2015

Non-coding RNAs of the Q fever agent, Coxiella burnetii

Indu Ramesh Warrier
The University of Montana

NON-CODING RNAS OF THE Q FEVER AGENT, *COXIELLA BURNETII*

By

INDU RAMESH WARRIER

M.Sc (Med), Kasturba Medical College, Manipal, India, 2010

Dissertation

presented in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
Cellular, Molecular and Microbial Biology

The University of Montana
Missoula, MT

August, 2015

Approved by:

Sandy Ross, Dean of The Graduate School
Graduate School

Michael F. Minnick, Chair
Division of Biological Sciences

Stephen J. Lodmell
Division of Biological Sciences

Scott D. Samuels
Division of Biological Sciences

Scott Miller
Division of Biological Sciences

Keith Parker
Department of Biomedical and Pharmaceutical Sciences

Warrier, Indu, PhD, Summer 2015 Cellular, Molecular and Microbial Biology

Non-coding RNAs of the Q fever agent, *Coxiella burnetii*

Chairperson: Michael F. Minnick

Coxiella burnetii is an obligate intracellular bacterial pathogen that undergoes a biphasic developmental cycle, alternating between a small cell variant (SCV) and a large cell variant (LCV). Despite the remarkable niche and life cycle of *C. burnetii*, little is known about its modes of regulation and the roles that non-coding RNAs play in its growth and development. One such element is the intervening sequence (IVS); a 444-nt RNA element that is inserted within helix 45 of *Coxiella*'s precursor 23S rRNA. *C. burnetii* may have acquired IVS through horizontal transfer, and it has been subsequently maintained by all strains of *C. burnetii* through vertical transfer. The IVS of *C. burnetii* contains an ORF that encodes a hypothetical ribosomal S23 protein (S23p). However, our data show that the S23p-encoding ORF is probably undergoing reductive evolution and therefore not expressed *in vivo*. Additionally, we observed that following RNase III-mediated excision, IVS RNA is degraded and levels of the resulting fragments of 23S rRNA differ significantly from each other and the 16S rRNA. Since the fragment of 23S rRNA that is lowest in quantity may dictate the number of mature ribosomes that are ultimately formed, we hypothesize that the biological role of IVS is to moderate *Coxiella*'s growth by fragmentation of its 23S rRNA thereby fostering *Coxiella*'s tendency towards slow growth and chronic infection. Further, we identified fifteen novel *Coxiella burnetii* sRNAs (CbSRs) using RNA-seq, which were verified using Northern analyses. Additionally, some of these CbSRs were upregulated in LCVs or during intracellular growth, suggesting adaptive roles in those contexts. Furthermore, we also identified and characterized the 6S RNA of *C. burnetii* and found that it accumulated during the SCV phase of the bacterium. The location of *ssrS* gene and the secondary structure of 6S RNA were similar to those of other eubacteria, indicating functionality. We also demonstrated that the 6S RNA of *C. burnetii* interacts specifically with RNA polymerase (RNAP). Finally, 6S RNA was highly expressed during intracellular growth of *C. burnetii* indicating that it probably regulates stress response by interacting with RNAP during transcription.

ACKNOWLEDGMENTS

Firstly, I would like to express my deepest gratitude to my mentor, Mike Minnick for instilling in me the qualities of being a good researcher. Without his guidance, patience and constant support this work would not have been possible. I am thankful to my committee members, Steve Lodmell, Scott Samuels, Scott Miller, and Keith Parker for their time and valuable suggestions. I would also like to thank my comprehensive exam committee chair, Brent Ryckman for his valuable advice regarding grant writing. A big shout-out to all of the current and past members of the Minnick Lab for creating an amazing environment to work especially, Jim Battisti for his assistance with 6S experiments and for all the invigorating discussions. I am indebted to our previous lab manager, Linda Hicks for her friendship and training to work with RNA. I would like to thank my dearest family and friends, especially my parents, parents-in-law, Mithila, Uncle Sudhir and Aunt Latha for their unconditional love and support. A special thanks to my wonderful husband, Sundaresh Shankar, for his constant support and inspiration that helped me to get through the hardships of graduate school with ease. I dedicate my work to my grandfather for his love and relentless belief in my abilities.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

COXIELLA BURNETII AND SMALL NON-CODING REGULATORY RNAS

A. *COXIELLA BURNETII*

I. Introduction

Coxiella burnetii is a Gram-negative obligate intracellular bacterium that causes Q fever in humans. *C. burnetii* is extremely infectious, with an ID_{50} of one to ten bacteria in the guinea pig model (1). The bacterium can exist in an endospore-like form, wherein it is stable in the environment for long periods. Due to its typical aerosol transmission to humans, its historic use as a bioweapon, and low ID_{50}, it is considered a potential bioterrorism agent. Hence, the CDC has classified *C. burnetii* as a category B select agent. This disease has gained greater notoriety due a recent outbreak in the Netherlands between 2007 and 2010, where more than 4,000 cases of acute Q fever were reported (4). The bacterium can be differentiated into phase I and II variants based on the length of the O-side chain of lipopolysaccharide (LPS). Strains with a phase I phenotype possess LPS with full-length O antigen and are fully virulent. In contrast, serial passage of phase I strains in embryonated chicken eggs or tissue culture media (i.e., immunologically incompetent backgrounds) causes truncation of the O antigen resulting in an avirulent, phase II phenotype and a deep rough LPS (2).

Until recently, *C. burnetii* was only propagated in eukaryotic host cells; a significant barrier to genetic manipulation. However, *C. burnetii* was “freed” from intracellular cultivation by development of an acidified citrate cysteine medium (ACCM) (5), which was later improved (ACCM2) to allow for isolation of single colonies on semisolid

medium (6). This advancement was followed by enhanced tools for genetic manipulation, including an improved *Himar1*-based transposon system (7), a stably maintained *ori*-based shuttle vector (8) and a site-directed mutagenesis strategy (9). These advances have provided new opportunities to gain insights into *C. burnetii*'s regulatory networks, intracellular parasitism and disease pathogenesis.

II. Q fever

Q fever is a zoonotic disease that is typically acquired by inhalation of aerosolized soil or animal products contaminated with *C. burnetii*. This disease has been reported worldwide with the exception of New Zealand. The most common reservoirs of this pathogen include wild animal species, arthropods and domestic animals (e.g., cattle, sheep and goats). Following exposure to the bacterium, the incubation periods varies from 2 to 6 weeks. In about half of cases the disease remains asymptomatic. In other instances, acute symptomatic Q fever presents with a flu-like illness accompanied by high fever, severe retro-orbital headache, malaise, myalgia, atypical pneumonia, and hepatitis for 1-2 weeks (10). Although the acute disease is self-limiting, in ~1-5% of the cases, it can develop into a chronic infection (11) characterized by endocarditis, hepatitis or a post-Q fever fatigue syndrome (12, 13). Endocarditis is the most severe manifestation of chronic Q fever, with a high mortality rate (up to 60%) even with timely antibiotic treatment (14).

Diagnosis of Q fever is difficult due to the generalized flu-like symptoms of the disease. Therefore, the diagnosis relies heavily on serology to detect the presence of *Coxiella* antigens or antibodies against them. An indirect immunofluorescence assay,

complement fixation or ELISAs are routinely used as diagnostic tests (10, 15). Additionally, the nature and titer of antibody can be useful to differentiate between acute and chronic infections. During acute infection, IgG antibodies against phase II antigens of the bacterium are higher, whereas titers against phase I antigens are higher during chronic infection, even though both antibody types are elevated [reviewed in (2)].

Although acute Q fever is most often self-resolving, doxycycline (100 mg twice daily) for 2-3 weeks is the first line of treatment to reduce severity of symptoms (11). However, chronic Q fever is difficult to treat and requires a prolonged treatment with doxycycline (100 mg twice daily) and hydroxychloroquine (600 mg once daily) for 18 months (16).

Since Q fever is a zoonotic disease, the chances of infection are high, especially in at-risk individuals like abattoir workers. However, a vaccine for Q fever is not widely available. Q-Vax, a formalin-inactivated whole-cell preparation is licensed for use in Australia (17). Additionally, a soluble, trichloroacetic acid extract of phase I bacteria is used as a vaccine in Czechoslovakia (18), and a chloroform-methanol residue vaccine is used by the US military (19).

III. Development cycle

C. burnetii undergoes a biphasic developmental cycle consisting of two distinct cellular forms (Fig. 1-1). Small cell variants (SCVs) are metabolically inert, highly infectious, endospore-like cells of the bacterium that are extremely resistant to harsh environmental conditions, including heat, UV light, desiccation, pressure and certain disinfectants (20, 21). Following aerosol transmission, the SCV passively enters host

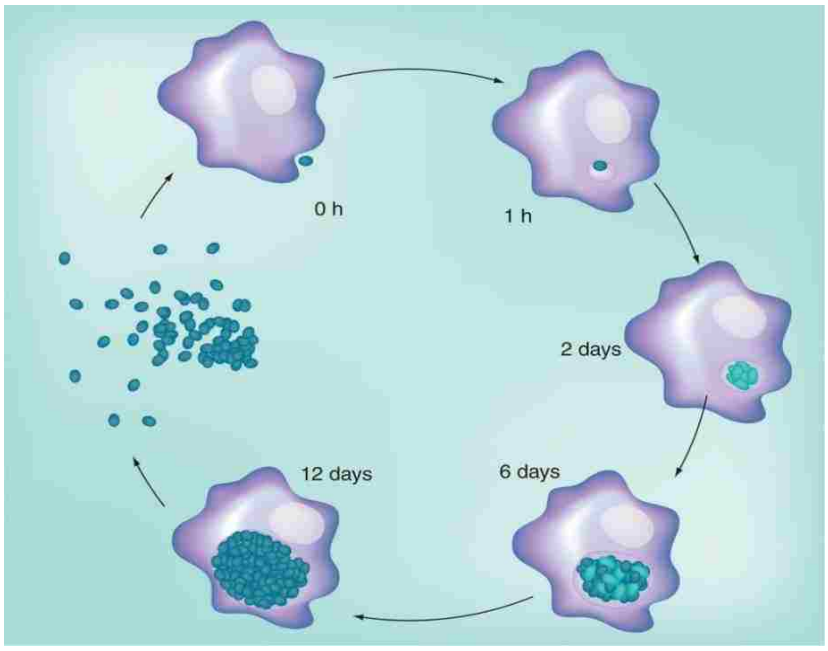


Figure 1-1. Developmental cycle of *C. burnetii*. A small cell variant (SCV) is typically internalized by an alveolar macrophage. Following acidification of the vacuole, SCVs metamorphose into large cell variants (LCVs) that are actively replicating by day 2. After 5-6 days, LCVs differentiate back to SCVs. By day 12, the PV fills the host cell cytoplasm and SCVs are released upon lysis of the host. Adapted from (2).

phagocytic cells (alveolar monocyte/macrophage) by actin-dependent phagocytosis (22, 23). Additionally, recent reports show that *C. burnetii* can invade non-phagocytic cells albeit by a different mechanism (24, 25). Phagocytosis of the bacterium results in a phagosome that matures through the endosomal cascade by recruiting RAB5 and RAB7, resulting in a phagolysosome-like compartment called a parasitophorous vacuole (PV) (26). During maturation, the PV undergoes acidification, increases in size and delays fusion with the lysosome to enhance the pathogen's survival. At this point, the PV interacts with autophagosome for nutrients, which along with the acidic pH (~4.5), results in the conversion of SCVs to metabolically-active, relatively fragile, large cell variants (LCVs) (26, 27). Eventually, the PV fuses with the lysosome; a step that is evident by the accumulation of lysosomal enzymes, including cathepsin D (CTSD) and lysosomal acid phosphatase (ACP2) (28). Interestingly, the PV membrane contains high amounts of cholesterol (twice that of a normal lysosomal compartment) such that inhibition of the host's cholesterol metabolism negatively affects PV membrane formation (29). After 5 to 6 days of infection, LCVs start to differentiate back into SCVs with drastic expansion of the PV's size to almost consume the host cell's volume (30). At this stage, the bacterium prolongs it host cell viability by inhibiting apoptotic signaling pathways and inducing pro-survival factors (31, 32). After an infection time of ~12 days, all the bacteria transform into SCVs that are released upon lysis of the infected cell (30).

IV. Parasitic genetic elements of *C. burnetii*

Genomes of bacteria are constantly changing by acquiring new DNA through horizontal gene transfer and losing by nucleotide deletion (33). These processes are more common in environmental bacteria where chances of horizontal gene transfer can

occur. On the other hand, host-associated bacteria have little or no chance of acquiring new genetic material. Although *Coxiella* lives within an acidic PV, its genome contains an unusually large number of parasitic genetic elements and pseudogenes, which the organism may have acquired during a free-living past (34, 35). Such elements include 31-59 insertion sequences spread across its genome, a putative intein in the C-terminal region of the replicative DNA helicase (DnaB) protein and two self-splicing group I introns (Cbu.L1917 and Cbu.L1951) along with an intervening sequence (IVS) within the 23S rRNA gene (Fig. 1-2) [reviewed in (2)]. All of these genetic elements are highly conserved in every genotype of *Coxiella*; however, their exact function is yet to be determined.

a. Intervening sequence

Although the 23S rRNA of bacteria is highly conserved, some are fragmented. This fragmentation is a result of an IVS (36). IVSs are inserted within the 23S and 16S rRNA of certain bacteria and are transcribed along with the rRNA primary transcript but are excised quickly thereafter. The size of IVS elements varies from 55 to 759 nt (37, 38). Although the sequences of IVSs are not conserved among bacteria, their location of occurrence is highly conserved. The mechanism by which IVS elements are inserted into the genes of certain bacteria and the reason for targeting rRNA genes are not known. Since the ends of these elements are complementary, similar to mobile genetic elements, it suggests that IVSs are either inserted into the genome by recombination or are footprints of an insertion and subsequent excision event of a transposable element (36). The broad, sporadic occurrence of IVS elements suggests that they were probably acquired by horizontal gene transfer. Researchers have been trying to understand the

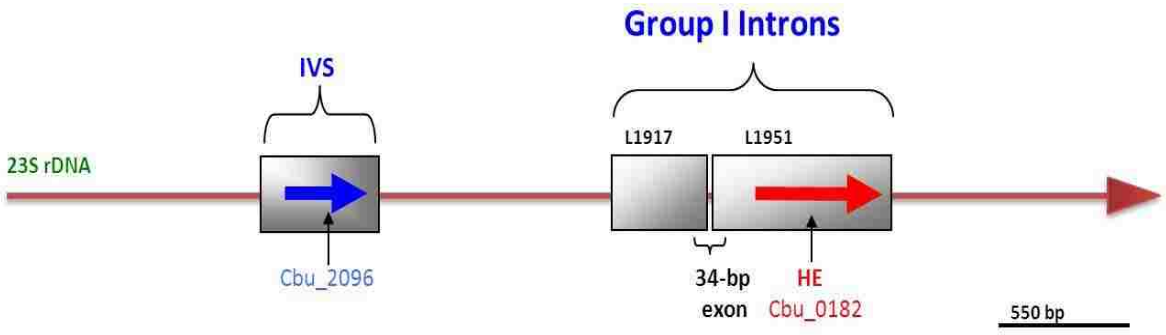


Figure 1-2. 23S rRNA gene linkage map of *C. burnetii*. The position of IVS is indicated along with a nested open reading frame (Cbu_2096). Two group I introns, L1917 and L1951 that flank a 34-bp exon is also shown. Adapted from (3)

biological function of this element in bacteria for some time; however, no specific role has been ascribed to date.

C. burnetii's IVS is a 444-nt element that is found inserted in the bacterium's 23S rRNA (Fig. 1-2), which is fragmented due to excision of the element (39). RNase III, a double-stranded specific endoribonuclease, has been shown to be involved in the excision of IVS in other bacteria (36-38). The IVS of *Leptospira* contains an ORF that encodes for a hypothetical protein that belongs to the S23p family of proteins. Although the function of this protein is not known, it is also found in *Xanthomonas* and *Brucella* (40). Interestingly, the IVS of *C. burnetii* also contains a similar ORF (39). The amino acid conservation of the S23p encoded in *Leptospira* and *Coxiella* IVSs suggest that it plays an important role, perhaps in mobility of the element. However, expression of the S23p protein in *Coxiella* is questionable, since no conserved Shine-Dalgarno sequence (SD) is observed preceding the start codon, indicating inefficient translation (39).

Several benefits of IVSs have been suggested to date. Fragmentation of 23S rRNA as a result of IVS excision could result in an increased efficiency of rRNA degradation during transition to the SCV phase. This possibility is based on precedence in *Salmonella*, where fragmentation of 23S rRNA resulted in increased degradation of the rRNA during stationary phase (41). Moreover, high levels of IVS during exponential phase could sequester ribonucleases, thereby regulating their intracellular levels. Further, IVS excision from 23S rRNA could also provide a point of regulation of 23S rRNA maturation. Another interesting possibility is since IVS elements are more common in bacteria that are closely associated with eukaryotic cells (42), they could conceivably promote communication between prokaryotes and eukaryotes. Finally,

since the presence of IVS elements has been correlated with pathogenicity in strains of *Yersinia enterocolitica* (43), the element might also play a role in somehow regulating pathogenesis.

B. SMALL NON-CODING REGULATORY RNAS

Small RNAs (sRNAs) are a heterogeneous group of short, non-coding RNAs ranging from 100-400 bases in length that perform a plethora of regulatory functions. They are either transcribed as a primary transcript or are formed as a result of processing of a larger transcript. sRNAs are known to modulate a variety of processes, including transcription, translation, mRNA stability, and DNA maintenance or silencing (44). However, most known sRNAs are post-transcriptional regulators that either inhibit or enhance translation (44). In contrast to protein-based regulation, sRNA-based regulation has multiple advantages. The most obvious is the energy cost, which is considerably lower for synthesizing these regulatory molecules. As a result of their small size and not having to undergo the extra step of translation, sRNAs are less “expensive” to produce than proteins. Additionally, sRNA-mediated regulation is much faster than proteins, and they can be used to respond in an extremely sensitive manner. A good example is the 5′ UTR element termed a riboswitch, which determines the fate of the downstream mRNA based on environmental cues (45). This type of regulation can be turned off easily and efficiently since sRNAs are much less stable than regulatory proteins (46). Moreover, based on the extent of complementarity with their respective targets, sRNAs can fine-tune regulation. For example, *trans*-encoded sRNAs that share incomplete complementarity with their targets can act on multiple targets and can differentially regulate translation based on their relative abundance and respective

binding affinities (46, 47). In the recent decade, research on discovery and characterization of sRNAs has increased greatly. Genome-wide analyses using high-density tiling arrays and RNA deep sequencing (RNA-seq) have revealed the transcriptomes of various pathogenic bacteria and exposed a high abundance of sRNAs. Many of these sRNAs are found to play a role in metabolism and virulence of these bacteria (48-50).

I. sRNA regulation in pathogenic bacteria

During survival in a hostile environment, bacteria must monitor the environment and alter gene expression accordingly. This is especially true for pathogenic bacteria thriving in an intracellular niche where nutrients are limited. In this environment, the bacterium is also under threat from the constant surveillance of the host's immune system, and it must withstand elevated concentrations of reactive oxygen species, degradative enzymes and various antibacterial constituents of lysosomes. In these instances, sRNAs act as excellent mediators that sense environmental cues and alter gene expression accordingly. sRNAs can be classified based on the mechanism by which they function in bacteria.

5' UTR elements are one of the best-studied classes of regulatory sRNAs and consist of two diverse groups: riboswitches and RNA thermometers. Basically, these elements are present in the 5' UTR of mRNAs and can perceive environmental signals, including temperature, pH, metabolite concentrations and stalled ribosomes [reviewed in (45)]. RNA thermometers can form secondary structures that are sensitive to temperature and affect the translation of the respective downstream mRNA. These

elements are very important in pathogenic bacteria. As an example, the 5' UTR of the *Listeria monocytogenes prfA* mRNA forms a secondary structure at lower temperature that masks the SD and prevents translation. Upon an increase in temperature to 37^0C, the secondary structure is relaxed, exposing the SD region to bring about translation of PrfA, which activates expression of several virulence genes in *Listeria monocytogenes* (51). Likewise, riboswitches can alternate between open and closed structures as a function of changing pH and metabolite concentrations, thereby interfering with translation of protein(s) encoded downstream [reviewed in (48, 49)].

Another broad class of regulatory sRNAs includes base-pairing sRNAs that are either *cis*- or *trans*-encoded. *Cis*-encoded sRNAs are antisense sRNAs that share complete complementarity with their target RNAs. The well-studied *cis*-encoded sRNA are expressed from plasmids, transposons or other mobile genetic elements (52). However, antisense sRNAs encoded on the chromosome are also being discovered. A few of these sRNA are also found to play a role in regulation of virulence gene expression (49). For instance, the *lesR-1* antisense sRNA encoded on the pSLT virulence plasmid of *Salmonella enterica* directly interacts with the 3' end of the bacterial *PSLT047* transcript (53). The PSLTO47 protein has been found to be important for the intracellular lifestyle of this pathogen (53).

On the other hand, *trans*-encoding sRNAs are encoded at a distal location on the chromosome in intergenic regions and thus share only limited complementarity with their target mRNAs. The majority of *trans*-encoded sRNAs are negative regulators that act by blocking the ribosome binding site, thereby bringing about translational inhibition, mRNA degradation or both (44). An excellent example of *trans*-encoded sRNA, with

relevance to *C. burnetii*, is *Chlamydia*'s *IhtA* sRNA that represses translation of the histone-like protein, Hc1. The increased expression of Hc1 during the late infectious phase, as a result of decreased transcription of *IhtA*, allows the pathogen to differentiate to a spore-like elementary body (EB) from a replicative, reticulate body (RB) (54). In contrast, *trans*-encoded sRNAs can also activate translation by base-pairing with an inhibitory secondary structure that sequesters the ribosome binding site. Such is the case with RNAIII of *Staphylococcus aureus* that activates α-hemolysis synthesis through base pairing with *hla* mRNA (55). Additionally, RNAIII represses translation of staphylococcal protein A and Rot transcription factor by binding to *spa* and *rot* mRNAs thereby controlling expression of virulence factors (55). Due to limited complementarity with their target RNAs, *trans*-encoded sRNAs usually associate with RNA chaperone protein, Hfq (44). Apart from acting as a platform to allow interactions between sRNAs and mRNAs, Hfq proteins also protect sRNAs from degradation in the absence of base pairing with mRNAs (44).

Another class of sRNAs control gene expression by associating with proteins. One of the best known examples of this is 6S RNA that interacts with σ^{70}-RNA polymerase (RNAP) and regulates transcription at specific promoters. In addition, a unique type of an adaptive microbial immune system has recently come to light called clustered regularly interspaced short palindromic repeats (CRISPR), which provide resistance against viruses and other invading genetic material (56). Recent reports have shown the relevance of crRNA in bacteria virulence [reviewed in (49)].

II. 6S RNA

6S RNA is a small noncoding RNA that is widely distributed among diverse bacteria. The 6S RNA is encoded by the *ssrS* gene and frequently found in the 5' UTR of *ygfA* (coding for 5-formyltetrahydrofolate cycloligase); a linkage that is widely conserved among α and γ Proteobacteria (57). 6S RNA has been found to be a global regulator of transcription by associating with RNA polymerase (RNAP). Bacterial RNAP is a multi-subunit core enzyme (α_2, β, β', ω) that associates with a sigma subunit (σ) to form a holoenzyme (Eσ). The holoenzyme form of RNAP is required for DNA promoter recognition and transcription initiation. 6S RNA specifically binds to housekeeping Eσ [i.e., $E\sigma^{70}$ in *Escherichia coli* (58) and $E\sigma^{A}$ in *Bacillus subtilis* (59)] but has weak or no binding with RNAP associated with alternate sigma factors σ^{38}, σ^{32} or the core RNAP enzyme (59, 60). Although 6S RNA is present throughout the growth phase of *E. coli*, it accumulates during stationary phase where over 75% of $E\sigma^{70}$ is bound to 6S RNA (58).

6S RNA has a characteristic secondary structure that consists of a single-stranded central bubble flanked by a closing stem and terminal loop (57). Although the nucleotide sequence of 6S RNA is not conserved, the overall secondary structure and certain base pairs and bulges are conserved and are important for its regulatory function (57). For example, mutants corresponding to the central bubble of the *E. coli* 6S RNA were unable to bind to RNAP, thus implicating this region in binding (59). The central bubble is reminiscent of the structure of DNA in an 'open promoter complex' when the DNA around the start site is melted during initiation of transcription. During its association with $E\sigma^{70}$, 6S RNA is lodged in the active site of RNAP. Studies have shown that the 4.2 region of σ^{70} that recognizes the -35 element of a DNA promoter is critical for its

interaction with 6S RNA. Specifically, eight positively-charged amino acid residues at the 4.2 region are essential for 6S RNA binding (61).

Studies in *E. coli* have shown that association of 6S RNA with $E\sigma^{70}$ results in decreased transcription from some σ^{70}-dependent promoters. Further, genome-wide transcriptome analyses of *ssrS* null mutants of *E. coli* showed downregulation of many genes, including a marked reduction in genes that encode the translational apparatus (62). In some cases, an extended -10 motif and a weak -35 element were identified as determinants of 6S RNA-sensitive promoters. However, this observation was not true in all cases. Additionally, 6S RNA was also found to activate several σ^{S}-dependent promoters (63). Due to its poor association with $E\sigma^{S}$, this effect may be indirect. Taken together, these observations indicate that 6S RNA is involved in regulating transcription during the transition from exponential to stationary phase during the growth cycle. In fact, 6S null mutants displayed reduced fitness during long-term stationary phase (63). In *Legionella pneumophila*, a close relative of *C. burnetii*, 6S RNA associates with $E\sigma^{70}$ and is important for intracellular multiplication in human macrophages (64). Moreover, *ssrS* null mutants showed a ten-fold reduction in intracellular growth compared to a wild-type strain, indicating an adaptive role for 6S RNA that may extend to other intracellular pathogens (64). Contrary to what was observed in *E. coli*, the 6S RNA of *L. pneumophila* mainly serves as a positive regulator affecting genes, including those involved in stress adaptation, amino acid metabolism and genes encoding Dot/Icm effectors (type IV secretion system substrates) that are important for survival of *Legionella* in its host (64). Furthermore, when stationary *E. coli* cells encounter new nutrient sources and re-enter growth, 6S RNA must be removed from $E\sigma^{70}$. This has

been shown to occur by the synthesis of pRNA, wherein RNAP acts as a RNA-dependent RNA polymerase by using 6S RNA as the template (65). In *B. subtilis*, formation of pRNA induces structural rearrangements that decrease 6S RNA affinity towards RNAP resulting in its release (66).

C. SPECIFIC AIMS, HYPOTHESES AND SIGNIFICANCE

As a result of *C. burnetii*'s extreme stability in the environment, its ubiquitous nature and typical aerosol route of transmission to humans, Q fever is a potential epidemic in waiting. The recent outbreak in the Netherlands (2007-2010) is clear evidence of the threat that this bacterium poses if sufficient diagnostic and therapeutic tools are not available. Understanding the biology of this pathogen will help us in developing effective strategies to counter this potential bioterrorism agent.

C. burnetii has a large number of parasitic genetic elements for an obligate intracellular bacterium [reviewed in (2)]. This is remarkable when one considers the reduced chances for horizontal gene transfer in the context of the intracellular niche. Moreover, all of these elements are highly conserved in *all* the genotypes of *C. burnetii*. One such genetic element is the intervening sequence (IVS) present within the 23S rRNA gene of *C. burnetii*. Due to the rare occurrence of IVS in bacteria and its conservation among *Coxiella* genotypes, this selfish genetic element might play a role in the pathogen's biology. Since IVSs are found in other bacteria that closely associate with eukaryotic cells, the results of our study may also be applicable to these bacteria.

Despite its remarkable niche and developmental cycle, little is known about *Coxiella*'s modes of regulation and what role non-coding RNAs play in growth and

development. At the start of *Coxiella*'s life cycle, the bacterium transitions from an extracellular environment (dust, milk, air, etc.) to an intracellular environment within its host. During this stage, *Coxiella* encounters various and sudden changes in environmental conditions, including a rapid upshift in temperature (~20^0C to ~37^0C), a change in osmotic pressure, a downshift in pH (pH 7.0 to pH 4.5) and an increase in the concentration of reactive oxygen species. These environmental cues are relevant to sRNA-mediated regulation, as they require a rapid response by the pathogen. Soon after phagocytosis by an alveolar macrophage, the typical initial host cell in humans, the SCV form of *Coxiella* differentiates into a metabolically-active LCV. sRNA-mediated regulation might also be advantageous during the pathogen's transition from one morphotype to another during development.

There are only a few reports to date on the role that regulatory sRNAs play in intracellular pathogens, e.g., *Chlamydia* and *Mycobacterium*, and nothing has been reported for *C. burnetii*. Hence, identifying the sRNAs of *C. burnetii* will help elucidate the regulatory networks of *C. burnetii* and improve our understanding of sRNA-mediated regulation in intracellular pathogens in general. Therefore, based on these observations ***we hypothesize that RNA elements are involved in regulation of C. burnetii's pathogenesis, metabolism and intracellular parasitism.***

Our first aim was to characterize the IVS element of *C. burnetii*. Although *Coxiella*'s IVS was identified two decades ago, its role in the pathogen's biology had not been explored. To this end, our first goal was to investigate the ORF present within the IVS that potentially encodes a S23p protein. Next, we sought to determine if the IVS RNA of *C. burnetii* is excised by RNase III, similar to what is presumed to occur in other bacteria

that have IVS within their 23S rRNA (36-38). The third goal was to determine the intrinsic stability of the IVS RNA, since an increased stability would suggest a possible adaptive role in *Coxiella*'s biology, akin to the two Group I introns encoded in the 23S rDNA (67). Finally, since *C. burnetii* must employ extensive and rapid decay of ribosomes during its dramatic transition from LCV to SCV morphotypes, we determined if IVS-mediated 23S rRNA fragmentation provides a selective advantage to the bacterium during stationary phase, thus establishing a biological function.

Since little is known about how *Coxiella* regulates its biphasic developmental cycle and infection of the host, our second aim was to identify sRNAs that may be involved in regulating these processes. Our first goal was to analyze sRNAs over the course of *Coxiella*'s development and identify sRNAs that are important during infection using RNA-seq. The next goal was to validate the existence of the identified sRNAs.

Preliminary analyses carried out in our lab identified the 6S RNA of *C. burnetii*. Therefore, our third aim was to characterize this sRNA. Since 6S RNA is highly expressed during stationary phase in *E. coli* and during intracellular growth in *L. pneumophila*, our first goal was to determine the expression pattern of 6S RNA as a function of *C. burnetii*'s growth time and culture conditions. Additionally, studies in *E. coli* and other bacteria have shown that 6S RNA associates with σ^{70}-RNAP. Therefore, we elucidated *C. burnetii*'s 6S RNA interaction with RNAP. 6S RNA down-regulates transcription from a subset of σ^{70} promoters while activating certain σ^{S} promoters, based upon work in *E. coli*. However, in *L. pneumophila*, a close relative of *C. burnetii*, it positively regulates several genes enhancing growth and virulence of the pathogen.

Thus, our final goal was to generate 6S-overexpression and 6S-knockdown mutants of *C. burnetii* to identify genes that are regulated by 6S RNA.

D. REFERENCES

1. **Tigertt WD, Benenson AS, Gochenour WS.** 1961. Airborne Q fever. Bacteriol Rev **25:**285-293.
2. **Minnick MF, Raghavan R.** 2011. Genetics of *Coxiella burnetii*: on the path of specialization. Future Microbiol **6:**1297-1314.
3. **Raghavan R, Miller SR, Hicks LD, Minnick MF.** 2007. The unusual 23S rRNA gene of *Coxiella burnetii*: two self-splicing group I introns flank a 34-base-pair exon, and one element lacks the canonical omegaG. J Bacteriol **189:**6572-6579.
4. **Enserink M.** 2010. Infectious diseases. Questions abound in Q-fever explosion in the Netherlands. Science **327:**266-267.
5. **Omsland A, Cockrell DC, Howe D, Fischer ER, Virtaneva K, Sturdevant DE, Porcella SF, Heinzen RA.** 2009. Host cell-free growth of the Q fever bacterium *Coxiella burnetii*. Proc Natl Acad Sci U S A **106:**4430-4434.
6. **Omsland A, Beare PA, Hill J, Cockrell DC, Howe D, Hansen B, Samuel JE, Heinzen RA.** 2011. Isolation from animal tissue and genetic transformation of *Coxiella burnetii* are facilitated by an improved axenic growth medium. Appl Environ Microbiol **77:**3720-3725.
7. **Beare PA, Howe D, Cockrell DC, Omsland A, Hansen B, Heinzen RA.** 2009. Characterization of a *Coxiella burnetii ftsZ* mutant generated by *Himar1* transposon mutagenesis. J Bacteriol **191:**1369-1381.

8. **Chen C, Banga S, Mertens K, Weber MM, Gorbaslieva I, Tan Y, Luo ZQ, Samuel JE.** 2010. Large-scale identification and translocation of type IV secretion substrates by *Coxiella burnetii.* Proc Natl Acad Sci U S A **107:**21755-21760.

9. **Beare PA, Larson CL, Gilk SD, Heinzen RA.** 2012. Two systems for targeted gene deletion in *Coxiella burnetii*. Appl Environ Microbiol **78:**4580-4589.

10. **Fournier PE, Marrie TJ, Raoult D.** 1998. Diagnosis of Q fever. J Clin Microbiol **36:**1823-1834.

11. **Maurin M, Raoult D.** 1999. Q fever. Clin Microbiol Rev **12:**518-553.

12. **Marmion BP, Shannon M, Maddocks I, Storm P, Penttila I.** 1996. Protracted debility and fatigue after acute Q fever. Lancet **347:**977-978.

13. **Marmion BP, Storm PA, Ayres JG, Semendric L, Mathews L, Winslow W, Turra M, Harris RJ.** 2005. Long-term persistence of *Coxiella burnetii* after acute primary Q fever. Qjm **98:**7-20.

14. **Raoult D, Marrie T, Mege J.** 2005. Natural history and pathophysiology of Q fever. Lancet Infect Dis **5:**219-226.

15. **Field PR, Mitchell JL, Santiago A, Dickeson DJ, Chan SW, Ho DW, Murphy AM, Cuzzubbo AJ, Devine PL.** 2000. Comparison of a commercial enzyme-linked immunosorbent assay with immunofluorescence and complement fixation tests for detection of *Coxiella burnetii* (Q fever) immunoglobulin M. J Clin Microbiol **38:**1645-1647.

16. **Raoult D, Houpikian P, Tissot Dupont H, Riss JM, Arditi-Djiane J, Brouqui P.** 1999. Treatment of Q fever endocarditis: comparison of 2 regimens containing doxycycline and ofloxacin or hydroxychloroquine. Arch Intern Med **159:**167-173.

17. **Kermode M, Yong K, Hurley S, Marmion B.** 2003. An economic evaluation of increased uptake in Q fever vaccination among meat and agricultural industry workers following implementation of the National Q Fever Management Program. Aust N Z J Public Health **27:**390-398.

18. **Camacho MT, Outschoorn I, Kovacova E, Tellez A.** 2000. Distribution of immunoglobulin G (IgG) and A (IgA) subclasses following Q fever vaccination with soluble phase I *Coxiella burnetii* extract. Vaccine **18:**1773-1777.

19. **Waag DM, England MJ, Tammariello RF, Byrne WR, Gibbs P, Banfield CM, Pitt ML.** 2002. Comparative efficacy and immunogenicity of Q fever chloroform:methanol residue (CMR) and phase I cellular (Q-Vax) vaccines in cynomolgus monkeys challenged by aerosol. Vaccine **20:**2623-2634.

20. **JC W.** 1991. Infectivity, virulence, and pathogenicity of *Coxiella burnetii* for various hosts, p 21–71. *In* JC W, HA T (ed), Q fever: The Biology of *Coxiella burnetii*. CRC Press, Boca Raton, FL, USA.

21. **TF M, T H, JC. W.** 1981. Ultrastructural and biological aspects of *Coxiella burnetii* under physical disruptions., p 267–280. *In* W B, RL A (ed), *Rickettsiae and Rickettsial Diseases.* Academic Press, NY, USA.

22. **Baca OG, Klassen DA, Aragon AS.** 1993. Entry of *Coxiella burnetii* into host cells. Acta Virol **37:**143-155.

23. **Tujulin E, Macellaro A, Lilliehook B, Norlander L.** 1998. Effect of endocytosis inhibitors on *Coxiella burnetii* interaction with host cells. Acta Virol **42:**125-131.

24. **Beare PA, Gilk SD, Larson CL, Hill J, Stead CM, Omsland A, Cockrell DC, Howe D, Voth DE, Heinzen RA.** 2011. Dot/Icm type IVB secretion system requirements for *Coxiella burnetii* growth in human macrophages. MBio **2:**e00175-00111.

25. **Carey KL, Newton HJ, Luhrmann A, Roy CR.** 2011. The *Coxiella burnetii* Dot/Icm system delivers a unique repertoire of type IV effectors into host cells and is required for intracellular replication. PLoS Pathog **7:**e1002056.

26. **Romano PS, Gutierrez MG, Beron W, Rabinovitch M, Colombo MI.** 2007. The autophagic pathway is actively modulated by phase II *Coxiella burnetii* to efficiently replicate in the host cell. Cell Microbiol **9:**891-909.

27. **Gutierrez MG, Vazquez CL, Munafo DB, Zoppino FC, Beron W, Rabinovitch M, Colombo MI.** 2005. Autophagy induction favours the generation and maturation of the *Coxiella*-replicative vacuoles. Cell Microbiol **7:**981-993.

28. **Howe D, Mallavia LP.** 2000. *Coxiella burnetii* exhibits morphological change and delays phagolysosomal fusion after internalization by J774A.1 cells. Infect Immun **68:**3815-3821.

29. **Howe D, Heinzen RA.** 2005. Replication of *Coxiella burnetii* is inhibited in CHO K-1 cells treated with inhibitors of cholesterol metabolism. Ann N Y Acad Sci **1063:**123-129.

30. **Coleman SA, Fischer ER, Howe D, Mead DJ, Heinzen RA.** 2004. Temporal analysis of *Coxiella burnetii* morphological differentiation. J Bacteriol **186:**7344-7352.

31. **Luhrmann A, Roy CR.** 2007. *Coxiella burnetii* inhibits activation of host cell apoptosis through a mechanism that involves preventing cytochrome c release from mitochondria. Infect Immun **75:**5282-5289.

32. **Voth DE, Howe D, Heinzen RA.** 2007. *Coxiella burnetii* inhibits apoptosis in human THP-1 cells and monkey primary alveolar macrophages. Infect Immun **75:**4263-4271.

33. **Ochman H, Moran NA.** 2001. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. Science **292:**1096-1099.

34. **Raghavan R, Hicks LD, Minnick MF.** 2008. Toxic introns and parasitic intein in *Coxiella burnetii*: legacies of a promiscuous past. J Bacteriol **190:**5934-5943.

35. **Seshadri R, Paulsen IT, Eisen JA, Read TD, Nelson KE, Nelson WC, Ward NL, Tettelin H, Davidsen TM, Beanan MJ, Deboy RT, Daugherty SC, Brinkac LM, Madupu R, Dodson RJ, Khouri HM, Lee KH, Carty HA, Scanlan D, Heinzen RA, Thompson HA, Samuel JE, Fraser CM, Heidelberg JF.** 2003. Complete genome sequence of the Q-fever pathogen *Coxiella burnetii*. Proc Natl Acad Sci U S A **100:**5455-5460.

36. **Burgin AB, Parodos K, Lane DJ, Pace NR.** 1990. The excision of intervening sequences from *Salmonella* 23S ribosomal RNA. Cell **60:**405-414.

37. **Evguenieva-Hackenberg E, Klug G.** 2000. RNase III processing of intervening sequences found in helix 9 of 23S rRNA in the alpha subclass of Proteobacteria. J Bacteriol **182:**4719-4729.

38. **Ralph D, McClelland M.** 1993. Intervening sequence with conserved open reading frame in eubacterial 23S rRNA genes. Proc Natl Acad Sci U S A **90:**6864-6868.

39. **Afseth G, Mo YY, Mallavia LP.** 1995. Characterization of the 23S and 5S rRNA genes of *Coxiella burnetii* and identification of an intervening sequence within the 23S rRNA gene. J Bacteriol **177:**2946-2949.

40. **Lin LY, Ching CL, Chin KH, Chou SH, Chan NL.** 2006. Crystal structure of the conserved hypothetical cytosolic protein Xcc0516 from *Xanthomonas campestris* reveals a novel quaternary structure assembled by five four-helix bundles. Proteins **65:**783-786.

41. **Hsu D, Shih LM, Zee YC.** 1994. Degradation of rRNA in *Salmonella* strains: a novel mechanism to regulate the concentrations of rRNA and ribosomes. J Bacteriol **176:**4761-4765.

42. **Baker BJ, Hugenholtz P, Dawson SC, Banfield JF.** 2003. Extremely acidophilic protists from acid mine drainage host Rickettsiales-lineage endosymbionts that have intervening sequences in their 16S rRNA genes. Appl Environ Microbiol **69:**5512-5518.

43. **Skurnik M, Toivanen P.** 1991. Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The IVSs in

Y. enterocolitica and *Salmonella typhimurium* have a common origin. Mol Microbiol **5:**585-593.

44. **Waters LS, Storz G.** 2009. Regulatory RNAs in bacteria. Cell **136:**615-628.

45. **Grundy FJ, Henkin TM.** 2006. From ribosome to riboswitch: control of gene expression in bacteria by RNA structural rearrangements. Crit Rev Biochem Mol Biol **41:**329-338.

46. **Mehta P, Goyal S, Wingreen NS.** 2008. A quantitative comparison of sRNA-based and protein-based gene regulation. Mol Syst Biol **4:**221.

47. **Mitarai N, Andersson AM, Krishna S, Semsey S, Sneppen K.** 2007. Efficient degradation and expression prioritization with small RNAs. Phys Biol **4:**164-171.

48. **Gripenland J, Netterling S, Loh E, Tiensuu T, Toledo-Arana A, Johansson J.** 2010. RNAs: regulators of bacterial virulence. Nat Rev Microbiol **8:**857-866.

49. **Oliva G, Sahr T, Buchrieser C.** 2015. Small RNAs, 5' UTR elements and RNA-binding proteins in intracellular bacteria: impact on metabolism and virulence. FEMS Microbiol Rev **39:**331-349.

50. **Ortega AD, Quereda JJ, Pucciarelli MG, Garcia-del Portillo F.** 2014. Non-coding RNA regulation in pathogenic bacteria located inside eukaryotic cells. Front Cell Infect Microbiol **4:**162.

51. **Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P.** 2002. An RNA thermosensor controls expression of virulence genes in *Listeria monocytogenes*. Cell **110:**551-561.

52. **Brantl S.** 2007. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol **10:**102-109.

53. **Gonzalo-Asensio J, Ortega AD, Rico-Perez G, Pucciarelli MG, Garcia-Del Portillo F.** 2013. A novel antisense RNA from the *Salmonella* virulence plasmid pSLT expressed by non-growing bacteria inside eukaryotic cells. PLoS One **8:**e77939.

54. **Tattersall J, Rao GV, Runac J, Hackstadt T, Grieshaber SS, Grieshaber NA.** 2012. Translation inhibition of the developmental cycle protein HctA by the small RNA IhtA is conserved across *Chlamydia*. PLoS One **7:**e47439.

55. **Toledo-Arana A, Repoila F, Cossart P.** 2007. Small noncoding RNAs controlling pathogenesis. Curr Opin Microbiol **10:**182-188.

56. **Sorek R, Kunin V, Hugenholtz P.** 2008. CRISPR--a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol **6:**181-186.

57. **Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR.** 2005. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. Rna **11:**774-784.

58. **Wassarman KM, Storz G.** 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell **101:**613-623.

59. **Trotochaud AE, Wassarman KM.** 2005. A highly conserved 6S RNA structure is required for regulation of transcription. Nat Struct Mol Biol **12:**313-319.

60. **Gildehaus N, Neusser T, Wurm R, Wagner R.** 2007. Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding, inhibition of in vitro transcription and synthesis of RNA-directed de novo transcripts. Nucleic Acids Res **35:**1885-1896.

61. **Klocko AD, Wassarman KM.** 2009. 6S RNA binding to Esigma(70) requires a positively charged surface of sigma(70) region 4.2. Mol Microbiol **73:**152-164.

62. **Neusser T, Polen T, Geissen R, Wagner R.** 2010. Depletion of the non-coding regulatory 6S RNA in *E. coli* causes a surprising reduction in the expression of the translation machinery. BMC Genomics **11:**165.

63. **Trotochaud AE, Wassarman KM.** 2004. 6S RNA function enhances long-term cell survival. J Bacteriol **186:**4978-4985.

64. **Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA.** 2010. *Legionella pneumophila* 6S RNA optimizes intracellular multiplication. Proc Natl Acad Sci U S A **107:**7533-7538.

65. **Wassarman KM, Saecker RM.** 2006. Synthesis-mediated release of a small RNA inhibitor of RNA polymerase. Science **314:**1601-1603.

66. **Beckmann BM, Hoch PG, Marz M, Willkomm DK, Salas M, Hartmann RK.** 2012. A pRNA-induced structural rearrangement triggers 6S-1 RNA release from RNA polymerase in *Bacillus subtilis*. Embo j **31:**1727-1738.

67. **Hicks LD, Warrier I, Raghavan R, Minnick MF.** 2011. Ribozyme stability, exon skipping, and a potential role for RNA helicase in group I intron splicing by *Coxiella burnetii*. J Bacteriol **193:**5292-5299.

CHAPTER TWO

The intervening sequence (IVS) of *Coxiella burnetii;* biological role and evolution of the RNA element

A. Abstract

The intervening sequence (IVS) of *Coxiella burnetii*, the agent of Q fever, is a 444-nt element inserted in helix 45 of the precursor 23S rRNA. The IVS element, in turn, contains an open reading frame (ORF) that encodes a hypothetical ribosomal S23 protein (S23p). Although the IVS is highly conserved among different strains of *Coxiella*, the region immediately upstream of the start codon is prone to change and the S23p-encoding ORF is apparently undergoing reductive evolution. Moreover, the S23p could not be expressed *in vitro*. Taken as a whole, data suggest that the RNA component of the IVS is more biologically significant than the S23p-encoding ORF. We observed that IVS excision was mediated by RNase III, and the cleavage sites were mapped to the middle of the stem of the predicted secondary structure of the element. Soon after excision from 23S rRNA, IVS RNA was found to be degraded. Levels of the two resulting 23S rRNA fragments (~1.2 kb and ~1.7 kb) were quantified in axenically-cultured *Coxiella* and found to significantly differ from those of the 16S rRNA originally derived from the same transcript. Specifically, quantities of the ~1.2-kb fragment (F1) were higher than 16S rRNA while those of the ~1.7-kb fragment (F2) were significantly lower. Thus, F2 may limit the number of mature ribosomes that are ultimately formed *in vivo*. We therefore hypothesize that the biological role of IVS is to modulate *Coxiella*'s growth by fragmentation of the 23S rRNA. We further hypothesize that IVS was initially

acquired through horizontal transfer and subsequently maintained by all strains of *Coxiella* through vertical transfer due to its adaptive role in fostering *Coxiella*'s tendency towards slow growth and chronic infection.

B. Importance

Intervening sequences (IVSs) are elements that occur in the 23S rRNA genes of many bacteria. Removal of these elements following transcription, results in fragmentation of the mature 23S rRNA. Although IVS elements were identified more than two decades ago and occur in a wide range of bacterial species, their biological function(s) remains elusive. Here, we report that, IVS-mediated fragmentation of 23S rRNA *in C. burnetii* results in two fragments that are differentially regulated, with one of them being significantly lower in concentration. We hypothesize that this event contributes to the markedly slow growth rate of *C. burnetii* and fosters its tendency for chronic infection of the host.

C. Introduction

Coxiella burnetii, the etiological agent of Q (query) fever, is an obligate intracellular pathogen that replicates in a parasitophorous vacuole (PV) (2). Human Q fever is a zoonotic disease that is generally acquired through inhalation of contaminated aerosols and is characterized by a flu-like illness accompanied by pneumonia and hepatitis (3). The bacterium undergoes a biphasic developmental cycle consisting of two cellular forms, including an infectious, dormant, small-cell variant (SCV) and a metabolically-active, large-cell variant (LCV). In spite of its intracellular nature, where chances for horizontal gene transfer are minimal, the genome of *C. burnetii* contains 31 insertion

sequence (IS) elements, a putative intein in the C-terminal region of the replicative DNA helicase (DnaB) protein and two self-splicing group I introns (Cbu_L1917 and Cbu_L1951) along with an intervening sequence (IVS) within the 23S rRNA gene (4).

IVSs are parasitic genetic elements that are found in 16S and 23S rRNA genes and are transcribed as part of the rRNA precursor. Excision of these elements occurs post-transcriptionally without ligation of the resulting fragments, which are held together by secondary and tertiary structures, thereby maintaining functional integrity of the 50S ribosome (5). IVSs have been identified in several bacteria, including *Salmonella* (6), *Leptospira* (5), *Yersinia* (7), *Campylobacter* (8), *Proteus* and *Providencia* (9). Within these bacteria, IVSs occur only in certain isolates, and in a given bacterial strain not all *rrn* operons contain the IVS (6, 9-11). This sporadic distribution of IVS suggests the possibility of horizontal transfer of the element among eubacteria. Although the distribution of IVS is random, its location within the 23S rRNA is relatively well-conserved. The most common sites for insertion of IVS include helices 9, 25 and 45 of the postulated secondary structure of *E. coli* 23S rRNA (12). IVS in both helix 25 and 45 are observed in *Salmonella* (13), *Helicobacter* (14), *Proteus* and *Providencia* (9). On the other hand, a single copy of IVS is found in helix 45 of *Leptospira* (5), *Yersinia* (7), *Campylobacter* (8) and in helix 9 of various alpha-Proteobacteria (15). Each of these helices originally consisted of a small tetraloop that was replaced by the extended stem loop structure of IVS. The conservation of the IVS insertion sites across species indicates that continuity of the 23S rRNA at these positions is probably not necessary for ribosomal function.

The IVS of *C. burnetii* is a 444-nt element that is inserted in the single-copy 23S rRNA gene (1). The element is bordered by complementary sequences that can form a stable stem-loop structure. IVSs of *Salmonella*, *Leptospira* and many species of alpha-Proteobacteria are known to be cleaved from their respective 23S rRNA by RNase III (5, 6, 15), an endoribonuclease that cleaves pre-23S and pre-16S rRNA from the 30S precursor during rRNA maturation (16). Therefore, we initially hypothesized that *Coxiella*'s IVS was also excised from its 23S rRNA precursor by RNase III. Additionally, IVS of *C. burnetii* contains an ORF that potentially encodes a ribosomal S23p family of hypothetical proteins. Although these proteins have no known function (4), the crystal structure of the S23p of *Xanthomonas campestris* has been solved and found to be a homopentamer comprised of four-helix bundles creating a toroidal (doughnut-shaped) structure (17). In addition to *Xanthomonas,* S23p orthologs are also encoded by IVS ORFs of *Leptospira* and *Brucella* spp. However, previous reports suggested that translation of the protein in *C. burnetii* may be inefficient since no conserved Shine-Dalgarno sequence (SD) occurs upstream of the start site of the respective ORF (1).

At this point, very little is known about the biological role of IVS in bacteria that possesses this element. The occurrence of IVS across bacterial species and its conservation in all strains of *C. burnetii* suggest that it might play an adaptive role. Interestingly, IVS elements appear to be most common in bacteria that form close associations with eukaryotes, possibly promoting communication between bacteria and their hosts (18). Such a role is conceivable in *Coxiella*, especially in view of its intracellular niche. A second possible role of IVS is in regulating the pathogenesis of *C. burnetii*. This possibility is supported by an observation in *Yersinia*, where all pathogenic

strains of *Y. enterocolitica* possessed IVS while non-pathogenic strains did not (7). A third possible role for IVS is in regulating the rate of rRNA degradation during *C. burnetii*'s development cycle. Studies in *Salmonella* have shown that the rate of degradation of 23S rRNA during stationary phase is directly related to the degree of IVS-mediated fragmentation of the 23S rRNA (19). Since *C. burnetii* transitions to an endospore-like SCV near the end of its life-cycle, this adaptation would clearly be advantageous.

The aim of this study was to fully characterize the IVS element of *C. burnetii* and to determine its role in the biology of the bacterium. During our investigation, we discovered that the IVS ORF of *C. burnetii* is undergoing reductive evolution with variation in the sequences immediately upstream of the start codon and deletional mutagenesis, in different strains of *C. burnetii*. We also discovered that, similar to other bacteria with IVSs in their 23S rRNA, IVS of *C. burnetii* is also excised by RNase III. Moreover, consistent with previous findings, *Coxiella*'s IVS RNA was found to be labile following excision from 23S rRNA. Together, these observations led us to hypothesize that a possible biological function of *C. burnetii*'s IVS is the physical fragmentation of 23S rRNA. Differential levels of the two resulting fragments of 23S rRNA, following IVS excision, may limit the number of mature ribosomes that are ultimately formed, thus ensuring a slow growth rate for this bacterium.

D. Materials and Methods

Axenic cultivation of *C. burnetii*. Cultures of *C. burnetii* Nine Mile phase II (RSA439; clone 4) were grown in ACCM-2 medium (20) using 0.2-µm-pore size filter-capped 500-

ml Erlenmeyer flasks containing 80 ml of medium. *C. burnetii* was inoculated at a concentration of 1.6 x 10^6 genome equivalents per ml (GE/ml). Cultures were incubated at 37^0C in a tri-gas incubator (2.5% O_2, 5% CO_2, 92.5% N_2) with continuous shaking at 75 RPM (20). Following the first week of growth in a tri-gas incubator, flasks were capped and moved to room temperature for generation of SCVs, as previously described (21).

Growth assays of *E. coli* expressing *Coxiella*'s IVS. This assay was performed essentially as previously described (22). Briefly, the IVS element and ~400 bp of flanking sequences were PCR amplified by standard protocol using IVSflank_For and _Rev primers (Table 2-1) and cloned into pCR2.1-TOPO per manufacturer's instructions (Invitrogen) to generate pIVS1. As a control, the IVS was cloned in opposite orientation relative to the vector's *lac* promoter to create pIVS2. *E. coli* (TOP10F′) was transformed with pIVS1 or pIVS2 as instructed by the manufacturer (Invitrogen). Resulting *E. coli* strains were grown overnight at 37^0C in Luria-Bertani (LB) broth containing 100 µg/ml ampicillin (LBamp). These cultures were used to inoculate fresh LBamp at a 1:10 (v/v) dilution, and the mixture was grown to mid-logarithmic phase (2 h) at 37^0C. Isopropyl-beta-D-thiogalactopyranoside (IPTG) was added to 1 mM and growth was assayed spectrophotometrically at 600 nm at hourly intervals for 7 h.

***In vitro* coupled transcription-translation (IVTT) of the IVS ORF.** Translation of the IVS ORF was attempted using an EasyXpress Protein Synthesis kit as instructed by the manufacturer (Qiagen). Templates included a PCR product containing the IVS ORF with an engineered T7 promoter (Table 2-1; IVS_For+T7 and IVS_Rev) or the same PCR product directionally cloned into pCR2.1-TOPO to generate pIVS3. The FluroTech™

Table 2-1. Primers and their respective targets in *C. burnetii

Target	Designation	Sequence
IVS	IVSflank_For+T7	TAATACGACTCACTATAGGGATAGCTGGTTCTCCTCG
	IVSflank_For	GATAGCTGGTTCTCCTCG
	IVSflank_Rev	CTTTTCCTGGAAGCGTGG
	IVS_For+T7	TAATACGACTCACTATAGGTAGCTCACTGGTCGAGTCG
	IVS_Rev	TGTCAGCATTCGCACTTC
	IVS_RACE_GSP1	GATTACGCCAAGCTTACCACACACGCATCTCATCTGCCGAAC
	IVS_RACE_GSP2	GATTACGCCAAGCTTTGGATTGGCAAGCCAAATCCGTCAAGCAAG
	IVS_RACE_NGSP1	GATTACGCCAAGCTTGAAACACTCTGCTTTCCAAACCCTTCAGC
	IVSprobe_F	TTTTTTTTTTTTTCATCAGAAGACAGATGAC
	IVSprobe_R+T7	TAATACGACTCACTATAGG TTTTTTTTTTTTTCATCAGAAGACAGATGGC
	IVS_qPCR_For	ATTGCAATGGGTTCGGCAGATGAG
	IVS_qPCR_Rev	ACAGATGGCAGAAGACTGAGGACA
RNase III	CbuRNaseIII_For	TTTTGGATCCAACCATCTTAACAAGTTA
	CbuRNaseIII_Rev	TTTTCCCGGGTCATTGGTCCCGCTCCGT
CBU_16S	16S_qPCR_For	TTCGGGAACCGAGTGACAGGTG
	16S_qPCR_Rev	TCGCTGGCAACTAAGGACGAGG
CBU_23S	23S Frag1_qPCR_For	AACACTGACTGGAGGCCCGAAC
	23S Frag1_qPCR_Rev	AGCCGAAACAGTGCTCTACCCC
	23S Frag2_qPCR_For	GGTGCTTGACTGCGAGAC
	23S Frag2_qPCR_Rev	GGTGGTATCAGCCTGTTATCC

*The T7 promoter sequence is underlined.

Green$_{Lys}$ *in vitro* Translation Labeling system (Promega) was included in the reaction for enhanced detection of translated proteins. Reactions were analyzed on 15% acrylamide (w/v) SDS-PAGE gels and visualized using a FLA 3000G Fluorescent Imager (Fujifilm).

RNase III assay. The *C. burnetii* RNase III gene (*rnc*, Cbu_1503; Primers: CbuRNaseIII_For and _Rev; Table 2-1) was cloned and purified as previously described for the RNA helicase (*rhlE*) gene (23). An RNase III assay substrate was prepared by T7 promoter-mediated *in vitro* transcription of a PCR product (from PCR primers IVSflank_For+T7 and IVSflank_Rev; Table 2-1) using a MAXIscript kit (Ambion), as instructed. The resulting RNA was electrophoresed in a 4% acrylamide (w/v) - 8M urea gel, and the substrate RNA was excised, eluted overnight at 37^{0}C into a buffer containing 0.5 M ammonium acetate, 1 mM EDTA, and 0.2% sodium dodecyl sulfate and then precipitated with 100% ethanol. RNase III assays were performed for 30 min at 37^{0}C using 200 nM substrate RNA in RNase assay buffer (50 mM NaCl, 10 mM Tris, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT, 450 mM KCl) in the presence of recombinant *C. burnetii* RNase III (30 µM). Positive controls were done using *E. coli* RNase III (Ambion) as per manufacturer's instructions. Products of the RNase III assays were analyzed on 4% acrylamide (w/v) - 8M urea gels.

Identification of 5′ and 3′ ends of IVS. Ends of the IVS RNA were determined by 5′ and 3′ Rapid Amplification of cDNA Ends (RACE) using a SMARTer® RACE 5'/3' kit (Clontech) as instructed. RNase III assay products were analyzed on 4% acrylamide (w/v) - 8M urea gels from which IVS was purified, before resuspending in RNase-free water. Prior to 3′ RACE, IVS RNA was polyadenylated at the 3′ end with *E. coli* Poly(A) polymerase as instructed (NEB). cDNA was generated from IVS RNA before

amplification by PCR using gene-specific primers (Table 2-1; IVS_RACE_GSP1, IVS_RACE_GSP2) and universal primers. Further, a second nested PCR was performed on the 5' RACE primary PCR product using IVS_RACE_NGSP1 (Table 2-1) and a universal short primer for increased resolution. PCR products were visualized on 2% agarose (w/v) gels, purified with a PCR clean-up and gel extraction kit as instructed (Clontech) and cloned into pRACE with an In-Fusion HD Cloning kit per manufacturer's instructions (Clontech). The plasmid content of individual clones was analyzed by automated Sanger sequencing at the Genomics Core Facility at the University of Montana.

RNase Protection Assay (RPA). The half-life of IVS RNA was determined using a RPA. *E. coli* TOP10F' (pIVS1) was cultured in LBamp for 16 h at 37^0C and then used to inoculate LBamp at a 1:20 dilution. Following growth for 2 h at 37^0C, IPTG was added to 1 mM and cultures were allowed to grow an additional 2 h. Rifampin was added to a final concentration of 160 µg/ml and 1-ml samples were taken at 5-min intervals for 35 min. RNA was purified using TRI Reagent (Ambion) and processed using a RNase Protection Assay (RPA) III kit (Ambion) (3 µg total RNA; 800pg IVS probe), as previously described (23). Biotinylated IVS RPA probes were prepared using a MEGAScript kit, biotin-labeled UTP (Bio 16-UTP; Ambion), and corresponding IVS primers (Table 2-1; IVSprobe_F and IVSprobe_R+T7). Relative levels of 23S and 16S rRNAs in *C. burnetii* over the course of its life cycle were also determined using a RPA. *C. burnetii* was cultured in ACCM-2, and DNA and RNA were isolated from the same flask at 3, 5, 7, 10, 12, 14, 17 and 21 d using a DNeasy Blood and Tissue kit (Qiagen) and Ribopure kit (Amibion), respectively, as per manufacturer's instructions. Resulting

RNA was treated with TURBO DNase (Ambion), precipitated with 100% ethanol and quantified by spectrophotometry. Biotinylated RNA probes for 16S, 23S fragment 1 (F1) and 23S fragment 2 (F2) were prepared as above, using corresponding primers (Table 2-1; 16S_qPCR_For and _Rev, 23S Frag1_qPCR_For and _Rev, 23S Frag2_qPCR_For and _Rev). RNA (250 ng) and corresponding probes (25 pg) were processed using RNase Protection Assay (RPA) III kit (Ambion), as previously described (23).

Quantitative real-time PCR (qRT-PCR). DNA and RNA were isolated from the same cultures at 0, 1, 7, 10, 13, 16, 19 and 22 d as described in 'RNase Protection Assay (RPA)' and 20 ng of RNA from each growth time-point was converted to cDNA using an iScript cDNA synthesis kit (Bio-Rad), as instructed by the manufacturer. Genome numbers were determined by quantitative PCR (qPCR) with a primer set specific to the *C. burnetii rpoS* gene (24). cDNA (diluted 100-fold) was used to perform quantitative real-time PCR (qRT-PCR) with corresponding primer sets (Table 2-1), as previously described (22). Amplified cDNA was normalized to respective genome numbers.

E. Results and Discussion

Evolution of the IVS ORF

Earlier attempts by our lab to clone the IVS of *C. burnetii* Nine Mile (NM) phase I (strain RSA493) in *E. coli* only generated clones whose insert was in opposite orientation to the vector's *lac* promoter, implying that the element was toxic to *E. coli*. However, a previous report showed that *E. coli* expressing the IVS of *S. typhimurium* did not display any growth defects (25). To clarify this disparity, the IVS of *C. burnetii*

NM phase II (strain RSA439) with ~400 flanking bases was cloned into pCR2.1-TOPO in both orientations relative to the vector's *lac* promoter to produce pIVS1 and pIVS2, respectively. *E. coli* strains harboring pIVS1 or pIVS2 were induced with IPTG and their growth rates monitored spectrophotometrically for 7 h. Contrary to what we previously observed, the growth rate of *E. coli* (pIVS1) was not significantly different from *E. coli* (pIVS2) or *E. coli* containing the pCR2.1-TOPO cloning vector alone (Fig. 2-1). These results suggested that the IVS elements of the two *C. burnetii* strains were distinct, with that of RSA493 possibly toxic and RSA439 innocuous to *E. coli*.

Interestingly, the *C. burnetii* IVS contains an ORF that encodes a 14.2-kDa protein that belongs to the ribosomal_S23 family of hypothetical proteins (S23p). Although IVS of *C. burnetii* was not found to be homologous to IVS of other bacteria, phylogenetic analyses performed using the predicted amino acid sequence of *Coxiella*'s S23p, showed that the protein is most closely related to a homolog encoded in the IVS of *Haemophilus* sp. with 71% identity (Fig. 2-2). As previously reported, a number of *C. burnetii* S23p homologs are also found in pathogenic *Leptospira* (1). These results prompted us to determine whether the *C. burnetii* S23p protein could be expressed *in vitro* using a coupled transcription-translation system from an *E. coli* S30 extract. However, we were unable to detect an insert-specific protein of ~14.2 kDa by this approach, indicating that the protein is not expressed *in vitro* (data not shown). These results were not surprising, since the absence of a conserved SD site upstream of the ORF was previously noted (1). In agreement with our data, the S23p protein was not identified in proteomic analyses of *C. burnetii* cell lysates and secreted molecules, implying that this protein is probably not expressed *in vivo* as well (26, 27).

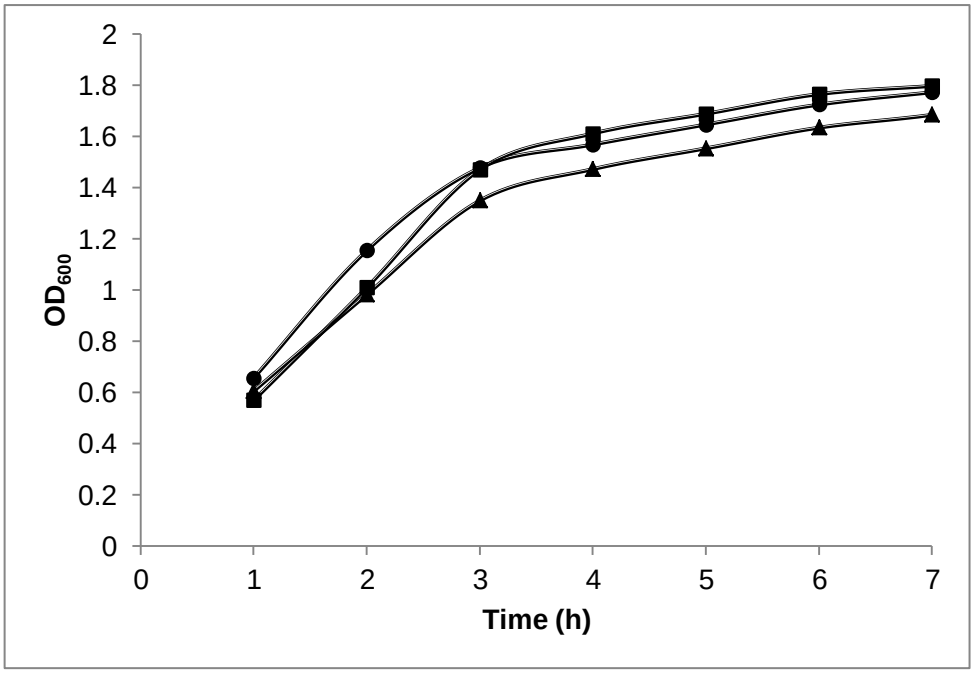


Figure 2-1. Effect of the *C. burnetii* (RSA439) IVS on *E. coli's* growth. *E. coli* (TOP10F′) expressing: IVS (square), control RNA (circle) or the empty vector (pCR2.1-TOPO; triangle) were induced with IPTG (1 mM) and assayed spectrophotometrically at 600 nm for growth at 37^{0}C. A typical representative of two assays is shown.

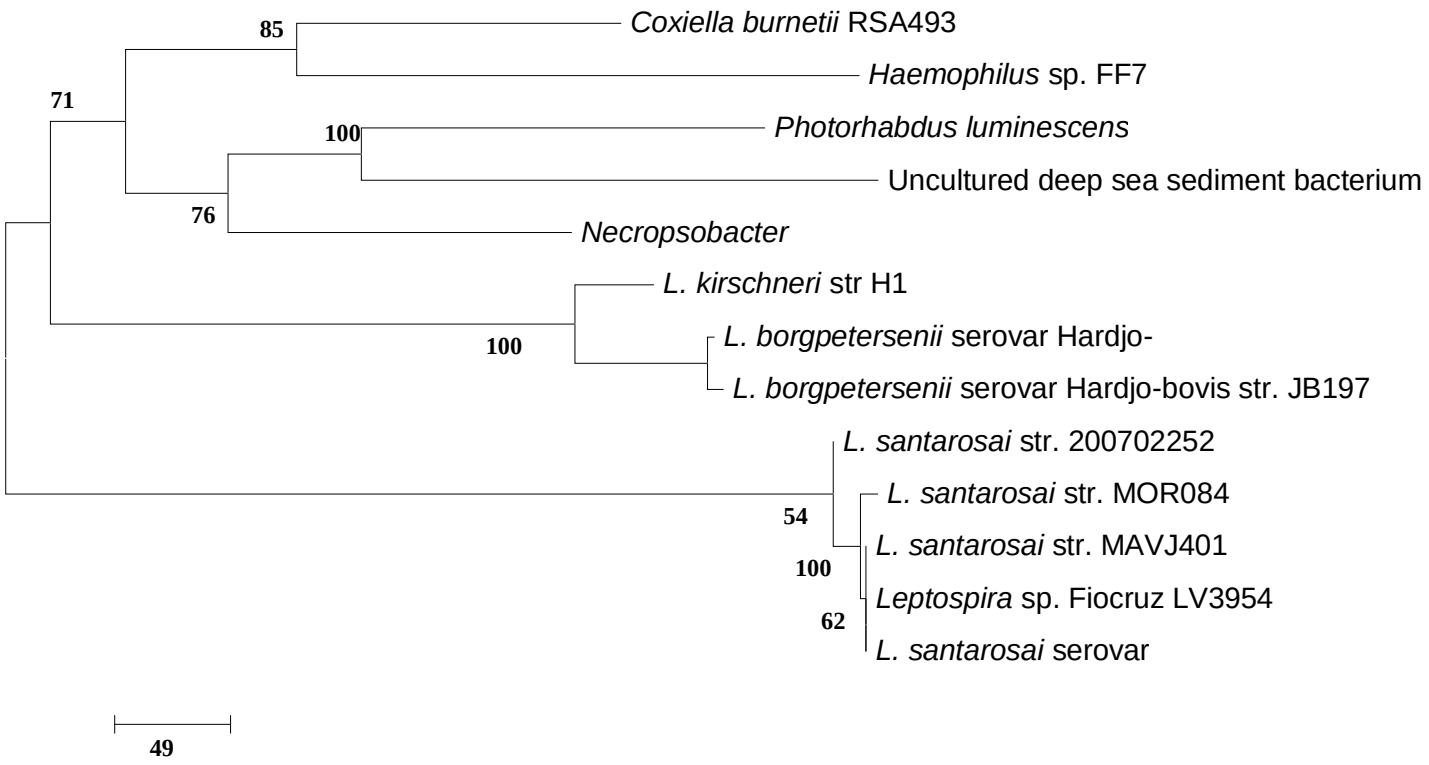


Figure 2-2. Phylogenetic analysis of the twelve top BLASTp hits to *C. burnetii*'s S23p protein. Neighbor-joining trees were built from predicted amino acid sequences of IVS ORF-encoded S23p superfamily proteins using MEGA5. Bootstrap values (1,000 replicates) are indicated at the nodes. Amino acid sequence identities ranged from 54 to 71% with *C. burnetii*'s S23p protein.

In silico analysis of the IVS regions of approximately 80 *C. burnetii* isolates was done to investigate the cloning disparity we observed when IVS elements were cloned from strains RSA493 and RSA439. Interestingly, results showed that some *C. burnetii* strains possess deletions in the sequence immediately upstream of the S23p ORF. As a rule, the 5′ region of the IVS consisted of a potential SD sequence of AGAAGA followed by three, tandem heptameric repeats (GACAGAT) upstream of the start codon of the S23p ORF (Fig. 2-3). There is an additional heptameric repeat with a degenerate nucleotide at its 3′ end (GACAGAA) prior to the predicted ATG of the S23p ORF (Fig. 2-3). Upon analysis of different strains of *C. burnetii*, a disparity in the number of heptameric repeats was observed. In Dugway, RSA331 and Q212 strains, three sets of heptameric repeat units were found, whereas in strains RSA493, Q154 and Q177 only two sets of the repeats were present. This disparity was also found to exist between RSA493 and RSA439, wherein RSA439 has three sets of heptameric repeat units (not shown) while RSA493 has only two sets (Fig. 2-3). This variation may explain why *E. coli* clones prepared from the IVS elements of the two strains behaved so differently. Specifically, those prepared from RSA439 (with 3 repeats) were non-toxic to *E. coli* and did not affect growth (Fig. 2-1). Apart from the SD region for this locus being poor, the presence of an extra repeat unit would move the start codon away from the putative AG-rich SD region, thus reducing the chance of translation of the ORF. Interestingly, another *C. burnetii* strain (3257) was found to have a large, 176-bp deletion its S23p ORF (not shown), implying that the protein-coding segment of the IVS is undergoing reductive evolution and is expendable. Taken together, these observations lead us to conclude that irrespective of the number of repeat units, the S23p ORF is probably not expressed

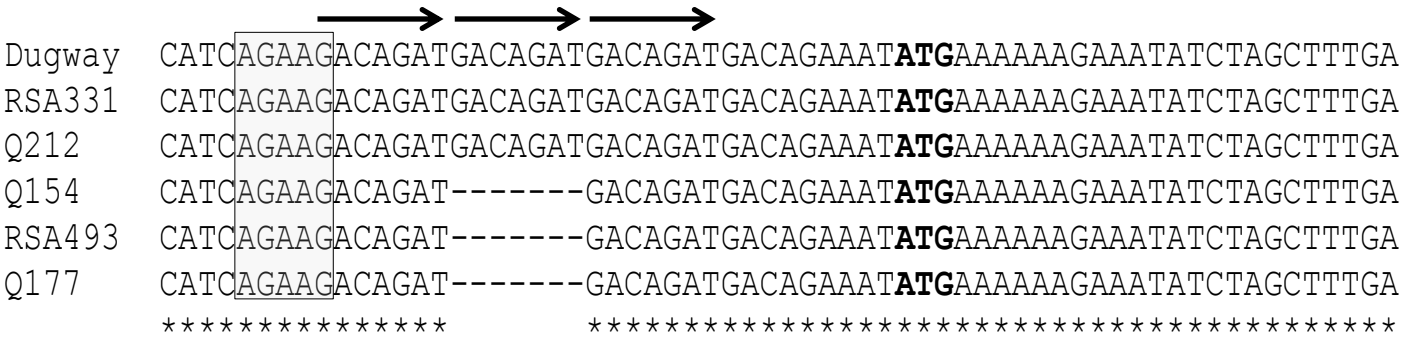


Figure 2-3. Multiple sequence alignment of the 5′ region of *C. burnetii*'s IVS. Corresponding sequences of selected *C. burnetii* strains (shown on the left) were aligned using ClustalX. Bases in grey represent a potential SD region. Arrows indicate a direct heptameric repeat, while a dashed line shows deletion of one of the repeats. ATG (in bold) represents the start codon of the S23p ORF, as predicted by Afseth et al. (1)

in vivo. Even though the IVS element is highly conserved among different strains of *C. burnetii* with uniform occurrence and 99% sequence identity, the region immediately upstream of the start codon is prone to minor changes and the protein S23p ORF is undergoing reductive evolution. These results lead us to hypothesize that the RNA component of the IVS is more biologically significant than the S23p protein.

***In vitro* processing of *C. burnetii* IVS by RNase III.**

The IVS of *C. burnetii* was first described by Afesth et al. as a 444-nt element that is excised to yield a mature, fragmented yet functional 23S rRNA (1). The IVS element is initially found inserted into helix 45 of *C. burnetii*'s precursor 23S rRNA. IVS RNA of *C. burnetii* (RSA493) forms a stable stem-loop secondary structure (ΔG= -144.36 kCal/mol) due to the presence of complementary sequences at its termini (Fig. 2-4). The terminal inverted repeats (IRs) form a 28-bp stem with complete complementarity except for a single mismatch and a bulge. The bulge results from a missing heptameric tandem repeat in this strain (Fig. 2-4). IVS elements of other bacteria are also associated with 23S rRNA genes and are excised by RNase III (5, 6, 15), an endoribonuclease that cleaves rRNA precursors during their maturation (28). To determine if this ribonuclease is responsible for excision of IVS from the 23S rRNA precursor of *C. burnetii*, we performed an *in vitro* assay using recombinant *C. burnetii* RNase III. A T7-transcribed *in vitro* RNase III substrate RNA containing IVS and ~400 nt of flanking sequences was used in the assay with either recombinant *C. burnetii* RNase III or commercial *E. coli* RNase III. Since RNase III cleavage is nonspecific at low ionic strength (29), we tested cleavage efficiencies at several ionic strengths (data not shown) and a final ionic concentration of 0.5 M (50 mM NaCl plus 450 mM KCl) was used in the reaction. When

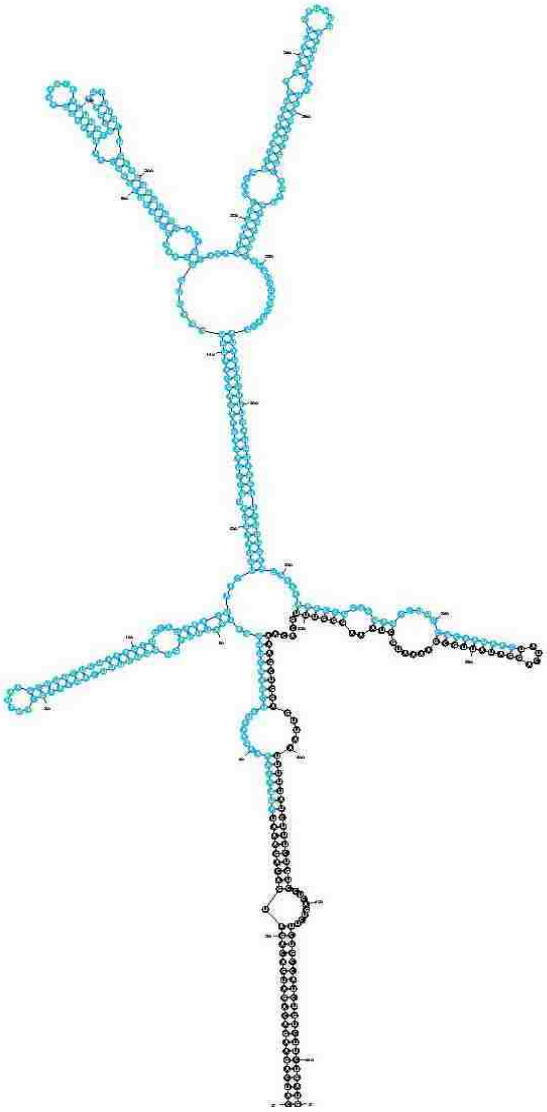


Figure 2-4. Predicted secondary structure of the *C. burnetii* IVS using Mfold. (ΔG = -144.36 kCal/mol). An ORF (CBU_2096) encoding a potential S23p protein is shown in blue.

products of the RNase III assay were analyzed on a polyacrylamide gel, three distinct RNA bands corresponding to the IVS (444 nt) and its flanking sequences (393 and 364 nts) were observed (Fig. 2-5). These results indicate that *C. burnetii* RNase III is capable of cleaving IVS *in vitro* and most probably does so *in vivo*. Identical RNA fragments were observed in reactions containing *E. coli* RNase III but were absent from the negative control (Fig. 2-5).

Characterization of the 5′ and 3′ ends of IVS.

Afesth et al. characterized the IVS by identifying the sequence that was not a portion of the 23S rRNA (1). To determine the exact ends of the IVS element, we performed 5′ and 3′ Rapid Amplification of cDNA Ends (RACE) analysis on IVS RNA isolated from the RNase III assay products. At least 7 clones were isolated from each reaction and sequenced. The sequence of the 5′ and 3′ ends of IVS RNA are shown with black dots indicating experimentally determined locations of the RNase III cleavage sites (Fig. 2-6A). While the 3′ end was uniform, the 5′ end showed a preferred cleavage site plus three minor, additional sites. When the sites of cleavage were mapped on the predicted secondary structure of IVS, it was found to be in the middle of the stem structure at about 9 bases from the 5′ end and about 15 bases from the 3′ end of the predicted IVS RNA sequence (Fig. 2-6B). Following cleavage, the remaining stem portion of the IVS could either be retained in the mature 23S rRNA that holds the rRNA pieces together, as observed in *Leptospira* (5, 8), or it could be trimmed to the correct size by exonucleases. One example of such a ribonuclease is RNase T, which plays a role in 3′ maturation of *E. coli*'s 23S rRNA and is also present in *C. burnetii* (16). However, an exoribonuclease involved in maturation of 5′ ends has not yet been identified (16).

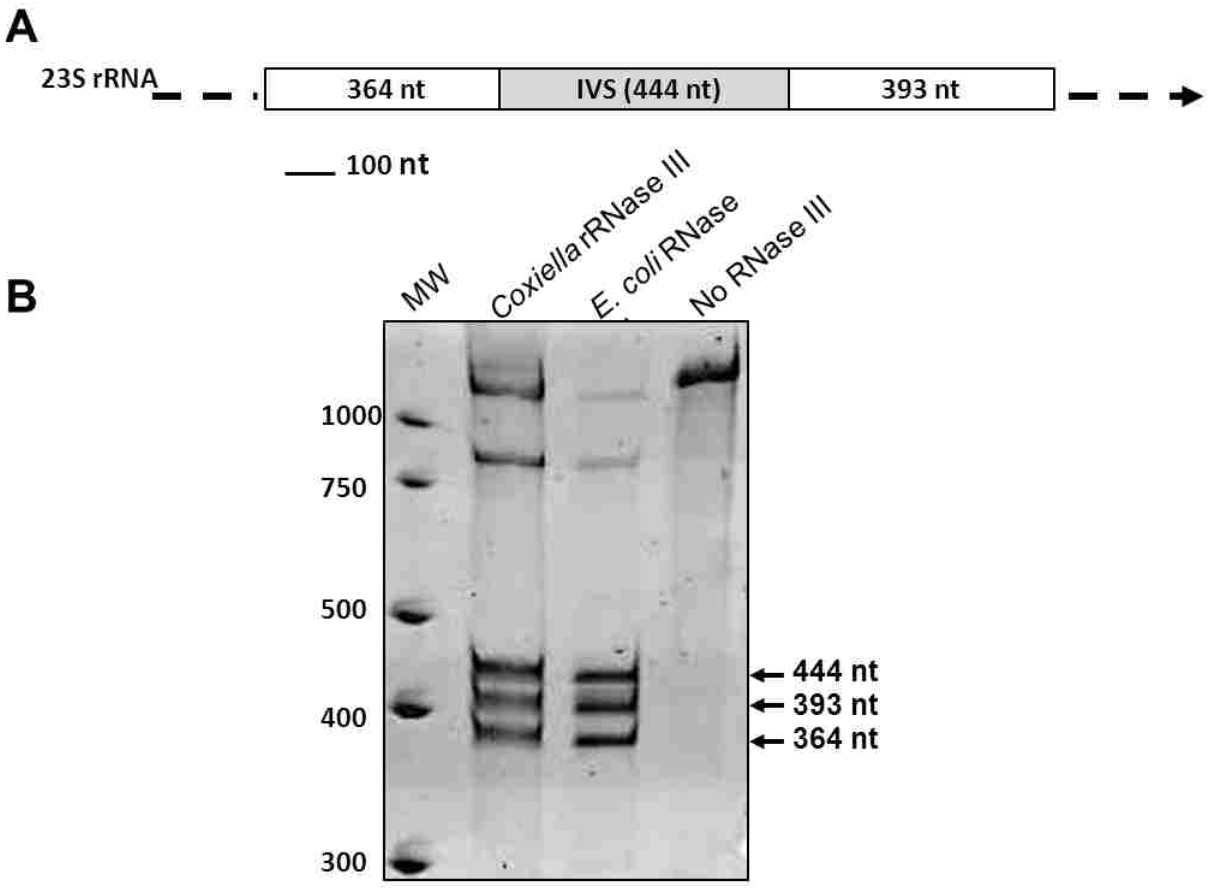


Figure 2-5. *In vitro* processing of the *C. burnetii* IVS by RNase III. (A) Map of the RNA substrate used in RNase III assays, consisting of a segment of 23S rRNA with the IVS and proximal flanking sequences (sizes shown). (B) Acridine orange-stained polyacrylamide gel [4% acrylamide (w/v)- 8M urea] showing cleavage products of substrate RNA following treatment with recombinant *C. burnetii* RNase III. Three discrete fragments of sizes corresponding to the substrate RNA (arrowed) were observed. Positive [*E. coli* RNase III (Ambion)] and negative (no RNase III) control reactions are shown for comparison.

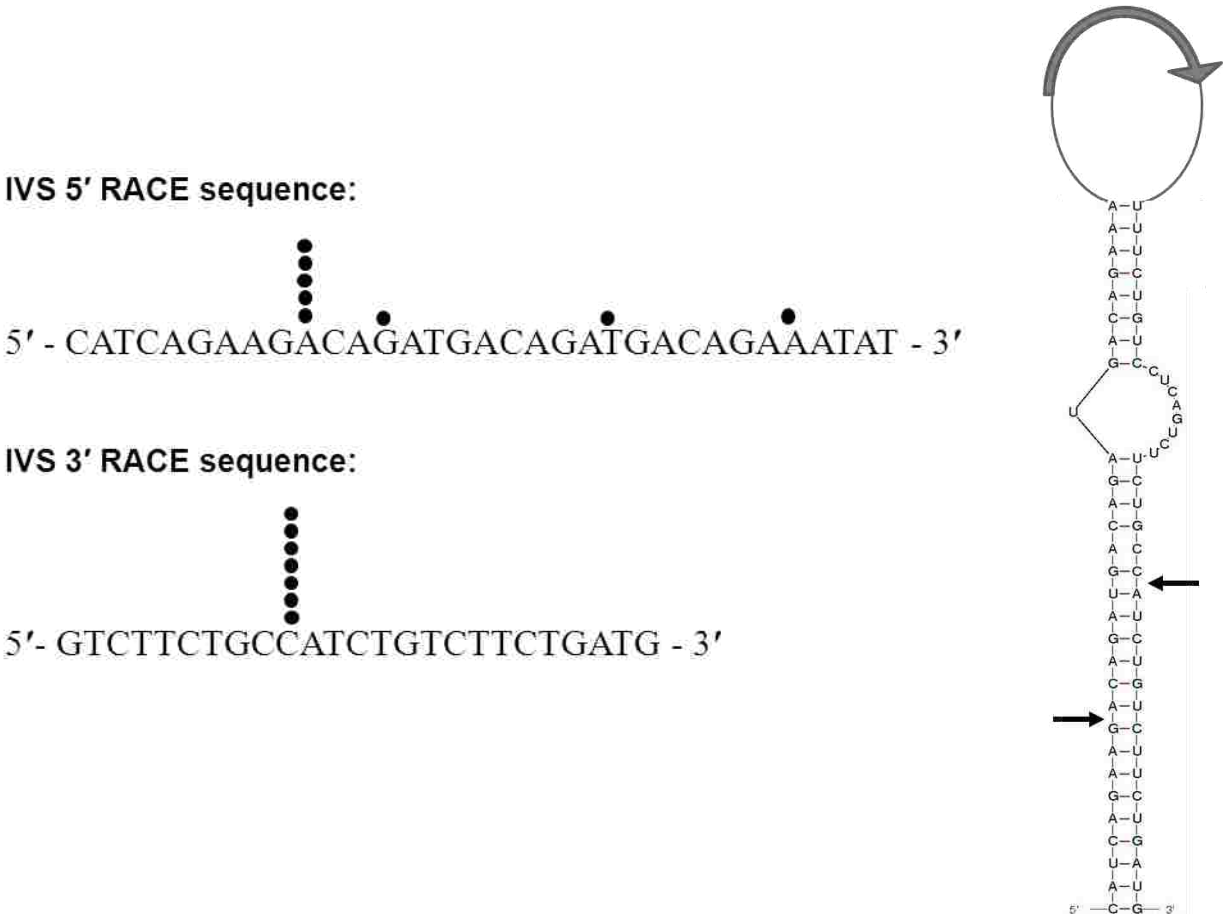


Figure 2-6. Characterization of *C. burnetii*'s IVS termini by 5′- and 3′-RACE analysis. (A) Sequence of the 5′ and 3′ ends of the IVS are shown. Each black dot indicates the last base in an individual sequence analysis of a cloned RACE PCR product. (B) Schematic representation of RNase III processing sites on the stem of the IVS RNA of *C. burnetii* (RSA493). Black arrows depict the processing sites as predicted by 5′- and 3′-RACE analyses. The relative location of the S23 ORF (CBU_2096) is indicated by a grey arrow.

***In vivo* stability of IVS RNA.**

Previous work by our laboratory showed that *Coxiella*'s introns, Cbu.L1917 and Cbu.L1951, are spliced from the precursor 23S rRNA and demonstrate remarkable stability (~11 min) in an *E. coli* background, suggesting a possible adaptive role in *Coxiella*'s biology (23). Since the IVS is uniformly present and highly conserved among all strains of *C. burnetii*, an adaptive role might also be attributed to IVS RNA. To determine the intrinsic stability of the IVS RNA, we performed an *in vivo* stability assay in an *E. coli* model. *E. coli* (pIVS1) was cultured, expression induced by IPTG, and cells sampled after transcription was stopped with rifampicin. RPAs performed on RNAs isolated from these samples showed the half-life of IVS RNA to be ~4 min (Fig. 2-7). Considering that the global average half-life of mRNA in *E. coli* is ~5.7 min (30), these results suggest that IVS RNA doesn't display extraordinary intrinsic stability like *Coxiella*'s intron RNAs. However, since the RNase activities of *E. coli* may be different from that of *C. burnetii*, we also quantified levels of IVS RNA in *C. burnetii* to determine its stability *in vivo*. *C. burnetii* was cultured in axenic medium for 22 d and IVS RNA levels were measured by performing qRT-PCR on total RNAs purified from the bacterium at 72-h intervals, using primers specific to the IVS (Table 2-1). qPCR was also performed on genomic DNA isolated at the same time-points to calculate the amount of IVS RNA on a per-genome basis. Results showed that IVS RNA was highest at 1 d (Fig. 2-8), or the beginning of the exponential phase of *Coxiella* (20). Since rRNAs are transcribed during this phase and mature 23S rRNA is formed following excision of IVS, the increased level correlates with increased synthesis of 23S rRNA. As the bacterium reached stationary phase (~4 d) (20), levels of IVS significantly declined

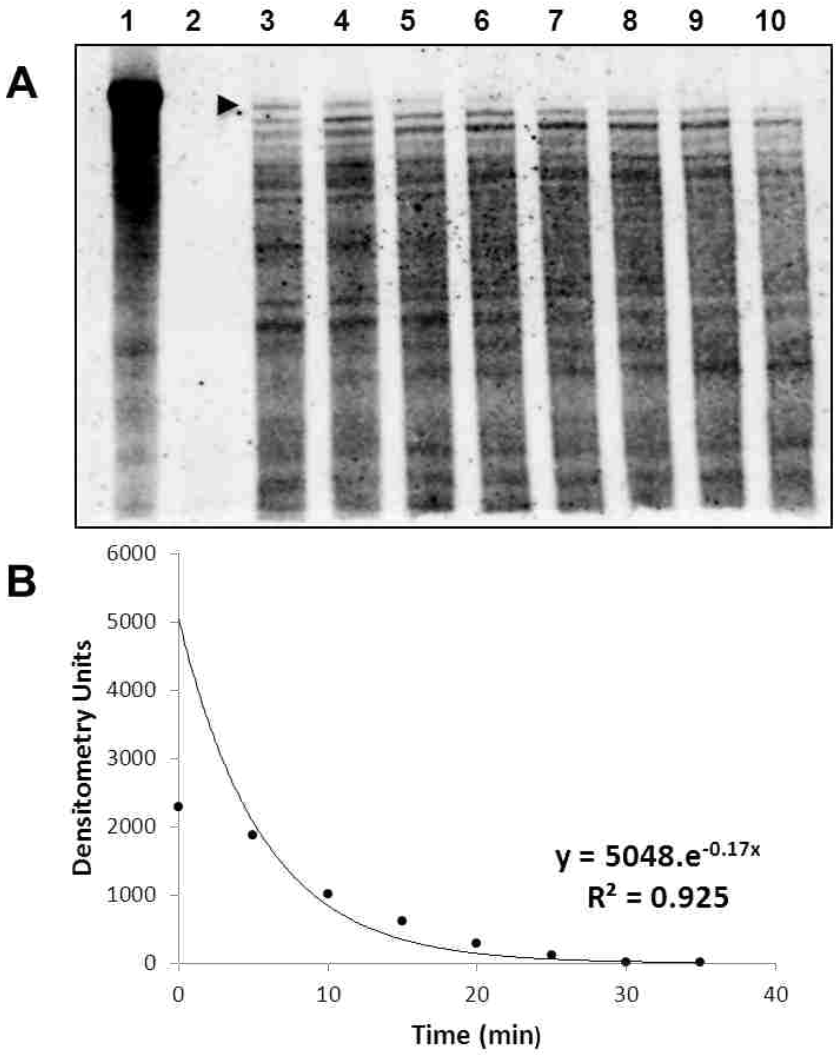


Figure. 2-7. *In vivo* half-life of *C. burnetii*'s IVS RNA in *E. coli*. A representative example of the *in vivo* IVS half-life as determined by RPA is shown. (A) RPA blot with the IVS band (508 bases) indicated by an arrowhead. (3 µg RNA per lane and 800 pg of IVS probe were used). Lanes 1 and 2 show controls with biotin-labeled IVS probe in the absence or presence of RNase A/T1, respectively; Lanes 3-10 show RNA isolated at 5-min intervals from 0-35 min following rifampin treatment. (B) Densitometric analysis of IVS bands in blot lanes 3-10 versus time of collection following rifampin treatment. The equation for the exponential curve and the R^2 value are inset and yields an IVS half-life of 3.9 min.

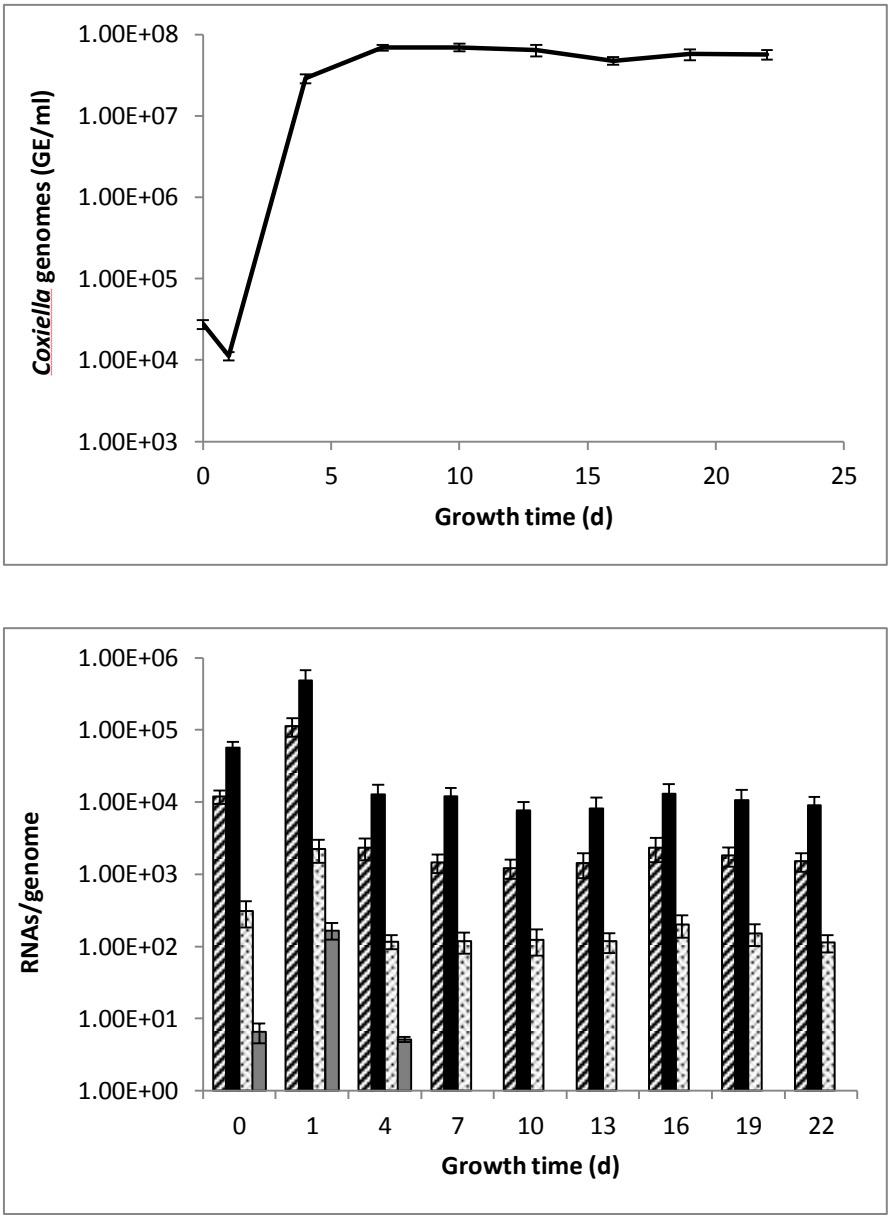


Figure 2-8. Levels of *C. burnetii* rRNAs and IVS RNA over time in culture. (A) q-PCR data showing *C. burnetii* genome equivalents per ml (GE/ml) during growth in axenic medium. (B) qRT-PCR data showing numbers of 16S rRNAs (striped bars), 23S F1 (black bars), 23S F2 (white, speckled bars) and IVS RNAs (grey bars) on a per genome basis. Data represent the means of six independent experiments ± S.D., except IVS data at 0 d, which represent the mean of three determinations ± S.D.

and subsequently became insignificant (Fig. 2-8). These results were consistent with previous reports for IVS elements of other bacteria, where IVS RNAs were not detectable by Northern analysis during growth, indicating reduced stability following excision from 23S rRNA (6, 8). Taken together, these results suggest that soon after excision from the 23S rRNA, IVS RNA is degraded *in vivo*, as observed *in vitro*.

Differential levels of 23S rRNA fragments following IVS excision.

The 23S rRNA of eubacteria is highly conserved, however, fragmentation caused by IVS elements does not appear to affect ribosome function (15). In fact, IVS elements occur within the 23S rRNA genes of several bacteria, including various *Leptospira* (5), *Salmonella* (6) and *Yersinia* (7). Although fragmentation does not affect the viability of these bacteria, it has been shown to affect the rate of 23S rRNA degradation in *Salmonella* as cells reach stationary phase (19). In fact, there appears to be a direct correlation between the degree of IVS-mediated fragmentation and the rate of 23S rRNA degradation, conferring a selective advantage to the bacterium. A similar adaptation could conceivably be advantageous to *C. burnetii* as it transitions to the small, endospore-like, SCV phase. To investigate this possibility and to help elucidate the biological function of IVS, we compared quantities of 16S rRNA to the two fragments of 23S rRNA by qRT-PCR. As a result of IVS excision, the 23S rRNA of *C. burnetii* is fragmented into ~1.2-kb and ~1.7-kb segments, i.e., fragment 1 (F1) and fragment 2 (F2), respectively. Results showed that on a per-genome basis, levels of 16S rRNA increased significantly at 1 d, i.e. at the beginning of log-phase ($P < 0.0001$), and fell significantly ($P < 0.0001$) as cells transitioned to stationary phase (4 d) (Fig. 2-8B).

Thereafter, levels of 16S rRNA remained stable with no significant changes from 4-22 d. Similar to what was observed with 16S rRNA, levels of F1 and F2 increased significantly (P<0.0002) at the beginning of log-phase (1 d), decreased significantly ($P < 0.0001$) as the cells approached stationary phase (4 d), and were stable between 4-22 d with no significant differences (Fig. 2-8B). However, the quantities of F1 and F2 appeared to be considerably different from one another. Since *C. burnetii*'s structural rRNAs are all encoded by a single-copy operon, they would be expected to be at parity, at least initially following transcription. Surprisingly, at each sampling time point, quantities of F1 were significantly higher (P<0.0007) than 16S rRNA (up to ~8 fold on day 7). On the other hand, levels of F2 were significantly lower ($P < 0.0001$) than 16S rRNA (up to ~50 fold on day 1). These results were corroborated by RPAs performed on total RNAs purified from *C. burnetii* isolated at similar time points (Fig. 2-9).

The differential stability of F1 and F2 was a surprising result, since studies in *Salmonella* have shown that following fragmentation, both segments of its 23S rRNA are equally prone to degradation near stationary phase. Since rRNA transcription, maturation and ribosome assembly are coupled in bacteria, it is conceivable that fragmentation of 23S rRNA might occur in the context of the ribosome. In such a case, following excision of IVS, F1 might be incorporated into the ribosome and thus protected from RNase degradation. Whereas, F2 might be more labile since it is relatively larger in size (~1.7 kb versus ~1.2 kb) and has two introns that must also be spliced out during maturation. These results show that following transcription and maturation each rRNA

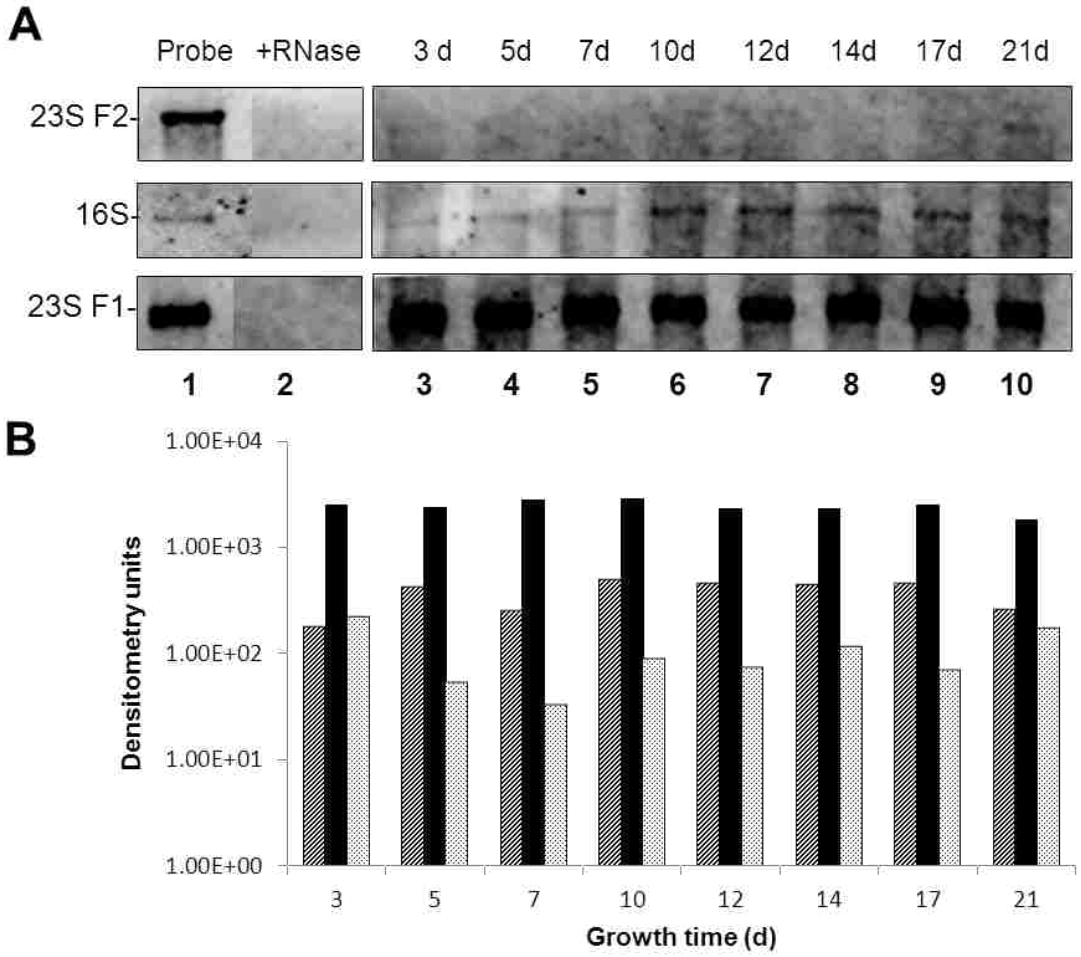


Figure. 2-9. Differential levels of *C. burnetii* rRNA species during growth. Levels of 16S rRNA and two fragments of 23S rRNA (23S F1 and 23S F2) were determined by RPA. (A) RPA blot showing F2, 16S and F1 bands. (In each case, 250 ng of RNA per lane and 25 pg of probe were used). Lanes 1 and 2 are control lanes with biotin-labeled probe in the absence or presence of RNase A/T1, respectively; Lanes 3-10 show RNA isolated over 3-21 d. (B) Densitometric analysis of 16S rRNA (striped bars), 23S F1 (black bars) and 23S F2 (white, speckled bars) in blot lanes 3-10 versus time of collection.

fragment is under the influence of a unique micro environment that determines its ultimate fate.

Considering that an equal number of each rRNA species (5S, 16S, 23S) is incorporated into a mature ribosome, the excess rRNA is undoubtedly degraded. F2, which is the lowest in relative quantity, would effectively limit the number of mature ribosomes that are ultimately formed. It is conceivable that this constraint could contribute to growth modulation in *C. burnetii*. We therefore hypothesize that IVS-mediated fragmentation of 23S rRNA is adaptive by virtue of the variable stability of the resulting 23S rRNA fragments, which contributes to the slow growth rate of *C. burnetii*. In turn, the slow growth rate fosters *Coxiella*'s tendency towards chronic infection and persistence in the host (22).

Phylogeny of *C. burnetii*'s IVS

Although 23S rRNAs are highly conserved among eubacteria, the occurrence of IVS within 23S rRNA is sporadic. In *Leptospira*, IVS occurs in only some species and a few strains of that species (5). While in *Salmonella*, with multiple *rrn* operons, IVS is present in only certain copies of the 23S rRNA genes (6). Additionally, in the LT2 strain of *S. typhimurium*, three distinct IVSs were found to occur in its 23S rRNA genes (11). This scattered distribution of IVS suggests the possibility of horizontal transfer. Moreover, Ralph and McClelland investigated the phylogeny of this element in species of *Leptospira* and showed evidence of natural horizontal transfer of IVS between species of this genus (31). Horizontal transfer is known to be an important process in evolution with well-established roles such as acquisition of antibiotic resistance genes (32) and

movement of group I and II introns between species (33). However, when the presence of IVS among different strains of *C. burnetii* was determined, it was found to be present in all strains with 99% sequence identity. In addition, the GC content of *C. burnetii*'s IVS is ~36% while that of its genome is ~42.5%. These observations suggest that modern *C. burnetii* strains may have acquired the element by vertical transfer from a common ancestor. Recently, Duron et al have shown that all known strains of *C. burnetii* have evolved from a *Coxiella*-like organism hosted by soft ticks of the *Ornithodoros* and *Argas* genera (34). They state that the maternally-inherited endosymbiont of ticks, a *Coxiella*-like organism, may have evolved into a highly-infectious pathogen of vertebrates, by acquisition of virulence genes through horizontal transfer from co-infecting pathogens (34). Keeping this in mind, it is conceivable that *C. burnetii* initially acquired its IVS through horizontal transfer before or during specialization to a vertebrate host. Thereafter, IVS was probably maintained within the genus by vertical transfer as a result of selective pressures imposed by the intracellular niche and the adaptive role of modulated growth imparted by the element.

F. Acknowledgements

We are grateful to Kip Barhaugh and Jeff Guccione for technical assistance. This work was supported by NIH grant R21AI078125 (to MFM).

G. References

1. **Afseth G, Mo YY, Mallavia LP.** 1995. Characterization of the 23S and 5S rRNA genes of *Coxiella burnetii* and identification of an intervening sequence within the 23S rRNA gene. J Bacteriol **177:**2946-2949.
2. **Voth DE, Heinzen RA.** 2007. Lounging in a lysosome: the intracellular lifestyle of *Coxiella burnetii*. Cell Microbiol **9:**829-840.
3. **Maurin M, Raoult D.** 1999. Q fever. Clin Microbiol Rev **12:**518-553.
4. **Minnick MF, Raghavan R.** 2011. Genetics of *Coxiella burnetii*: on the path of specialization. Future Microbiol **6:**1297-1314.
5. **Ralph D, McClelland M.** 1993. Intervening sequence with conserved open reading frame in eubacterial 23S rRNA genes. Proc Natl Acad Sci U S A **90:**6864-6868.
6. **Burgin AB, Parodos K, Lane DJ, Pace NR.** 1990. The excision of intervening sequences from *Salmonella* 23S ribosomal RNA. Cell **60:**405-414.
7. **Skurnik M, Toivanen P.** 1991. Intervening sequences (IVSs) in the 23S ribosomal RNA genes of pathogenic *Yersinia enterocolitica* strains. The IVSs in *Y. enterocolitica* and *Salmonella typhimurium* have a common origin. Mol Microbiol **5:**585-593.
8. **Konkel ME, Marconi RT, Mead DJ, Cieplak W, Jr.** 1994. Identification and characterization of an intervening sequence within the 23S ribosomal RNA genes of *Campylobacter jejuni*. Mol Microbiol **14:**235-241.

9. **Miller WL, Pabbaraju K, Sanderson KE.** 2000. Fragmentation of 23S rRNA in strains of *Proteus* and *Providencia* results from intervening sequences in the rrn (rRNA) genes. J Bacteriol **182:**1109-1117.

10. **Hsu D, Zee YC, Ingraham J, Shih LM.** 1992. Diversity of cleavage patterns of *Salmonella* 23S rRNA. J Gen Microbiol **138:**199-203.

11. **Mattatall NR, Sanderson KE.** 1996. *Salmonella typhimurium* LT2 possesses three distinct 23S rRNA intervening sequences. J Bacteriol **178:**2272-2278.

12. **Noller HF.** 1984. Structure of ribosomal RNA. Annu Rev Biochem **53:**119-162.

13. **Pabbaraju K, Miller WL, Sanderson KE.** 2000. Distribution of intervening sequences in the genes for 23S rRNA and rRNA fragmentation among strains of the *Salmonella* reference collection B (SARB) and SARC sets. J Bacteriol **182:**1923-1929.

14. **Hurtado A, Clewley JP, Linton D, Owen RJ, Stanley J.** 1997. Sequence similarities between large subunit ribosomal RNA gene intervening sequences from different Helicobacter species. Gene **194:**69-75.

15. **Evguenieva-Hackenberg E, Klug G.** 2000. RNase III processing of intervening sequences found in helix 9 of 23S rRNA in the alpha subclass of Proteobacteria. J Bacteriol **182:**4719-4729.

16. **Deutscher MP.** 2009. Maturation and degradation of ribosomal RNA in bacteria. Prog Mol Biol Transl Sci **85:**369-391.

17. **Lin LY, Ching CL, Chin KH, Chou SH, Chan NL.** 2006. Crystal structure of the conserved hypothetical cytosolic protein Xcc0516 from *Xanthomonas campestris*

reveals a novel quaternary structure assembled by five four-helix bundles. Proteins **65:**783-786.

18. **Baker BJ, Hugenholtz P, Dawson SC, Banfield JF.** 2003. Extremely acidophilic protists from acid mine drainage host Rickettsiales-lineage endosymbionts that have intervening sequences in their 16S rRNA genes. Appl Environ Microbiol **69:**5512-5518.
19. **Hsu D, Shih LM, Zee YC.** 1994. Degradation of rRNA in *Salmonella* strains: a novel mechanism to regulate the concentrations of rRNA and ribosomes. J Bacteriol **176:**4761-4765.
20. **Omsland A, Beare PA, Hill J, Cockrell DC, Howe D, Hansen B, Samuel JE, Heinzen RA.** 2011. Isolation from animal tissue and genetic transformation of *Coxiella burnetii* are facilitated by an improved axenic growth medium. Appl Environ Microbiol **77:**3720-3725.
21. **Sandoz KM, Sturdevant DE, Hansen B, Heinzen RA.** 2014. Developmental transitions of *Coxiella burnetii* grown in axenic media. J Microbiol Methods **96:**104-110.
22. **Raghavan R, Hicks LD, Minnick MF.** 2008. Toxic introns and parasitic intein in *Coxiella burnetii*: legacies of a promiscuous past. J Bacteriol **190:**5934-5943.
23. **Hicks LD, Warrier I, Raghavan R, Minnick MF.** 2011. Ribozyme stability, exon skipping, and a potential role for RNA helicase in group I intron splicing by *Coxiella burnetii*. J Bacteriol **193:**5292-5299.

24. **Coleman SA, Fischer ER, Howe D, Mead DJ, Heinzen RA.** 2004. Temporal analysis of *Coxiella burnetii* morphological differentiation. J Bacteriol **186:**7344-7352.

25. **Gregory ST, O'Connor M, Dahlberg AE.** 1996. Functional *Escherichia coli* 23S rRNAs containing processed and unprocessed intervening sequences from *Salmonella typhimurium*. Nucleic Acids Res **24:**4918-4923.

26. **Coleman SA, Fischer ER, Cockrell DC, Voth DE, Howe D, Mead DJ, Samuel JE, Heinzen RA.** 2007. Proteome and antigen profiling of *Coxiella burnetii* developmental forms. Infect Immun **75:**290-298.

27. **Skultety L, Hajduch M, Flores-Ramirez G, Miernyk JA, Ciampor F, Toman R, Sekeyova Z.** 2011. Proteomic comparison of virulent phase I and avirulent phase II of *Coxiella burnetii*, the causative agent of Q fever. J Proteomics **74:**1974-1984.

28. **Dunn JJ, Studier FW.** 1973. T7 early RNAs and *Escherichia coli* ribosomal RNAs are cut from large precursor RNAs in vivo by ribonuclease 3. Proc Natl Acad Sci U S A **70:**3296-3300.

29. **Dunn JJ.** 1976. RNase III cleavage of single-stranded RNA. Effect of ionic strength on the fideltiy of cleavage. J Biol Chem **251:**3807-3814.

30. **Bernstein JA, Khodursky AB, Lin PH, Lin-Chao S, Cohen SN.** 2002. Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. Proc Natl Acad Sci U S A **99:**9697-9702.

31. **Ralph D, McClelland M.** 1994. Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. J Bacteriol **176:**5982-5987.

32. **Krawiec S, Riley M.** 1990. Organization of the bacterial chromosome. Microbiol Rev **54:**502-539.

33. **Dujon B.** 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. Gene **82:**91-114.

34. **Duron O, Noel V, McCoy KD, Bonazzi M, Sidi-Boumedine K, Morel O, Vavre F, Zenner L, Jourdain E, Durand P, Arnathau C, Renaud F, Trape JF, Biguezoton AS, Cremaschi J, Dietrich M, Leger E, Appelgren A, Dupraz M, Gomez-Diaz E, Diatta G, Dayo GK, Adakal H, Zoungrana S, Vial L, Chevillon C.** 2015. The Recent Evolution of a Maternally-Inherited Endosymbiont of Ticks Led to the Emergence of the Q Fever Pathogen, *Coxiella burnetii*. PLoS Pathog **11:**e1004892.

CHAPTER THREE

Identification of Novel small RNAs of *Coxiella burnetii*

As published in: PLoS One. 2014 Jun 20;9(6):e100147. doi:

10.1371/journal.pone.0100147

A. Abstract

Coxiella burnetii, an obligate intracellular bacterial pathogen that causes Q fever, undergoes a biphasic developmental cycle that alternates between a metabolically-active large cell variant (LCV) and a dormant small cell variant (SCV). As such, the bacterium undoubtedly employs complex modes of regulating its lifecycle, metabolism and pathogenesis. Small RNAs (sRNAs) have been shown to play important regulatory roles in controlling metabolism and virulence in several pathogenic bacteria. We hypothesize that sRNAs are involved in regulating growth and development of *C. burnetii* and its infection of host cells. To address the hypothesis and identify potential sRNAs, we subjected total RNA isolated from *Coxiella* cultured axenically and in Vero host cells to deep-sequencing. Using this approach, we identified fifteen novel *C. burnetii* sRNAs (CbSRs). Fourteen CbSRs were validated by Northern blotting. Most CbSRs showed differential expression, with increased levels in LCVs. Eight CbSRs were upregulated (≥2-fold) during intracellular growth as compared to growth in axenic medium.

B. Introduction

During infection, pathogenic bacteria must adapt to diverse and dynamic environments imposed by their host and regulate synthesis of a variety of molecules (DNA, RNA and proteins) needed to colonize, replicate and persist. This kind of regulation must be rapid, metabolically inexpensive and efficient. There is growing evidence that post-transcriptional control mediated by small RNAs (sRNAs) plays a significant role in bacterial regulation (1, 2). In pathogenic bacteria, sRNAs are known to coordinate virulence gene expression and also stress responses that are important for survival in the host (3, 4). Bacterial sRNAs are typically 100-400 bases in length and are categorized as *cis*-encoded sRNAs and *trans*-encoded sRNAs. Most *cis*-encoded sRNAs are located within 5′ untranslated regions (UTRs) of mRNAs and are transcribed in the antisense orientation to the corresponding mRNA. *Cis*-encoded sRNAs can expose or block a ribosome-binding site (RBS) by adopting different conformations in response to various environmental cues, thereby regulating translation. On the other hand, *trans*-encoded sRNAs are located in intergenic regions (IGRs). They share only limited complementarity with their target RNAs and are thought to regulate translation and/or stability of these RNAs (2). sRNAs can interact with mRNA or protein in order to bring about regulation, but a majority of them function by binding to mRNA targets. An example of a widely distributed and well-studied sRNA is 6S RNA. 6S RNA binds to RNA polymerase (RNAP)-σ^{70} complex and regulates transcription by altering RNAP's promoter specificity during stationary phase (5, 6).

Coxiella burnetii, the causative agent of Q fever, is classified as a Gram-negative obligate intracellular γ-proteobacterium. Human Q fever is generally a zoonosis acquired

by inhalation of contaminated aerosols and can present either as an acute or chronic disease. An acute case of Q fever typically ranges from an asymptomatic infection to an influenza-like illness accompanied by high fever, malaise, atypical pneumonia, myalgia and hepatitis. In approximately 2-5% of cases, chronic Q fever occurs and manifests as endocarditis, especially in patients with predisposing valvular defects (7). The pathogen's biphasic developmental cycle consists of two cellular forms. An infectious, dormant small cell variant (SCV) is spore-like and can endure adverse environmental conditions such as heat, pressure, UV light and desiccation. Following inhalation, *Coxiella* enters alveolar macrophages by endocytosis and generates a phagolysosome-like vacuole termed a parasitophorus vacuole (PV). The PV interacts with autophagosomes for bacterial nutrition (8). At approximately 8 h post-infection, SCVs metamorphose to form metabolically active LCVs in the PV, with a doubling time of approximately 11 hours (9, 10). Following 6-8 days of intracellular growth, the PV reaches maturity and occupies almost the entire volume of the cell, and it is filled with a mixture of LCVs and SCVs. By approximately 12 days, the entire bacterial population has transformed into SCVs that are eventually released upon lysis of the host cell (10).

C. burnetii encounters various and sudden changes in environmental conditions during its life cycle, including a rapid upshift in temperature upon transmission from contaminated aerosols to the human lung, and a downshift in pH and an increase in reactive oxygen intermediates (ROIs) in the PV. All of these events are relevant to rapid, sRNA-mediated regulation (2). Recent reports have identified sRNAs in a variety of pathogenic bacteria, including *Legionella pneumophila* (11) and *Streptococcus pyogenes* (12). Reports have also shown the involvement of sRNAs in the pathogenesis

of *Streptococcus pneumoniae*, *Salmonella spp., Yersinia pseudotuberculosis* and *Listeria monocytogenes* (13-16).

The sRNAs of intracellular bacterial pathogens are poorly characterized, and there are no reports on sRNAs of *C. burnetii*. Thus, the aim of our study was to identify sRNAs associated with the bacterium's developmental cycle and host cell infection. Here, we describe a set of 15 novel *Coxiella* sRNAs identified by high-throughput sequencing of RNA (RNA-seq) isolated from distinct life stages and culture conditions.

C. Materials and Methods

Cultivation of *C. burnetii*. *C. burnetii* Nine Mile phase II (strain RSA439, clone 4) was propagated in African green monkey kidney (Vero) fibroblast cells (CL-81; American Type Culture Collection) grown in RPMI medium (Invitrogen Corp.) supplemented with 10% fetal bovine serum at 37^{o}C in a 5% CO_2 atmosphere. Bacteria were purified from host cells using differential centrifugation, as previously described (17). LCVs were harvested at 72 h post-infection from infected cells using digitonin (18). SCVs were harvested and prepared at 21 days post-infection (dpi), as previously described (19), and used to infect Vero cell monolayers for the production of synchronized bacterial cultures. *C. burnetii* was also cultivated axenically in ACCM2 at 37^{o}C in a tri-gas incubator (2.5% O_2, 5% CO_2, 92.5% N_2) with continuous shaking at 75 RPM (20). To generate LCVs, ACCM2 was inoculated with 10-d-old ACCM2-cultured bacteria, incubated 72 h, and isolated by centrifugation (10,000 x g for 20 min at 4^{o}C), as previously described (9). SCV generation in ACCM2 was identical to LCVs except

bacteria were grown for 7 d, and then flask lids were tightened and cultured an addition at 14 d on the lab bench (~25^{o}C) without replenishing the medium (21).

RNA isolation and deep-sequencing. To isolate *C. burnetii* RNA from infected Vero cells, LCVs were prepared as above and treated with 40 µg/ml RNase A in RNase A digestion buffer [10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 5 mM EDTA] to reduce host cell RNA contamination. SCVs were prepared as above and used directly. Total RNA used in deep-sequencing was purified from LCVs and SCVs with a Ribopure kit (Ambion). Resulting RNAs were treated with excess DNase I to remove trace amounts of residual DNA using a DNA-*free* kit, as instructed by the manufacturer (Applied Biosystems). RNA was precipitated with 100% ethanol and enriched for bacterial RNAs by sequential use of MICROB*Enrich* (Ambion), MICROB*Express* (Ambion) and Ribo-Zero (Epicentre) kits to increase the relative level of *C. burnetii* RNA derived from Vero cell-propagated organisms and to exclude rRNAs, respectively. RNA from *C. burnetii* cells cultured in ACCM2 was done as for infected Vero cells, however, the MICROB*Enrich* (Ambion) step was omitted. RNA was quantified using a NanoPhotometer (Implen) and checked for integrity using a 2100 Bioanalyzer (Agilent Technologies). Sequencing libraries were prepared with a TruSeq RNA sample preparation kit (Illumina). Libraries were sequenced on an Illumina HiSeq 2000 (76 cycles) at the Yale Center for Genome Analysis (West Haven, CT). Two independent samples were sequenced from all conditions, and sequencing statistics are given in Table 3-1. Deep sequencing data were submitted to the Sequence Read Archive (SRA) database, NCBI, and assigned the accession number SRP041556.

Table 3-1. Sequencing statistics.

Sample	Reads mapped	Total reads	% mapped
ACCM2-LCV1	31,890,682	32,945,220	97
ACCM2-LCV2	34,064,618	35,157,535	97
ACCM2-SCV1	22,937,399	23,682,545	97
ACCM2-SCV3	29,168,247	30,027,641	97
VERO-LCV2	61,191,633	81,206,707	75
VERO-LCV3	45,315,665	65,102,022	70
VERO-SCV1	33,270,940	47,656,046	70
VERO-SCV2	48,317,348	59,069,599	82

Mapping of sequencing reads and quantification of transcripts. Sequencing reads were mapped on the *C. burnetii* Nine Mile Phase I (RSA 493) genome (NC_002971.3) using BWA software (22). The algorithm was set to allow for two mismatches between 76-nt reads and the genome sequence. Coverage at each nucleotide position was visualized using Artemis software (23). Expression values for each genomic location were calculated by determining the number of reads overlapping that region and normalizing it to the total number of reads in each library and the region's length. The average expression values obtained from two independent libraries per time point were denoted as Mean Expression Values (MEVs). Transcripts were qualified as sRNAs if they were 50-400 nt in length, had an MEV ≥ 5 times that of the flanking 50 nucleotides and did not correspond exactly to an annotated open reading frame (ORF). The presence of σ^{70} consensus promoters and rho-independent terminators was predicted using BPROM (24) and TranstermHP (25) software, respectively.

Northern blot analysis. Northern blots were carried out using a NorthernMax kit (Ambion) as per manufacturer's instructions. Briefly, total RNAs of *C. burnetii* grown in Vero cells or ACCM2 were isolated by sequential use of Ribopure (Ambion) and DNA-*free* (Applied Biosystems) kits and then precipitated with 100% ethanol. For quality control purposes, RNA samples were occasionally analyzed on denaturing acrylamide gels to check for RNA integrity. RNA degradation was not observed in samples used in the study (data not shown). RNA (3 µg per lane, except CbSR 2, where 1.7 µg RNA was used) was electrophoresed through 1.5% agarose-formaldehyde gels and blotted onto positively-charged BrightStar-Plus nylon membranes (Ambion). Membranes were then UV-cross-linked or chemically cross-linked by 1-ethyl-3-(3-dimethylaminopropyl)

carbodiimide (EDC) (Sigma-Aldrich), as previously described (26). Hybridizations were carried out using single-stranded RNA probes specific to each sRNA. RNA probes were generated by T7 promoter-mediated *in vitro* transcription of PCR products using a MEGAscript kit as instructed (Ambion), in the presence of biotin-labeled UTP (Bio-16-UTP; Ambion). Finally, membranes were developed with a BrightStart BioDetect kit (Ambion) following the manufacturer's protocol, and visualized using a LAS-3000 imaging system (Fujifilm). Densitometry was performed using ImageJ software (27). [Please see Table 3-2 and 3-3 for primers and probe details].

D. Results

Identification of *C. burnetii* sRNAs

To investigate the transcriptome profile of *C. burnetii* and to identify potential sRNAs, we first isolated RNA from LCVs and SCVs co-cultured in Vero cells as well as those cultured axenically in ACCM2 medium. cDNAs prepared from these RNAs were subjected to Illumina sequencing. This deep sequencing analysis resulted in roughly 23 to 32 million reads from RNA isolated from *C. burnetii* cultured axenically, and ~47 to 81 million reads from total RNA isolated from Vero co-cultures. On the whole, sequencing reads obtained from *C. burnetii* cultured in ACCM2 mapped well to the genome (97%). On the other hand, sequencing reads from total RNA isolated from bacteria cultured in Vero host cells mapped 76% and ~72.5% to the genome, respectively (Table 3-1). By analyzing the sequencing reads, we identified a total of 15 novel sRNAs, which will hereafter be referred to as CbSRs (*Coxiella burnetii* small RNAs) (Table 3-4).

Table 3-2. PCR primers used to make CbSR probes*.

Primers	**Sequence 5' to 3'**
CbSR 1_Forward	CTTTCTGAAGAGGTAATCACGAAG
CbSR 1_Reverse	TAATACGACTCACTATAGGGTCCCTACCAAGCAGTTCTGTC
CbSR 2_ Forward	GATGCTGTTCTTCGTAGGC
CbSR 2_Reverse	TAATACGACTCACTATAGGGCGGCTATCGCTTCTTTGC
CbSR 3_ Forward	AAACAAACCTTGATAGAAAGCG
CbSR 3_Reverse	TAATACGACTCACTATAGGGACTACTTGATATTCCCTCTTTACC
CbSR 4_ Forward	GATAGCGTGGTGGGAATCGGTTAC
CbSR 4_Reverse	TAATACGACTCACTATAGGGGGGTTTGGTGCGGTCAAGTGG
CbSR 5_ Forward	CGAAATGAAGAAAAGCAACTC
CbSR 5_Reverse	TAATACGACTCACTATAGGGGCTTGGACTTCCTCTAAATG
CbSR 6_ Forward	GTAACTACGGGCATTCCATCG
CbSR 6_Reverse	TAATACGACTCACTATAGGGCGGTACTAACGGTTTCTCAAGC
CbSR 7_ Forward	CTTACTAGGGGATTTTTTTTACTCG
CbSR 7_Reverse	TAATACGACTCACTATAGGGTCGTTAGTTGAAACATTGAACG
CbSR 8_ Forward	CGAGGTGCTTTAGCCATTG
CbSR 8_Reverse	TAATACGACTCACTATAGGGAGATTGTAACGAGACGATGAAG
CbSR 9_ Forward	GAGTACCGTTATAAACATGGATAC
CbSR 9_Reverse	TAATACGACTCACTATAGGGTGAAGATGGGTGGAAAGCC
CbSR 10_ Forward	TCTTTTAATGAAGCGGGAATGG
CbSR 10_Reverse	TAATACGACTCACTATAGGGGCAACATTGGCACGATGG
CbSR 11_ Forward	GACATAACTAGACATCAGGTG
CbSR 11_Reverse	TAATACGACTCACTATAGGGGATTGGCTGCTGTAATGG
CbSR 12_ Forward	TAGCTGAGGTCTCTAGGATCTTG
CbSR 12_Reverse	TAATACGACTCACTATAGGGGACCTTAGACTACCTCATTACGTTTAG
CbSR 13_ Forward	GTCGTTCCCGTGCGTAGG
CbSR 13_Reverse	TAATACGACTCACTATAGGGGCCGTCTGCTGTAGTATTGAAG
CbSR 14_ Forward	GCTTTGGGAGATGACCTTCGC
CbSR 14_Reverse	TAATACGACTCACTATAGGGGACCGTGAGGACAGCAGTTTG
CbSR 15_ Forward	GAATCGGGAGAAACACCAC
CbSR 15_Reverse	TAATACGACTCACTATAGGGTGCCTTTAGGAGGTCAGTC

*The T7 promoter sequence is underlined

Table 3-3. Probes used in Northern blots and RPAs.

Probes	Sequence 5' to 3'
CbSR 1	CTTTCTGAAGAGGTAATCACGAAGTTAGGAAACTTTATCTCATGGAGAGAGAA AGGTCTAAGGACAGAACTGCTTGGTAGGGA
CbSR 2	GATGCTGTTCTTCGTAGGCGGCTTGGCCGCTTGCCAACAACAGGATCGCACC CATCAGGATAGACAAAATGGCGCTGTTACAACGCCCCACCCCGATGATCAAC AAAAAGATAATATGCAAAGAAGCGATAGCCG
CbSR 3	AAACAAACCTTGATAGAAAGCGCGCTAGATTCCCGCCTGCGCGCATAGACGT CAACTTAAGGAGCGAGGTAAAGAGGGAATATCAAGTAGT
CbSR 4	GATAGCGTGGTGGGAATCGGTTACGGAAGATCGTTCCTGCCATTCTGTCGCA AATGTAGCCACTTGACCGCACCAAACCC
CbSR 5	CGAAATGAAGAAAAGCAACTCTTGGCCTGTCAAATGCTGCAAGCACAGATCG ATCCCTTGTTGATTGCGAAACTAACAGGACTTCATTTAGAGGAAGTCCAAGC
CbSR 6	GTAACTACGGGCATTCCATCGCGGGTAGATTTAATCAGTCCTTTAAAAAAATTT TGTTGCGCTGCTTGAGAAACCGTTAGTACCG
CbSR 7	CTTACTAGGGGATTTTTTTTACTCGTTTTCAATTCTATTGAACCGTTCAATGTTT CAACTAACGA
CbSR 8	CGAGGTGCTTTAGCCATTGGGCAACCCGTTTATTGGACTAGTATGGATAGATT AAATCAAAACAAATTGAGCATGCGAACTGAACAAAGAGATAAAAGGCCACGTC AAATTCCTAAGTGAGAAAGAATGTTTTGCAGCTCGTTTCAGGTCTTGAAGAAG CTTCATCGTCTCGTTACAATCT
CbSR 9	GAGTACCGTTATAAACATGGATACCCACTAGGTTGATGTTAGGCATTTTTAAAA AGAAATTCTTATTCGAAAAAACGATTTCTTCATCCATGCCGAACTGCGAAGGC TTTCCACCCATCTTCA
CbSR 10	TCTTTTAATGAAGCGGGAATGGTTGCCATTGGAAACGTGGAGATGCCAGTAG ATATGATGCTGGGAGGACCAATGACGTCGGTAACAGGCCCATCGTGCCAATG TTGC
CbSR 11	GACATAACTAGACATCAGGTGTAACCAAACAATCACGGAGGATCGACAAATGA GAAAAGCACTTGCTAGCGTTGTTGTAATTATATTTGCAGGCTTACTTTGCACG AGTTTACTGACAATCTTTACAGGCAACTCATCAGGTAACCATTACAGCAGCCA ATC
CbSR 12	TAGCTGAGGTCTCTAGGATCTTGGTGGACAAGGAAGTCCTCGGTGTACTCTA AACGTAATGAGGTAGTCTAAGGTC

CbSR 13	GTCGTTCCCGTGCGTAGGCCGTCTATCGATAGACAAAGAATCGAGGGCTTTC ACGGCCGAAGTCACGGGAATCCCGGCTAAGAGGGGCTTGAAGAACACTAAC GGTGTTTTTCTTAGCTCCTTAATCTGGGTCCCCCCGACTCGGCCGTGAAGGTT TTGTATTCTTCAATACTACAGCAGACGGC
CbSR 14	GCTTTGGGAGATGACCTTCGCTGTGAAAATGGGGGTACCGTTACAGGAGTTA CTCATAAGAATTGCGGGCTAACCTGGGAGACTAATCCTCCCGAATTCAATTTT TGCCCAAAGGGCCAAACAACAAACTGCTGTCCTCACGGTC
CbSR 15	GAATCGGGAGAAACACCACTCTTAAAATTACTATTGAAACGATTACCCAAGCA ATCTTCCTTGAAGAAAGTAGCAAGTAATCCTCACTATATCGAGATGACTGACC TCCTAAAGGCA

Table 3-4. Novel *C. burnetii* sRNAs (CbSRs) identified by RNA-seq*.

sRNA	Left End[a]	Right End[a]	Size (nt)	Strand	Axenic LCV	Axenic SCV	Vero LCV	Vero SCV
					Mean Expression Value (MEV)			
CbSR 1	12005	12117	113	F	15444.89	257.96	59710.85	711.00
CbSR 2	75261	75503	243	R	7381.43	782.49	15719.76	5885.67
CbSR 3	481609	481806	198	R	577.92	757.68	1703.55	1225.17
CbSR 4	544387	544582	196	R	1392.97	644.14	4163.36	3433.46
CbSR 5	702095	702304	210	R	3302.63	293.32	584.04	338.12
CbSR 6	727878	728097	220	R	1776.92	332.35	778.38	401.85
CbSR 7	657100	657198	99	F	808.27	158.84	103.52	176.22
CbSR 8	866381	866666	286	R	8384.18	338.06	1867.89	508.76
CbSR 9	973800	974012	213	F	370.36	276.08	763.43	484.25
CbSR 10	1090006	1090228	223	R	7138.08	914.10	5713.77	1768.32
CbSR 11	1327797	1328052	256	F	6428.52	922.77	4258.87	4309.17
CbSR 12	1403153	1403300	148	R	4423.01	1477.31	434177.60	35288.85
CbSR 13	1816997	1817305	309	F	3716.12	1186.83	1621.83	1206.80
CbSR 14	1838698	1838886	189	R	790.92	451.29	3719.02	691.73
CbSR 15	1878295	1878543	249	F	1054.67	298.45	205.01	396.01

[a]Numbering according to NC_002971.3

All 15 CbSRs were present in both LCVs and SCVs cultured in axenic medium as well as in Vero cells. Comparison of the MEVs of LCVs and SCVs indicates that most CbSRs are present at higher levels in LCVs regardless of culture conditions. CbSRs could be classified into three groups based on the relative location of their coding sequence on the genome. Specifically, group I includes sRNAs encoded entirely within an IGR; group II consists of sRNAs situated antisense to identified ORFs (antisense sRNA), and group III includes sRNAs that are ORF-derived (Fig. 3-1). A majority (eight of fifteen) of the identified sRNAs were antisense sRNAs. Sizes of the CbSRs ranged from 99-309 nt with a minimum MEV of ~104 and a maximum MEV of ~434,178 in at least one growth condition. BLAST analyses showed that all sRNAs were found only in *Coxiella* and most sRNAs were highly conserved within six available *C. burnetii* genomes (RSA 493, RSA 331, Dugway 5J108-111, Cb175 Guyana, CbuK Q154 and CbuG Q212) with ≥97% sequence identity. The exception was CbSR 8, which was only found in *C. burnetii* strains RSA493, RSA331 and Cb175 Guyana. Regions immediately upstream of all sRNAs possessed predicted σ^{70} consensus promoters (Table 3-5), and intrinsic (Rho-independent) terminators were predicted just downstream of seven sRNAs (Table 3-6), suggesting that these are bona fide sRNAs.

Verification of sRNA candidates

The fifteen sRNA candidates identified by RNA-seq were further analyzed by Northern blotting of total RNAs from *C. burnetii* LCV or SCV morphotypes (Fig. 3-2). Northern blots were probed using strand-specific biotinylated RNA oligonucleotides specific to each sRNA. In each case, CbSRs produced distinct bands on the Northern blots, validating their existence in the transcriptome as well as the strand of origin in the

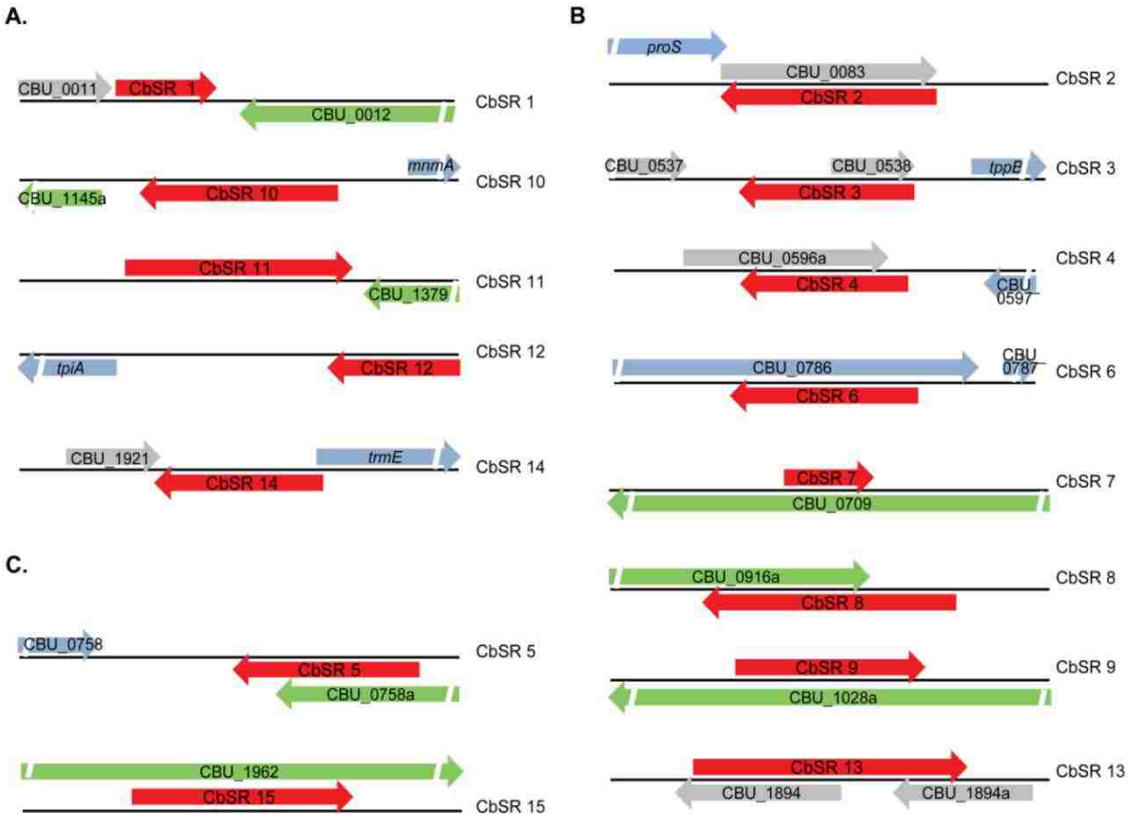


Figure 3-1. Linkage maps showing CbSR loci on the *C. burnetii* chromosome (black line). Red arrows indicate CbSRs and their relative orientation. Blue, grey and green arrows represent annotated, hypothetical ORFs and pseudogenes, respectively. CbSRs are classified into three groups based on their location relative to adjacent genes: **A.** Group I: CbSRs encoded within IGRs, **B.** Group II: CbSRs located antisense to identified ORFs and **C.** Group III: CbSRs that are ORF-derived.

Table 3-5. Putative σ^{70} promoters of CbSRs identified upstream of sRNA coding sequences using BPROM (24).

sRNA	-35 box	-10 box	sRNA	-35 box	-10 box
CbSR 1	TTTATA	GATTGT	CbSR 9	TTTAAT	TACACT
CbSR 2	TTTAAA	TATATT	CbSR 10	TTGTCT	TATAAT
CbSR 3	TTCTAA	CAGGAT	CbSR 11	TTTCAA	TATCTT
CbSR 4	TTGAGA	TAGTCT	CbSR 12	TTGTTA	TATATT
CbSR 5	TTATCA	TGAAAT	CbSR 13	TTGGAG	TATAAT
CbSR 6	TGGCCA	TATAAT	CbSR 14	TTGCTA	TAAAAA
CbSR 7	TTCACA	GATAAT	CbSR 15	TTATCA	GATAAT
CbSR 8	TGGCCA	TATAAT			

Table 3-6. Rho-independent terminators of CbSRs identified using TranstermHP (25)*.

sRNA	**Predicted terminator sequence**
CbSR 1	AGGGATCACCAACCCGGGGTGGTTATAGCAACCACCCCTTTTTTTTATTATTA
CbSR 2	CGCCTCAGTATGAAAGAAATCTCGGCCGTTGATGTCCGAGATTTCTTCATCTAAACACAG
CbSR 3	AAAGCCTAAGAAAAGCGCCATCGGTGTTTTTCTTAGCCCCC
CbSR 10	ATCTACGTAAACAAAGCAGGCAAAATCCTCGAATCGGATCTGCCTGCTTTTTTTTGAAGAAA
CbSR 11	TGATTATTTCCCCCAGCCTAGTCTGTCCGTTGTAAAACGGCAGCTAGGCTGCTTTCATTCCAGG
CbSR 12	TTGTACTAATAAAGAGGACCGCTTTTGCGGTCCTTTTTTTTTCTCACTT
CbSR 13	GAGGGGCTTGAAGAACACTAACGGTGTTTTTCTTAGCTCCT

*underlined are the sRNAs overlapping with terminator sequences.

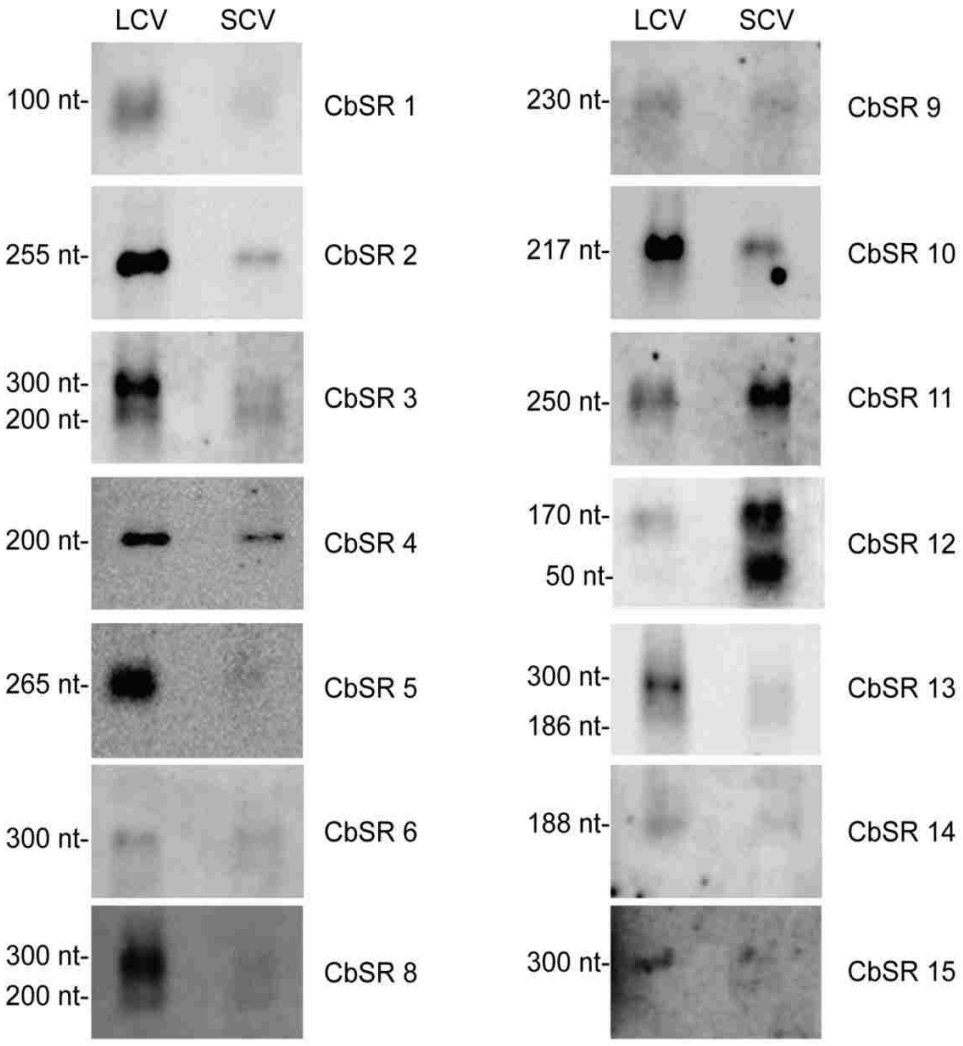


Figure 3-2. Northern blot detection of CbSRs. RNA was isolated from LCVs (3 dpi) and SCVs (21 dpi) grown in ACCM2. Hybridizations were performed at high stringency using biotinylated oligonucleotide probes specific to each CbSR. 3 µg RNA was used for all lanes. Apparent sizes of the CbSRs, as calculated from Northern blots, are indicated. (Note: intensity of bands is not comparable between panels, since exposure times for each panel have not been optimized).

C. burnetii genome. However, CbSR 7 was not observed. We believe this was due to relatively low CbSR 7 transcript quantity that was undetectable by Northern analysis (28).

For most of the CbSRs, estimated band sizes on Northern blots corresponded to the sRNA lengths predicted by RNA-seq analysis. However, four out of fourteen CbSRs showed multiple bands on blots (e.g., CbSR 3, CbSR 8, CbSR 12 and CbSR 13). First, in the case of CbSR 3, a longer transcript (~300 nt) was observed, which could represent a primary transcript that is cleaved to give a mature sRNA of ~200 nt, as obtained by RNA-seq. A similar processing has been previously described for sRNAs of other bacteria (29, 30). Second, CbSR 8, CbSR 12 and CbSR 13 Northern blots revealed two bands in which the larger bands corresponded to sizes obtained by RNA-seq, suggesting that the upper band is the actual sRNA. In CbSR 12, the molar ratios of the two different-sized bands observed varied between the LCV and SCV stages, possibly indicating different RNA processing, similar to what occurs with the SroF sRNA of *E. coli* (30). When transcript levels of the fifteen CbSRs were compared on blots, most had increased expression during the metabolically-active LCV phase with exceptions like CbSR 9 which was present in seemingly equal amounts in both morphotypes.

Northern blot signal intensity of most CbSRs corresponded to the MEVs obtained by RNA-seq (Table 3-4), with a few anomalies like CbSR 11, CbSR 12 and CbSR 13. Although the larger band (~300 nt) of CbSR 3 doesn't correspond well with RNA-seq MEV ratios, the lower (~200 nt) band does. Furthermore, signal intensities of CbSR 11 and CbSR 12 bands on Northern blots, as a function of morphotype, were consistently

reversed relative to their deep sequencing MEVs with increased transcript level in SCVs rather than LCVs. Further investigation is required to determine the basis of these discrepancies.

sRNAs up-regulated during intracellular growth

To search for sRNA regulators that are significantly up-regulated during a host cell infection, we compared expression levels of CbSRs from *C. burnetii* cultured in Vero host cells to those cultured axenically in ACCM2 (Table 3-4). Results showed eight CbSRs with increased MEVs in host cells (i.e., at least 2-fold higher) relative to axenic medium, including CbSR 1, CbSR 2, CbSR 3, CbSR 4, CbSR 9, CbSR 11, CbSR 12 and CbSR 14. Northern hybridizations were performed on each of these CbSRs to confirm their existence and determine their levels under different growth conditions (Fig. 3-3). The results observed were consistent with RNA-seq data. CbSR 12 showed a marked 24-fold higher level in Vero-grown *C. burnetii* suggesting a possible role in regulating a bacterial response related to intracellular survival. Other CbSRs that were markedly increased during intracellular growth included CbSR 2 and CbSR 4, which were 8-fold and 5-fold higher by MEV, respectively, compared to values obtained from axenically-grown *C. burnetii*.

E. Discussion

Although *C. burnetii* is an obligate intracellular parasite in nature, its life cycle includes a endospore-like, dormant SCV morphotype that enables the bacterium to persist and survive outside of host cells. Given the disparate physical conditions encountered by *Coxiella* in the context of the environment and host, it is highly likely

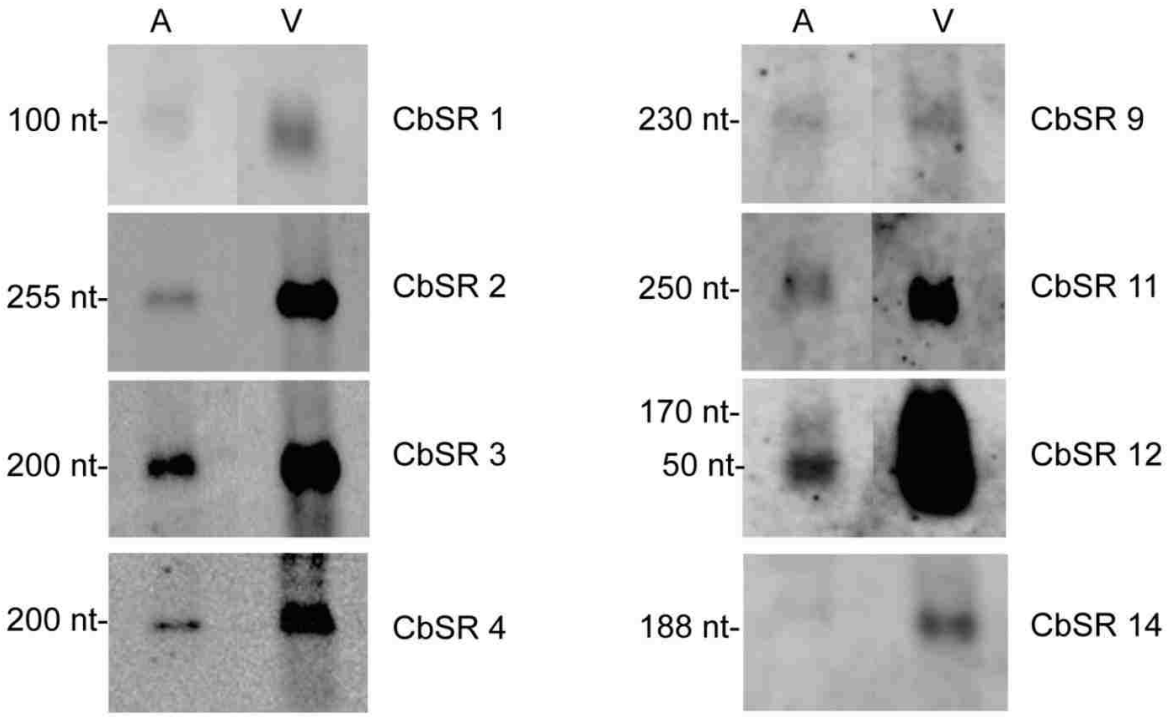


Figure. 3-3. Northern blots showing CbSRs up-regulated (≥ 2 fold) in host cells relative to ACCM2. RNA was isolated from SCVs (3 dpi) grown in ACCM2 (A) and in Vero host cells (V). Hybridizations were performed at high stringency using biotinylated oligonucleotide probes specific to each CbSR. 3 µg RNA was used for all lanes. Apparent sizes of the CbSRs, as calculated from the Northern blots, are indicated. (Note: intensity of bands is not comparable between panels, since exposure times for each panel have not been optimized).

that the bacterium employs a rapid and efficient means of regulation to withstand the changing, harsh conditions. Recently, sRNAs have become increasingly recognized as modulators of gene expression, and their role in controlling stress response and virulence, directly or indirectly, has been shown in several bacteria (4, 31). Here, we describe a deep sequencing-based identification of sRNAs in *C. burnetii*. RNA-seq has been used previously on several other organisms to identify novel non-coding RNAs (32-34), but this is the first experimental evidence for, and identification of, sRNAs in *C. burnetii*.

Analysis of sRNA libraries generated from total RNA isolated from *C. burnetii* grown in Vero cells and in axenic medium led to the identification of fifteen novel sRNAs, referred to as CbSRs 1-15. To ensure that the identified sRNAs were authentic, we experimentally verified their existence using Northern blot analyses and identified their strand of origin. However, CbSR 7, although detected by RNAseq, was not detectable by Northern blot analysis (28). The lengths of most of the CbSRs estimated from Northern blots were in fairly good agreement with that determined by RNA-seq. Moreover, the CbSRs are unique to *C. burnetii,* and with the exception of CbSR 8, highly conserved among six strains of the bacterium. All CbSRs were independently detected in both morphotypes of *Coxiella* isolated from both Vero cells and ACCM2, but their levels changed as a function of growth conditions. These results strongly suggest that CbSRs play a regulatory role in the physiology of *Coxiella.* Not surprisingly, transcript levels of most CbSRs increased during growth phase (LCV) as compared to stationary phase (SCV). A similar observation has been reported in *S. pyogenes*, where transcript levels of most sRNAs are abundant at exponential and early stationary phase

as compared to late stationary phase (12). Based on these observations one could predict that CbSRs help regulate genes that are involved in metabolic functions.

When we compared the transcript levels of CbSRs obtained from *Coxiella* grown in host cells versus axenic medium, eight sRNAs were found to be at higher levels during intracellular growth. Of these, CbSR 12 is particularly striking with regards to its up-regulation in the host cell, and is a current focus of research in our lab. The role of sRNAs in controlling pathogenesis and virulence has been reported in a number of bacteria, including *L. monocytogenes* (35), *Salmonella typhimurium* (36), and *Vibrio cholerae* (37). We hypothesize that these eight CbSRs are involved in regulating the bacterium's stress response in the intracellular niche. In addition to identification of 15 novel sRNAs, we also identified the bacterium's RNase P RNA (encoded by *rnpB*), tmRNA (encoded by *ssrA*) and 6S RNA (encoded by *ssrS*) by RNA-seq. RNase P RNA and tmRNA are well-studied sRNAs that are conserved among all bacteria. Studies have shown that RNase P RNA is the ribozyme component of RNase P that is involved in processing of 4.5S RNA and tRNA precursor molecules (38). On the other hand, the tmRNA rescues stalled ribosomes during translation and tags incompletely translated proteins for degradation (39). 6S RNA is known to associate with RNA polymerase and interfere with transcription, especially during stationary phase (40).

Interestingly, using in silico analysis we also discovered that *C. burnetii* lacks an apparent Hfq; an RNA chaperone that modulates translation of many mRNAs and also stabilizes interactions of sRNAs with target RNAs. Previous reports have shown that *hfq* null mutants of pathogens that normally possess Hfq show decreased growth rates, increased sensitivity to stress conditions and impaired virulence (41, 42). The

significance of this observation in *C. burnetii* is unclear. Either *Coxiella*'s sRNAs are Hfq-independent, similar to many Gram-positive bacteria (43), or the bacterium possess an atypical Hfq, as reported for *Borrelia burgdorferi* (44).

In the past few years, sRNAs have been identified in many bacteria; however, there are few reports on characterization of their target(s). Various computational and experimental approaches have been employed in order to identify these targets (45). With the aim of predicting potential roles for the identified CbSRs, we used TargetRNA2 software (46) to predict mRNA targets that could base pair with the sRNAs. However, none showed significant binding to a specific and prominent mRNA target. We can speculate on the roles of some of the sRNAs based on the location of their coding sequence relative to neighboring genes. In the case of antisense sRNAs, the RNA that they regulate could be the corresponding mRNAs. Further, most of the antisense and ORF-derived sRNAs are less abundant than intergenic sRNAs indicating that they preferably base-pair with mRNAs encoded nearby. Unfortunately, since most of these genes are pseudogenes or encode hypothetical proteins, their regulatory function is difficult to predict based on location alone. Interestingly, CbSR 14 is transcribed antisense to the 5' UTR of *trmE*, a bacterial tRNA modification GTPase that has been implicated in ribosome assembly and other cellular processes including stress response, sporulation and pathogenesis (47). Since these functions would likely be advantageous to *C. burnetii*, CbSR 14 possibly regulates *trmE*, however, this hypothesis must be experimentally validated.

In conclusion, this study is the first step towards elucidating sRNA-mediated regulation of *C. burnetii*'s physiology and pathogenesis. Further investigations are

required to determine the exact role played by each CbSR to help *C. burnetii* transition between the two different cell morphotypes and adapt to the intracellular niche.

F. Acknowledgments

We thank Drs. Karen M. Wassarman and Jeffrey E. Barrick for helpful discussions, technical assistance and unpublished data. We are also grateful to Karen M. Wassarman for the generous gift of RNAP-specific antibodies used in this study. This work was supported by NIH grant R15AI103511.

G. References

1. **Papenfort K, Vogel J.** 2010. Regulatory RNA in bacterial pathogens. Cell Host Microbe **8:**116-127.
2. **Waters LS, Storz G.** 2009. Regulatory RNAs in bacteria. Cell **136:**615-628.
3. **Mandin P, Gottesman S.** 2010. Integrating anaerobic/aerobic sensing and the general stress response through the ArcZ small RNA. Embo j **29:**3094-3107.
4. **Romby P, Vandenesch F, Wagner EG.** 2006. The role of RNAs in the regulation of virulence-gene expression. Curr Opin Microbiol **9:**229-236.
5. **Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA.** 2010. *Legionella pneumophila* 6S RNA optimizes intracellular multiplication. Proc Natl Acad Sci U S A **107:**7533-7538.
6. **Wassarman KM, Storz G.** 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell **101:**613-623.
7. **Maurin M, Raoult D.** 1999. Q fever. Clin Microbiol Rev **12:**518-553.

8. **Romano PS, Gutierrez MG, Beron W, Rabinovitch M, Colombo MI.** 2007. The autophagic pathway is actively modulated by phase II *Coxiella burnetii* to efficiently replicate in the host cell. Cell Microbiol **9:**891-909.

9. **Coleman SA, Fischer ER, Howe D, Mead DJ, Heinzen RA.** 2004. Temporal analysis of *Coxiella burnetii* morphological differentiation. J Bacteriol **186:**7344-7352.

10. **McCaul TF, Williams JC.** 1981. Developmental cycle of *Coxiella burnetii*: structure and morphogenesis of vegetative and sporogenic differentiations. J Bacteriol **147:**1063-1076.

11. **Weissenmayer BA, Prendergast JG, Lohan AJ, Loftus BJ.** 2011. Sequencing illustrates the transcriptional response of *Legionella pneumophila* during infection and identifies seventy novel small non-coding RNAs. PLoS One **6:**e17570.

12. **Tesorero RA, Yu N, Wright JO, Svencionis JP, Cheng Q, Kim JH, Cho KH.** 2013. Novel regulatory small RNAs in *Streptococcus pyogenes*. PLoS One **8:**e64021.

13. **Gong H, Vu GP, Bai Y, Chan E, Wu R, Yang E, Liu F, Lu S.** 2011. A *Salmonella* small non-coding RNA facilitates bacterial invasion and intracellular replication by modulating the expression of virulence factors. PLoS Pathog **7:**e1002120.

14. **Koo JT, Alleyne TM, Schiano CA, Jafari N, Lathem WW.** 2011. Global discovery of small RNAs in *Yersinia pseudotuberculosis* identifies Yersinia-specific small, noncoding RNAs required for virulence. Proc Natl Acad Sci U S A **108:**E709-717.

15. **Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P.** 2007. Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets. Nucleic Acids Res **35:**962-974.

16. **Mann B, van Opijnen T, Wang J, Obert C, Wang YD, Carter R, McGoldrick DJ, Ridout G, Camilli A, Tuomanen EI, Rosch JW.** 2012. Control of virulence by small RNAs in *Streptococcus pneumoniae*. PLoS Pathog **8:**e1002788.

17. **Williams JC, Peacock MG, McCaul TF.** 1981. Immunological and biological characterization of *Coxiella burnetii*, phases I and II, separated from host components. Infect Immun **32:**840-851.

18. **Cockrell DC, Beare PA, Fischer ER, Howe D, Heinzen RA.** 2008. A method for purifying obligate intracellular *Coxiella burnetii* that employs digitonin lysis of host cells. J Microbiol Methods **72:**321-325.

19. **Raghavan R, Hicks LD, Minnick MF.** 2008. Toxic introns and parasitic intein in *Coxiella burnetii*: legacies of a promiscuous past. J Bacteriol **190:**5934-5943.

20. **Omsland A, Cockrell DC, Howe D, Fischer ER, Virtaneva K, Sturdevant DE, Porcella SF, Heinzen RA.** 2009. Host cell-free growth of the Q fever bacterium *Coxiella burnetii*. Proc Natl Acad Sci U S A **106:**4430-4434.

21. **Sandoz KM, Sturdevant DE, Hansen B, Heinzen RA.** 2014. Developmental transitions of *Coxiella burnetii* grown in axenic media. J Microbiol Methods **96:**104-110.

22. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics **26:**589-595.

23. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944-945.

24. **Solovyev V, Salamov A.** 2011. Automatic Annotation of Microbial Genomes and Metagenomic Sequences. Nova Science Publishers.

25. **Kingsford CL, Ayanbule K, Salzberg SL.** 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol **8:**R22.

26. **Pall GS, Hamilton AJ.** 2008. Improved northern blot method for enhanced detection of small RNA. Nat Protoc **3:**1077-1084.

27. **Schneider CA, Rasband WS, Eliceiri KW.** 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods **9:**671-675.

28. **Sahr T, Rusniok C, Dervins-Ravault D, Sismeiro O, Coppee JY, Buchrieser C.** 2012. Deep sequencing defines the transcriptional map of *L. pneumophila* and identifies growth phase-dependent regulated ncRNAs implicated in virulence. RNA Biol **9:**503-519.

29. **Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S.** 2001. Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. Curr Biol **11:**941-950.

30. **Vogel J, Bartels V, Tang TH, Churakov G, Slagter-Jager JG, Huttenhofer A, Wagner EG.** 2003. RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. Nucleic Acids Res **31:**6435-6443.

31. **Toledo-Arana A, Repoila F, Cossart P.** 2007. Small noncoding RNAs controlling pathogenesis. Curr Opin Microbiol **10:**182-188.

32. **Liu JM, Livny J, Lawrence MS, Kimball MD, Waldor MK, Camilli A.** 2009. Experimental discovery of sRNAs in *Vibrio cholerae* by direct cloning, 5S/tRNA depletion and parallel sequencing. Nucleic Acids Res **37:**e46.

33. **Raghavan R, Groisman EA, Ochman H.** 2011. Genome-wide detection of novel regulatory RNAs in *E. coli*. Genome Res **21:**1487-1497.

34. **Yan Y, Su S, Meng X, Ji X, Qu Y, Liu Z, Wang X, Cui Y, Deng Z, Zhou D, Jiang W, Yang R, Han Y.** 2013. Determination of sRNA expressions by RNA-seq in *Yersinia pestis* grown in vitro and during infection. PLoS One **8:**e74495.

35. **Mraheil MA, Billion A, Mohamed W, Mukherjee K, Kuenne C, Pischimarov J, Krawitz C, Retey J, Hartsch T, Chakraborty T, Hain T.** 2011. The intracellular sRNA transcriptome of *Listeria monocytogenes* during growth in macrophages. Nucleic Acids Res **39:**4235-4248.

36. **Padalon-Brauch G, Hershberg R, Elgrably-Weiss M, Baruch K, Rosenshine I, Margalit H, Altuvia S.** 2008. Small RNAs encoded within genetic islands of *Salmonella typhimurium* show host-induced expression and role in virulence. Nucleic Acids Res **36:**1913-1927.

37. **Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL.** 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell **118:**69-82.

38. **Kazantsev AV, Pace NR.** 2006. Bacterial RNase P: a new view of an ancient enzyme. Nat Rev Microbiol **4:**729-740.

39. **Moore SD, Sauer RT.** 2007. The tmRNA system for translational surveillance and ribosome rescue. Annu Rev Biochem **76:**101-124.

40. **Wassarman KM.** 2007. 6S RNA: a regulator of transcription. Mol Microbiol **65:**1425-1431.

41. **Christiansen JK, Larsen MH, Ingmer H, Sogaard-Andersen L, Kallipolitis BH.** 2004. The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. J Bacteriol **186:**3355-3362.

42. **Sonnleitner E, Hagens S, Rosenau F, Wilhelm S, Habel A, Jager KE, Blasi U.** 2003. Reduced virulence of a hfq mutant of *Pseudomonas aeruginosa* O1. Microb Pathog **35:**217-228.

43. **Romby P, Charpentier E.** 2010. An overview of RNAs with regulatory functions in gram-positive bacteria. Cell Mol Life Sci **67:**217-237.

44. **Lybecker MC, Abel CA, Feig AL, Samuels DS.** 2010. Identification and function of the RNA chaperone Hfq in the Lyme disease spirochete *Borrelia burgdorferi*. Mol Microbiol **78:**622-635.

45. **Vogel J, Wagner EG.** 2007. Target identification of small noncoding RNAs in bacteria. Curr Opin Microbiol **10:**262-270.

46. **Tjaden B.** 2008. TargetRNA: a tool for predicting targets of small RNA action in bacteria. Nucleic Acids Res **36:**W109-113.

47. **Verstraeten N, Fauvart M, Versees W, Michiels J.** 2011. The universally conserved prokaryotic GTPases. Microbiol Mol Biol Rev **75:**507-542, second and third pages of table of contents.

CHAPTER FOUR

Investigation of *Coxiella burnetii*'s 6S RNA and examination of its regulatory role in growth and development

As published in part in: PLoS One. 2014 Jun 20;9(6):e100147. doi: 10.1371/journal.pone.0100147

A. Introduction

6S RNA was first identified in *E. coli*, where it was found to accumulate during stationary phase and interact with the RNA polymerase (RNAP)-σ^{70} complex (2). 6S RNA forms an extended hairpin structure that is highly conserved among eubacteria, consisting of a single-stranded central bubble that resembles an open promoter, flanked by a closing stem and terminal loop (4). As a result, 6S RNA associates with RNAP and occupies the active site, thereby interfering with formation of a transcription initiation complex (4) and blocking transcription. Apart from its role in transcription, functions of 6S RNA also include: upregulation of genes involved in stress response and nutrient acquisition (5), long-term survival during stationary phase (6) and regulation of *relA* and ppGpp synthesis during the stringent response (7). In *E. coli*, 6S RNA is a major regulator of transcription; negatively affecting 148 genes while activating transcription of 125 genes (8). In *L. pneumophila*, a close relative of *C. burnetii*, 6S RNA positively regulates several genes, many of which enhance intracellular growth and virulence (5). Early work by our lab showed a prominent RNA of ~200 nt in length that was more prominently expressed in SCVs compared to LCVs and eventually identified as 6S RNA (Fig. 4-1). We mapped the *ssrS* gene by in silico analysis and unpublished 6S RNA

Figure 4-1. *C. burnetii* total RNA separated on a denaturing acrylamide gel. RNA isolated from *C. burnetii* LCVs (3 dpi) and SCVs (14 dpi) grown in Vero host cells was separated on a denaturing 8 M urea - 8% acrylamide gel and stained with ethidium bromide (5 µg RNA per lane). Arrow indicates the position of 6S RNA at ~200 nucleotides. The number of nucleotides in the RNA size standards (Std) is indicated to the left.

sequence data [kindly provided by Ronald Breaker (4)]. The *ssrS* gene is located in the 5′ untranslated region (UTR) of *C. burnetii*'s *ygfA* locus (encoding formyl tetrahydrofolate cyclo-ligase; CBU_0066) (Fig. 4-2), a linkage that is highly conserved among α and γ-proteobacteria (4). Here, we characterize the 6S RNA of *C. burnetii.* We found that 6S RNA specifically binds to *C. burnetii*'s RNAP, reaches its highest concentration in SCVs, and its expression is markedly increased during intracellular versus axenic growth.

B. Materials and Methods

Cultivation of *C. burnetii.* *C. burnetii* Nine Mile phase II (strain RSA439, clone 4) was propagated in African green monkey kidney (Vero) fibroblast cells (CL-81; American Type Culture Collection) grown in RPMI medium (Invitrogen Corp.) supplemented with 10% fetal bovine serum at 37^0C in a 5% CO_2 atmosphere. Bacteria were purified from host cells using differential centrifugation, as previously described (9). LCVs were harvested at 72 h post-infection from infected cells using digitonin (10). SCVs were harvested and prepared at 21 days post-infection (dpi), as previously described (11), and used to infect Vero cell monolayers for the production of synchronized bacterial cultures. *C. burnetii* was also cultivated axenically in ACCM2 at 37^0C in a tri-gas incubator (2.5% O_2, 5% CO_2, 92.5% N_2) with continuous shaking at 75 RPM (12). To generate LCVs, ACCM2 was inoculated with 10-d-old ACCM2-cultured bacteria, incubated 72 h, and isolated by centrifugation (10,000 x g for 20 min at 4^0C), as previously described (11). SCV generation in ACCM2 was identical to LCVs except bacteria were grown for 7 d, and then flask lids were tightened and cultured an additional 14 d on the lab bench (~25^0C) without replenishing the medium (13).

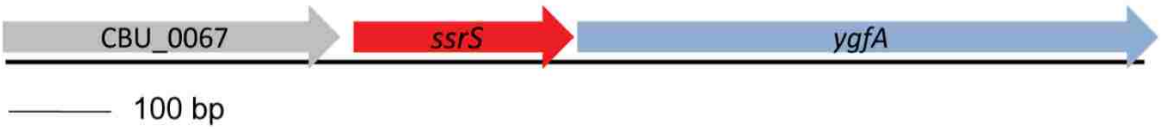


Figure 4-2. Linkage map showing the location of *C. burnetii*'s 6S RNA gene (*ssrS*). *ssrS* is encoded in the 5' untranslated region (UTR) of *ygfA* (encoding formyl tetrahydrofolate cyclo-ligase; CBU_0066). The gene immediately upstream (CBU_0067) encodes a hypothetical protein.

Northern Blot Analysis. Northern blots were carried out using a NorthernMax kit (Ambion) as per manufacturer's instructions. Briefly, total RNAs of *C. burnetii* grown in Vero cells or ACCM2 were isolated by sequential use of Ribopure (Ambion) and DNA-free (Applied Biosystems) kits and then precipitated with 100% ethanol. RNA (3 µg per lane) was electrophoresed through 1.5% agarose formaldehyde gels and blotted onto positively-charged BrightStar Plus nylon membranes (Ambion). Membranes were then UV cross-linked or chemically cross-linked by 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (Sigma-Aldrich), as previously described (14). Hybridizations were carried out using single-stranded RNA probes specific to 6S RNA (Table. 4-1). 6S RNA probes were generated by T7 promoter-mediated *in vitro* transcription of PCR products (PCR primers, Table 4-2) using a MEGAscript kit, as instructed (Ambion), in the presence of biotin-labeled UTP (Bio-16-UTP; Ambion). Finally, membranes were developed with a BrightStart BioDetect kit (Ambion) following the manufacturer's protocol, and visualized using a LAS-3000 imaging system (Fujifilm). Densitometry was performed using ImageJ software (15).

Quantitative PCR (qPCR) and quantitative real-time PCR (qRT-PCR). qPCR was performed as previously described (16) using a primer set specific to *C. burnetii*'s *rpoS* gene (Table 4-3) for generation of a growth curve showing genome numbers as a function of time (11). Primers specific to *C. burnetii*'s 6S RNA encoding gene (*ssrS*) (Table 4-3) were designed using Beacon Designer 7.5 software (Biosoft International). qRT-PCR data were obtained with a 6S RNA primer set and normalized to corresponding *C. burnetii* genome numbers.

Table 4-1. Probes used in Northern blots and RPAs.

Probe	Sequence
6S RNA	AATATAAGTGTATCCTCTGTGACTCGTGGCAAGGACCACATATTTGAACCGATAC GAATATGATAGGGAATTGGCTGTGGTCACACTGTTGAGCAAGCCCGTTTTCGGG GTCTCATAACC
5S RNA	CCACCTGATTCCATTCCGAACTCAGAAGTGAAAACGCTTAGCGCCGATGATAGT GTGGGTCTCCCATGCGAAAGT

Table 4-2. PCR primers and their sequence*

Primers	Sequence
6SRNA_Forward	AATATAAGTGTATCCTCTGT
6SRNA_Reverse	TAATACGACTCACTATAGGGGTTATGAGACCCCGAAAAC
5SRNA_Forward	CCACCTGATTCCATTCCGAACTCAGAAG
5SRNA_Reverse	TAATACGACTCACTATAGGGACTTTCGCATGGGAGACC

*The T7 promoter sequence is underlined.

Table 4-3. qPCR and qRT-PCR primers.

Primers	Sequence 5' to 3'	Source or Reference
6S_Forward	ACCCCTAAGGGAAGCCTGAA	This study
6S_Reverse	TTGAACCCAAAGGCTCAAGTG	This study
RpoS_Forward	CGCGTTCGTCAAATCCAAATA	(11)
RpoS_Reverse	GACGCCTTCCATTTCCAAAA	(11)

***C. burnetii* extract preparation.** A mixed population (11 dpi) of *C. burnetii* grown in Vero cells was pelleted by centrifugation (10,000 x g for 10 min at 4^{o}C) and resuspended in 250 µl Net2 buffer [50 mM Tris (pH 7.4), 150 mM NaCl, 0.05% NP-40 (triton X-100)] supplemented with protease inhibitor (Complete Mini Protease inhibitor cocktail tablets used as instructed; Roche). RNasin Plus (Promega) was added to a final concentration of 1 U/µl and bacteria were lysed by five alternating freeze-thaws cycles in liquid nitrogen and a 37^{o}C water bath (5 min each). The resulting lysate was clarified by centrifugation (10,000 x g for 10 min at 4^{o}C), and the supernatant was used for further analysis.

Immunoprecipitation (IP). Protein A Sepharose (PAS) beads (CL-4B; GE Healthcare) were swelled (2 mg PAS in 100 µl Net2 buffer) for 30 min at room temperature and washed three times with 400 µl cold Net2 buffer followed by centrifugation (400 x g for 30 sec). IPs were carried out using rabbit anti-*Escherichia coli* RNAP core polyclonal antibody (a generous gift from Dr. Karen Wassarman, University of Wisconsin-Madison), a corresponding rabbit pre-immune serum or rabbit anti-*Coxiella* Com1 polyclonal antibody. Antibodies were incubated with 100 µl PAS-Net2 at a 1:50 dilution for 16 h at 4^{o}C with gentle agitation. PAS-antibody conjugates were then washed five times with 400 µl cold Net2 buffer, as above. *C. burnetii* extract (25 µl) was added to each PAS-antibody conjugate and incubated for 2 h at 4^{o}C with rocking. IP reactions were separated by centrifugation, and PAS beads and supernatants were retained for further analysis. PAS beads were washed five times as above, and the final pellet resuspended in Net2 buffer (200 µl). Approximately 20% of the IP suspension was used for protein analysis and 80% for RNA analysis.

Protein analysis. IP beads and supernatants were mixed with equal volumes of 2X Laemmli sample buffer, boiled for 5 min and centrifuged 1 min at 16,000 x g. The resulting supernatants were resolved on a 10% acrylamide SDS-PAGE gel. The gel was immediately blotted onto a nitrocellulose membrane (0.45 µm pore size) and blocked for 2 h at room temperature in blocking buffer [5% nonfat dry milk in TBS-T (25 mM Tris-HCl, pH 8.0; 125 mM NaCl; 0.1% Tween 20)] with rocking. Blots were subsequently probed for 16 h with a 1:2000 dilution of anti-RNAP antibody in antibody binding buffer (TBS-T containing 1% nonfat dry milk) followed by 5 washes of 10 min each in TBS-T. Blots were then incubated for 1 h at room temperature with rocking in a 1:2000 dilution of peroxide-conjugated goat anti-rabbit IgG antibodies (Sigma) in antibody binding buffer, followed by 5 washes (10 min each) in TBS-T. Finally, blots were developed using a chemiluminescent substrate as instructed by the manufacturer (SuperSignal West Pico kit, Thermo Scientific) and visualized using a LAS-3000 imaging system (Fujifilm).

RNA extraction and RNase Protection Assay (RPA). Total RNA from IP beads and supernatant was isolated by extraction with phenol:chloroform:isoamyl alcohol [25:24:1 ; v/v; (pH 8-8.3)] (Invitrogen) followed by ethanol precipitation. Purified RNA was processed using an RNase Protection assay (RPA) III kit (Ambion) as per manufacturer's instructions. Specifically, 43 ng of RNA and 4.3 pg of probe were used in each reaction, except in the IP from the anti-Com1 antibody, where 22.8 ng RNA was used. The 6S RNA probe prepared for Northern blot analysis was also used in RPAs. RPA reactions were resolved on gels (5% acrylamide; 8 M urea), transferred to BrightStar-Plus nylon membrane (Ambion) and UV-cross-linked. RPA blots were

developed using a BrightStar BioDetect kit as instructed (Ambion) and visualized with a LAS-3000 imaging system (Fujifilm) [Please see Table 4-1 for probe details].

Construction of 6S-overexpression and knockdown vectors

The pKM244 plasmid (provided by Jim Samuel, Texas A&M University), a derivative of pKM230 (17), was used to generate the 6S-knockdown plasmid, p6Skd. An IPTG-inducible *lac* promoter (Plac) and *lac* operator (lacO) were PCR amplified from pUC19 plasmid using Plac_lacO+EcoRI Forward and Plac_lacO+NcoI Reverse primers Table 4-4) and *ssrS* was PCR amplified from NMII genomic DNA using 6S antisense For_XmaI and 6S antisense Rev_NcoI. The amplified *ssrS* gene of *C. burnetii* was ligated to Plac_lacO, in the opposite orientation with respect to the *lac* promoter, to generate a Plac_lacO+6Santisense amplicon. The p6Skd plasmid was generated by ligating Plac_lacO+6Santisense and pKM244 following digestion with EcoRI and XmaI thereby cloning the *ssrS* gene into the multiple cloning site of pKM244 to take advantage of its *tac* promoter. The 6S-overexperssion plasmid, p6Sovrexp, was generated by Linda D. Hicks of our lab, using the same principle except that the *ssrS* gene was ligated to Plac_lacO in the same orientation with respect to the *lac* promoter to generate Plac_lacO+6Ssense, which was ligated into pKM244 as described above.

Isolation of 6S-overexpression and knockdown clones

C. burnetii was axenically cultured in ACCM2 from an inoculum of 2 X 10^6 genome equivalents per ml for 7 d at 37^0C in a 2.5% O_2 and 5% CO_2 environment, as previously described (12). Bacterial cells were washed five times in cold 10% glycerol and resuspended in the same, at an approximate concentration of 10^{11} cells/ml. *C. burnetii*

Table 4-4. Primers used for cloning *ssrS*.

Primers	**Sequence**
Plac_lacO+EcoRI Forward	TTTTGAATTCTTTACACTTTATGCTTCC
Plac_lacO+NcoI Reverse	TTTCCATGGAATTGTTATCCGCTCAC
6S antisense For_XmaI	TTTCCCGGGAATATAAGTGTATCCTCTGT
6S antisense Rev_NcoI	TTTCCATGGTTGAATACATAAAACCCTCT
6S sense For_NcoI	TTTCCATGGAATATAAGTGTATCCTCTGT
6S sense Rev_XmaI	TTTCCCGGGTTGAATACATAAAACCCTCT

was mixed with 3.78 µg of p6Skd or pKM244 and electroporated at 1.8 kV, 400Ω and 25 µF as previously described (12). *C. burnetii* transformants were rescued in 6 ml ACCM-2 containing 1% fetal bovine serum (FBS) and incubated overnight before addition of chloramphenicol (final concentration 3 µg/ml). Transformants were cultured for 7 d and a 10 µl aliquots were plated on chloramphenicol (3 µg/ml) containing ACCM-2 agarose. The clonality of the transformants was confirmed by performing two rounds of serial dilution followed by plating on chloramphenicol (5 µg/ml) containing ACCM2. At each stage, the expression of mCherry was monitored by microscopy. Plasmid preparations were performed on *C. burnetii* NMII transformed with p6Skd or p6Sovrexp and were sequenced to confirm plasmid content and orientation of *ssrS*

IPTG induction experiments using *C. burnetii* NMII *ssrS* clones

C. burnetii strains harboring p6Skd, p6Sovrexp or pKM244 were used to examine the IPTG inducibility of the plasmids. ACCM2 (20 ml) was inoculated with 10^6 GE/ml of each clone and cultured for 3 d at 37^0C in 2.5% O_2 and 5% CO_2 with shaking, as previously described (12). Culture volumes were then divided and one half supplemented with 1 mM IPTG. Cells were isolated at 24 h and 48 h post-induction by centrifugation at 16,000 x g for 16 min at 4^0C. DNA and RNA were isolated from the cells using a DNeasy Blood and Tissue kit (Qiagen) and Ribopure kit (Amibion), respectively, as per manufacturer's instructions. Resulting RNA was treated with TURBO DNase (Ambion), precipitated with 100% ethanol and quantified by spectrophotometry. 50 ng of RNA from each growth time-point was converted into cDNA using an iScript cDNA synthesis kit (Bio-Rad), as instructed by the manufacturer. Genome numbers were determined by qPCR with a primer set specific to the *C. burnetii rpoS* gene (11). cDNA was used to

perform qRT-PCR with *ssrS* primers (Table 4-3), as previously described (16). Amplified cDNA was normalized to the respective genome numbers.

IPTG induction experiment using *E. coli ssrS* clones

E. coli (TOP10F') was transformed with p6Skd, p6Sovrexp and pKM244 as instructed by the manufacturer (Invitrogen). Resulting *E. coli* strains were grown overnight at 37^0C in Luria-Bertani (LB) broth containing 5 µg/ml chloramphenicol (LBcam). These cultures were used to inoculate fresh LBcam at a 1:10 (v/v) dilution, and the mixture was grown to mid-logarithmic phase (2 h) at 37^0C. IPTG was added to 1 mM and cells were isolated following 2 h growth at 37^0C with shaking. DNA and RNA were isolated using DNeasy Blood and Tissue kit (Qiagen) and Ribopure kit (Ambion), respectively, as per manufacturer's instructions. Resulting RNA was treated with TURBO DNase (Ambion), precipitated with 100% ethanol and quantified by spectrophotometry. RNA (500 ng) was converted to cDNA using an iScript cDNA synthesis kit (Bio-Rad), as instructed by the manufacturer. Genome numbers were determined by quantitative PCR (qPCR) with a primer set specific to the *E. coli* 16S rDNA. cDNA was used to perform qRT-PCR with the *ssrS* primer set (Table 4-3), as previously described (16). Amplified cDNA was normalized to the respective genome numbers.

C. Results

Characterization of *C. burnetii*'s 6S RNA

The *ssrS* gene of *C. burnetii* was mapped to the 5′ UTR of *ygfA* (Fig. 4-2). A linkage of *ssrS* and *ygfA* is conserved among many bacterial species (4). Also, the predicted secondary structure of the *C. burnetii* 6S RNA was found to be highly similar to the

published consensus structure of 6S RNA, consisting of a single-stranded central bubble, including two conserved G-C base pairs surrounding the bubble on both sides, flanked by a closing stem and terminal loop (Fig. 4-3) (4). The central bubble mimics the structure of a DNA template in an open promoter complex and also occupies the active site of the RNAP. These observations suggest that the 6S RNA of *C. burnetii* is functional.

To confirm the identity of the presumed 6S RNA band, we performed Northern blot analyses of RNA isolated from both morphotypes of *C. burnetii* cultured in Vero cells and in ACCM2. Northern blot analyses were also performed on total RNA isolated from SCVs at 14 dpi and 21 dpi to compare 6S RNA levels at early and late stationary phase. Blots were then probed with a biotinylated RNA designed from the 5′ UTR of *C. burnetii*'s *ygfA* locus. The resulting Northern blot validated the identity of 6S RNA and the size was observed to be ~185 nt, which we later confirmed by RNA-seq (not shown). Furthermore, 6S RNA was found to accumulate in SCVs relative to LCVs, irrespective of growth conditions (Fig. 4-4). This is similar to the 11-fold increase reported for *E. coli* 6S RNA during stationary phase versus exponential phase (2). Levels of 6S RNA in *C. burnetii* cultured in Vero cells were ~9-fold higher in SCVs at 14 dpi compared to LCVs at 3 dpi when blots were analyzed by densitometry. Levels of 6S RNA dropped ~2 fold between 14 dpi and 21 dpi (Fig. 4-4, lane 1 compared to lanes 2 and 3). On the other hand, the transcript level of 6S RNA in *C. burnetii* cultured axenically was ~2-fold higher in SCVs at 14 dpi compared to LCVs at 3 dpi and then remained stable through 21 dpi (Fig. 4-4, lane 4 compared to lanes 5 and 6). To further analyze the increased transcript level of 6S RNA during the SCV phase, we used qRT-

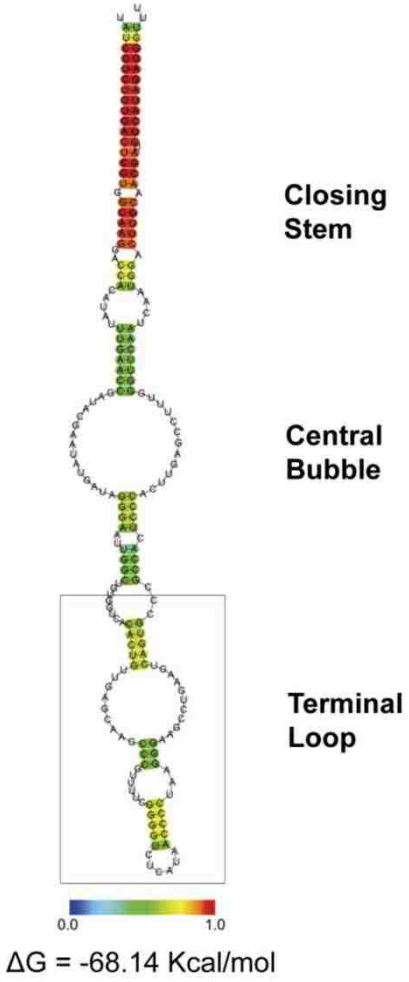


Figure 4-3. Predicted secondary structure of *C. burnetii*'s 6S RNA as determined by Centroidfold (1). The color scale at the bottom represents a heat color gradation from blue to red, corresponding to base-pairing probability from 0 to 1, respectively. The free energy of the structure is also shown.

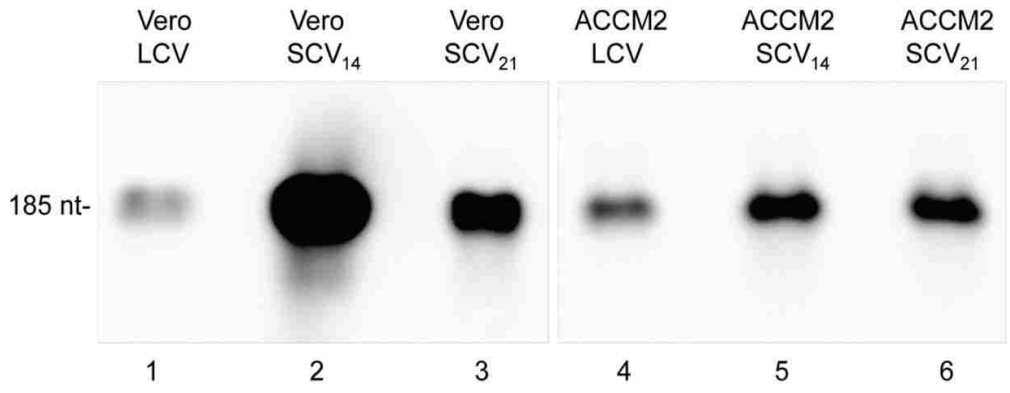


Figure 4-4. Northern blots showing 6S RNA levels of *C. burnetii*. RNA was isolated from LCVs at 3 days post-infection (dpi) and SCVs (SCV_{14}, 14 dpi; SCV_{21}, 21 dpi) grown in Vero host cells and ACCM2, respectively. Hybridizations were performed at high stringency using a 6S RNA-specific biotinylated oligonucleotide probe. 3 µg RNA was used for all lanes. The size of the signal is indicated to the left.

PCR to quantify and compare *C. burnetii* genome numbers to 6S RNA levels over a 14-d infection period in Vero cells (Fig. 4-5A). Results showed the greatest increase (~6-fold) in 6S RNA at 14 d as compared to 0 d of the infection period (Fig. 4-5B). When 6S RNA levels were compared between SCVs isolated from infected Vero cells versus axenic cultures, a ~7-fold higher transcript level was observed in SCVs isolated from Vero cells (Fig. 4-4, lanes 2 versus 5) indicating a potential role for 6S RNA during intracellular growth. A similar observation has been reported for *L. pneumophila*, where 6S RNA was shown to be important for optimal expression of genes during intracellular growth (5).

6S RNA Co-immunoprecipitates with RNAP

Previous studies with *E. coli* (2), *Bacillus subtilis* (18) and *L. pneumophila* (5) have shown a physical interaction between 6S RNA and RNAP. To investigate whether this interaction also exists in *C. burnetii*, we carried out IP studies using a *C. burnetii* lysate and antibodies that recognize *E. coli*'s core RNAP subunits (a generous gift from Dr. Karen Wassarman, University of Wisconsin-Madison). When IP products were analyzed on western blots, two protein bands (~154 kDa and ~43 kDa) were observed that likely correspond to β/β´ and α subunits of *C. burnetii*'s RNAP, respectively (Fig 4-6A, lane 5), based on previous observations in *E. coli* IPs using the same antibody (2). These two bands were not observed in IPs carried out without antibody, irrelevant antibody (anti-*Coxiella* Com1) or the corresponding pre-immune rabbit serum (Fig 4-6A, lanes 2-4, respectively) indicating that the antibody specifically recognizes *C. burnetii*'s RNAP. RPAs on RNA prepared from IP samples showed that 6S RNA was present in IPs prepared using anti-RNAP antibody (Fig. 4-6B, lane 9) and was absent in controls,

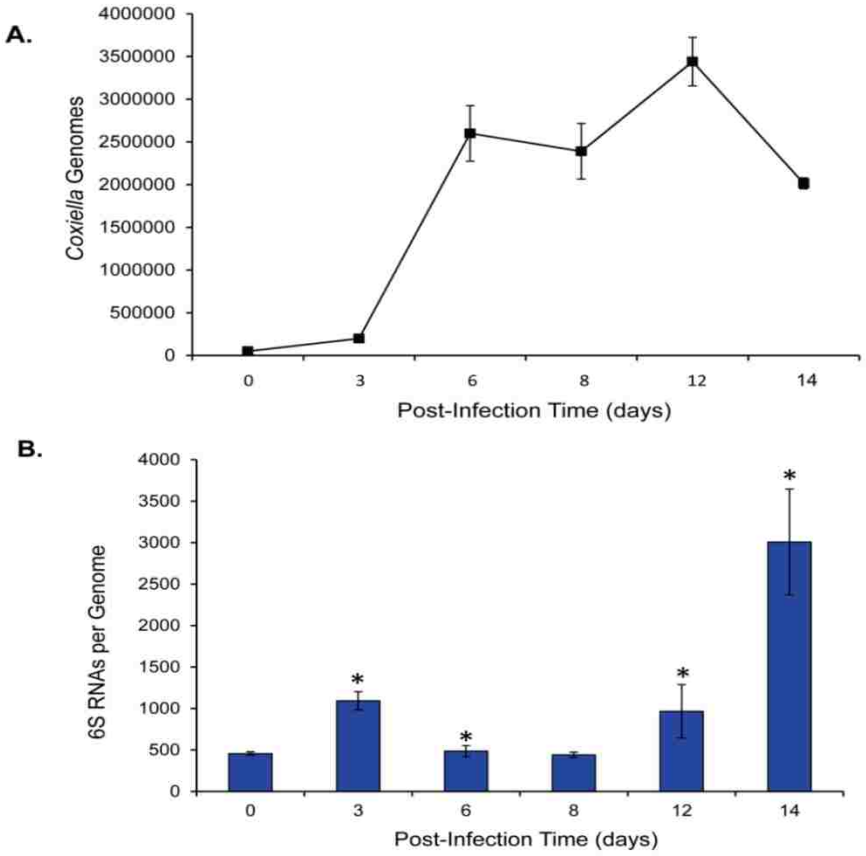
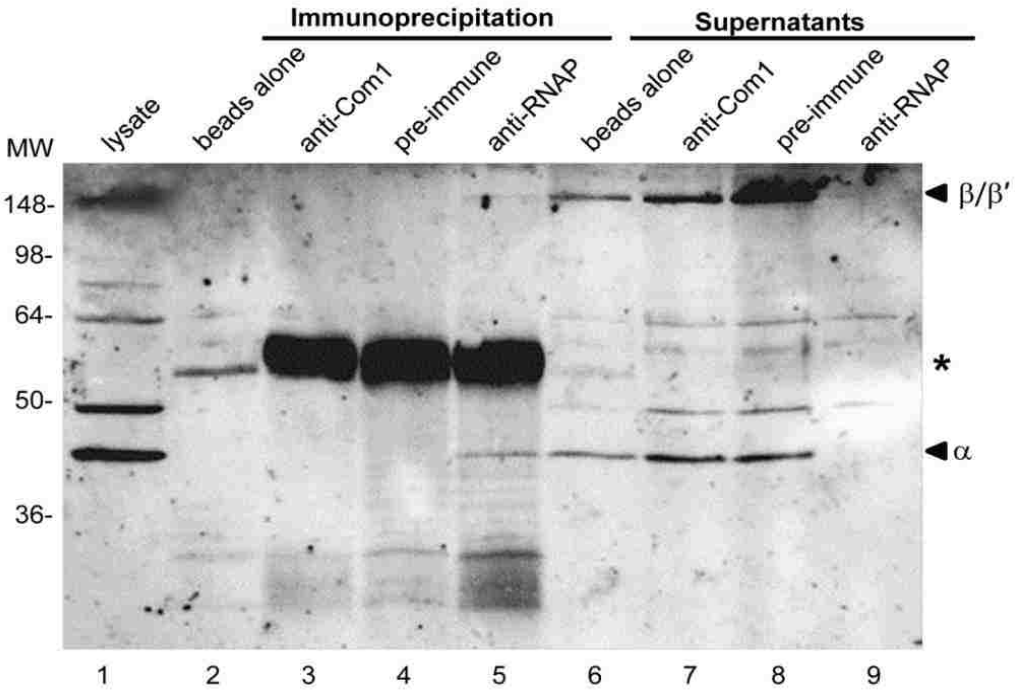


Figure 4-5. *C. burnetii* 6S RNA copies per genome over a 14-d infection period. A. Number of *C. burnetii* genomes over a 14-d infection of Vero cells, as determined by qPCR with a primer set specific to *rpoS*. Values represent the means ± S.D. of the results of 6 independent determinations. **B.** Average number of copies of *C. burnetii* 6S RNAs per genome over a 14-d infection of Vero cells. The number of 6S RNA copies was determined by qRT-PCR using primers specific for 6S RNA and 1 µg total RNA from each time point using the same source cultures as in panel A. Values represent the means ± S.D. of the results of 6 independent determinations. Asterisks denote a significant difference relative to the 0-d sample ($p<0.05$ by student's t test)

A.


Immunoprecipitation
Supernatants
MW
lysate
beads alone
anti-Com1
pre-immune
anti-RNAP
beads alone
anti-Com1
pre-immune
anti-RNAP
148-
98-
64-
50-
36-
β/β′
*
α
1 2 3 4 5 6 7 8 9

B

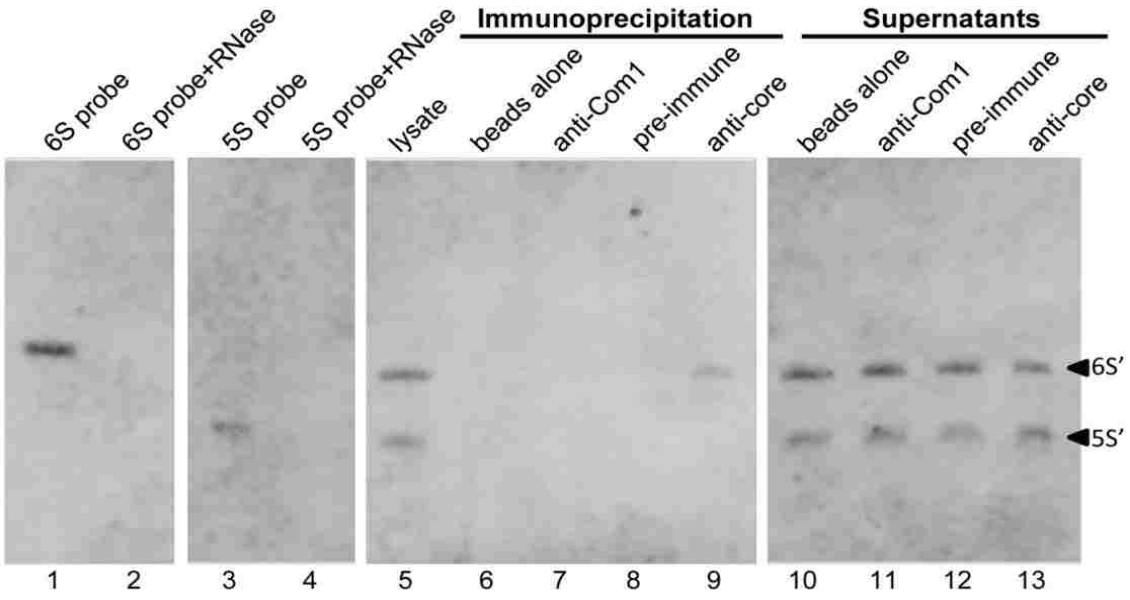

6S probe
6S probe+RNase
5S probe
5S probe+RNase
lysate
Immunoprecipitation
beads alone
anti-Com1
pre-immune
anti-core
Supernatants
beads alone
anti-Com1
pre-immune
anti-core
6S′
5S′
1 2 3 4 5 6 7 8 9 10 11 12 13

Figure 4-6. 6S RNA co-immunoprecipitates with *C. burnetii* RNAP. A. Immunoblot showing IP reactions of a *C. burnetii* lysate and corresponding supernatant samples using various antibodies. IPs were performed with no antibody (lanes 2 and 6), rabbit anti-*Coxiella* Com1 antibody (lanes 3 and 7), pre-immune rabbit serum from the rabbit used to generate anti-RNAP antibodies (lanes 4 and 8) and rabbit anti-RNAP antibody (lanes 5 and 9). The presumed β/β′ and α subunits of RNAP are indicated to the right. Molecular weight values from standards are given to the left in kDa. An asterisk indicates the IgG heavy chain band in lanes 2-5. **B.** RPAs performed on IP samples. Specific biotinylated probes were used to detect samples containing 6S RNA and 5S RNA. 43 ng of RNA and 4.3 pg probe were used in each RPA reaction, except IP-anti-Com1, where 22.8 ng RNA was used. Lanes 1 and 3 contain untreated 6S RNA and 5S RNA probes, respectively, while lanes 2 and 4 contain 6S RNA and 5S RNA probes plus RNase, respectively. The RNase-protected portions of the 6S and 5S RNAs (6S' and 5S'; respectively) are arrowed to indicate the presence or absence of corresponding signals in lanes 5–13

indicating that 6S RNA co-immunoprecipitates with the core RNAP. Further, a 5S RNA control was detected in IP supernatants but was absent in IP samples, indicating that binding of 6S RNA to RNAP is specific.

Overexpression and knockdown of *C. burnetii*'s 6S RNA

To determine the role of 6S RNA in regulating *C. burnetii*'s transcriptome, we generated *ssrS* knockdown and overexpression strains of *C. burnetii*. The pKM244 plasmid, a derivative of pKM230 (17), contains a mCherry gene (mChe) and chloramphenicol resistance marker cassette (CAT) and was used to generate 6S-knockdown (p6Skd) (Fig. 4-7A) and 6S-overexpression (p6Sovrexp) (Fig. 4-7B) vectors. The expression of sense or anti-sense 6S RNA in the plasmid is driven by an IPTG-inducible *tac* promoter (*Ptac*), while repression is maintained by a *lac* operator (*lacO*) and repressor (*lacIq*). These plasmids were first transformed into *E. coli* to test the IPTG-inducibility of the system. *E. coli* transformed with pKM244 was used as a vector-only control. Following IPTG induction of 6S-knockdown and 6S-overexpression strains of *E. coli*, genomic DNA and *Coxiella* 6S sense/antisense RNAs were quantified using qPCR and qRT-PCR, respectively. Results showed that following IPTG-induction, levels of 6S sense/antisense RNA significantly ($p < 0.02$) increased, on a per genome basis (Fig. 4-8). On comparison of induced verses un-induced, the amount of 6S RNA increased up to ~4-fold in *E. coli* (p6Skd) and about ~8-fold in *E. coli* (p6Sovrexp) (Fig. 4-8B), confirming the IPTG-inducibility of the constructs. Since the 6S primers used in qRT-PCR are specific to *C. burnetii* and not *E. coli*, no significant reads were obtained from *E. coli* transformed with pKM244 alone.

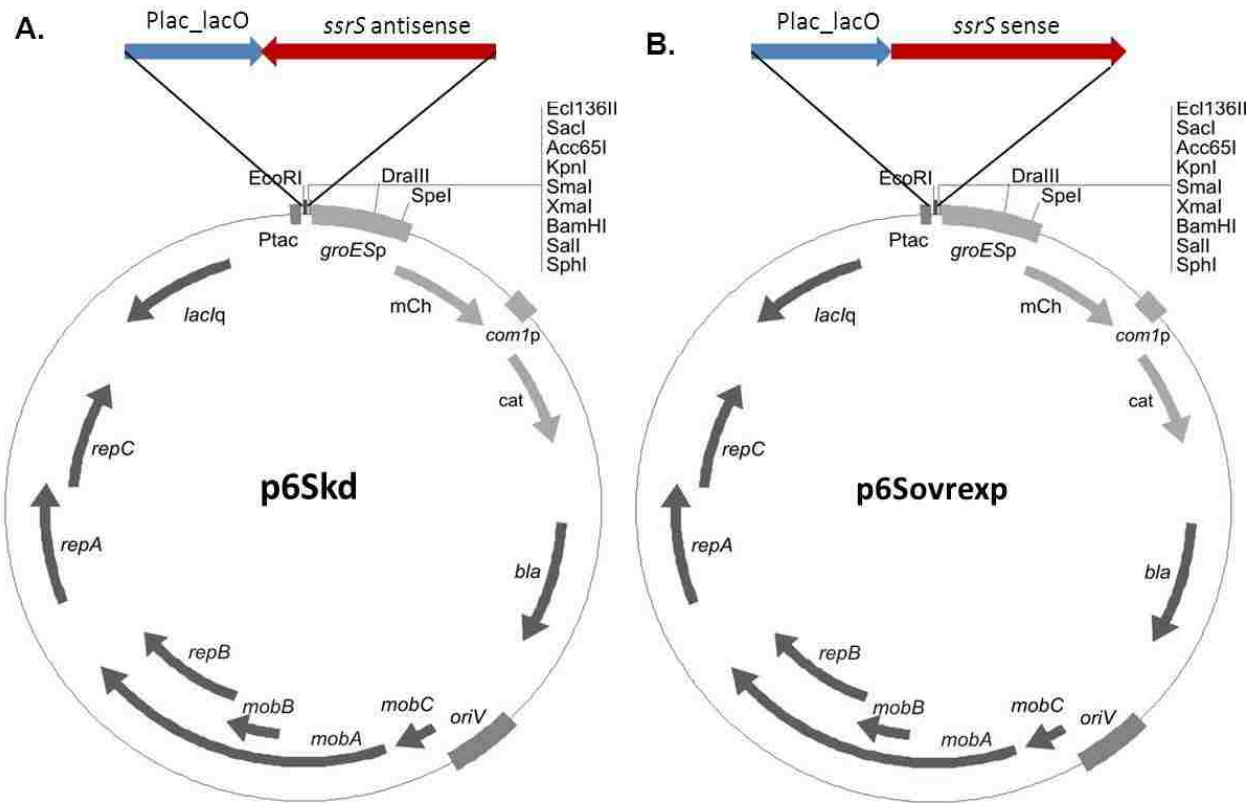


Figure 4-7. Schematic maps of the constructed shuttle vectors to: (A) knock down or (B) overexpress 6S RNA levels in *C. burnetii*.

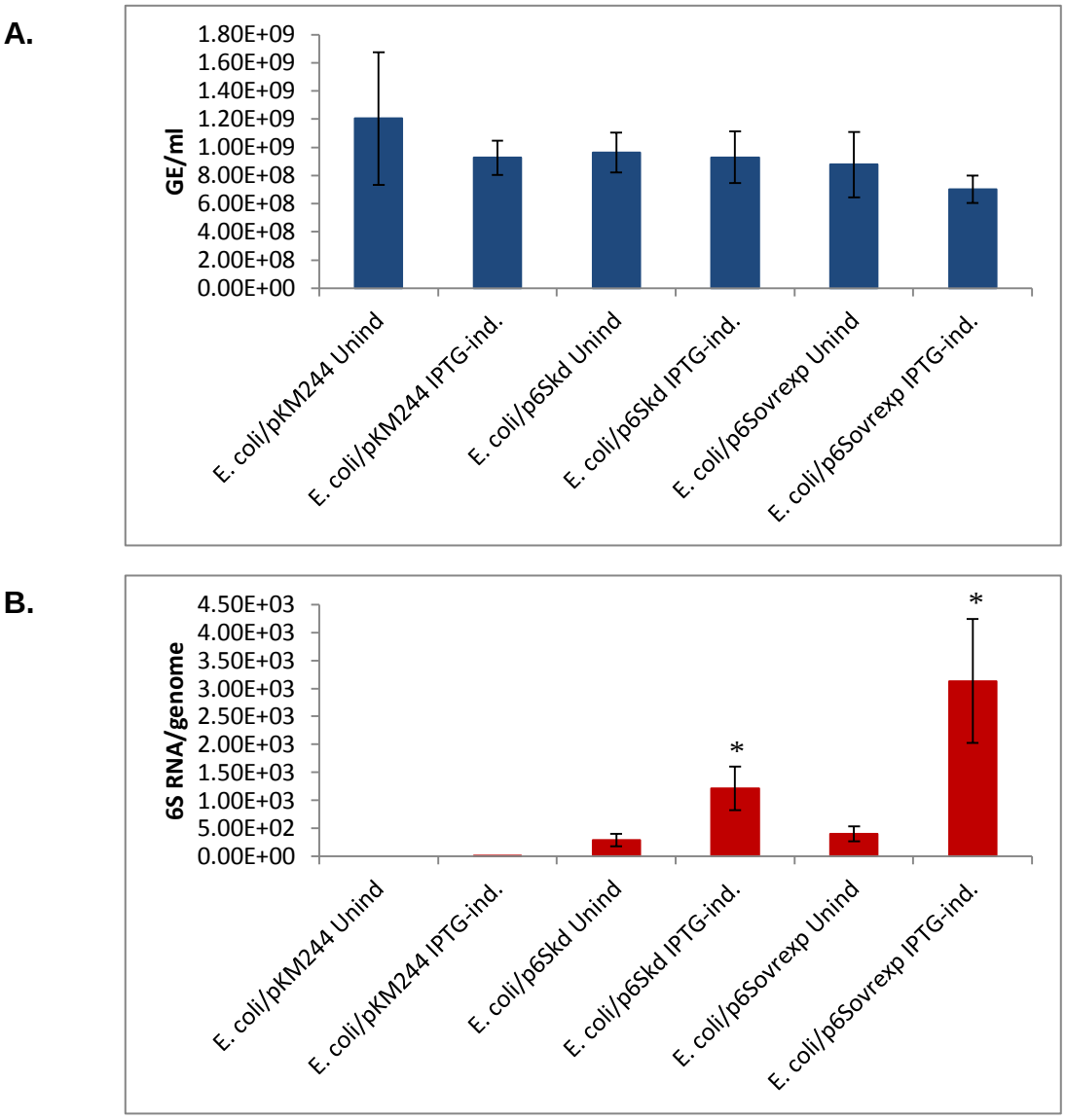


Figure 4-8. Levels of 6S sense/antisense RNAs following IPTG-induction of *E. coli* clones. (A) q-PCR data showing *E. coli* genome equivalents per ml (GE/ml). **(B)** qRT-PCR data showing numbers of 6S sense/antisense RNAs per genome in the presence or absence of 1 mM IPTG. Data represent the means of six independent experiments ± S.D. Asterisks denote statistically significant increases in 6S RNA levels in IPTG-induced clones as compared to their un-induced counterparts ($P < 0.02$).

To identify the set of genes that are regulated by 6S RNA, p6Skd, p6Sovrexp and pKM244 plasmids were transformed into axenically-cultured *C. burnetii*. Individual clones of these strains were isolated by multiple rounds of serial dilution followed by plating on ACCM2-agarose plates containing chloramphenicol (3 µg/ml). Verification of each clone's plasmid content was performed by PCR, mCherry expression (by fluorescence microscope) and a chloramphenicol-resistant phenotype. Plasmid preparations from these strains were also sequenced to confirm plasmid content and proper orientation of *ssrS* with respect to the *lac* promoter. Wild-type (WT) *C. burnetii* and *C. burnetii* (p6Skd, p6Sovrexp and pKM244) strains were cultured for 3 d, induced with 1 mM IPTG and samples were isolated at 24 h and 48 h post-induction. Unfortunately, results of qPCR and qRT-PCR showed highly variable levels of 6S RNA, on a per genome basis, with poor correlation of plasmid content and IPTG induction (Fig. 4-9). Furthermore, to determine if IPTG-inducibility of the *C. burnetii* strains was affected by time in culture, clones were also cultured for 1 d (log phase) and then induced with IPTG. DNA and RNA were purified from samples isolated at 24 h and 48 h post-induction and q(RT)-PCR was performed. However, the levels of 6S RNA were again inconsistent (data not shown), suggesting that these plasmids are not useful for this particular application in *C. burnetii*.

D. **Discussion**

6S RNA is widely distributed among bacteria, and its biology has been under investigation since its identification in 1976 (19). Studies in *E. coli* (2), *B. subtilis* (18) and *L. pneumophila* (5) have shown that 6S RNA specifically associates with RNAP and regulates transcription. For example, some functions of 6S RNA include

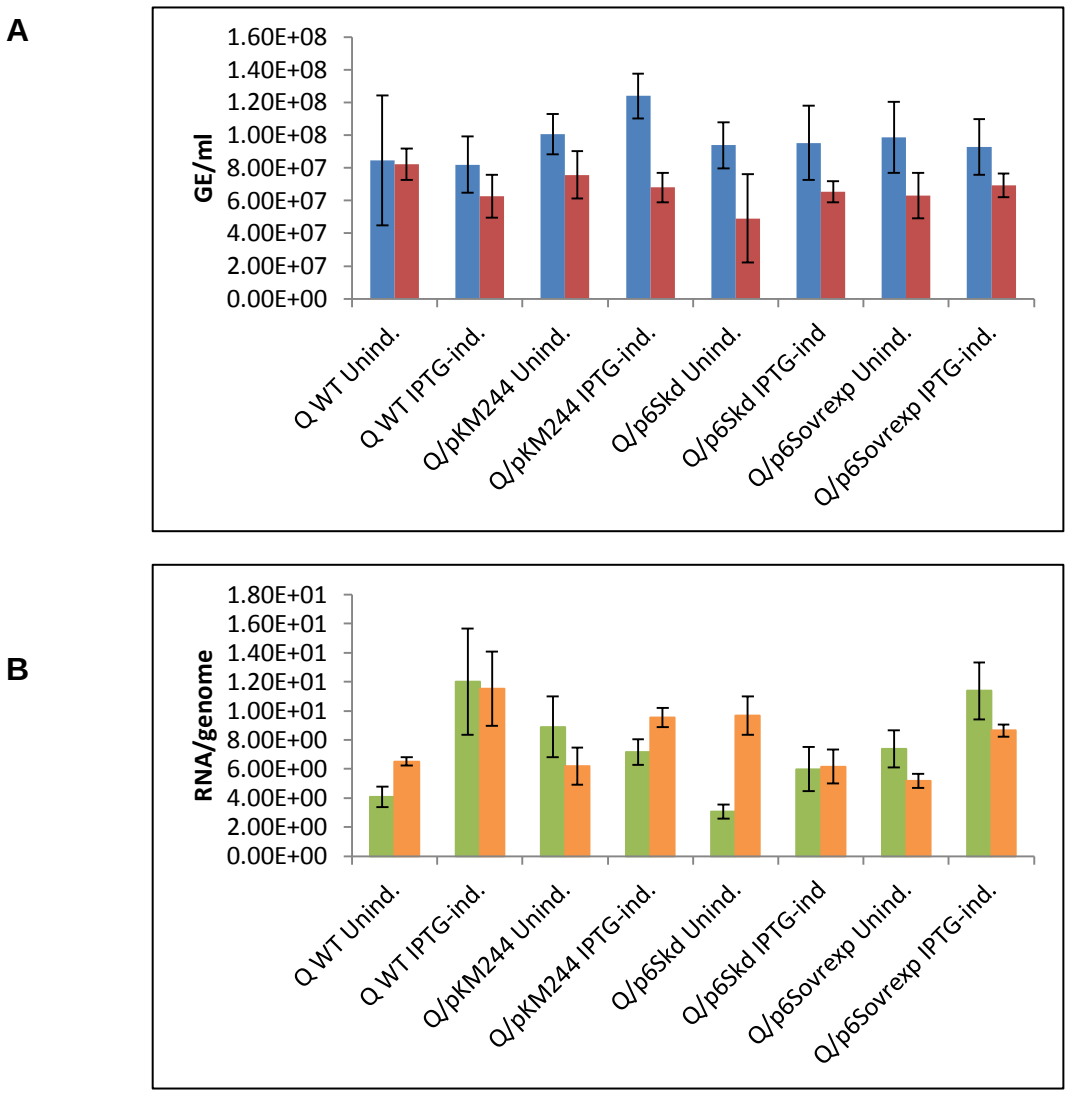


Figure 4-9. Levels of *C. burnetii* 6S sense/antisense RNAs following IPTG-induction. (A) q-PCR data showing *C. burnetii* genome equivalents per ml (GE/ml) at 24 h (blue) and 48 h post-induction (red). **(B)** qRT-PCR data showing numbers of 6S sense/antisense RNAs following IPTG-induction for 24 h (green) and 48 h (orange). Levels of 6S RNA in the absence of IPTG and the cloning vector alone (pKM244) are also shown as a control. Data represent the means of three independent experiments ± S.D.

upregulation of genes involved in stress response and nutrient acquisition (5) , long-term survival during stationary phase (6) and regulation of *relA* and ppGpp synthesis during the stringent response (7). Considering *C. burnetii's* intracellular niche, these functions would be clearly beneficial. This encouraged us to further investigate the biology of 6S RNA in *C. burnetii*. The *ssrS* gene of *C. burnetii* was mapped to the 5′ UTR of *ygfA* (Fig. 4-2). A linkage of *ssrS* and *ygfA* is conserved among many bacterial species (4). When we examined the transcript levels of 6S RNA in *C. burnetii* using Northern blot analysis, we found that it was present at much higher levels in the SCV stage of the bacterium, irrespective of growth conditions (Fig. 4-4). These results were also confirmed by qRT-PCR (Fig. 4-5). This increase is similar to what has been observed in other bacteria, where 6S RNA reaches its highest abundance during stationary phase (2, 5). Interestingly, a ~7-fold higher transcript level was observed in SCVs during intracellular versus axenic growth (Fig. 4-4, compare lanes 2 and 5). A similar observation has been reported for *L. pneumophila*, a bacterium that is closely related to *C. burnetii*. In fact, deletion of the *ssrS* gene of *L. pneumophila* reduced intracellular growth in host cells by ~10-fold, while there was no effect on the mutant's growth in axenic medium (5). A recent study in another pathogenic bacterium, *Y. pestis*, also showed increased transcript levels of 6S RNA *in vivo* (20). Also, the transcript level of 6S RNA in *C. burnetii* grown in Vero cells increased ~9-fold by 14 dpi (SCV), compared to 3 dpi (LCV), and then dropped ~2 fold at 21 dpi (SCV) (Fig. 4-4). However, this drop was not observed in *C. burnetii* grown axenically. It is possible that in Vero co-cultures, some SCVs are still intracellular at 14 d (i.e. some Vero cells are extant), whereas by 21 d all the host cells are dead, SCVs are extracellular and 6S RNA falls to

a background level. Taken as a whole, our observations are suggestive of 6S RNA's involvement in regulating genes related to *C. burnetii*'s stress response during intracellular growth.

Our studies have also shown that, similar to other bacteria, *C. burnetii's* 6S RNA associates specifically with RNAP. This was demonstrated by IP experiments using a *C. burnetii* lysate and an antibody that recognizes core RNAP. Western blotting was performed to confirm that the antibodies were specific to *C. burnetii's* RNAP (Fig. 4-6A). An RPA was also performed on RNA isolated from the IP samples using 6S RNA- and 5S RNA-specific biotinylated probes. Results clearly showed that 6S RNA was present exclusively in IP samples where RNAP was present (Fig. 4-6B). This confirms a physical association between 6S RNA and RNAP. Based on these observations and previous research on other bacteria, we predict that 6S RNA alters transcription in *C. burnetii* by associating with its RNAP. The 6S RNA of *E. coli* is known to bind to all forms of RNAP, however, it preferentially interacts with RNAP-σ^{70} (21).

RpoS (σ^S) is classically the major starvation/stationary phase sigma factor, but it serves as the dominant sigma factor during exponential growth of *C. burnetii* (22). The choice of σ^S is thought to be due to the stressful conditions encountered by *Coxiella* in the context of the PV. With this in mind, we were curious about the potential targets of *Coxiella*'s 6S RNA. Eight positively-charged amino acids have been shown to create a surface that is required for binding of 6S RNA to the 4.2 region of *E. coli*'s RpoD (σ^{70}) (3). Analysis of *Coxiella*'s RpoS and RpoH 4.2 regions indicates that they each possess only five positively-charged amino acid residues (Fig. 4-10). This suggests that *Coxiella*'s 6S RNA interactions with RNAP-RpoS and RNAP-RpoH would be minimal or

Figure 4-10. The 4.2 region of *E. coli* RpoD in comparison to predicted, homologous regions of *C. burnetii* sigma factors. *E. coli* (*Ec*) and *C. burnetii* (*Cb*) 4.2 regions of sigma factors RpoD, RpoS and RpoH are shown. Positively-charged amino acids of the *E. coli* sigma factor RpoD 4.2 region involved in binding 6S RNA (3) are shown in red. Positively-charged residues in the predicted 4.2 region of *C. burnetii* sigma factors are shown in green. ClustalW alignment results are shown on the bottom line, where an asterisk indicates perfect identity, a colon indicates similar amino acids with conservation and a period indicates weakly similar amino acids with conservation to the *E. coli* sequence.

absent. In contrast, the 4.2 region of *Coxiella*'s RpoD (σ^{70}) shares 100% identity with 30 amino acid residues of the *E. coli* σ^{70} 4.2 region with all eight positively-charged amino acids present (Fig. 4-10). Taken together, these analyses suggest that *Coxiella*'s 6S RNA would mainly interact with RNAP-σ^{70}. Nevertheless, the dominant role of RpoS in the log-phase growth of *C. burnetii* suggests the potential for an atypical mechanism of 6S RNA-mediated gene regulation that warrants additional research. .

Early work with *E. coli* demonstrated that 6S RNA binding to RNAP-σ^{70} during stationary phase inhibits polymerase binding to certain σ^{70}-dependent promoters, thus selectively regulating transcription [reviewed in (23)]. Later, it was revealed that the 6S RNA of *E. coli* also activates certain σ^{s}-dependent promoters [reviewed in (23)]. In contrast, the 6S RNA of *L. pneumophila* was found to serve mainly as a positive regulator of genes involved in amino acid metabolism, stress adaptation, DNA repair/replication and detoxification (5). Based upon these observations, we predict that *C. burnetii*'s 6S RNA acts as both a positive and negative regulator as cells approach stationary phase (SCV stage). To identify the genes that are regulated by 6S RNA we generated 6S-knockdown and overexpression strains of *C. burnetii* using a pKM244 shuttle vector. The transcription of 6S sense/antisense RNA was under the control of a *tac* promoter and subject to IPTG induction. These 6S-knockdown and overexpression plasmids were positively tested for their IPTG-inducibility in *E. coli*, where they showed significantly increased 6S sense/antisense RNA levels upon induction (Fig. 4-8). However, when IPTG-inducibility of these plasmids was tested in *C. burnetii*, the levels of 6S sense/antisense RNAs following induction were found to be highly variable with little correlation to plasmid content or IPTG treatment (Fig. 4-9). One possible

explanation for this unfortunate observation is that since *C. burnetii*'s genome does not contain any constituents of the *E. coli lac* operon, IPTG induction may not be functional or is suboptimal in this bacterium. However, this explanation runs counter to a previous report of successful IPTG induction of a fusion protein encoded on this vector (24). Despite our attempts to use several time points in the growth cycle for IPTG induction of sense/antisense RNA, it is possible that the concentration and timing of IPTG induction may need to be further adjusted to optimize the system. An alternative is to generate different constructs where expression is under the control of one of the heat shock promoters of *C. burnetii*.

E. References

1. **Sato K, Hamada M, Asai K, Mituyama T.** 2009. CENTROIDFOLD: a web server for RNA secondary structure prediction. Nucleic Acids Res **37:**W277-280.
2. **Wassarman KM, Storz G.** 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell **101:**613-623.
3. **Klocko AD, Wassarman KM.** 2009. 6S RNA binding to Esigma(70) requires a positively charged surface of sigma(70) region 4.2. Mol Microbiol **73:**152-164.
4. **Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR.** 2005. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. Rna **11:**774-784.
5. **Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA.** 2010. *Legionella pneumophila* 6S RNA optimizes intracellular multiplication. Proc Natl Acad Sci U S A **107:**7533-7538.

6. **Trotochaud AE, Wassarman KM.** 2004. 6S RNA function enhances long-term cell survival. J Bacteriol **186:**4978-4985.

7. **Cavanagh AT, Chandrangsu P, Wassarman KM.** 2010. 6S RNA regulation of relA alters ppGpp levels in early stationary phase. Microbiology **156:**3791-3800.

8. **Neusser T, Polen T, Geissen R, Wagner R.** 2010. Depletion of the non-coding regulatory 6S RNA in *E. coli* causes a surprising reduction in the expression of the translation machinery. BMC Genomics **11:**165.

9. **Williams JC, Peacock MG, McCaul TF.** 1981. Immunological and biological characterization of *Coxiella burnetii*, phases I and II, separated from host components. Infect Immun **32:**840-851.

10. **Cockrell DC, Beare PA, Fischer ER, Howe D, Heinzen RA.** 2008. A method for purifying obligate intracellular *Coxiella burnetii* that employs digitonin lysis of host cells. J Microbiol Methods **72:**321-325.

11. **Coleman SA, Fischer ER, Howe D, Mead DJ, Heinzen RA.** 2004. Temporal analysis of *Coxiella burnetii* morphological differentiation. J Bacteriol **186:**7344-7352.

12. **Omsland A, Beare PA, Hill J, Cockrell DC, Howe D, Hansen B, Samuel JE, Heinzen RA.** 2011. Isolation from animal tissue and genetic transformation of *Coxiella burnetii* are facilitated by an improved axenic growth medium. Appl Environ Microbiol **77:**3720-3725.

13. **Sandoz KM, Sturdevant DE, Hansen B, Heinzen RA.** 2014. Developmental transitions of *Coxiella burnetii* grown in axenic media. J Microbiol Methods **96:**104-110.

14. **Pall GS, Hamilton AJ.** 2008. Improved northern blot method for enhanced detection of small RNA. Nat Protoc **3:**1077-1084.

15. **Schneider CA, Rasband WS, Eliceiri KW.** 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods **9:**671-675.

16. **Raghavan R, Hicks LD, Minnick MF.** 2008. Toxic introns and parasitic intein in *Coxiella burnetii*: legacies of a promiscuous past. J Bacteriol **190:**5934-5943.

17. **Chen C, Banga S, Mertens K, Weber MM, Gorbaslieva I, Tan Y, Luo ZQ, Samuel JE.** 2010. Large-scale identification and translocation of type IV secretion substrates by *Coxiella burnetii*. Proc Natl Acad Sci U S A **107:**21755-21760.

18. **Trotochaud AE, Wassarman KM.** 2005. A highly conserved 6S RNA structure is required for regulation of transcription. Nat Struct Mol Biol **12:**313-319.

19. **Hindley J.** 1967. Fractionation of 32P-labelled ribonucleic acids on polyacrylamide gels and their characterization by fingerprinting. J Mol Biol **30:**125-136.

20. **Yan Y, Su S, Meng X, Ji X, Qu Y, Liu Z, Wang X, Cui Y, Deng Z, Zhou D, Jiang W, Yang R, Han Y.** 2013. Determination of sRNA expressions by RNA-seq in *Yersinia pestis* grown in vitro and during infection. PLoS One **8:**e74495.

21. **Gildehaus N, Neusser T, Wurm R, Wagner R.** 2007. Studies on the function of the riboregulator 6S RNA from *E. coli*: RNA polymerase binding, inhibition of in vitro transcription and synthesis of RNA-directed de novo transcripts. Nucleic Acids Res **35:**1885-1896.

22. **Seshadri R, Samuel JE.** 2001. Characterization of a stress-induced alternate sigma factor, RpoS, of *Coxiella burnetii* and its expression during the development cycle. Infect Immun **69:**4874-4883.

23. **Wassarman KM.** 2007. 6S RNA: a regulator of transcription. Mol Microbiol **65:**1425-1431.

24. **Weber MM, Chen C, Rowin K, Mertens K, Galvan G, Zhi H, Dealing CM, Roman VA, Banga S, Tan Y, Luo ZQ, Samuel JE.** 2013. Identification of *Coxiella burnetii* type IV secretion substrates required for intracellular replication and *Coxiella*-containing vacuole formation. J Bacteriol **195:**3914-3924.

CHAPTER FIVE

CONCLUSIONS

Small RNAs (sRNAs) are a group of non-coding RNAs that regulate a variety of processes in bacteria, including transcription, translation, and mRNA stability (1). The roles of these molecules are especially important in pathogenic bacteria where they regulate processes that are involved in controlling virulence and adapting to harsh environmental conditions (2). Since *C. burnetii* is an obligate intracellular bacterium in nature, these adaptations would be clearly beneficial. Therefore, the overarching goal of my research was to identify and characterize such regulatory RNA elements of *C. burnetii* and to elucidate the role of some of these RNA elements in *C. burnetii*'s biology.

The RNA that we were initially interested in studying was the IVS element inserted in the 23S rRNA gene of *C. burnetii* (3). Similar to other parasitic genetic elements of *C. burnetii*, IVS is highly conserved among all the genotypes of this bacterium. Removal of the IVS element results in fragmentation of the 23S rRNA without affecting its function. In spite of its occurrence in one of the most conserved genes of bacteria, IVS distribution is sporadic, indicating that it is likely acquired by horizontal gene transfer (4). Due to *C. burnetii*'s intracellular niche, it should have been shielded from such an event. However, a recent study has shown that the pathogenic *C. burnetii* evolved from a *Coxiella*-like progenitor hosted by ticks (5). Evolution from a tick symbiont to a virulent pathogen of vertebrates probably involved the acquisition of virulence factors through horizontal gene transfer from co-infecting pathogens (5). In the light of this scenario, it is possible that *C. burnetii* initially acquired the IVS element through horizontal transfer

before or during its transition to a vertebrate host. Thereafter IVS was possibly maintained within the genus by vertical transfer due to its adaptive role in growth modulation.

Similar to *Leptospira,* the IVS of *C. burnetii* contains an ORF that potentially encodes a hypothetical S23p protein (6). Additionally, orthologs of this protein are also encoded by the IVS ORFs of *Haemophilus*, *Brucella* and *Xanthomonas* (6). In spite of several attempts and various strategies, we were unable to express the *C. burnetii* S23p protein *in vitro* using a coupled transcription-translation system. Surprisingly, when the sequence of the IVS was investigated among several *C. burnetii* strains, we observed variability in nucleotide sequence, but maintenance of the stem portion of the element involved in fragmentation. This is indicative of reductive evolution of the ORF, akin to the pseudogenes of *C. burnetii.* In fact, the ORF could eventually be eliminated by this bacterium. These observations led us to predict that IVS is not expressed *in vitro* and probably not *in vivo*. Therefore, we directed our attention to the IVS RNA.

The predicted secondary structure of the IVS was found to be a stem-loop structure held together by complementary sequences present on the ends. We showed that, similar to other bacteria, IVS in *C. burnetii* is excised by RNase III, an endoribonucleases that usually cleaves 23S and 16S from their rRNA precursor. Additionally, we were able to map the ends of the IVS following its cleavage by RNase III. To deduce the *in vivo* role of IVS RNA, we determined the *in vivo* half-life of IVS in an *E. coli* model and in *C. burnetii*. An extended half-life might indicate an adaptive role in *C. burnetii* (7). However, IVS was found to be degraded as soon as it was excised from the 23S rRNA in *C. burnetii* as well as in an *E. coli* model. These results are

consistent with previous observations in other bacteria where IVS was not detected by Northern analysis (8, 9).

Although fragmentation of 23S rRNA due to IVS removal doesn't affect its function, the degradation of these fragments might be altered. Studies in *Salmonella* have shown that increased fragmentation of 23S rRNA results in its increased degradation during stationary phase (10). However, when we analyzed the fragments of 23S rRNA of *C. burnetii* and compared them to 16S rRNA, the two fragments of 23S rRNA (F1 and F2) were differentially regulated. In fact, the levels of smaller 23S rRNA fragment F1 were significantly higher than 16S rRNA while that of larger 23S rRNA fragment F2 was significantly lower than 16S rRNA. These results show that F2 is more labile than F1, which may be as a result of its large size and the presence of two introns that need to be spliced out during maturation. Considering that equal numbers of rRNAs are incorporated into the ribosomes, F2 levels might dictate the number of mature ribosomes that are formed. It is conceivable that this limitation could contribute to growth modulation in *C. burnetii*. Therefore, differential stability of 23S rRNA, as a result of IVS-mediated excision, could contribute to slow growth rate and promote *C. burnetii*'s tendency towards chronic infection.

Survival of *C. burnetii* within the intracellular environment and transitioning between the two development forms require quick and efficient regulation that is probably aided by sRNAs. In the recent decade, research on discovery and characterization of sRNAs has increased greatly. Genome-wide analyses using high-density tiling arrays and RNA-seq have revealed the transcriptomes of various pathogenic bacteria and exposed a high abundance of sRNAs. Although these RNAs might initially appear as products of

spurious transcriptional events, recent reports have shown that they are functional and can be important regulators of gene expression (11). Using RNA-seq, we identified 15 novel sRNAs (CbSR 1-15) of *C. burnetii* and confirmed their presence, size and strand of origin using Northern analyses. Most of the CbSRs were found to be elevated in the metabolic phase of the life cycle, indicating that they might play a role in physiology. The genes that are expressed during the LCV phase are probably involved in survival in the hostile environment. In many bacteria, sRNAs have been identified to be key components of regulatory cascades that manage environmental changes caused by changes in pH, concentration of reactive oxygen species and metabolic-by-products (12). Therefore, sRNAs that are highly expressed during the LCV stage might be involved in regulating such processes. Interestingly, when expression of CbSRs was compared between *C. burnetii* cultured in Veros and axenic medium, eight of them were found to be increasingly expressed during intracellular growth. The role of sRNAs in controlling pathogenesis had been reported in several bacteria (2). Therefore, these eight sRNAs could be involved in regulating *Coxiella*'s virulence and stress response during *Coxiella*'s survival within the hostile intracellular environment. The obvious next step is to determine the target(s) of these sRNAs to identify the genes that they regulate. One of the approaches that can be used to identify the target of these novel sRNAs is by computational prediction approaches using algorithms like RNApredator (13) and IntaRNA (14). The general approaches also include high-throughput screening of targets using genomics- and proteomics-based methods (15). Another probable method of target identification is monitoring the phenotypic changes resulting from experimental manipulation of sRNA transcript levels.

In addition to identification of the novel sRNAs, we also identified the 6S RNA of *C. burnetii*. The *ssrS* gene was mapped to the 5' UTR of the *yfgA* locus. Interestingly, the secondary structure of 6S RNA of *C. burnetii* was found to consist of a central bubble flanked by a closing stem and terminal loop. The functional importance of this structure, especially the central bubble, has been reported for *E. coli* (16, 17). When the levels of 6S RNA was analyzed over the growth cycle of *C. burnetii*, it was found to be highest during the SCV phase. This observation is similar to other bacteria where 6S RNA was highest during stationary phase of the organism (18).

To identify the role it plays in *C. burnetii* biology, we determined whether 6S RNA interacts with RNAP. Immunoprecipitation with antibodies that identify core RNAP showed that 6S RNA efficiently and specifically binds to the RNAP of *C. burnetii*. In *E. coli*, 6S RNA associates with σ^{70}-RNAP holoenzyme and affects transcription by negatively regulating genes expressed from σ^{70} dependent promoters and positively affecting genes expressed from σ^{S} dependent promoters (19). When the 4.2 regions of the sigma factors of *C. burnetii* (σ^{70}, σ^{S}, σ^{32}) were compared for the presence of eight positively-charged amino acids that are known to be required for 6S RNA binding, σ^{70} was found as the most probable candidate. However, contrary to other bacteria, σ^{S} is the dominant sigma factor during exponential growth of *C. burnetii* (20). Based on these observations we conclude that although 6S RNA of *C. burnetii* probably associates with σ^{70}-RNAP, similar to other bacteria, its mode of regulation might be very different.

Additionally, 6S RNA of *L. pneumophila* and *Y. pestis* were found to be comparatively higher during intracellular growth (21, 22). In fact, 6S null mutants of *L. pneumophila* displayed reduced fitness during intracellular growth indicating that 6S

RNA functions in a way that is essential for surviving in a eukaryotic host (21). Due to the intracellular nature of *C. burnetii*, such as role would be clearly beneficial. Further, to identify the genes that are regulated by 6S RNA, we generated 6S-overexpression and 6S-knockdown mutants of *C. burnetii* using a pKM244 shuttle vector. 6S RNA was sequentially cloned in either sense or antisense orientation with respect to a *lac* promoter to generate overexpression and knockdown constructs, respectively. The IPTG inducibility of these plasmids was first tested successfully in *E. coli*. These plasmids were then electroporated into *C. burnetii* and clones were isolated based on chloramphenicol resistance, mCherry expression and PCR confirmation of plasmid content. When the IPTG inducibility of these plasmids was tested in *C. burnetii*, the levels of sense and antisense 6S RNA expression was highly variable. Similar results were observed on repeating the experiment by altering the timing of IPTG induction of the strains indicating that these plasmids are not useful for this particular application in *C. burnetii*. Further, adjusting the concentration of IPTG might be required to optimize this system. Alternatively, a different construct could be generated where the expression of *ssrS* is under the control of a *C. burnetii* promoter such as CBU1169 that encodes a small heat shock protein Hsp20 or CBU0311, an outer membrane porin P1 promoter (23).

In conclusion, we have identified and confirmed 15 novel sRNAs of *C. burnetii*. Additionally, we have characterized two non-coding RNAs: IVS, a parasitic genetic element, and 6S RNA, a sRNA that regulates transcription. This is the first step towards elucidating the regulatory RNA networks of *C. burnetii*. Further work on these sRNAs

will reveal details of transcriptional regulation that will further our understanding of the physiology and virulence of this mysterious pathogen.

A. References

1. **Waters LS, Storz G.** 2009. Regulatory RNAs in bacteria. Cell **136:**615-628.
2. **Gripenland J, Netterling S, Loh E, Tiensuu T, Toledo-Arana A, Johansson J.** 2010. RNAs: regulators of bacterial virulence. Nat Rev Microbiol **8:**857-866.
3. **Afseth G, Mo YY, Mallavia LP.** 1995. Characterization of the 23S and 5S rRNA genes of Coxiella burnetii and identification of an intervening sequence within the 23S rRNA gene. J Bacteriol **177:**2946-2949.
4. **Ralph D, McClelland M.** 1994. Phylogenetic evidence for horizontal transfer of an intervening sequence between species in a spirochete genus. J Bacteriol **176:**5982-5987.
5. **Duron O, Noel V, McCoy KD, Bonazzi M, Sidi-Boumedine K, Morel O, Vavre F, Zenner L, Jourdain E, Durand P, Arnathau C, Renaud F, Trape JF, Biguezoton AS, Cremaschi J, Dietrich M, Leger E, Appelgren A, Dupraz M, Gomez-Diaz E, Diatta G, Dayo GK, Adakal H, Zoungrana S, Vial L, Chevillon C.** 2015. The Recent Evolution of a Maternally-Inherited Endosymbiont of Ticks Led to the Emergence of the Q Fever Pathogen, Coxiella burnetii. PLoS Pathog **11:**e1004892.
6. **Minnick MF, Raghavan R.** 2011. Genetics of Coxiella burnetii: on the path of specialization. Future Microbiol **6:**1297-1314.

7. **Hicks LD, Warrier I, Raghavan R, Minnick MF.** 2011. Ribozyme stability, exon skipping, and a potential role for RNA helicase in group I intron splicing by Coxiella burnetii. J Bacteriol **193:**5292-5299.

8. **Burgin AB, Parodos K, Lane DJ, Pace NR.** 1990. The excision of intervening sequences from Salmonella 23S ribosomal RNA. Cell **60:**405-414.

9. **Konkel ME, Marconi RT, Mead DJ, Cieplak W, Jr.** 1994. Identification and characterization of an intervening sequence within the 23S ribosomal RNA genes of Campylobacter jejuni. Mol Microbiol **14:**235-241.

10. **Hsu D, Shih LM, Zee YC.** 1994. Degradation of rRNA in Salmonella strains: a novel mechanism to regulate the concentrations of rRNA and ribosomes. J Bacteriol **176:**4761-4765.

11. **Lybecker M, Bilusic I, Raghavan R.** 2014. Pervasive transcription: detecting functional RNAs in bacteria. Transcription **5:**e944039.

12. **Repoila F, Majdalani N, Gottesman S.** 2003. Small non-coding RNAs, co-ordinators of adaptation processes in Escherichia coli: the RpoS paradigm. Mol Microbiol **48:**855-861.

13. **Eggenhofer F, Tafer H, Stadler PF, Hofacker IL.** 2011. RNApredator: fast accessibility-based prediction of sRNA targets. Nucleic Acids Res **39:**W149-154.

14. **Busch A, Richter AS, Backofen R.** 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics **24:**2849-2856.

15. **Vogel J, Wagner EG.** 2007. Target identification of small noncoding RNAs in bacteria. Curr Opin Microbiol **10:**262-270.

16. **Barrick JE, Sudarsan N, Weinberg Z, Ruzzo WL, Breaker RR.** 2005. 6S RNA is a widespread regulator of eubacterial RNA polymerase that resembles an open promoter. Rna **11:**774-784.

17. **Trotochaud AE, Wassarman KM.** 2005. A highly conserved 6S RNA structure is required for regulation of transcription. Nat Struct Mol Biol **12:**313-319.

18. **Wassarman KM, Storz G.** 2000. 6S RNA regulates E. coli RNA polymerase activity. Cell **101:**613-623.

19. **Wassarman KM.** 2007. 6S RNA: a regulator of transcription. Mol Microbiol **65:**1425-1431.

20. **Seshadri R, Samuel JE.** 2001. Characterization of a stress-induced alternate sigma factor, RpoS, of Coxiella burnetii and its expression during the development cycle. Infect Immun **69:**4874-4883.

21. **Faucher SP, Friedlander G, Livny J, Margalit H, Shuman HA.** 2010. Legionella pneumophila 6S RNA optimizes intracellular multiplication. Proc Natl Acad Sci U S A **107:**7533-7538.

22. **Yan Y, Su S, Meng X, Ji X, Qu Y, Liu Z, Wang X, Cui Y, Deng Z, Zhou D, Jiang W, Yang R, Han Y.** 2013. Determination of sRNA expressions by RNA-seq in Yersinia pestis grown in vitro and during infection. PLoS One **8:**e74495.

23. **Beare PA, Gilk SD, Larson CL, Hill J, Stead CM, Omsland A, Cockrell DC, Howe D, Voth DE, Heinzen RA.** 2011. Dot/Icm type IVB secretion system requirements for Coxiella burnetii growth in human macrophages. MBio **2:**e00175-00111.