagenesis of USE.

As an alternative to expressing *lacZ* transcriptional fusions in *M. mycoides* subsp. *capri*, the possibility of using *E. coli* as an expression host was considered. The original plasmid constructs (pXL-USE$_{ctpA}$ – P$_{ctpA}$ *lacZ*) were used to transform *E. coli* Stbl2 cells. Because the plasmids are self-replicating, we felt that the issue of genome disruption would be eliminated and the multiple copy number of plasmid should enhance the expression and, in turn, increase the sensitivity of detection. All *E. coli* clones but one formed blue colonies on LB agar + X-gal plates within 24 hours. Unlike the expression of *lacZ* in *M. mycoides* subsp. *capri*, in *E. coli*, *lacZ* expression was strong and prominently visible. The formation of blue colonies indicates that *ctpA* promoter successfully drove the expression of *lacZ* gene in *E. coli*; this is consistent with previous reports regarding the ability of *E. coli* RNA polymerase to recognize mycoplasma promoters (Taschke et al. 1988).

The complete and truncated versions of USE and promoter regions of *ctpA* produced a similar pattern of *lacZ* expression to that was observed in *M. mycoides* subsp. *capri*; that is, successive deletions in the region upstream *lacZ* fusion did not affect its expression until the deletion of the promoter region, which abolished the expression of *lacZ*. This suggests that the promoter region may be the minimum regulatory sequence of *ctpA* required for expression of the downstream gene.

The next step was to determine if the transcription of *lacZ* under the control of *ctpA* USE and promoter in *E. coli* showed a similar response to pH and temperature shifts as did *ctpA* transcription in *M. mycoides* subsp. *capri* . Simultaneously, the link

between the USE and the effect of the pH and temperature on *ctpA* transcription was investigated. The presence of clones of *lacZ* gene under the control of a complete and deleted upstream sequence element (USE) made it possible to characterize the significance of USE in response to shifts in pH and temperature. *E. coli lacZ* gene under the control of its own cognate promoter was used as an internal reference. The rationale for this control is that in *E. coli*, *lacZ* gene driven by its native promoter is regulated by lactose and subjected to catabolite repression by glucose but it is influenced by neither pH nor temperature (Wilson et al. 2007).

Only with the presence of upstream nucleotide sequence (USE) was the expression of *lacZ* affected by the shift in pH and temperature, while the *ctpA* promoter alone (i.e. without upstream nucleotides sequence) elicited a constitutive expression of *lacZ* regardless of change in pH or temperature. Therefore, while the upstream nucleotide sequence is not required for the expression of the downstream gene, it may play a role in its regulation in response to pH and temperature. Interestingly, in *E. coli*, the upstream nucleotide sequence increased *lacZ* expression in response to acidic pH and temperature shock while decreased it at alkaline pH. This pattern of gene expression in *E. coli* coincides with the one observed with *ctpA* transcription in *M. mycoides* subsp. *capri*.

This similarity regarding pH- and temperature-regulated *ctpA* expression in two phylogenetically distant bacteria has many implications. First, it signifies the conservation of the regulatory mechanism. Second, since mycoplasmas have only one sigma factor and no specific activators or repressors, the *ctpA* regulatory region driving *lacZ* expression in *E. coli* is very likely to depend only on the basal transcriptional

machinery and the conventional Sigma factor (σ^{70}). Third, an importantly practical corollary for the conservation of the mechanism is to afford us the study of *ctpA* regulation in *E. coli*. There is a wealth of information about genetics, physiology and metabolism of *E. coli*. In addition, tools, protocols, sensitive assays, and a variety of well defined media have been developed for *E. coli*. These features facilitate investigating gene regulation mechanisms in relatively easy, well controlled experiments and without confounding factors. In contrast, mycoplasmas are fastidious, genetically difficult to manipulate, lack necessary cloning vectors and molecular tools, and their genetics is poorly defined. For these reasons, genetics and gene regulation studies are plagued with many confounding issues and hurdles to overcome. Therefore, unraveling the regulatory mechanism(s) may be more amenable in *E. coli*; however, ultimately it will have to be confirmed or disproved in Mycoplasmas.

Taken together, the data from these experiments indicate that the expression of *ctpA* is not constitutive; rather it is sensitive to the intracellular milieu as well as modified in response to environmental/host stimuli. These experiments have also shown that the expression of *ctpA* has been influenced through modulating the level of mRNA. This suggests that *ctpA* is regulated primarily at the transcription level. Moreover, this *ctpA* regulation is more likely to be quantitative rather than a switch-type mechanism. This coincides with recent findings in *M. capricolum*, *M. genitalium*, *Ureaplasma urealyticum* , and *M. pneumoniae* (Kannan et al. 2010, Muto and Ushida 2002, Glass et al. 2000, Himmelreich et al. 1996, Bork et al. 1995, Fraser et al. 1995) and is consistent with the paucity of repressors and activators in mycoplasma genomes (Bork et al. 1995, Fraser et al. 1995, Glass et al. 2000, Himmelreich et al. 1996, Glass et al. 2006).

In conclusion, the expression of *ctpA* was quantitatively regulated in response to environmental or host stimuli, including pH and temperature. The level of gene regulation is primarily at the transcription step. This does not, however, exclude other levels of regulation, especially posttranslational regulation. The mechanism of *ctpA* transcription is likely to be conserved, dependent on the basal transcriptional machinery, and influenced by the nucleotide sequence upstream of the *ctpA* promoter.

Table 4-1. Primers used for PCR.

Primers	5' →3' sequence
BamHI-USE$_{ctpA}$ For	GGATCCGTTTATAAAAATGGAGAATTTGCACAA
BamHI-Trunc1 For	GGATCC CATTTTTATACAAATGTTTTTTTCTTTGG
BamHI-Trunc2 For	GGATCC GGATTTTTTTATTTAAAAAACGTCTAAA
BamHI-Trunc3 For	GGATCC ATTTAAAAAACGTCTAAAAAACCAGTA
BamHI-Promoter$_{ctpA}$ For	GGATCC CGTCTAAAAAACCAGTATAGATCTTGTATAA
BamHI-RBS$_{ctpA}$ For	GGATCC AAATAAAGCAAAAAGGAGAAAAA
SpeI lacZ For	ACTAGTTATAGATCCCGTCGTTTTACAACGT
lacZ-Rev	TAGTTATTGCTCAGCGGTGG
qRT-PCR_ctpA For	CACTTTTTATAGACGTGGTGAAGAAG
qRT-PCR_ctpA Rev	GTACTAATGGTAAATACAAGTCATTTTG
qRT-PCR_RiboL13 For	GGAGATCACGTTATTATAATTAATGCTG
qRT-PCR_RiboL13 Rev	GCATCTAGTTCTCTTTGAACTTCAAC
qRT-PCR_RiboL09 For	GAAGTTAGTGACGGATATGCAAG
qRT-PCR_RiboL09 Rev	GCTAGTTCTTTTTCTTCTTGATCAG
qRT-PCR_TEF For	AGAAGTTGTTAGACCTAAAGTAATTGAAG
qRT-PCR_TEF Rev	AGCTTGTCTGTTTCTAGCAGCATC
Tn4001tSeq	GTACTCAATGAATTAGGTGGAAGACCGAGG

Table 4-2. Tn916 (J8 only) and Tn4001T random insertional mutants of *M. mycoides capri* with altered proteolytic phenotype.

Gene locus	Mutant #	Zone (mm)	Relative activity	Gene	Protein ID
MMCAP_0241	J8	2.5	↓↓	*ctpA* (upstream)	ACU79406.1; GI:256384837
MMCAP_0343	2	2.0	↓↓	CHP	ACU79457.1; GI:256384888
MMCAP_0038	7	3.0	↓	CHP, presumed *ftsH*	ACU79266.1; GI:256384697
MMCAP_0169	9	3.0	↓	Putative lipoprotein	ACU79312.1; GI:256384743
MMCAP_0157	10	4.5	↑	*mgtE*	ACU79055.1; GI:256384486
MMCAP_0792	38	3.0	↓	*atp*	ACU79340.1; GI:256384771
MMCAP_0241	152	0	None	S41 peptidase *(ctpA)*	ACU79406.1; GI:256384837
MMCAP_0144	153	3.0	↓	CHP	ACU79095.1; GI:256384526
MMCAP_0902	154	3.0	↓	*vlcH* (upstream)	ACU79765.1; GI:256385196
MMCAP_0223	155	2.0	↓↓	*nox*	ACU79270.1; GI:256384701

The reference proteolytic zone size for *M. mycoides capri* GM12 wild type control was 4.0 mm. Tn916 was used to obtain mutant J8 only; all other mutants were obtained using Tn4001T. CHP denotes conserved hypothetical protein. MMCAP_0241 is currently annotated as an S41 peptidase; however this will likely be changed to *ctpA* (J. Glass, pers. communication). CHP denotes conserved hypothetical protein. The full genome of *M. mycoides capri* GM12 is available via GenBank as well as the Molligen database (http://cbi.labri.fr/outils/molligen/).

Table 4-3. Genomic location of genes and site of disruption of the coding sequence by Tn4001T random insertion.

Mutant	Mnemonic	CD nt	Insertion Site nt	Strand	Gene/Product
2	MMCAP2_0343	429776…430471	::428632	+	CHP
7	MMCAP2_0038	53623…54906	::54359	+	CHP
9	MMCAP2_0169	213034…216135	::213252	+	Putative lipoprotein
10	MMCAP2_0157	196489…197892	::197650	-	*mgtE*
38	MMCAP2_0792	931805…933382	::932755	-	*atpA*
152	MMCAP2_0241	302759…304714	::304151	+	*ctpA*
153	MMCAP2_0144	178992…179492	::179286	+	CHP
154	MMCAP2_0902	1077603…1078082	::1077510	+	*vlcH*
155	MMCAP2_0223	282548…283912	::282766	+	*nox*

CD denotes coding sequence; nt denotes nucleotide; CHP denotes conserved hypothetical protein. :: denotes insertion site of Tn4001T. The full genome of *M. mycoides capri* GM12 is available via GenBank as well as the Molligen database (http://cbi.labri.fr/outils/molligen/).

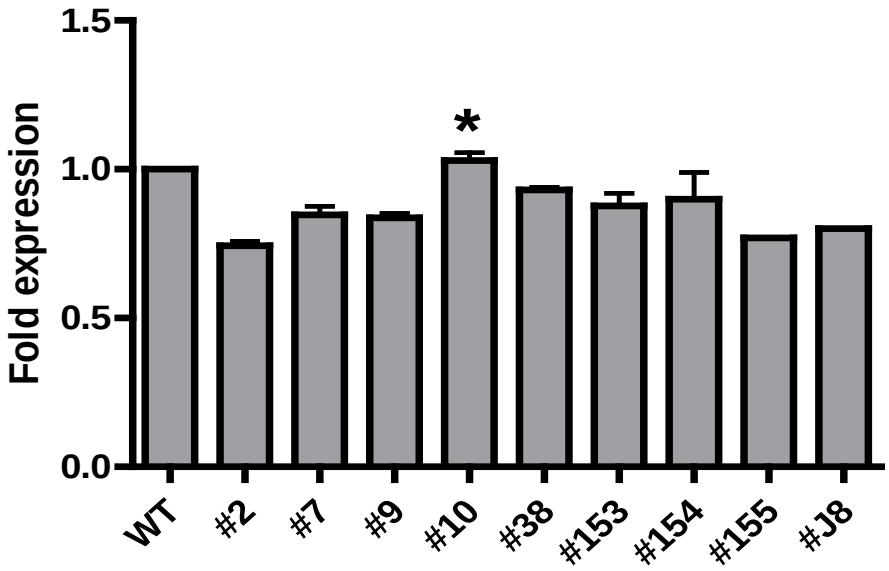

Figure 4-1. Measurement of *ctpA* gene expression in WT *M. mycoides* and Tn4001T, Tn916 generated mutants by real-time RT-PCR. Bacterial cultures were grown to log phase at 37^0 C. For RT-PCR reactions, RNA from Ribosome L013 was used as an internal reference. The threshold cycle number for each experimental sample was converted to fold expression. Fold expressions were normalized to that of WT. Values represent the mean fold change ± SD of 3 biological replicates. ANOVA confirmed that there was statistically significant difference among groups ($P \leq 0.0001$). Fishers PLSD test revealed that all mutants except 10 (*) displayed a significantly different fold change in ctpA expression when compared to wild type *M. mycoides* ($P < 0.02$).

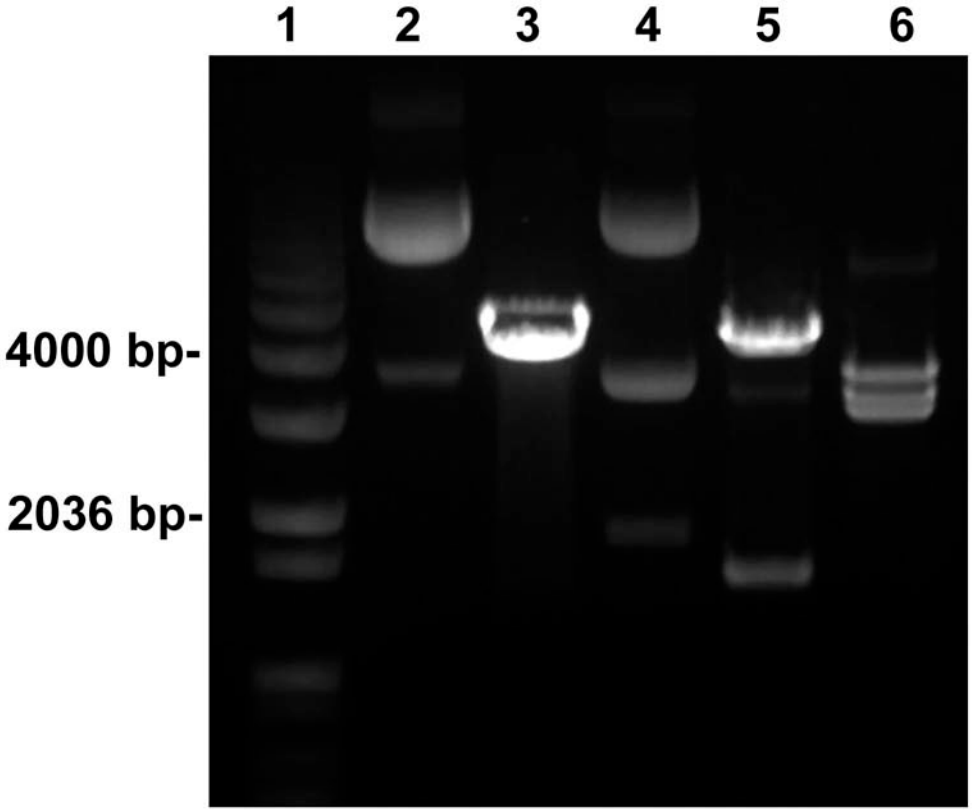

Figure 4-2. Preparation of transcriptional fusion constructs of *lacZ* reporter to *ctpA* gene promoter and upstream elements**.** The gel is demonstrates the key features in the preparation of the modified pXL-USE*ctpA-LacZ* constructs. Lane 1: 1kb DNA marker (Invitrogen); Lane 2: pXL-USE$ctpA_p$*lacZ* plasmid template; Lane 3: PCR of Full Promoter$_{ctpA}$-*lacZ* fragment; Lane 4; intact pXL-Promoter$_{ctpA}$-*lacZ*; Lane 5: verification of cloned pXL-Promoter$_{ctpA}$-*lacZ*, intact plasmid was cut with BamHI and NcoI; Lane 6: confirmation of desired orientation of the cloned fragment, pXL- Promoter$_{ctpA}$-*lacZ* cut with XbaI and NdeI.

Intact USE_{ctpA} + $Promoter_{ctpA}$

ATAATATTAAAACATTTGTTTATAAAAATGGAGAATTTGCACAAGAATAAAAACATTTT
TATACAAATGTTTTTTTCTTTGGATTTTTTTATTTAAAAAA CGTCTAAAAAACCAGTAT
AGAT CTTGTAT AAGCAAAAAAATGGAC TAATAT TTAAATAAAGCAAAAA GGAG AAAAAG
ATT ATGAAACTA *LacZ coding sequence*

USE_{Trunc1} + $Promoter_{ctpA}$

ATAATATTAAAACATTT GTTTATAAAAATGGAGAATTTGCACAAGAATAAAAACATTTTT
ATACAAATGTTTTTTTCTTTGGATTTTTTTATTTAAAAAA CGTCTAAAAAACCAGTATAG
AT CTTGTAT AAGCAAAAAAATGGAC TAATAT TTAAATAAAGCAAAAA GGAG AAAAAGATT
ATGAAACTA *LacZ coding seque nce*

USE_{Trunc2} + $Promoter_{ctpA}$

ATAATATTAA AACATTTGTTTATAAAAATGGAGAATTTGCACAAGAATAAAAA CATTTT
TATACAAATGTTTTTTTCTTTGGATTTTTTTATTTAAAAAA CGTCTAAAAAACCAG TAT
AGAT CTTGTAT AAGCAAAAAAATGGAC TAATAT TTAAATAAAGCAAAAA GGAG AAAAAG
ATT ATGAAACTA *LacZ coding sequence*

USE_{Trunc3} + $Promoter_{ctpA}$

ATAATATTAAAACATTTGTTTATAAAAATG GAGAATTTGCACAAGAATAAAAACATTTT
TATACAAATGTTTTTTTCTTTGGATTTTTTT ATTTAAAAAA CGTCTAAAAAACCAG TAT
AGAT CTTGTAT AAGCAAAAAAATGGAC TAATAT TTAAATAAAGCAAAAA GGAG AAAAAG
ATT ATGAAACTA *LacZ coding sequence*

$Promoter_{ctpA}$

ATAATATTAAAACATTTGTTTATAAAAATGGAGAATTTGCACAAGAATAAAAACATTTT
TAT ACAAATGTTTTTTTCTTTGGATTTTTTTATTTAAAAAA CGTCTAAAAAACCAG TAT
AGAT CTTGTAT AAGCAAAAAAATGGAC TAATAT TTAAATAAAGCAAAAA GGAG AAAAAG
ATT ATGAAACTA *LacZ coding sequence*

RBS_{ctpA}

AT AATATTAAAACATTTGTTTATAAAAATGGAGAATTTGCACAAGAATAAAAACATTTT
TATACAAATGTTTTTTTCTTTGGATTTTTTT ATTTAAAAA ACGTCTAAAAAACCAGTAT
AGAT CTTGTAT AAGCAAAAAAATGGAC TAATAT TT AAATAAAGCAAAAA GGAG AAAAAG
ATT ATGAAACTA *LacZ coding sequence*

Figure 4-3. Sequence of *lacZ* deletion constructs. The deleted sequences for each construct are shown in gray font. The -35 region of the promoter (CTTGTAT) is shown in yellow highlight. The -35 region of the promoter (CTTGTAT) is shown in yellow highlight. The -10 region of the promoter (TAATAT) is shown in green highlight. The ribosomal binding site (GGAG) is shown in red highlight. The codons of the *ctpA* for the first three amino acids (MLK) are shown in turquoise highlight. The *lacZ* coding sequence that was cloned in frame is shown in pink highlight.

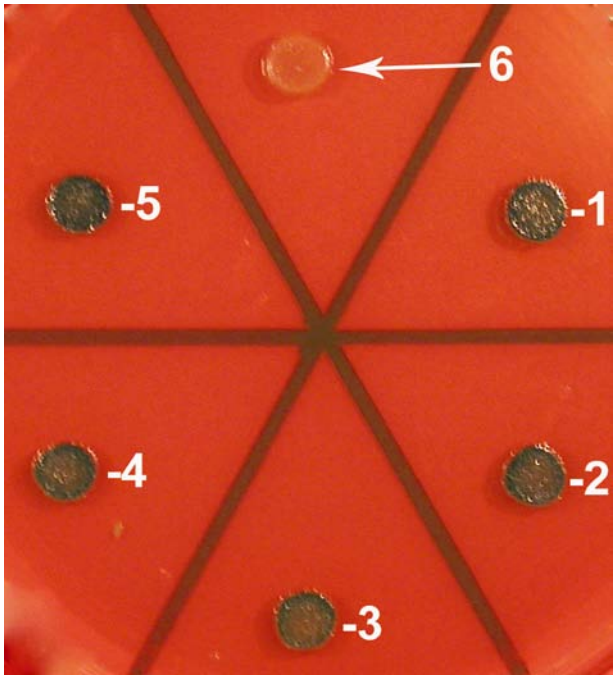

Figure 4-4. Expression of *lac*Z in *E. coli* under the control of USE and promoter of *ctpA*. Deletion of the promoter region (6) abrogated *lacZ* expression**,** whereas successive deletions of upstream nucleotides sequences (1-4) of *ctpA* did not affect *lacZ* expression. See Figure 4-3 for nucleotide sequences that were deleted. 1 = USE_{ctpA} + $Promoter_{ctpA}$; 2 = USE_{Trunc1} + $Promoter_{ctpA}$; 3 = USE_{Trunc2} + $Promoter_{ctpA}$; 4 = USE_{Trunc3} + $Promoter_{ctpA}$; 5 = $Promoter_{ctpA}$; 6 = $Ribosomal\ Binding\ Site_{ctpA.}$

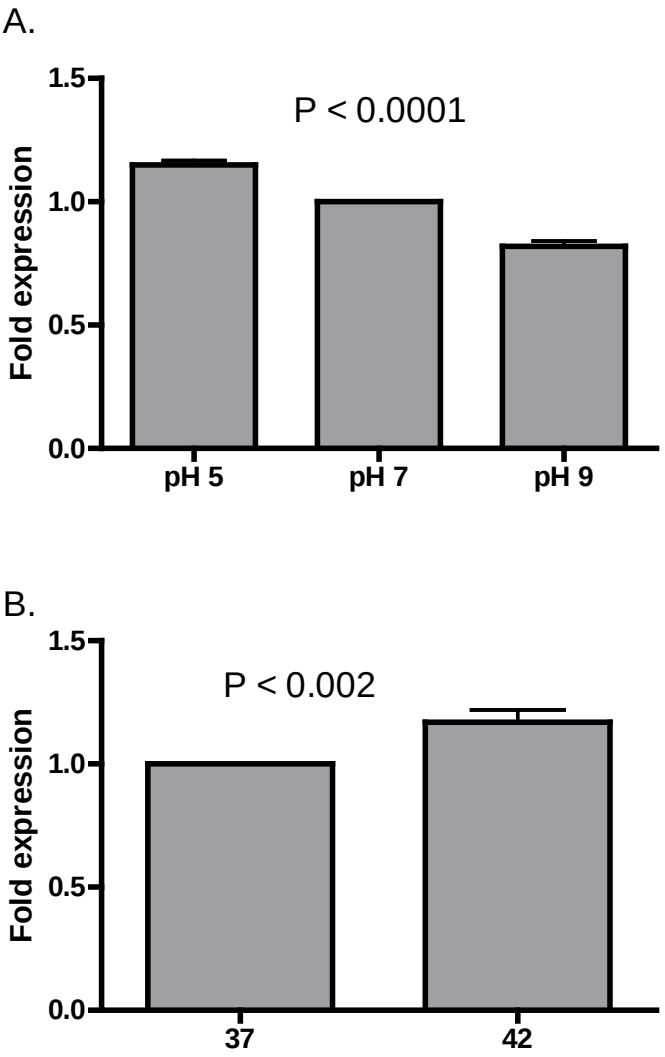

Figure 4-5. Measurement of *ctpA* gene expression in *M. mycoides capri* wild type by real-time RT-PCR. Bacterial cultures were grown to log phase then subjected to varying pH (A) or varying temperature (B). For RT-PCR reactions, mRNA for *tef* gene was used as an internal reference for pH experiment while mRNA for 50S Ribosome L09 was used as an internal reference for heat shock experiment. The threshold cycle number for each experimental sample was converted to fold expression. Fold expression values were normalized to pH 7 (graph A) or 37^0 C (graph B). Values represent the mean fold change ± SD of 5 replicates. ANOVA confirmed that there was a statistical significance among groups (P≤ 0.0001). Fisher's PLSD test (P valued shown in graph A) confirmed that the fold change observed in *ctpA* expression at either pH 5 or 9 was significantly different from expression at pH 7. Unpaired student's t test showed there was a statistically significant fold change in *ctpA* expression in cultures grown at 42^0 C.

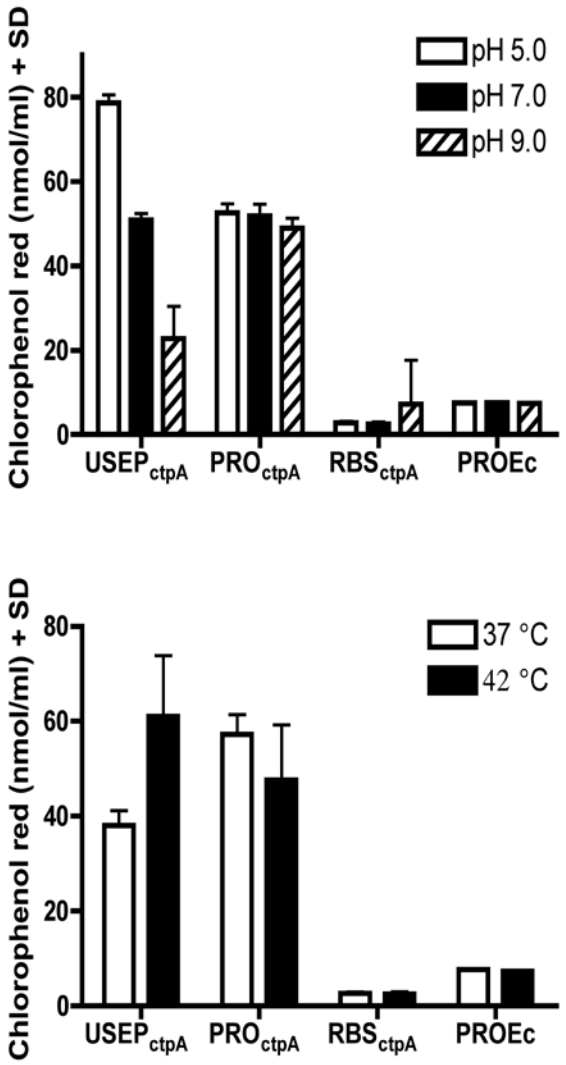

Figure 4-6. Effect of pH and heat shock on ctpA USE and P driven *lacZ* expression in *E. coli*. Upper panel: effect of USE on the *lacZ* expression in response to pH change. *LacZ* expressed from its own native promoter (PROEc) was used as an internal reference for all experiments. *LacZ* under the control of *ctpA* (RBS_{ctpA}) alone without the promoter was a negative control for all experiments. Five replicas were used for each experiment. Complete USE and P region ($USEP_{ctpA)}$ increased *lacZ* expression in response to acidic pH, while *LacZ* expression was reduced at alkaline pH. The *ctpA* promoter (PRO_{ctpA}) alone resulted in constitutive expression of *lacZ* and rendered it nonresponsive to shift in pH. Lower panel: effect of USE on *lacZ* expression in response to heat shock. An increase in *lacZ* expression in response to heat shock to 42^{o}C was seen with the complete USE and P region. Deletion of USE caused *lacZ* expression to be irresponsive to heat shock.

CHAPTER 5
SUMMARY AND FUTURE DIRECTIONS

Mycoplasmas, which have the smallest genomes known among bacteria, are thought to represent the minimum gene set required for sustaining an independent and self-replicating organism (Fadiel et al. 2007). The genetics of these microbes and their use to create a synthetic bacterium have attracted significant attention in recent years (Gibson et al. 2010, Lartigue et al. 2009, Gibson et al. 2008). Studies aimed at analysis of the whole genome or the global gene expression have revealed surprising degrees of complexity and versatility at many levels of gene expression. Despite the accumulation of data about mycoplasma genomes and their expression profiles, a mechanistic view of how they manage and regulate their gene expression is still lacking. In order to unravel the mycoplasma gene regulatory mechanisms, many hurdles have to be overcome. Among the factors impeding progress are the lack of effective genetic tools to target and manipulate specific genes in mycoplasmas as well as the lack of good models to study gene regulation in controlled and precise experimental settings. These factors are particularly challenging in mycoplasmas because of their fastidious growth requirements and complex medium components. In addition, very few genes can offer a discernible phenotype that would facilitate investigating a gene response, and screening or identifying mutants of interest. In this study, I identified a target gene with an identifiable phenotype that allowed me to develop improved methods to create targeted gene disruptions and to investigate gene regulation.

Mycoplasma mycoides subsp. *capri*, a member of the mycoides cluster, is characterized by its ability to hydrolyze casein on an agar plate. The gene responsible for this proteolytic phenotype has been identified through random mutagenesis. The

sequence of the gene revealed that it encodes for an S41 peptidase protein belonging to carboxyl-terminal proteases and shares key conserved motifs with CtpA proteases of several bacterial species (Hara et al. 1991). Expression of this gene seems to be coordinated with the growth phase, reaching its maximum at stationary phase. The discernible phenotype and the consistent response to the growth phase made the *ctpA* gene a good target for developing a specific gene disruption method and a model for studying gene regulation in mycoplasma.

Major Findings

In Chapter 2, I describe the development of a new technique for targeted mutagenesis through homologous recombination in *M. mycoides* subsp. *capri*. A novel method was developed for targeted mutagenesis that substantially increased the efficiency of targeted gene disruption. This method is a major improvement over existing methods in terms of both consistency and efficiency. By including the *E. coli RecA* gene in frame behind the promoter and RBS of the *ctpA* gene, the consistency and efficiency of obtaining the desired mutation in *M. mycoides* subsp. *capri* were significantly increased.

Although the decision to include *E. coli RecA* might not be intuitively obvious, the rationale was based on knowledge of mycoplasmal systems. The complete pathway of homologous recombination is missing in mycoplasma genomes (Rocha et al., 2005).

In addition, initiator systems and resolving complexes do not seem to be conserved among mycoplasmas (Ingleston et al. 2002). These findings have led to the belief that mycoplasmas are deficient or incapable of performing homologous recombination. However, when homologous recombination events have been reported (Dybvig and Woodard 1992), they were dependent on the presence of RecA. The exact

mechanism for the increased efficiency of homologous recombination in mycoplasmas when *E. coli* RecA was included remains to be determined. Importantly, to the best of our knowledge, the successful and consistent disruption of *ctpA* represents one of the very few documented, targeted, double crossover homologous recombination events in Mollicutes, and the only case achieved deliberately by experimental design.

Creation of an isogenic mutant of *M. mycoides* subsp. *capri* made it possible to investigate the consequences of *ctpA* disruption on the expression profile in the mutant. In Chapter 3, I used differential proteome profiling to investigate the impact of *ctpA* gene disruption on *M. mycoides* subsp. *capri*. The disruption of *ctpA* had a pleiotropic effect on the expression profile. Particularly noticeable was its impact on the glycolysis/gluconeogenesis pathway in carbon and energy metabolism. Disruption of *ctpA* seems to cause a shift toward glycerol metabolism and downregulation of lactate dehydrogenase genes. Both of these effects may precipitate an increase in oxidative stress. A corollary of increased glycerol metabolism is an elevated production of H_2O_2 which has been considered a virulence factor in the closely related *M. mycoides* subsp. *mycoides* Small Colony (Pilo et al., 2007). Disruption of the *ctpA* gene seems to increase the sensitivity of the mutant to heat shock. It is not clear whether there is a link between the shift toward glycerol metabolism and possible increase of H_2O_2 production on one hand and the temperature sensitivity on the other.

The role of CtpA as part of the stress response is also consistent with its elevated expression at the stationary phase and its response to acidic stress. Because mycoplasmas lack cytochromes and therefore depend on the proton motive potential as a source of energy (Pollack et al. 1981), an increasing concentration of H^+ ions can also

be a source of oxidative stress. Since the *ctpA* gene seems to be part of the stress response, determining genes affected by *ctpA* disruption in the mutant may help us understand how mycoplasmas cope with stress conditions.

The proteomic profiling results in Chapter 3 indicated that CtpA is likely a component of the stress response, particularly acidic and heat stress. Therefore, in Chapter 4, I investigated the effect of pH and temperature on the expression of *ctpA*, and the step at which *ctpA* expression is controlled. I used qRT-PCR to confirm that *ctpA* is quantitatively regulated at the transcriptional level by pH and temperature as well as certain genes –other than *ctpA*- whose disruptions altered the proteolytic phenotype. Interestingly, transcriptional levels of *ctpA* in these mutants were impacted. The common factor among these disrupted genes is that they are all involved in the maintenance, homeostasis and in fundamental biochemical and metabolic processes.

It was noted in a previous study that random insertion of Tn916 into the noncoding upstream (USE) region of ctpA resulted in decreased transcription and marked reduction of proteolytic activity (Rosentel 2003). Therefore, in Chapter 4, I also used mutation analysis to address the role of this noncoding upstream sequence (USE) region in transcriptional regulation of *ctpA.* Initially, *lacZ* transcriptional fusion constructs were created and tested in *M. mycoides* subsp. *mycoides.* However, weak β-galactosidase enzymatic activity precluded effective use of the constructs in *M. mycoides* subsp*. capri*.

In order to study the effects of these mutations in USE, I developed *lacZ* transcriptional fusion constructs that could be cloned and used in *E. coli*. Using this model, I confirmed by β-galactosidase enzymatic assay that the impacts of temperature

and pH in *E. coli* were similar to those observed in *M. mycoides* subsp. *mycoides.* These results suggest that the mechanism of regulation is conserved and is dependent on the basal transcriptional machinery; this in turn made it possible to further investigate the transcription mechanism in *E. coli*.

With this approach, I demonstrated that the *ctpA* promoter alone could drive transcription. However, the promoter region alone was not responsive to environmental stimuli such as pH and temperature. However, if the upstream nucleotide sequence (USE) was present, then *ctpA* transcription was responsive to these environmental signals. Based on the current literature, this is the first description of the quantitative regulation of a mycoplasmal gene that is mediated by the noncoding upstream nucleotides. This is also the first report that transcriptional regulation in mycoplasmas can be studied in *E. coli* using reporter gene fusion constructs.

Future Directions

Despite the fact that the suicide plasmid that was described in Chapter 2 significantly improved the ability to perform targeted mutagenesis in *M. mycoides* subsp. *capri*, further improvements could be made. For example, one approach is to place *E. coli recA* under the control of a stronger promoter such as the spiralin promoter (Cordova et al. 2005), which also has the advantage of being constitutively expressed. Another possibility is to transiently express some of the components of the initiation or resolution complexes that are missing in mycoplasmas (Rocha et al. 2005), preferably, in combination with *E. coli recA.* Finally, improved constructs should be tested in other mycoplasmal species.

Further investigations of the network of proteins involved in the stress response are warranted. Future experiments could address the effect of carbon source on metabolism, H_2O_2 production, and tolerance for heat shock.

Finally, a comprehensive study of the upstream elements could provide further insights into the mechanism by which mycoplasmal genes are quantitatively regulated in response to environmental signals. Determination of the specific nucleotides of USE region required for transcriptional regulation as well as identification of proteins that might bind to, and interact with, USE could help elucidate the specific mechanism(s) involved in transcriptional regulation.

Mutagenesis could be used to determine the minimal nucleotide sequence and critical bases responsible for regulation of *ctpA* expression. The identities of potential DNA binding proteins could be investigated using molecular biology assays such as the DNaseI Footprinting assay, the filter binding assay, and the gel mobility shift assay.

There have been few genetic tools developed for mycoplasmas including transposon Tn4001T. The attempt to use Tn4001T as a cloning vector has been encumbered with instability of the cloned DNA fragment. In addition, the expression and detection of *lacZ* gene from *E. coli* in *M. mycoides* subsp. *capri* was found to be too inefficient and weak to use for quantitative analysis.

Developing a better reporter gene that functions well in mycoplasmas would permit addressing regulatory questions directly in mycoplasmas. Expression of the *LacZ* in *M. mycoides* subsp. *capri* was disappointing in my study. However, improving the translation efficiency of *E. coli lacZ* gene expression in mycoplasmas through codon optimization is one potential approach. Another approach may be to include alternative

reporter genes that are more suitable for mycoplasmas in lieu of *lacZ*. For example, one candidate is the sialidase gene that has been reported in many mycoplasma species (May et al. 2007). Sialidase activity can be quantitatively assayed using a fluorogenic substrate such as 2′-(4-methylumbelliferyl)-α-D-*N*-acetylneuraminic acid (MUAN) and spectrofluorometry (Hunt and Brown 2007). Therefore, activity would be quantifiable and codon optimization would require no or minimal effort.

In conclusion, my study has provided two significant tools to better manipulate and investigate gene regulation in mycoplasmas. Both the ability to target specific genes for disruption by homologous recombination and the development of a model system to study mycoplasmal gene regulation in *E. coli* are significant contributions to the field. Finally, my studies have revealed that quantitative regulation of *ctpA* occurs by mechanisms other than those associated with antigenic variation (Citti et al. 2005) or stochastic processes (Simmons et al. 2007). The importance of the noncoding upstream nucleotides in the quantitative regulation of a gene in response to environmental stimuli has provided new insights into the regulation of mycoplasmal genes.

LIST OF REFERENCES

Abu-Groun, E. A. M., Taylor, R. R., Varsani, H., Wadher, B. J., Leach, R. H. & Miles, R. J. (1994). Biochemical diversity within the *Mycoplasma mycoides* "cluster".
Microbiology **140,** 2033-42.

Abouhamad, W. N., Manson, M., Gibson, M. M. & Higgins, C. F. (1991). Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella typhimurium*: characterization of the dipeptide permease (Dpp) and the dipeptide-binding protein. *Mol Microbiol* **5,** 1035-47.

Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M. & Adler, L. (1997). The two isoenzymes for yeast NAD+ -dependent glycerol 3-phosphate dehydrogenase encoded by GPD1 and GPD2 have distinct roles in osmoadaptation and redox regulation. *EMBO J* **16,** 2179 – 2187.

Barré, A., de Daruvar, A. & Blanchard, A. (2004). MolliGen, a database dedicated to the comparative genomics of Mollicutes. *Nucleic Acids Res* **32,** D307-10.

Baumler, A. J., Kusters, J. G., Stojiljkovic, I. & Heffron, F. (1994). *Salmonella typhimurium* loci involved in survival within macrophages. *Infect Immun* **62,** 1623-30.

Beebe, K. D., Shin, J., Peng, J., Chaudhury, C., Khera, J. & Pei, D. (2000). Substrate recognition through a PDZ domain in tail-specific protease. *Biochemistry* **39,** 3149-3155.

Bischof, D. F., Janis, C., Vilei, E. M., Bertoni, G. & Frey J. (2008). Cytotoxicity of *Mycoplasma mycoides* subsp*. mycoides* small colony type to bovine epithelial cells. *Infect Immun* **76,** 263-9.

Boesen, T., Emmersen, J.,Baczynska, A., Birkelund, S. & Christiansen, G. (2004). The vaa locus *of Mycoplasma hominis* contains a divergent genetic islet encoding a putative membrane protein. *BMC Microbiol* **4,** 37.

Bork, P., Ouzounis, C., Casari, G., Schneider, R., Sander, C., Dolan, M., Gilbert, W. & Gillevet, P. M. (1995). Exploring the *Mycoplasma capricolum* genome: a minimal cell reveals its physiology. *Mol Microbiol* **16,** 955-67.

Burgos, R., Pich, O. Q., Querol, E. & Pinol, J. (2008). Deletion of the *Mycoplasma genitalium* MG_217 gene modifies cell gliding behaviour by altering terminal organelle curvature. *Mol Microbiol* **69,** 1029-40.

Burgos, R., Pich, O. Q., Querol, E. & Pinol, J. (2007). Functional analysis of the *Mycoplasma genitalium* MG312 protein reveals a specific requirement of the MG312 N-terminal domain for gliding motility. *J Bacteriol* **189,** 7014-23.

Carvalho, F. M., Fonseca, M. M., De Medeiros, S. B., Scortecci, K. C., Blaha, C. A. & Agnez-Lima, L. F. (2005). DNA repair in reduced genome: the Mycoplasma model. *Gene* **360,** 111-9.

Cecchini, K. R., Gorton, T. S. & Geary, S. J. (2007). Transcriptional responses of *Mycoplasma gallisepticum* strain R in association with eukaryotic cells. *J Bacteriol* **189,** 5803-7.

Chopra-Dewasthaly, R., Citti, C., Glew, M. D., Zimmermann, M.,Rosengarten, R. & Jechlinger, W. (2008). Phase-locked mutants of *Mycoplasma agalactiae*: defining the molecular switch of high-frequency Vpma antigenic variation. *Mol Microbiol* **67,** 1196-210.

Chopra-Dewasthaly, R., Marenda, M., Rosengarten, R., Jechlinger, W. & Citti, C. (2005). Construction of the first shuttle vectors for gene cloning and homologous recombination in *Mycoplasma agalactiae*. *FEMS Microbiol Lett* **253,** 89-94.

Chopra-Dewasthaly, R., Zimmermann, M., Rosengarten, R. & Citti, C. (2005). First steps towards the genetic manipulation of *Mycoplasma agalactiae* and *Mycoplasma bovis* using the transposon Tn4001mod. *Int J Med Microbiol* **294,** 447-53.

Cirillo, V. P. (1993). Transport systems in mycoplasmas. *Subcell Biochem* **20,** 293-310.

Cocks, B. G., Brake, F. A., Mitchell, A. & Finch, L. R. (1985). Enzymes of intermediary carbohydrate metabolism in *Ureaplasma urealyticum* and *Mycoplasma mycoides* subsp. *mycoides*. *J Gen Microbiol* **131,** 2129-35.

Coll, J. L., Heyde, M. & Portalier, R. (1994). Expression of the *nmpC* gene of *Escherichia coli* K-12 is modulated by external pH. Identification of cis-acting regulatory sequences involved in this regulation. *Mol Microbiol* **12,** 83-93.

Cordova, C. M., Lartigue, C., Sirand-Pugnet, P., Renaudin, J., Cunha, R. A. & Blanchard, A. (2002). Identification of the origin of replication of the *Mycoplasma pulmonis* chromosome and its use in oriC replicative plasmids. *J Bacteriol* **184,** 5426-35.

Cottew, G. S., Breard, A., DaMassa, A. J., Ernø, H., Leach, R. H., Lefevre, P. C., Rodwell, A. W. & Smith, G. R. (1987). Taxonomy of the *Mycoplasma mycoides* cluster. *Isr J Med Sci* **2,** 632-5.

Cox, M. M. (2007). Motoring along with the bacterial RecA protein. *Nat Rev Mol Cell Biol* **8,** 127-38.

DaMassa, A. J., Brooks, D. L. & Adler, H. E. (1983). Caprine mycoplasmosis: widespread infection in goats with *Mycoplasma mycoides subsp mycoides* (large-colony type). *Am J Vet Res* **44,** 322-5.

Dandekar, T., Huynen, M., Regula, J. T., Ueberle, B., Zimmermann, C. U., Andrade, M. A., Doerks, T., Sánchez-Pulido, L., Snel, B., Suyama, M., Yuan Y. P., Herrmann, R. & Bork, P. (2000). Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res* **28,** 3278-88.

Deutscher, J., Galinier, A. & Martin-Verstraete, I. (2002). Carbohydrate uptake and metabolism. In *Bacillus Subtilis and Its Closest Relatives*, pp. 129-150. Edited by A. L. Sonenshein and others. Washington, DC: American Society for Microbiology.

Dhandayuthapani, S., Blaylock, M. W., Bebear, C. M., Rasmussen, W. G. & Baseman, J. B. (2001). Peptide methionine sulfoxide reductase (MsrA) is a virulence determinant in *Mycoplasma genitalium*. *J Bacteriol* **183,** 5645-50.

Dhandayuthapani, S., Rasmussen, W. G. & Baseman, J. B. (1999). Disruption of gene mg218 of *Mycoplasma genitalium* through homologous recombination leads to an adherence-deficient phenotype. *Proc Natl Acad Sci USA* **96,** 5227-32.

Diner, B. A., Ries, D. F., Cohen, B. N. & Metz, J. G. (1988). COOH-terminal processing of polypeptide D1 of the photosystem II reaction center of *Scenedesmus obliquus* is necessary for the assembly of the oxygen-evolving complex. *J Biol Chem* **263,** 8972-8980.

Duret, S., A. Andre, A. & Renaudin, J. (2005). Specific gene targeting in *Spiroplasma citri*: improved vectors and production of unmarked mutations using site-specific recombination. *Microbiology* **151,** 2793-803.

Duret, S., Berho, N., Danet, J. L., Garnier, M. & Renaudin, J. (2003). Spiralin is not essential for helicity, motility, or pathogenicity but is required for efficient transmission of *Spiroplasma citri* by its leafhopper vector *Circulifer haematoceps*. *Appl Environ Microbiol* **69,** 6225-34.

Duret, S., Danet, J. L., Garnier, M. & Renaudin, J. (1999). Gene disruption through homologous recombination in *Spiroplasma citri*: an scm1-disrupted motility mutant is pathogenic. *J Bacteriol* **181,** 7449-56.

Dybvig, K. & Voelker, L. L. (1996). Molecular biology of mycoplasmas. *Annu Rev Microbiol* **50**, 25-57.

Dybvig, K. & Woodard, A. (1992). Construction of recA mutants of *Acholeplasma laidlawii* by insertional inactivation with a homologous DNA fragment. *Plasmid* **28,** 262-6.

Ekici, O. D., Paetzel, M. & Dalbey, R. E. (2008). Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* **17,** 2023-37.

Ezraty, B., Aussel, L. & Barras, F. (2005). Methionine sulfoxide reductases in prokaryotes. *Biochim Biophys Acta* **17,** 221-9.

Fadiel, A., Eichenbaum, K. D., Semary, N. El. & Epperson, B. (2007). Mycoplasma genomics: tailoring the genome for minimal life requirements through reductive evolution. *Front Biosci* **12,** 2020-8.

Forage, R. G. & Lin, E. C. (1982). DHA system mediating aerobic and anaerobic dissimilation of glycerol in *Klebsiella pneumoniae* NCIB 418. *J Bacteriol* **151,** 591−599.

Fraser, C. M. , Gocayne, J. D.,White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, R. D., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J., Nguyen, D., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J. F., Dougherty, B. A., Bott, K. F., Hu, P. C., Lucier, T. S. , Peterson, S. N., Smith, H. O., Hutchison 3rd, C. A. & Venter, J. C. (1995). The minimal gene complement of *Mycoplasma genitalium*. *Science* **270,** 397-403.

French, C. T., Lao, P., Loraine, A. E., Matthews, B. T., Yu, H. & Dybvig, K. (2008). Large-scale transposon mutagenesis of *Mycoplasma pulmonis*. *Mol Microbiol* **69,** 67-76.

Frey, J. (2002). Mycoplasmas of animals. In *Molecular Biology and Pathogenicity of mycoplasmas*, pp. 73-90. Edited by S. Razin & R. Herrmann. New York: Plenum.

Gamper, M. & Kas, P. (2005). Strategy for chromosomal gene targeting in RecA-deficient *Escherichia coli* strains. *Biotechniques* **38,** 405-8.

Garcia-Alles, L. F., Siebold, C., Nyffeler, T. L., Flükiger-Brühwiler, K., Schneider, P., Bürgi, H. B., Baumann, U. & Erni, B. (2004). Phosphoenolpyruvate- and ATP-dependent dihydroxyacetone kinases: covalent substrate-binding and kinetic mechanism. *Biochemistry* **19,** 13037-45.

Gardner, S. W. & Minion, F. C. (2010). Detection and Quantification of Intergenic Transcription in *Mycoplasma hyopneumoniae*. *Microbiology* 2010 Apr 29. [Epub ahead of print]

Gawron-Burke, C. & Clewell, D. B. (1984). Regeneration of insertionally inactivated streptococcal DNA fragments after excision of transposon Tn916 in *Escherichia coli*: strategy for targeting and cloning of genes from gram-positive bacteria. *J Bacteriol* **159,** 214-21.

Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A. 3rd, Smith, H. O. & Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329,** 52-6

Gibson, D. G., Benders, G. A., Andrews-Pfannkoch, C., Denisova, E. A., Baden-Tillson, H., Zaveri, J., Stockwell, T. B., Brownley, A., Thomas, D. W., Algire, M. A., Merryman, C., Young, L., Noskov, V. N., Glass, J. I., Venter, J. C., Hutchison III, C. A. & Smith, H. O. (2008). Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319,** 1215–1220.

Glass, J. I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M. R., Maruf, M.,Hutchison III, C. A., Smith, H. O. & Venter, J. C. (2006). Essential genes of a minimal bacterium. *Proc Natl Acad Sci USA* **103,** 425–430.

Glass, J. I., Lefkowitz, E. J., Glass, J. S., Heiner, C. R., Chen, E. Y. & Cassell, G. H. (2000). The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature* **407,** 757-62.

Gollin, D. J., L. E. Mortenson, L. E. & Robson, R. L. (1992). Carboxyl-terminal processing may be essential for production of active NiFe hydrogenase in *Azotobacter vinelandii*. *FEBS Lett* **309,** 371-375.

Goryshin, I. Y., Jendrisak, J., Hoffman, L. M., Meis, R. & Reznikoff, W. S. (2000). Insertional transposon mutagenesis by electroporation of released Tn5 transposition complexes. *Nat Biotechnol* **18,** 97-100.

Goryshin, I. Y. & Reznikoff, W. S. (1998). Tn5 in vitro transposition. *J Biol Chem* **273,** 7367-74.

Greenberg, E. P. (2000). Bacterial genomics. Pump up the versatility. *Nature* **31,** 947.

Güell, M., van Noor,t V., Yus, E., Chen, W. H., Leigh-Bell, J., Michalodimitrakis, K., Yamada, T., Arumugam, M., Doerks, T., Kühner, S., Rode, M., Suyama, M., Schmidt, S., Gavin, A. C., Bork, P. & Serrano L. (2009). Transcriptome complexity in a genome-reduced bacterium. *Science* **326,** 1268-71.

Gutknecht, R., Beutler, R., Garcia-Alles, L. F., Baumann, U. & Erni, B. (2001). The dihydroxyacetone kinase of *Escherichia coli* utilizes a phosphoprotein instead of ATP as phosphoryl donor. *EMBO J* **15,** 2480-6.

Halbedel, S. & Stulke, J. (2007). Tools for the genetic analysis of Mycoplasma. *Int J Med Microbiol* **297,** 37-44.

Halbedel, S., Eilers, H., Jonas, B., Busse, J., Hecker, M., Engelmann, S. & Stülke, J. (2007). Transcription in *Mycoplasma pneumoniae*: analysis of the promoters of the ackA and ldh genes. *J Mol Biol* **371,** 596-607.

Hames, C., Halbedel, S., Hoppert, M., Frey, J. & Stülke, J. (2009). Glycerol metabolism is important for cytotoxicity of *Mycoplasma pneumoniae*. *J Bacteriol* **191,** 747-53.

Hara, H., Yamamoto, Y., Higashitani, A., Suzuki, H. & Nishimura, Y. (1991). Cloning, mapping, and characterization of the *Escherichia coli* prc gene, which is involved in C-terminal processing of penicillin-binding protein 3. *J Bacteriol* **173,** 4799-813.

Hatchikian, E. C., Magro, V., Forget, N., Nicolet, Y. & Fontecilla-Camps, J. C. (1999). Carboxy-terminal processing of the large subunit of [Fe] hydrogenase from *Desulfovibrio desulfuricans* ATCC 7757. *J Bacteriol* **181,** 2947-2952.

Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B. C. & Herrmann, R. (1996). Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res* **24,** 4420-49.

Hoffman, L. M., Jendrisak, J. J., Meis, R. J., Goryshin, I. Y. & Reznikof, S. W. (2000). Transposome insertional mutagenesis and direct sequencing of microbial genomes. *Genetica* **108,** 19-24.

Holms, H. (1996). Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol Rev* 19, 85 - 116.

Horii, T., Ozawa, N., Ogawa, T. & Ogawa, H. (1992). Inhibitory effects of N- and C-terminal truncated *Escherichia coli* recA gene products on functions of the wild-type recA gene. *J Mol Biol* **5,** 105-14.

Hunt, M. E. & Brown, D. R. (2007). Role of sialidase in *Mycoplasma alligatoris*-induced pulmonary fibroblast apoptosis. *Vet Microbiol* **121,** 73-82.

Hutchison, C. A., Peterson, S. N., Gill, S. R., Cline, R. T., White, O., Fraser, C. M., Smith, H. O. & Venter, J. C. (1999). Global transposon mutagenesis and a minimal Mycoplasma genome. *Science* **286,** 2165-9.

Ingleston, S. M., Dickman, M. J., Grasby, J. A., Hornby, D. P., Sharples, G. J. & Lloyd, R. G. (2002). Holliday junction binding and processing by the RuvA protein of *Mycoplasma pneumoniae*. *Eur J Biochem* **269,** 1525-33.

Islam, M. R., Grubb, J. H. & Sly, W. S. (1993). C-terminal processing of human beta-glucuronidase. The propeptide is required for full expression of catalytic activity, intracellular retention, and proper phosphorylation. *J Biol Chem* **268,** 22627-22633.

Jaffe, J. D., Berg, H. C., & Church, G. M. (2004). Proteogenomic mapping as a complementary method to perform genome annotation. *Proteomics* **4,** 59-77.

Janis, C., Bischof, D., Gourgues, G., Frey, J., Blanchard, A. & Sirand-Pugnet, P. (2008). Unmarked insertional mutagenesis in the bovine pathogen *Mycoplasma mycoides* subsp. *mycoides* SC: characterization of a lppQ mutant. *Microbiology* **154,** 2427-36.

Janis, C., Lartigue, C., Frey, J., Wroblewski, H., Thiaucourt, F., Blanchard, A. & Sirand-Pugnet, P. (2005). Versatile use of oriC plasmids for functional genomics of *Mycoplasma capricolum* subsp. *capricolum*. *Appl Environ Microbiol* 7**1,** 2888-93.

Jarhede, T. K., Le Henaff, M. & Wieslander, A. (1995). Expression of foreign genes and selection of promoter sequences in *Acholeplasma laidlawii*. *Microbiology* **141,** 2071-9.

Jores, J., Meens, J., Buettner, F. F., Linz, B., Naessens, J. & Gerlach, G. F. (2009). Analysis of the immunoproteome of *Mycoplasma mycoides* subsp. *mycoides* small colony type reveals immunogenic homologues to other known virulence traits in related Mycoplasma species.*Vet Immunol Immunopathol* **131,** 238-45.

Kannan, T. R., Musatovova, O., Balasubramanian, S., Cagle, M., Jordan, J. L., Krunkosky, T. M., Davis, A., Hardy, R. D. & Baseman, J. B. (2010). *Mycoplasma pneumoniae* CARDS toxin expression reveals growth phase and infection-dependent regulation. *Mol Microbiol* 2010 Feb 28. [Epub ahead of print]

Kannan, T. R., & J. B. Baseman, J. B. (2006). ADP-ribosylating and vacuolating cytotoxin of *Mycoplasma pneumoniae* represents unique virulence determinant among bacterial pathogens. *Proc Natl Acad Sci USA* **103,** 6724-9.

Karnauchov, I., Herrmann, R. G., Pakrasi, H. B. & Klosgen, R. B. (1997). Transport of CtpA protein from the cyanobacterium *Synechocystis* 6803 across the thylakoid membrane in chloroplasts. *Eur J Biochem* **249,** 497-504.

Keiler, K. C., Waller, P. R. H. & Sauer, R. T. (1996). Role of a peptide tagging system in degredation of proteins synthesized from damaged messenger RNA. *Science* **271,** 990-3.

Krishnakumar, R., Assad-Garcia, N., Benders, G. A., Phan, Q., Montague, M. G. & Glass, J. I. (2010). Targeted Chromosomal Knockouts in *Mycoplasma pneumoniae*. *Appl Environ Microbiol* 2010 Jun 11. [Epub ahead of print]

Kühner, S., van Noort, V., Betts, M. J., Leo-Macias, A., Batisse, C., Rode, M., Yamada, T., Maier, T., Bader, S., Beltran-Alvarez, P., Castaño-Diez, D., Chen, W. H., Devos, D., Güell, M., Norambuena, T., Racke, I., Rybin, V., Schmidt, A., Yus, E., Aebersold, R., Herrmann, R., Böttcher, B., Frangakis, A. S., Russell, R. B., Serrano, L., Bork, P. & Gavin, A. C. (2009). Proteome organization in a genome-reduced bacterium. *Science* **326,** 1235-40.

Lad, S. P., Yang, G., Scott, D. A., Wang, G., Nair, P., Mathison, J., Reddy, V. S. & Li, E. (2007). Chlamydial CT441 is a PDZ domain-containing tail-specific protease that interferes with the NF-kappaB pathway of immune response. *J Bacteriol* **189,** 6619-25.

Lartigue, C., Vashee, S., Algire, M. A., Chuang, R. Y., Benders, G. A., Ma, L., Noskov, V. N., Denisova, E. A., Gibson, D. G., Assad-Garcia, N., Alperovich, N., Thomas, D. W., Merryman, C., Hutchison 3rd, C. A., Smith, H. O., Venter, J. C. & Glass, J. I. (2009). Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science* **325,** 1693-6.

Lartigue, C., Blanchard, A., Renaudin, J., Thiaucourt, F. & Sirand-Pugnet, P. (2003) . Host specificity of mollicutes oriC plasmids: functional analysis of replication origin. *Nucleic Acids Res* **31,** 6610-8.

Lavery, P. E. & Kowalczykowski, S. C. (1992). Enhancement of recA protein-promoted DNA strand exchange activity by volume-occupying agents. *J Biol Chem* **267,** 9307-14.

Leach, R. H., Erno, H. & MacOwan, K. J. (1993). Proposal for designation of F38-type caprine mycoplasmas as *Mycoplasma capricolum* subsp. *capripneumoniae* subsp. nov. and consequent obligatory relegation of strains currently classified as *M. capricolum* (Tully, Barile, Edward, Theodore, and Erno 1974) to an additional new subspecies, M. *capricolum* subsp. *capricolum* subsp. nov. *Int J Syst Bacteriol* **43,** 603-5.

Lee, C. D. & Wang, T. F. (2009). The N-terminal domain of *Escherichia coli* RecA have multiple functions in promoting homologous recombination. *J Biomed Sci* **1,** 16:37.

Letunic, I., Doerks, T. & Bork, P. (2009). SMART 6: recent updates and new developments. *Nucleic Acids Res* **37,** D229-32.

Liao, D. I., Qian, J., Chisholm, D. A., Jordan, D. B. & Diner, B. A. (2000). Crystal structures of the photosystem II D1 C-terminal processing protease. *Nat Struct Biol* **7,** 749-53.

Madsen, M. L., Puttamreddy, S., Thacker, E. L., Carruthers, M. D. & Minion, F. C. (2008). Transcriptome changes in *Mycoplasma hyopneumoniae* during infection. *Infect Immun* **76,** 658-63.

Madsen, M. L., Nettleton, D., Thacker, E. L., Edwards, R. & Minion, F. C. (2006). a Transcriptional profiling of *Mycoplasma hyopneumoniae* during heat shock using microarrays. *Infect Immun* **74,** 160-6.

Madsen, M. L., Nettleton, D., Thacker, E. L. & Minion, F. C. (2006). b Transcriptional profiling of *Mycoplasma hyopneumoniae* during iron depletion using microarrays. *Microbiology* **152,** 937-44.

Manso-Silván, L., Vilei, E. M., Sachse, K., Djordjevic, S. P., Thiaucourt, F. & Frey, J. (2009). *Mycoplasma leachii* sp. nov. as a new species designation for Mycoplasma sp. bovine group 7 of Leach, and reclassification of *Mycoplasma mycoides* subsp. *mycoides* LC as a serovar of Mycoplasma mycoides subsp. capri. *Int J Syst Evol Microbiol* **59,** 1353-8.

Marais, A., Bove, J. M. & Renaudin, J. (1996). *Spiroplasma citri* virus SpV1-derived cloning vector: deletion formation by illegitimate and homologous recombination in a spiroplasmal host strain which probably lacks a functional recA gene. *J Bacteriol* **178,** 862-70.

Markham, P. F., Kanci, A., Czifra, G., Sundquist, B., Hains, P. & Browning, G. F. (2003). Homologue of macrophage-activating lipoprotein in *Mycoplasma gallisepticum* is not essential for growth and pathogenicity in tracheal organ cultures. *J Bacteriol* **185,** 2538-47.

Masui, R., Mikawa, T. & Kuramitsu, S. (1997). Local folding of the N-terminal domain of *Escherichia coli* RecA controls protein-protein interaction. *J Biol Chem* 31, 27707-15.

Mattman, L. H. (2001). Cell wall Deficient forms: Stealth Pathogens. Boca Raton, Florida: CRC Press LLC.

May, M., Kleven, S. H. & Brown, D. R. (2007). Sialidase activity in *Mycoplasma synoviae*. *Avian Dis* **51,** 829-33.

Mazmanian, S. K., Ton-That, H. & Schneewind, O. (2001). Sortase-catalysed anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. *Mol Microbiol* **40,** 1049-1057.

Menon, N. K., Robbins, J., Der Vartanian, M., Patil, D., Peck, H. D., Menon, A. L., Robson, R. L. & Przybyla, A. E. (1993). Carboxy-terminal processing of the large subunit of [NiFe] hydrogenases. *FEBS Lett* **331,** 91-95.

Michel, B. (1999). Illegitimate recombination in bacteria. In *Organization of prokaryotic genome*, pp. 129-150. Edited by R. L. Charlebois. Washington DC: American Society for Micorbiology

Miles, R. J. (1992). Catabolism in mollicutes. *J Gen Microbiol* **138,** 1773-83.

Mitchell, S. J. & Minnick, M. F. (1997). A carboxy-terminal processing protease gene is located immediately upstream of the invasion-associated locus from *Bartonella bacilliformis*. *Microbiology* **143,** 1221-1233.

Monnet, V. (2003). Bacterial oligopeptide-binding proteins. *Cell Mol Life Sci* **60,** 2100-14.

Muto, A. & Ushida, C. (2002). Transcription and Translation. p323-345. In *Molecular biology and pathogenicity of mycoplasmas*, pp. 323-345. Edited by S. Razin & R. Herrmann. New York: Plenum.

Muto, A., Yamao, F. & Osawa, S. (1987). The genome of *Mycoplasma capricolum*. *Prog Nucleic Acid Res Mol Biol* **34,** 29-58.

Muto, A. (1987). The genome structure of *Mycoplasma capricolum*. *Isr J Med Sci* **23,** 334-41.

Nagasawa, H., Sakagami, Y., Suzuki, A., Suzuki, H., Hara, H. & Hirota, Y. (1989). Determination of the cleavage site involved in C-terminal processing of penicillin-binding protein 3 of *Escherichia coli*. *J Bacteriol* **171,** 5890-5893.

Nicholas, R. A. & Bashiruddin, J. B. (1995). *Mycoplasma mycoides* subspecies *mycoides* (small colony variant): the agent of contagious bovine pleuropneumonia and member of the "*Mycoplasma mycoides* cluster". *J Comp Pathol* **113,** 1-27.

Nixon, P. J., Trost, J. T. & Diner, B. A. (1992). Role of the carboxy terminus of polypeptide D1 in the assembly of a functional water-oxidizing manganese cluster in photosystem II of the cyanobacterium *Synechocystis* sp. PCC 6803: assembly requires a free carboxyl group at C-terminal position 344. *Biochemistry* **31,** 10859-10871.

Ogasawara, N., Moriya, S. & Yoshikawa, H. (1991). Initiation of chromosome replication: structure and function of oriC and DnaA protein in eubacteria. *Res Microbiol* **142,** 851-9.

Oneal, M. J., Schafer, E. R., Madsen, M. L. & Minion, F. C. (2008). Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to norepinephrine. *Microbiology* **154,** 2581-8.

Ostberg, Y., Carroll, J. A., Pinne, M., Krum, J. G., Rosa, P. & Bergström, S. (2004). Pleiotropic effects of inactivating a carboxyl-terminal protease, CtpA, in *Borrelia burgdorferi*. *J Bacteriol* **186,** 2074-84.

Paetzel, M. & Dalbey, R. E. (1997). Catalytic hydroxyl/amine dyads within serine proteases. *Trends Biochem Sci* **22,** 28-31.

Pelicic, V., Reyrat, J. M. & Gicquel, B. (1996). Expression of the *Bacillus subtilis* sacB gene confers sucrose sensitivity on mycobacteria. *J Bacteriol* **178,** 1197-9.

Pilo, P., Frey, J. & Vilei, E. M. (2007). Molecular mechanisms of pathogenicity of *Mycoplasma mycoides* subsp. *mycoides. SC Vet J* **174,** 513-21.

Pilo, P., Vilei, E. M., Peterhans, E., Bonvin-Klotz, L., Stoffel, M. H., Dobbelaere, D. & Frey, J. (2005). A metabolic enzyme as a primary virulence factor of *Mycoplasma mycoides* subsp. *mycoides* small colony. *J Bacteriol* **187,** 6824-31.

Pinto, P. M., Chemale, G., de Castro, L. A., Costa, A. P., Kich, J. D., Vainstein, M. H., Zaha, A. & Ferreira, H. B. (2007). Proteomic survey of the pathogenic *Mycoplasma hyopneumoniae* strain 7448 and identification of novel post-translationally modified and antigenic proteins. *Vet Microbiol* **121,** 83-93.

Pollack, J. D. (2002). Central carbohydrate pathways: metabolic flexibility and the extra role of some “housekeeping” enzymes. In *Molecular biology and pathogenicity of mycoplasmas*, pp. 163-199. Edited by S. Razin & R. Herrmann. New York: Plenum.

Pollack, J. D., Williams, M. V. & McElhaney R. N. (1997). The comparative metabolism of the mollicutes (Mycoplasmas): the utility for taxonomic classification and the relationship of putative gene annotation and phylogeny to enzymatic function in the smallest free-living cells. *Crit Rev Microbiol* **23,** 269-354.

Pollack, J. D., Merola, A. J., Platz, M. & Booth Jr., R. L. (1981). Respiration-associated components of Mollicutes. Journal of Bacteriology, 146 (3), pp. 907-913.

Razin, S., D. Yogev, D. & Naot, Y. (1998). Molecular biology and pathogenicity of mycoplasmas. *Microbiol Mol Biol Rev* **62,** 1094-156.

Renaudin, J. & Lartigue, C. (2005). OriC plasmids as gene vectors for mollicutes. In *Mycoplasmas: Pathogenesis, Molecular Biology, and Emerging Strategies for Control*, pp. 3-30. Edited by A. Blanchard & G. Browning. Norwich, UK: Horizon Scientific Press.

Renaudin, J. (2002). Extrachromosomal elements and gene transfer. In *Molecular biology and pathogenicity of mycoplasmas,* pp. 347-371. Edited by S. Razin & R. Herrmann. New York: Plenum.

Rocha, E., Sirand-Pugnet, P. & Blanchard, A. (2005). Genome analysis: Recombination, Repair, and Recombinational Hotspots. In *Mycoplasmas: Molecular biology, Pathogenicity, and Strategies for control*, pp. 31-73. Edited by A. Blanchard & G. Browning. UK: Horizon Bioscience.

Rocha, E. P., and A. Blanchard, A. (2002). Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution. *Nucleic Acids Res* **30,** 2031-42.

Rodreguez, J. L., Proveda, J. B., Oros, J., Herraez, P., Sierra, M. A. & Fernandez, A. (1995). High mortality in goats associated with the isolation of a strain of *Mycoplasma mycoides* subsp*. mycoides* (Large Colony type). *Zentralble-Veterinarmed-B* **42,** 587-593.

Rosentel, J. K. (2003). Extracellular proteases: potential virulence factor of *Mycoplasma mycoides* subsp. *mycoides*. Ph.D. dissertation. University of Florida. Florida.

Rossmann, R., Sauter, M., Lottspeich, F., & Bock, A. (1994). Maturation of the large subunit (HYCE) of *Escherichia coli* hydrogenase 3 requires nickel incorporation followed by C-terminal processing at Arg537. *Eur J Biochem* **220,** 377-384.

Samuelsson, T., Elias, P., Lustig, F. & Guindy, Y. S. (1985). Cloning and nucleotide sequence analysis of transfer RNA genes from *Mycoplasma mycoides*. *Biochem J* **232,** 223-8.

Schafer, E. R., Oneal, M. J., Madsen, M. L. & Minion, F. C. (2007). Global transcriptional analysis of *Mycoplasma hyopneumoniae* following exposure to hydrogen peroxide. *Microbiology* **153,** 3785-90.

Schmidl, S. R. , Gronau, K., Hames, C., Busse, J., Becher, D., Hecker, M. & Stülke, J. (2010). The stability of cytadherence proteins in *Mycoplasma pneumoniae* requires activity of the protein kinase PrkC. *Infect Immun* **78,** 184-92.

Schumann, W. (1999). FtsH--a single-chain charonin? *FEMS Microbiol Rev* **23,** 1-11.

Schurmann, M., & Sprenger, G. A. (2001). Fructose-6-phosphate aldolase is a novel class I aldolase from *Escherichia coli* and is related to a novel group of bacterial transaldolases. *J Biol Chem* **276,** 11055-11061.

Shahram, M., Nicholas, R. A., Miles, R. J., Wood, A. P. & Kelly, D. P. (2009). Kinetics of substrate oxidation and hydrogen peroxide production by *Mycoplasma mycoides* subsp. *mycoides* Large Colony (LC) type and *Mycoplasma mycoides* subsp. *capri*. *Res Vet Sci* **87,** 364-6.

Shilov, I. V., Seymour, S. L., Patel, A. A., Loboda, A., Tang, W. H., Keating, S. P., Hunter, C. L., Nuwaysir, L. M. & Schaeffer, D. A. (2007). The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics* **6,** 1638-1655.

Shirvan, M. H. & Rottem, S. (1993). Ion pumps and volume regulation in mycoplasma. *Subcell Biochem* **20,** 261-92.

Silber, K. R., Keiler, K. C. & Sauer, R. T. (1992). Tsp: a tail-specific protease that selectively degrades proteins with nonpolar C termini. *Proc Natl Acad Sci* **89,** 295-299.

Simmons, W. L., Bolland, J. R., Daubenspeck, J. M. & Dybvig, K. (2007). A stochastic mechanism for biofilm formation by *Mycoplasma pulmonis. J Bacteriol* **189,** 1905-13.

Smits, W. K., Dubois, J. Y., Bron, S.,van Dijl, J. M. & Kuipers, O. P. (2005). Tricksy business: transcriptome analysis reveals the involvement of thioredoxin A in redox homeostasis, oxidative stress, sulfur metabolism, and cellular differentiation in *Bacillus subtilis*. *J Bacteriol* **187,** 3921-30.

Stintzi, A., D. Marlow, D., Palyada, k., Naikare, H., Panciera, R., Whitworth, L. & Clarke, C. (2005). Use of genome-wide expression profiling and mutagenesis to study the intestinal lifestyle of *Campylobacter jejuni*. *Infect Immun* **73,** 1797-810.

Stone, K. L. & Williams, K. R. (1997). Digestion of Proteins in Gels for Sequence Analysis. (ed.) In *Current protocols in protein science,* vol. 2 pp. 11.3.1 – 11.3.12. Edited by J. E. Coligan and others. NJ: Wiley & Sons, Inc.

Story, R. M. & Weber, I. T. & Steitz T. A. (1992). The structure of the *E. coli* recA protein monomer and polymer. *Nature* **355**, 318-25.

Su, Y. A., He, P. & Clewell, D. B. (1992). Characterization of the tet(M) determinant of Tn916: evidence for regulation by transcription attenuation. *Antimicrob Agents Chemother* **36,** 769-78.

Taschke, C. & Herrmann, R. (1988). Analysis of transcription and processing signals in the 5' regions of the two *Mycoplasma capricolum* rRNA operons. *Mol Gen Genet* **212,** 522-30.

Tryon, V. V. & Baseman, J. B. (1992). Pathogenic determinants and mechanisms. In *Mycoplasmas: molecular biology and pathogenesis,* pp. 457- 471. Edited by J. Maniloff. Washington, DC: American Society for Microbiology.

Tully, J. G., Rose, D. L., Whitcomb, R. F. & Wenzel, R. P. (1979). Enhanced isolation of *Mycoplasma pneumoniae* from throat washings with a newly-modified culture medium. *J Infect Dis* **139,** 478-82.

Ueberle, B., Frank, R. & Herrmann, R. (2002). The proteome of the bacterium *Mycoplasma pneumoniae*: comparing predicted open reading frames to identified gene products. *Proteomics* **2,** 754-64.

Van Melderen, L., Thi, M. H. D., Lecchi, P., Gottesman, S., Couturier, M. & Maurizi, M.R. (1996). ATP-dependent degradation of CcdA by Lon protease - effects of secondary structure and heterologous subunit interactions. *J Biol Chem* **271,** 27730 -27738.

Vilei, E. M., Abdo, E. M., Nicolet, J., Botelho, A., Gonçalves, R. & Frey, J. (2000). Genomic and antigenic differences between the European and African/Australian clusters of *Mycoplasma mycoides* subsp. *mycoides* SC. *Microbiology* **146,** 477-86.

Voelker, L. L. & Dybvig, K. (1998). Transposon mutagenesis. *Methods Mol Biol* **104,** 235-8.

Waldo, R. H., Popham, P. L., Romero-Arroyo, C. E.,Mothershed, E. A., Lee, K. K. & Krause, D. C. (1999). Transcriptional analysis of the hmw gene cluster of *Mycoplasma pneumoniae*. *J Bacteriol* **181,** 4978-85.

Wasinger, V. C., Pollack, J. D., & Humphery-Smith, I. (2000). The proteome of *Mycoplasma genitalium*. Chaps-soluble component. *Eur J Biochem* **267,** 1571-82.

Weiner, J. 3rd, Zimmerman, C. U., Göhlmann, H. W. & Herrmann, R. (2003). Transcription profiles of the bacterium *Mycoplasma pneumoniae* grown at different temperatures. *Nucleic Acids Res* **31,** 6306-20.

Weiner, J. 3rd, Herrmann, R. & Browning, G. F. (2000). Transcription in *Mycoplasma pneumoniae*. *Nucleic Acids Res* **28,** 4488-96.

Wilson, C. J., Zhan, H., Swint-Kruse, L. & Matthews, K. S. (2007). The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. Cell Mol Life Sci **64**, 3-16.

Yamao, F., Iwagami, S., Azumi, Y., Muto, A., Osawa, S., Fujita, N. & Ishihama, A. (1988). Evolutionary dynamics of tryptophan tRNAs in *Mycoplasma capricolum*. *Mol Gen Genet* **212,** 364-9.

Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y. & Osawa, S. (1985). UGA is read as tryptophan in *Mycoplasma capricolum*. *Proc Natl Acad Sci USA* **82,** 2306-9.

Ye, F., Renaudin, J., Bové, J. M. & Laigret, F. (1994). Cloning and sequencing of the replication origin (oriC) of the *Spiroplasma citri* chromosome and construction of autonomously replicating artificial plasmids. *Curr Microbiol* **29,** 23-9.

Yus, E., Maier, T., Michalodimitrakis, K., van Noort, V., Yamada, T., Chen, W. H., Wodke, J. A., Güell, M., Martínez, S., Bourgeois, R., Kühner, S., Raineri, E., Letunic, I., Kalinina, O. V., Rode, M. , Herrmann, R., Gutiérrez-Gallego, R., Russell, R. B., Gavin, A. C., Bork, P. & Serrano, L. (2009). Impact of genome reduction on bacterial metabolism and its regulation. *Science* **326,** 1263-8.

Zeller, T. & G. Klug. (2006). Thioredoxins in bacteria: functions in oxidative stress response and regulation of thioredoxin genes. *Naturwissenschaften* **93,** 259-66.

Zhong, G., Fan, P., Ji, H., Dong, F. & Huang, Y. (2001). Identification of a chlamydial protease-like activity factor responsible for the degradation of host transcription factors. *J Exp Med* **193,** 935-42.

Zou, N. & Dybvig, K. (2002). DNA replication, repair, and stress response. In *Molecular Biology and pathogenicity of Mycoplasmas*, pp. 303-321. Edited by S. Razin & R. Herrmann. New York: Plenum.

BIOGRAPHICAL SKETCH

Ayman Allam was born in the City of Cairo, Egypt. He received his B.SC (Botany/Chemistry) from Ain Shams University/ Cairo, Egypt. As a one year compulsory military service, he served in the army as a chemist specialized in rocket fuel analysis. He then joined the Plant Taxonomy and Natural Products Unit at the National Research Center (NRC) – Cairo, Egypt.

At NRC he began working on identification and isolation of plant phytoalexins, antimicrobial substances synthesized de novo by plants that rapidly accumulate at the site of pathogen infection. He won the National ICGEB (the International Center for Genetic Engineering and Biotechnology – Trieste, Italy) fellowship to study molecular biology, DNA recombinant technology and biotechnology. He studied at Department of Biochemistry and Molecular Biology, Bari University, Bari – Italy. His “Diploma of Specialization” was in Plant Molecular Biology. Upon returning to NRC in Cairo he joined the Microbial Biotechnology Unit working on metal chelation by algae and fungi. He won a scholarship from Georgia State University where he studied and obtained a M.SC on “Molecular microbial Genetics of *Bacillus stearothermophilus*”.

He then joined KBI (Kinetic Biosystems Inc.), a research company located in Advanced Technology Development Center (ATDC) on the campus of the Georgia Institute of Technology. He worked as a microbiologist to study the possibility of using and/or genetically adapting *P. putida* to treat the nitrate- contaminated groundwater.

He was accepted as a Ph.D. student at the Department of Biochemistry and Molecular Biology, College of Medicine, University of Florida and then transferred to the College of Veterinary Medicine, “Department of Infectious Diseases and Pathology”.