12-2009

Discovery and Validation of New Regulatory RNA Elements in Chlamydia trachomatis

Yasser Mohammed Elsayed Metwally AbdelRahman
University of Tennessee Health Science Center

Discovery and Validation of New Regulatory RNA Elements in Chlamydia trachomatis

Abstract

Chlamydia trachomatis is an obligate intracellular bacterium that exhibits a unique biphasic developmental cycle that can be disrupted by growth in the presence of IFN-g and b-lactams, giving rise to an abnormal growth state termed persistence. Relatively little is known about the regulatory mechanisms that control temporal gene expression during the developmental cycle or the control of persistence and reactivation. Here we have examined the expression of a newly defined family of non-coding RNAs (ncRNAs) that are differentially expressed during the developmental cycle and the induction of persistence and reactivation (Using IFNγ and Carbenicillin). Non-coding RNAs were initially identified using an intergenic tiling microarray and were confirmed by Northern blotting. A group of 10 ncRNAs were mapped and characterized and compared to the previously described chlamydial ncRNAs (*IhtA*, pCHL antisense transcripts). The 5' and 3' ends of the ncRNAs were determined using an RNA circularization procedure. Promoter predictions indicated that all ncRNAs were expressed from s^{66} promoters and 9 ncRNAs contained non-templated 3' poly-A or poly-AG additions.

Expression of ncRNAs was studied by Northern blotting during i) the normal developmental cycle, ii) IFN-g-induced persistence, and iii) carbenicillin-induced persistence. Differential temporal expression during the developmental cycle was seen for all ncRNAs and distinct differences in expression were seen during IFN-g and carbenicillin-induced persistence and reactivation.

Two of the studied ncRNAs were *cis* acting antisense molecules (*CTIG270* and *CTIG153*). Expression of *CTIG270* in a surrogate *E. coli* system along with its target gene *ftsI* proved that it was an antisense RNA, and it effectively controlled *ftsI* availability. A screening system was developed to determine the targets of potential *trans*-acting ncRNAs. The screen was composed of two compatible plasmids, a high copy effector plasmid expressing the ncRNA and a low/medium copy target plasmid expressing chlamydial library. Target plasmid contained translational fusion between chlamydial library and a positive/negative selection fusion system (*ccdB/CAT* fusion/TEV protease). The selection process to identify potential targets for ncRNAs was carried out over two stages. Stage one was to eliminate self ligated plasmids and non translational fusions, in this phase, only CAT^{+} *ccdB* resistant cells survived chloramphenicol selection *i.e.* cells representing true translational fusions survived antibiotic selection. Plasmids enriched for by stage one selection were subjected to stage two selection. In stage two *ccdB* sensitive cells were double transformed with both target plasmid and effector plasmid. Only in instances when ncRNA inhibits translation of *ccdB* would cells survive this selection. Stage two selection enriched for target(s) of ncRNA under test. We show in details the construction of this screening system and its functional aspect.

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Microbiology, Molecular Biology and Biochemistry

Research Advisor

Robert Belland, Ph.D.

Keywords

Chlamydia trachomatis, Developmental cycle, non-coding RNAs, Carbenicillin persistence, IFN-gamma persistence, RNA Circularization

Subject Categories

Medical Microbiology | Medical Molecular Biology | Medical Sciences | Medicine and Health Sciences

Comments

One year embargo expired December 2010.

Discovery and Validation of New Regulatory RNA Elements in *Chlamydia trachomatis*

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Yasser Mohammed Elsayed Metwally AbdelRahman
December 2009

To my Father, who taught me how to succeed.
To my Mother, who gave up all life luxuries for our family.
To my Wife, who proved to be the ROCK of our house.

ACKNOWLEDGEMENT

I would like to thank my mentor, Dr. Robert Belland, for offering me the opportunity to pursue my Ph.D. study. I especially thank him not only for his extensive scientific knowledge and teachings, but also for his helpful advice and guidance in many other areas during my graduate studies. His unmatched patience, trust, support and positive attitude exemplified role model for mentor.

I also thank my committee members Dr. Yan Cui, Dr. Richard Lee, Dr. Thomas Hatch, for their valuable suggestions and positive critique of my research. I would like to give very worm thanks for Dr Gerald Byrne for being the first one to introduce me to *Chlamydia* during my rotation. I additionally thank Dr. Tina Henkin at Ohio State University for her invaluable advice about the role of ncRNAs. I would also like to thank Dr. Susan Gottesman Head, Biochemical Genetics Section Laboratory Co-Chief for her great advice and encouragement. Special thanks to Leighton LeGros Research Compliance Officer at the Memphis VA Medical Center. He taught me many technical tricks about cloning and general lab protocols. I thank all Belland lab members, not to mention Loren Rose for doing an outstanding job at designing the array, and optimizing the hybridization protocols also special thanks for Nirun Naher for keeping our supplies of *Chlamydia* replenished, she also helped me greatly with the Western blotting experiments. In addition, I thank Scot Ouellette for giving me the first lessons in dealing with *Chlamydia.*

I deeply thank my Father for teaching me how to love academia and how to love my work, I deeply thank my Mother and Father in law for taking care of my kid while I am at lab (day and night) Their prayers always paved the way for me and made it possible for me to attain this goal. I cannot find words to express my appreciation for my amazing wife who, besides her duties as a Pharmacist, bore the main responsibility of caring for our children throughout our time in Memphis; her never ending support and understanding helped me accomplish a project of that magnitude. I thank my friends Walid Awad, Mohammed Nooh, Ramy Aziz, Shehab Ismail, Hossam Abd-Elsamed, Ali Ellbidy, Mohammed Elmeligy for their support during my stay in Memphis; they are the brothers I never had.

Last but not least I would like to pay my greatest respect and thanks to the Department of Microbiology and Immunology in the Faculty of Pharmacy Cairo University. I am deeply honored to mention two professors that passed away during the course of my Ph.D., Dr Hussien Shoeb former head of the department and Dr AbdelHameed. Both professors supervised my masters and both tought me a lot during the course of my study. I am in deep gratitude to all members of the department of Microbiology and Immunology. I single out professors: Dr AbdelGawad Hashim the current head of the department, Dr. Ossama Eltayeb, Dr. Magdy Amin, Dr Ala Shawki, Dr Fathi'ya, Dr Aysha, Dr Manal, and all members of the department that I forgot to mention or new members that I am yet to meet.

ABSTRACT

Chlamydia trachomatis is an obligate intracellular bacterium that exhibits a unique biphasic developmental cycle that can be disrupted by growth in the presence of IFN-γ and β-lactams, giving rise to an abnormal growth state termed persistence. Relatively little is known about the regulatory mechanisms that control temporal gene expression during the developmental cycle or the control of persistence and reactivation. Here we have examined the expression of a newly defined family of non-coding RNAs (ncRNAs) that are differentially expressed during the developmental cycle and the induction of persistence and reactivation (Using IFN-γ and Carbenicillin). Non-coding RNAs were initially identified using an intergenic tiling microarray and were confirmed by Northern blotting. A group of 10 ncRNAs were mapped and characterized and compared to the previously described chlamydial ncRNAs (*IhtA*, pCHL antisense transcripts). The 5' and 3' ends of the ncRNAs were determined using an RNA circularization procedure. Promoter predictions indicated that all ncRNAs were expressed from σ^{66} promoters and 9 ncRNAs contained non-templated 3' poly-A or poly-AG additions.

Expression of ncRNAs was studied by Northern blotting during i) the normal developmental cycle, ii) IFN-γ-induced persistence, and iii) carbenicillin-induced persistence. Differential temporal expression during the developmental cycle was seen for all ncRNAs and distinct differences in expression were seen during IFN-γ and carbenicillin-induced persistence and reactivation.

Two of the studied ncRNAs were *cis* acting antisense molecules (*CTIG270* and *CTIG153*). Expression of *CTIG270* in a surrogate *E. coli* system along with its target gene *ftsI* proved that it was an antisense RNA, and it effectively controlled *ftsI* availability. A screening system was developed to determine the targets of potential *trans*-acting ncRNAs. The screen was composed of two compatible plasmids, a high copy effector plasmid expressing the ncRNA and a low/medium copy target plasmid expressing chlamydial library. Target plasmid contained translational fusion between chlamydial library and a positive/negative selection fusion system (*ccdB/CAT* fusion/TEV protease). The selection process to identify potential targets for ncRNAs was carried out over two stages. Stage one was to eliminate self ligated plasmids and non translational fusions, in this phase, only CAT^{+} *ccdB* resistant cells survived chloramphenicol selection *i.e.* cells representing true translational fusions survived antibiotic selection. Plasmids enriched for by stage one selection were subjected to stage two selection. In stage two *ccdB* sensitive cells were double transformed with both target plasmid and effector plasmid. Only in instances when ncRNA inhibits translation of *ccdB* would cells survive this selection. Stage two selection enriched for target(s) of ncRNA under test. We show in details the construction of this screening system and its functional aspect.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION[1]

1.1 *CHLAMYDIA*

The order Chlamydiales encompasses a large group of bacteria characterized by their obligate growth in eukaryotic cells. The last decade has seen a rapid expansion of the number of organisms within the order, including organisms from numerous animal and environmental sources. This has led to the proposed division of the family Chlamydiaceae to include two genera, *Chlamydia* and *Chlamydophila* and the inclusion of non-Chlamydiaceae such as Simkaniaceae, Waddliaceae and Parachlamydiaceae (Everett *et al*, 1999). The taxonomy of the Chlamydiales is controversial at this time and many feel the genera divisions are unnecessary. I have chosen to use the terminology described in Schachter *et al*, 2001 *i.e.* the use of a single genus, *Chlamydia*.

Chlamydiaceae are the etiological agents of many important human and animal diseases. In humans, the genital serovars of *Chlamydia trachomatis* (*C. trachomatis*) are the most prevalent cause of sexually transmitted disease worldwide (Gerbase *et al*, 1998), while the ocular serovars result in blinding trachoma in developing countries (Schachter, 1978). *Chlamydia pneumoniae* (*C. pneumoniae*) is a widespread respiratory pathogen (Kuo *et al*, 1995) and chronic infections are associated with an enhanced risk of developing atherosclerotic (Saikku *et al*, 1988 and Kuo *et al*, 1993), cerebrovascular (Balin *et al*, 1998), and chronic lung disease (Hahn *et al*, 1991). Pathogenic chlamydial isolates have been characterized from a number of animal hosts including birds, cats, rodents, cattle and pigs (*e.g.* Meijer, 2002).

Chlamydia alternates between two morphological forms, the elementary body (EB) and the reticulate body (RB) (Moulder, 1991). EBs are extra-cellular, metabolically inert forms, responsible for dissemination of infection by their ability to attach to and invade susceptible cells. Upon infection, EBs are internalized in membrane bound vacuoles termed inclusions. EBs differentiate into metabolically active forms, termed RBs, and undergo repeated cycles of binary fission leading to secondary differentiation back to EBs. The host cell then lyses, releasing EBs, which infect neighboring cells. Under stressful growth conditions, imposed by immunological responses, antibiotics, or nutrient deprivation, the developmental cycle is disrupted, resulting in the appearance of large, aberrant RBs (reviewed in Hogan *et al*, 2004). This altered growth scheme appears to be associated with continued expression of genes associated with DNA replication but not with those genes involved with bacterial cell division (Belland *et al*, 2003a).

[1] *Adapted by permission. Abdelrahman YM, Belland RJ (2005) The chlamydial developmental cycle. FEMS Microbiol Rev* ***29:*** *949-959.*

1.2 THE DEVELOPMENTAL CYCLE

1.2.1 ELEMENTARY BODY

The term "elementary body" (EB) refers to the small (ca. 0.3 μm), round, electron dense, infectious form of the organism (*e.g.* Matsumoto, 1973and Eb *et al*, 1976). EBs' nucleotide is highly compacted due to the condensation of nuclear material by the bacterial histone-like proteins HctA and HctB (Barry *et al*, 1992 and Brickman *et al*, 1993). Chlamydial EBs are unusual in that little or no peptidoglycan is present in the cell wall. Structural rigidity is thought instead to be due to the highly cross-linked nature of the outer-membrane complex. Inter and intramolecular cystine bonds exist between the cysteine rich proteins of the outer envelope including OmpA, OmcB, and OmcA (reviewed in Hatch, 1999).

Recent studies indicate the interaction of EBs with host cells occurs in a two-stage process. The initial reversible attachment occurs through electrostatic interactions of the bacteria with heparan sulfate containing glycosaminoglycans (Su *et al*, 1996; Su and Caldwell, 1998; Stephens *et al*, 2001) and (Taraktchoglou *et al*, 2001); and a second, irreversible, binding stage (Carabeo and Hackstadt, 2001). Clifton *et al*, 2004 have recently shown that, immediately following the irreversible binding step, a type III secretion system secretion system (TTSS) exported protein is delivered into the host cell. This protein, termed Tarp (Translocated actin-recruiting phosphoprotein, *CT456*), is rapidly phosphorylated at tyrosine residues and phosphorylation correlates spatially and temporally with actin recruitment (Figure 1-1).

1.2.2 PRIMARY DIFFERENTIATION

The chlamydial developmental cycle has a singular enigmatic stage, the differentiation of the infecting EB to a metabolically active RB. The process of differentiation can be blocked by the addition of antibiotic inhibitors of transcription or translation, suggesting that de novo protein expression is required to begin intracellular growth (Scidmore *et al*, 1996). This finding is at odds with the generally held understanding that the presence of the bacterial histone-like proteins HctA and HctB render the EB transcriptionally incompetent. This paradox has been at least partially addressed by the findings of Grieshaber *et al*, 2004 chlamydial histone–DNA interactions are disrupted upon germination by a small metabolite in the non-mevalonate pathway (MEP) pathway of isoprenoid biosynthesis. The metabolite is thought to be 2-C-methylerythritol 2, 4-cyclodiphosphate and is involved in functional antagonism of HctA. Findings showed that an *E. coli* HctA expression system was rescued from the lethal effects of HctA by chlamydial *ispE*. *ispE* is an intermediate enzyme in the MEP pathway, paradoxically; experimental data showed that *E. coli* orthologue of *ispE* could not protect *E. coli* from the lethal effects of HctA expression.

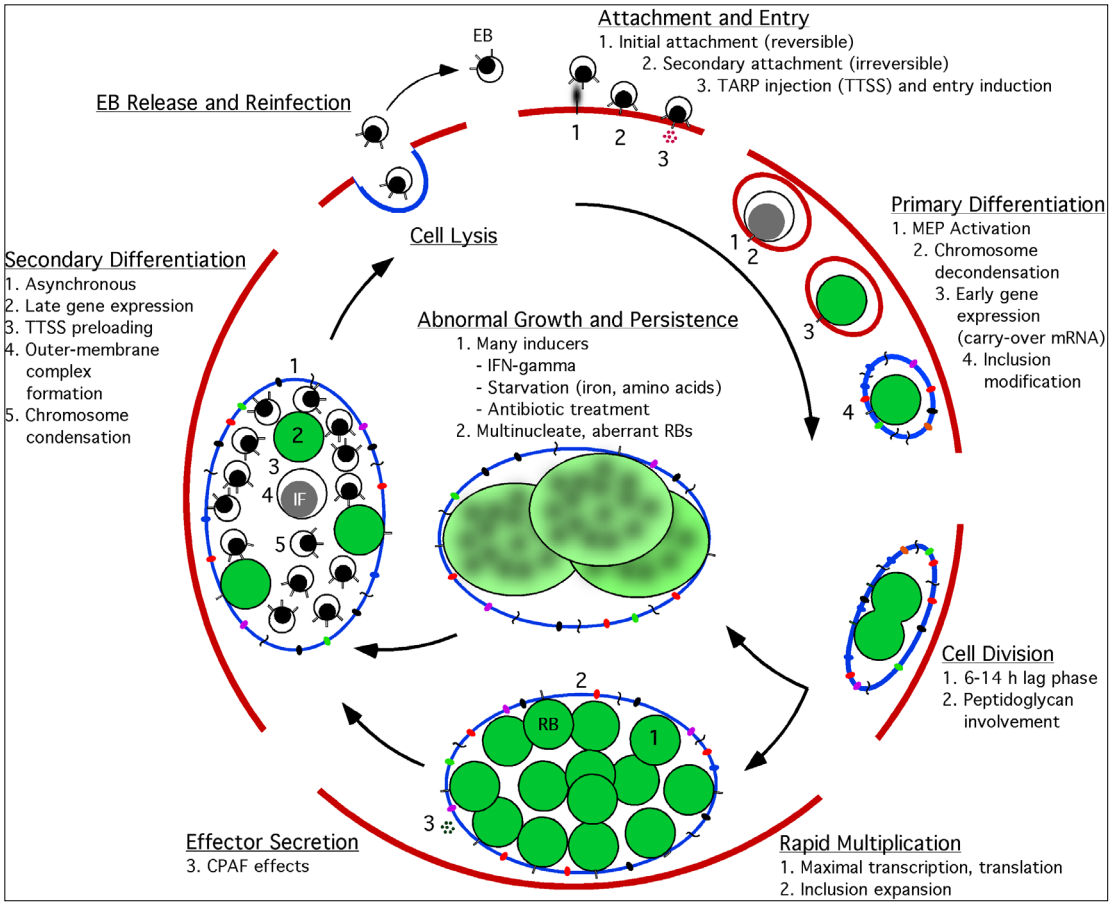

Figure 1-1 A schematic representation of the chlamydial developmental cycle. The host cell cytoplasmic membrane (red line) is shown to depict the interactions of chlamydial EBs and the origin of the inclusion membrane. The major events in the developmental cycle are categorized as described in the text to allow an overview of the processes and mechanisms associated with acute and persistent chlamydial growth. TARP (Translocated Actin-Recruiting Phosphoprotein, *CT456*), TTSS (Type III Secretion System), MEP (non-Mevalonate Pathway), CPAF (Chlamydial Protease/proteasome-like Activity Factor).

The question arises whether it is special enzyme kinetics attributed to chlamydial IspE or the transcriptional activity of the chlamydial *ispE* in the presence of HctA, or another unknown function that chlamydial *ispE* has that helps in the rescue from HctA lethality. *Chlamydia* possess another histone-like protein HctB, it is still unknown what antagonizes its function during the differentiation step (Grieshaber *et al*, 2004).

Transcription begins within the differentiating RB almost immediately following internalization. New protein expression can be detected within 15 min PI using intrinsic labeling procedures (Plaunt and Hatch, 1988) and newly synthesized RNA can be detected in "host-free" *Chlamydia* by 1 h PI (Crenshaw *et al*, 1990). Wichlan and Hatch (Wichlan and Hatch, 1993) identified the first early gene (*euo*, Early Upstream open reading frames [ORF]) using "host-free" early RNA to select clones from a genomic library. Transcription of *euo* was confirmed by Northern blotting and primer extension analysis. Euo appears to be specific to *Chlamydia* in that it is highly conserved within the genus but has no homology to other bacterial proteins. Despite its early identification, the function of Euo has not been clearly defined to date.

Following the sequencing of the serovar D genome, a number of studies used RT-PCR to identify genes expressed early in the cycle. Shaw *et al*, 2000, identified a number of inclusion-associated protein genes (*inc*) as early genes and proposed the generally accepted temporal classes of developmentally expressed genes as (i) early (1–2 h PI), (ii) mid-cycle (3–18 h PI), (iii) late (20–48 h PI). More recently, genome-wide DNA microarrays have been used to determine the transcriptome throughout the cycle. Belland *et al*, 2003b identified 29 genes expressed by *C. trachomatis* serovar D at 1 h PI in HeLa 229 cells. Seven of these genes had been previously characterized as early genes, including the previously mentioned *euo* and *inc* genes.

1.2.3 RETICULATE BODY

The "reticulate body" (RB) arises from the internalized EB following primary differentiation, as discussed above. RBs are larger than EBs (ca. 1 μm) and the cytoplasm appears granular with diffuse, fibrillar nucleic acids, in contrast with the highly condensed nucleic acid content of the EB (reviewed in Matsumoto, 2000). RBs are non-infectious and are bounded by an inner and outer-membrane, resembling other, Gram-negative, eubacteria. RBs undergo binary fission throughout the middle part of the developmental cycle.

1.2.4 GENE EXPRESSION

Microarray analyses of gene transcription during the developmental cycle have been reported for *C. trachomatis* serovar L2 (Nicholson *et al*, 2003) and serovar D (Belland *et al*, 2003b). While a number of technical and experimental design differences were used for the analyses, both reports indicate that by 6–8 h PI the developing RBs were highly transcriptionally active. Belland *et al* (2003b) classified temporal gene

expression groups into early, mid-cycle and late categories with a subgrouping of genes expressed at 1 h PI as immediate-early. While this study focused on the immediate-early and late genes, the complete listing of transcriptional activities (Belland *et al*, 2003b) throughout the cycle indicated that the period of intense transcriptional activity (16–24 h PI) correlated with the rapid growth and division of RBs. Furthermore, they reported that virtually every gene in the organism is expressed at some point in the cycle, indicating that *C. trachomatis* has virtually no facultative capacity, as might be expected for an obligate intracellular organism with a "minimal" genome.

1.2.5 TYPE III SECRETION MACHINERY

Genomic data reveals that all chlamydial species possess a complement of genes encoding TTSS (Stephens *et al*, 1998; Kalman *et al*, 1999) and (Read *et al*, 2000). TTSS genes in *Chlamydia* show temporal regulation through out the developmental cycle (Fields and Hackstadt, 2000; Fields *et al*, 2003; Slepenkin *et al*, 2003).

1.2.6 CELL DIVISION

Cell division takes place during the RB stage. Genome sequencing indicated that *Chlamydia* lack an identifiable *fts*Z orthologue, which encodes a protein centrally involved in bacterial cell division and found in all other sequenced eubacteria. Another surprising finding in the genomic analyses was the presence of a complete set of genes for the synthesis, assembly, and degradation of peptidoglycan (PG) (Chopra *et al*, 1998; Stephens *et al*, 1998; Read *et al*, 2000;) and (Read *et al*, 2003) Numerous studies had reported that *Chlamydia* lacked peptidoglycan, with a single study reporting trace amounts in EBs (Su H., 1985). Similarly, attempts to identify peptidoglycan in RBs were unsuccessful (Barbour *et al*, 1982). Somewhat surprisingly, the production of infectious EBs is highly sensitive to inhibitors of PG synthesis, including β-lactam antibiotics and D-cycloserine (Matsumoto and Manire, 1970; Tamura and Manire, 1968; Weiss, 1950). Treatment of infected cells with these agents inhibits cell division and leads to the formation of large, aberrant RB that cannot differentiate to EB (Moulder, 1993). These studies indicate that chlamydial PG synthesis may be required for chlamydial cell division and proper differentiation. This paradoxical situation has been referred to as the "chlamydial anomaly" (Ghuysen and Goffin, 1999). McCoy *et al*, 2003 have shown that the chlamydial MurA orthologue (UDP-*N*-acetylglucosamine enopyruvyl transferase), which catalyzes the first committed step in peptidoglycan biosynthesis, is functional in *E. coli*. The chlamydial MurA was found to encode a fosfomycin-resistant form of the enzyme and this resistance was imparted to *E. coli* expressing the chlamydial enzyme. Expression studies indicated that *Chlamydia* are naturally fosfomycin resistant and that *mur*A is expressed at the point in the developmental cycle immediately preceding cell division (Belland *et al*, 2003b) and (McCoy *et al*, 2003), providing circumstantial evidence for the involvement of peptidoglycan synthesis in cell division. Brown and Rockey (Brown and Rockey, 2000) have identified an antigen named the septum antigen (SEP antigen) localized to a ring-like structure at the apparent septum within dividing

chlamydial reticulate bodies (RB). Antisera directed against SEP show similar patterns of antigen distribution in *C. trachomatis* and *C. psittaci* RB. Treatment of *chlamydiae* with inhibitors of peptidoglycan synthesis or culture of *chlamydiae* in medium lacking tryptophan leads to the formation of nondividing, aberrant RB. Staining of aberrant RB with anti-SEP reveals a marked redistribution of the antigen. The reversal of aberrant formation results in the production of normal RB and a redistribution of SEP to the apparent plane of bacterial division.

1.2.7 SECONDARY DIFFERENTIATION AND LATE GENE EXPRESSION

Following the period of rapid cell division, RBs begin to redifferentiate to EBs, here termed "secondary differentiation". The signal for this process is unknown and efforts to identify quorum sensing type pathways are not been supported by genomic analyses.

Expression of a number of late-cycle genes occurs during secondary differentiation (reviewed in Hatch, 1999), including genes that encode components of the outer-membrane complex (*e.g.*, OmcA and B) and proteins involved in the condensation of the chromosome (*e.g.*, HctA and B). Microarray analyses have extended the number of genes expressed during this stage of development. Belland *et al*, 2003b identified 26 late genes including a number of genes previously characterized as late genes based on other molecular biological studies (*e.g.*, *omc*AB, *hct*AB, *ltu*B, *lcr*H.1) and a number of new genes that encode proteins with speculative and unknown function. Interestingly, expression of some late cycle genes may be directed towards the arming of EBs with proteins necessary during the early stages of the next infectious cycle (*e.g.*, the TTSS protein LcrH.1), hence "late" genes may encode "early" proteins.

1.3 PROMOTER ELEMENTS AND REGULATION

1.3.1 PROMOTER ELEMENTS IN E. coli.

The bacterial core RNA polymerase complex, consists of five subunits ($\beta\beta'\alpha_2\omega$), these subunits are sufficient for transcription elongation and termination but are not able to initiate transcription. Transcription initiation from prokaryotic promoter requires another dissociable subunit (σ^{70} factor) (Burgess *et al*, 1969), which reversibly associates with the core RNA polymerase complex to form a holoenzyme RNAP. (reviewed in Paget and Helmann, 2003).

An RNAP-promoter complex capable of transcription initiation is formed through interactions that involve several regions of RNAP and that span 70–80 base pairs (bp) of promoter DNA (~−60 to +20) with respect to the transcription start site (TSS). The majority of the sequence-specific interactions occur with the σ subunit. In *E. coli*, sequence-specific interactions with σ^{70} occur at the −10 element (regions 2.3–2.4 in σ^{70}),

the extended −10 element (σ^{70} region 3.0), the −35 element (σ^{70} region 4.2), and the discriminator element immediately downstream of the −10 hexamer (σ^{70} region 1.2), Figure 1-2; and (Haugen *et al*, 2008). In addition, the C-terminal domain of α subunit can interact sequence-specifically with the UP element, located upstream of the −35 hexamer (Gourse *et al*, 2000). Other interactions also take place that are not sequence specific (Korzheva and Mustaev, 2001; Murakami *et al*, 2002; Vassylyev *et al*, 2007). (Reviewed in Ross and Gourse 2009).

1.3.2 CHLAMYDIAL GENE REGULATION

Control of the developmental cycle may have important translational and post-translational components but the bulk of the experimental evidence suggests the primary mechanism involved is transcriptional regulation. Genomic studies have shown that all species sequenced to date (Stephens *et al*, 1998; Read *et al*, 2000; Read *et al*, 2003) have a highly conserved multisubunit RNA polymerase that has many shared features compared to other eubacterial RNA polymerases (*i.e.* the core RNA polymerase consists of two α subunits and β and β' subunits). Three sigma factors (σ) that combine with the core polymerase to form the holoenzyme have also been found in all sequenced species. The chlamydial-specific sigma factors are the major sigma factor σ^{66} (Koehler *et al*, 1990; Mathews and Stephens, 1999a; Douglas and Hatch, 2000) (orthologous to σ^{70} in *E. coli*) and the minor sigma factor σ^{28} (Mathews *et al*, 1999b; Yu and Tan, 2003; Shen *et al*, 2006) (orthologous to σ^{28} in *E. coli*) and the alternative sigma factor σ^{54} (Mathews and Timms, 2000) (orthologous to σ^{54} in *E. coli*). Functional evidence has been demonstrated for several σ^{66} promoters and one σ^{28} promoter using mutational analysis and *in vitro* transcription assays (*e.g.* Ricci *et al*, 1993; Fahr *et al*, 1995; Douglas and Hatch, 1996; Tan *et al*, 1996; Tan *et al*, 1998; Ochiai *et al*, 1999; Schaumburg and Tan, 2003; Yu *et al*, 2003; Yu *et al*, 2006). Chlamydial promoters share homology with each of the eubacterial consensus sequences with certain exceptions, including the presence of a *cis*-acting, A/T spacer region in the majority of σ^{66} promoters analyzed (Schaumburg and Tan, 2000). In addition, the stringency of consensus in chlamydial promoters appears to

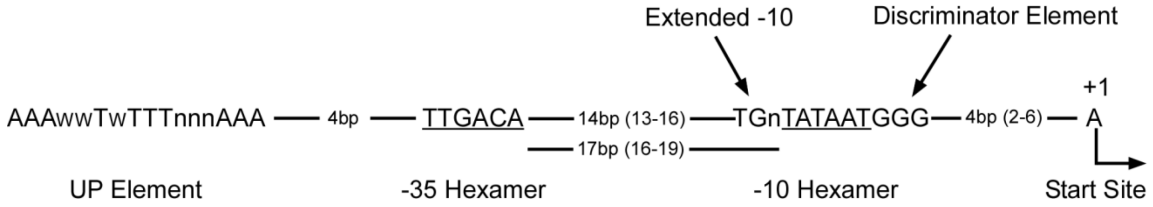

Figure 1-2 Consensus sequences for $E\sigma^{70}$ recognition elements and RNAP regions that interact with these elements. UP Element with α subunit C-terminal domains, −35 hexamer with σ^{70} 4.2, extended −10 element with σ^{70} 3.0, −10 hexamer with σ^{70} 2.3–2.4, and discriminator region element with σ^{70} 1.2. The TSS is indicated as (+1). The most common spacing between recognition elements, in bp, with (range) is indicated. Both the spacing of the −35 element with respect to the extended −10 element (14bp), and with respect to the −10 element (17bp) are shown. *Reprint by permission: Ross W, Gourse RL (2009) Analysis of RNA polymerase-promoter complex formation. Methods* ***47:*** *13-24.*

be more relaxed than found in other eubacteria (Mathews and Sriprakash, 1994), allowing for changes in the consensus -10 and -35 regions of the promoters that interact with the major sigma factor. This lack of stringency has hampered in silico efforts to identify promoters from complete genome sequences.

Several transcriptional regulatory mechanisms have been found in *Chlamydia* including *cis*-acting DNA elements (Schaumburg *et al*, 2000), phosphorylation cascades (Koo and Stephens, 2003) and transcription factors (activators and repressors). Transcriptional regulators have been found that repress transcription. HrcA has been shown to specifically bind a chlamydial CIRCE (Controlling Inverted Repeat for Chaperone Expression) element upstream of *dna*K (Wilson and Tan, 2002), and a degenerate CIRCE element upstream of *gro*ELS (Wilson *et al*, 2005) and repress transcription of the *dna*K and *gro*ELS genes.

A distant relative of the *E. coli fur* gene has been identified in *Chlamydia* (termed *dcr*A) by complementation of an *E. coli fur* mutant and the ability to bind a 19 bp consensus operator sequence (or Fur box) found in promoters of iron-regulated genes (Wyllie and Raulston, 2001). The repressor-operator binding kinetics of DcrA differ from the *E. coli* Fur protein suggesting that DcrA may have additional response cues than its *E. coli* counterpart or that the chlamydial operator sequence may differ from the consensus Fur box. Other regulators include IHF (Zhong *et al*, 2001), TrpR (Akers and Tan, 2006), ArgR (Schaumburg and Tan, 2006) and ChxR (Koo *et al*, 2006).

1.4 ALTERNATIVE GROWTH MODELS AND PERSISTENCE

Many chlamydial diseases are associated with a long term or chronic infectious state. In most cases it is difficult to establish whether chronic or recurrent infections arise through the inability of the host to resolve the infection or the occurrence of repeated infections with similar species or genotypes. Despite the unresolved nature of the disease etiology, persistence models of chlamydial infection have been studied to provide insight into the nature of chronic disease (reviewed in Hogan *et al*, 2004). Persistence is defined as a long-term association between *Chlamydia* and their host cell in which these organisms remain in a viable but culture-negative state. The *in vitro* persistence systems often share altered chlamydial growth characteristics for example, many studies have described enlarged, and pleomorphic RBs that neither undergo binary fission, nor differentiate to EBs, but nevertheless continue to replicate their chromosomes. These changes are generally reversible upon removal of the growth inhibitory factor (Hogan *et al*, 2004). Persistent *in vitro* infections have been induced by penicillin treatment (Matsumoto *et al*, 1970), amino acid starvation (Coles *et al*, 1993), iron deficiency (Raulston, 1997), IFN-γ exposure (Beatty *et al*, 1993), monocyte infection (Koehler *et al*, 1997), phage infection (Hsia *et al*, 2000), continuous culture (Kutlin *et al*, 2001). This subject has been thoroughly reviewed in Hogan *et al*, 2004.

IFN-γ mediated inhibition of intracellular chlamydial replication occurs by depletion of the essential amino acid tryptophan, via the induction of indoleamine-2,3-

dioxygenase (IDO) (Beatty *et al*, 1994). The effect of IFN-γ on chlamydial infection could be reversed by addition of tryptophan (Byrne *et al*, 1986). Belland *et al.* have studied the induction of persistence with IFN-γ and the subsequent reactivation, using microarray analysis. Persistent growth, characterized by large aberrant RBs, led to the up-regulation of genes involved in tryptophan utilization, DNA repair and recombination, phospholipid biosynthesis and translation. Up-regulation of the repressible *trp* BA operon (Wood *et al*, 2003) confirms the previous observations that IFN-γ treatment reduces intracellular concentrations of tryptophan. In addition, a number of early genes were up-regulated, particularly the *euo* gene (30-fold increase) which encodes a DNA-binding protein that has been shown to bind to a late gene promoter region (*i.e. omc*AB (Zhang *et al*, 1998). Down-regulation of genes involved in RB to EB differentiation (late genes such as *hct*AB and *omc*AB), proteolysis and peptide transport, and cell division were seen during persistent growth. The transcriptional analyses were consistent with the biological properties associated with aberrant RBs in that cells were blocked in cytokinesis and the developmental cycle was arrested at a point preceding late gene expression. Removal of IFN-γ and supplementation with added tryptophan led to a rapid reactivation from persistent growth. During reactivation the expression differences rapidly returned to control levels, *i.e. euo* expression dropped 20-fold in 12 h. This coordinated biological response appears to have evolved to allow the organism to rapidly respond to immunological pressure in a manner that allows for a period of resistance followed by rapid recovery after the waning of the host response.

1.5 REGULATORY ncRNAs IN BACTERIAL PATHOGENS

1.5.1 GENERAL

Bacterial RNAs have long been known to encode the information for protein expression (mRNAs) and the structural components of the translational machinery (tRNAs and rRNAs). Numerous studies have proved that there are other type(s) of RNAs with other functions. For example, a number of abundant and frequently stable RNAs were identified by metabolic labeling and direct analysis by various fractionation procedures. These included: 1) 4.5S RNA, part of the secretion machinery (Ribes *et al*, 1990); 2) the catalytic part of the RNase P Ribozyme (M1RNA) (Robertson *et al*, 1972), that was found to be a ubiquitous endonuclease catalyzing the maturation of the 5' end of tRNA (reviewed in Kazantsev and Pace, 2006); 3) tmRNA, which serves an important role in translational quality control (Komine *et al*, 1994); 4) 6S RNA, which modulates RNA polymerase activity (Narayan *et al*, 1966; Barrick *et al*, 2005); and 5) Spot 42, an ncRNA that acts as an anti-sense regulator of the gal operon (Ikemura and Dahlberg, 1973; Moller *et al*, 2002). Not surprisingly, this sort of regulation is not confined to *E. coli*. Phages, plasmids and transposons have long been recognized to use anti-sense RNA regulators involved in controlling copy number of plasmids, transposons, and the life cycle of bacteriophages (reviewed in Delihas, 1995; Wagner and Brantl, 1998; Wagner *et al*, 2002). In eukaryotic cells, microRNAs and RNAi parallel in many ways bacterial ncRNA regulators; confirming that this level of regulation is widespread and is as central

to creating a working organism as are the more highly characterized protein regulators of transcription (Carrington and Ambros, 2003; McManus and Sharp, 2003).

1.5.2 FUNCTIONAL ASPECTS

Since the discovery of this new category of ncRNAs, several exhaustive searches have taken place. This lead to the identification of approximately 80 ncRNAs in *E. coli* between 50 and 400 bases in size (1%–2% of the number of protein-coding genes) (Argaman *et al*, 2001; Gottesman *et al*, 2001; Rivas and Eddy, 2001; Wassarman *et al*, 2001; Kawano *et al*, 2005). Mechanistically, ncRNAs act through base-pairing interactions with other RNAs, thus inducing premature termination of message (Novick *et al*, 1989), mediating the destabilization of target mRNAs (Masse *et al*, 2003; Morita *et al*, 2005), promoting efficient translation of target mRNA (Majdalani *et al*, 1998), negatively regulating translation of target mRNA (Grieshaber *et al*, 2006) by mimicking the structure of other RNAs and DNAs (Willkomm and Hartmann, 2005), or as integral parts of protein-RNA complexes (Ribes *et al*, 1990) (reviewed in Storz *et al*, 2004; Gottesman, 2005 and Waters and Storz, 2009).

Regulatory ncRNAs can be broadly divided into *cis* and *trans*-acting molecules. *Cis*-acting ncRNAs (also termed anti-sense RNAs) generally overlap the target gene and form extensive complementary stretches of double stranded RNA (dsRNA) that are degraded by RNase III (Vogel *et al*, 2004), or cause premature termination (attenuation) of transcription (Novick *et al*, 1989). *Trans*-acting ncRNAs interact with non-overlapping target mRNAs and the resulting RNA hybrids are generally degraded by RNase E (Masse *et al*, 2003), or the binding interaction interferes with recognition of the translational start portion of the mRNA (reviewed in Storz *et al*, 2004), or act as anti-anti-sense where they remodel target RNA in favor of a translational permissive secondary structure (Majdalani *et al*, 1998). *Trans*-acting ncRNA function is often difficult to predict since the anti-sense interaction generally occurs through a limited number of base pairs with the target mRNA, often using non-contiguous regions of the molecule folded into hairpin loops. Many *trans*-acting ncRNAs have been shown to bind multiple target sequences; *e.g.* RyhB that binds to the mRNA of several iron metabolizing genes and possibly other targets as well (Masse *et al*, 2005).

1.5.3 DETECTION METHODS

Identification of bacterial ncRNAs has increased dramatically with the advances in genomic technology and the appreciation of the importance of these molecules in regulatory pathways. Experimental approaches to identifying ncRNAs (reviewed in Huttenhofer and Vogel, 2006; Vogel and Sharma, 2005) have been diverse and include: i) direct metabolic labeling of abundant ncRNAs and detection by gel electrophoresis, ii) genetic screens for regulatory phenotypes that mapped to non-coding regions, iii) cloning of size fractionated RNAs, iv) bioinformatic identification of conserved intergenic regions, v) co-purification with RNA-binding proteins, and vi) microarray analysis of

intergenic regions. Although each approach has strengths and weaknesses, a microarray analysis has become the method of choice for initial identification. Traditionally, microarray design has been limited to regions coding for proteins, thereby omitting transcriptional measurements in intergenic regions (IGRs) where the majority of ncRNAs are encoded. Recent studies with *E. coli* have utilized high-density microarrays that contain overlapping 25'mers oligonucleotide probes for all mRNAs, tRNAs, and rRNAs and IGRs (Selinger *et al*, 2000). Although initially used to validate technical aspects of microarray profiling, a number of ncRNAs were confirmed using this procedure. Subsequent studies, specifically directed at ncRNA detection, have shown the approach to be particularly useful in analyzing ncRNA expression under different growth conditions (Wassarman *et al*, 2001; Tjaden *et al*, 2002). An alternative approach has been used to analyze intergenic expression patterns in *Staphylococcus aureus* that involved PCR amplification of selected regions and arraying the dsDNAs on glass slides (Pichon and Felden, 2005).

1.5.4 KNOWN EXAMPLES OF ncRNAs

Recent studies have shown that, in addition to the plethora of protein repressors and activators, ncRNAs act as major regulators of adaptive responses in bacterial pathogens. These ncRNAs control the global response of *E. coli* to oxidative stress (Altuvia *et al*, 1997), iron starvation (Masse *et al*, 2005), glucose starvation (Vanderpool and Gottesman, 2004), and the expression of outer membrane proteins and porins (Rasmussen *et al*, 2005; Udekwu *et al*, 2005). Several pathogens utilize a cell-density based system (quorum sensing) to control the expression and secretion of virulence factors. Non-coding RNAs were found to be the main effectors of quorum sensing systems in some organisms, *e.g.* in *Vibrio cholerae*, the expression of four redundant Qrr RNAs (quorum regulatory RNAs) is dependent on the quorum sensing protein (phosphorylated LuxO at low cell density) to block expression of the downstream effector, HapR (Lenz *et al*, 2004). This leads to the up-regulation of a large number of virulence factors. In *Staphylococcus aureus*, the expression of virulence factors is controlled by the expression of RNAIII that acts as the intracellular effector of the agr regulon (Novick and Jiang, 2003). In *Streptococcus pyogenes*, the pel locus encodes a pathogenesis-related RNA that forms part of a signal transduction cascade that is dependent on the growth phase of the organism (Mangold *et al*, 2004). The multifaceted effects of ncRNAs can be illustrated by the *S. aureus* RNAIII (Novick *et al*, 2003) molecule, which controls the switch between the expression of surface proteins (colonization) and secreted toxins (pathogenicity). The RNAIII molecule uses different structural domains to act in three different ways: 1) as a translational activator for hla (hemolysin δ) (Morfeldt *et al*, 1995); 2) as a translational inhibitor of spa mRNA; and 3) as a hemolysin mRNA (Novick *et al*, 2003). RNAIII is also thought to regulate the expression of approximately 110 other genes, but it is not known if this is a direct effect on transcription or the result of post-transcriptional regulation of other transcription factors (Dunman *et al*, 2001). Recently, Grieshaber *et al*, 2006 have shown that heterologous expression of a ncRNA from *C. trachomatis* can block translation of HctA (a chlamydial histone) in *E. coli*, thereby rescuing *chlamydiae* from the lethal effects of

HctA expression. In addition, they demonstrated that the ncRNA (ihtA) was developmentally expressed in *C. trachomatis* at times in the cycle when HctA was not expressed.

CHAPTER 2. MATERIALS AND METHODS

2.1 MICROARRAY DESIGN AND CONSTRUCTION

The custom microarray was designed in conjunction with the Affymetrix Design Team and purchased through support of the University of Tennessee Health Sciences Center and the UT Center of Excellence for Genomics and Bioinformatics. The custom microarray is termed the "Multi-Pathogen Array, University of Tennessee number 1" or MPAUT-1, as it has representations of four chlamydial genomes of four different species. The species represented are *C. trachomatis* D (Stephens *et al*, 1998), *C. pneumoniae* AR-39 (Read *et al*, 2000), *C. muridarum* (Read *et al*, 2000), and *C. caviae* (Read *et al*, 2003) and the complete genomic sequences are available on the TIGR website (www.tigr.org).

The custom array format is a "CustomExpress Advantage 49-7875" that has approximately 253,000 addressable features (*i.e.* defined positions for 25'mer oligonucleotides termed "probes"). Sequences of the complete ORFs for each species were submitted to the Affymetrix Design Team as FASTA files and probes were selected based on their evaluation process. This process looks at all possible 25'mer probes and assigns probe scores based on a number of criteria including; thermodynamic properties, hybridization potential, synthesis steps, and cross-reactivity. Cross-reactivity was addressed by a series of steps involving "hard" pruning (elimination of probes that cross-react with abundant RNAs, *i.e.* rRNAs) and "soft" pruning (elimination of probes that cross-react with other sequences on the microarray and internal control sequences). Each ORF is represented by a number of 25'mer oligonucleotides (mean value, 11 25'mers per gene) that are species-specific. Each probe (termed Perfect Match or PM) has a corresponding mismatched probe (MM, altered at position 13, the middle of the 25'mer) to control for specificity. The values reported for ORF analysis are therefore the mean value of the 11 measurements *i.e.* $\sum PM - MM/11$.

2.2 *C. trachomatis* GROWTH AND CELL CULTURE

C. trachomatis serovar D (strain UW-3/Cx) was grown in HeLa 229 cells cultivated at 37^0 C with 5% CO_2 in high glucose-containing DMEM (Cellgro, Mediatech) supplemented with 10% heat-inactivated FBS. EBs were purified on density gradients of RenoCal-76 (Bracco Diagnostics, NJ) as previously described (Belland *et al*, 2003b). Monolayers were pre-treated with DEAE-Dextran (30 μg/ml) and infected at a multiplicity of infection (MOI) of 1. PI timing began immediately following the addition of infectious EBs to the monolayers. Growth in the presence of IFN-γ was done as described previously (Belland *et al*, 2003a). 5ng per ml IFNγ was used to induce persistence. The amount of IFN-γ used was determined based on a titration experiment for the minimum amount of IFN-γ that induce persistence and allowed for reactivation. Carbenicillin-persistence was induced by adding 2 μg per ml antibiotic to the media at the time of infection. Similar to the IFN-γ experiments, the amount of carbenicillin used was determined by a titration experiment that allowed for reactivation.

2.3 RNA PURIFICATION

Total RNA was purified from infected HeLa 229 cultures (9 x 10^5 cells per well cultivated as monolayers in 6 well culture plates) at various times PI. At the designated times PI, the culture media was discarded and cells lysed in 1 ml per well lysis buffer containing Proteinase K (0.167 µg/µl) (MasterPure™, Epicentre) and total nucleic acid was isolated. An aliquot of the lysate (300µl) was treated with RNase A (0.0167 µg/µl) and used to purify DNA. The remainder of the lysate was treated with DNase I and used for total RNA preparation according to manufacturer's instructions (MasterPure, Epicentre). Purified DNA was quantified by qPCR and used for RNA normalization to calculated *C. trachomatis* genome numbers as previously described (Ouellette *et al*, 2006). RNA for microarray analysis was further purified from residual contaminating DNA by treating with DNase I (0.3U per 1 µg RNA) at 37°C for 1 hour (TurboDNAFree, Ambion). RNA was then precipitated using sodium acetate (0.1 volume) and ethanol (3 volumes) overnight at -20°C. RNA was then pelleted by centrifugation, washed with 80% ethanol and re-suspended in DPEC water (1µg/µl). RNA (250 µg) was then enriched using the MICROB*Enrich*™ (Ambion) and MICROB*Express*™ (Ambion) protocols, resulting in the removal of a significant portion of HeLa cell polyadenylated mRNA, 18S and 28S rRNAs in addition to bacterial 16S and 23S rRNA.

RNA was reverse transcribed, hybridized, washed, stained, and scanned according to the Prokaryotic sample and array processing section in the GeneChip Expression Analysis Technical Manual (Affymetrix, CA). In short, cDNA was generated from enriched RNA preparations using random primers (Invitrogen, CA) and Superscript RT II (Invitrogen, CA). Following cDNA generation, RNA was removed by adding NaOH (1 N) and incubating for 30 min at 65°C followed by neutralization in HCl (1 N). cDNA was then purified using MinElute columns (Qiagen, CA). cDNA was then fragmented using DNase I (Amersham, UK) using 0.6 U per 1µg of cDNA generated. Fragmented cDNA products were then end-labeled using GeneChip DNA labeling Reagent (Affymetrix, CA) and Terminal Deoxynuceotidyl Transferase (Promega, WI). Fragmented and labeled cDNA were analyzed using 4-20% TBE gels (Invitrogen, CA) using a gel shift assay using NeutrAvidin (NeutrAvidin, Pierce). The hybridization cocktail, consisting of labeled cDNA, control B2 oligo (Affymetrix, CA), herring sperm DNA (Promega, WI), and BSA (Invitrogen, CA) was then hybridized to MPAUT-1 GeneChip and incubated for ca. 16 h in the Affymetrix Hybridization Oven 640 (Affymetrix, CA) at 45°C rotating at 60 rpm. Samples were then washed and stained on the Affymetrix Fluidics Station 450 (Affymetrix, CA) using Streptavidin (Pierce, IL), Anti-Streptavidin antibody (Vector Laboratories, CA), and R-Streptavidin Phycoerythrin (Molecular Probes, CA) and using fluidics protocol ProKGE-WS2_450 in GCOS 1.1 (Affymetrix, CA). After washing and staining, completed chips were scanned using the Genechip Scanner 3000 (Affymetrix, CA). Data was scaled using GCOS 1.1 (Affymetrix, CA) to a mean intensity of 1000 using the *C. trachomatis* ORF's as a scaling mask. The pivot file for the sample was then saved in GCOS as a text tab delimited file and the data was imported into GeneSpring 7.2 (Agilent, CA) for data analysis.

2.4 QUANTITATIVE RT-PCR

Primer/probe sets have been designed for *C. trachomatis ompA* using Primer Express software (Applied Biosystems, CA). Standard curves were performed using purified chromosomal template DNA at concentrations ranging from 10 - 0.001 ng/ml (data not shown). Assays were performed (Universal PCR System, Applied Biosystems) using DNA preparations of a portion of the sample for RNA purifications that were used for Northern Blotting. Bacterial genome copy numbers were estimated by converting mean critical threshold (Ct) values to DNA concentrations (using Standard curves) and converting concentrations to copy numbers using the calculated molecular mass of the bacterial genome (7700 Sequence Detector, ABI Prism).

2.5 RNA CIRCULARIZATION, AND 5'/3' END SEQUENCING

The procedure for RNA 5'/3' end sequencing has been described previously (Vogel and Hess, 2001) and was used with some modifications. A total of 8 µg of DNA-free total RNA (DNased twice) prepared from 24 h PI cultures was treated with ten units tobacco acid pyrophosphatase (Epicentre, WI) for 30 min at 37°C to convert 5' triphosphate groups of primary transcripts to 5' monophosphates. Following organic extraction, RNA was treated with 40 U of T4 RNA ligase (New England Biolabs, MA) overnight at 17°C. RNA concentration during ligation was kept at 0.5 µg/50µl; this dilution allowed the reaction to proceed in the direction of intramolecular ligation (circularization) rather than intermolecular ligation (concatemerization). Following organic extraction and ethanol precipitation, 1.5 µg of self-ligated RNA was converted to cDNA using gene specific reverse primes (Table 2-1) under 5'3' SEQ application heading) and Superscript III (200 units) reverse transcription kit (Invitrogen, CA) in a 20 µl reaction. Incubation was carried out in 20-minute intervals at 42 °C, 50 °C, 55 °C, and 60 °C. After heat inactivation of the reverse transcriptase for 5 min at 85 °C, cDNA was treated with one unit of RNase H (New England Biolabs, MA) at 37 °C for 20 min. A total of 1 µl of the reaction served as template in a subsequent standard 25 µl PCR reaction using TITANIUM Taq DNA Polymerase, TaqStart Antibody (1.1 µg/µl), 1x Advantage 2 SA PCR Buffer (Tris-HCl 1 mM (pH 8.5), KCl 5 mM, $MgCl_2$ 0.2 mM), and 0.2 mM dNTP (Clontech, CA) and primer pairs designed to amplify products representing successful self-ligated ncRNA transcripts. PCR products were separated by 10% PAGE TBE gel electrophoresis (BIORAD, CA), and fragments of the expected size were excised, incubated overnight in 10mM Tris-HCl buffer then cloned into the pCR 2.1-TOPO vector (Invitrogen, CA). Positive colonies were picked, grown overnight in LB media with carbenicillin (50 µg/ml) and plasmid were prepare using MiniPlasmid Prep (Qiagen, CA). Plasmids were sequenced using a standard M13 Forward primer.

Table 2-1 List of primers, organized according to application.

Application/ Use	Primer Name	Sequence
CTIG153 Northern	*CTIG153*_FWD_T7	TAATACGACTCACTATAGGGAGATTCTCCGCGATGGCACCTTAAT
	*CTIG153*_REV	TGTGATAGCCTGAGCGTGTAGATG
	*CTIG153*band2_FWD_T7	TAATACGACTCACTATAGGTAGCTCTTGCAGAGAAGCATCGTG
	*CTIG153*band2_REV	CCGTCTCAAATGTCTACTGTCGCT
CTIG153 5’3’ SEQ	*CTIG153* FWD 5'3' SEQ	AGGTTCCCTTGAAAGACACGGTCA
	CTIG153 REV 5'3'SEQ	CCGTCTCAAATGTCTACTGTCGCTGC
CTIG153 5’RACE	*CTIG153*_GSP1	AGGTTCCCTTGAAAGACACGGTCA
	*CTIG153*_GSP2	ACTACCATCCGGGAATGATTAGACGCTTCCGT
CTIG241 Northern	*CTIG241*_FWD	CCTTTGGAGTATATTTGAGGCGG
	*CTIG241*_REV_T7	TAATACGACTCACTATAGGGGGAGAACTCAGAGGG
CTIG241 5’3’ SEQ	*CTIG241* FWD 5'3'SEQ	CCCTCTGAGTTCTCCCCCGAA
	CTIG241 REV 5'3'SEQ	GGGTGCTGGTCACTAGCAAAAACTAGA
CTIG241 5’RACE	*CTIG241*_GSP1	GGGAGAACTCAGAGGG
	*CTIG241*_GSP2	GGGTGCTGGTCACTA
CTIG270 Northern	*CTIG270*_FWD	CTACGTCATCTGTCAGAGAAAGAGC
	*CTIG270*_REV_T7	TAATACGACTCACTATAGGAGGAATGGAATCGCAAATAGTGAG
CT270 Northern	*CT270*_FWD	ACACGCCTAGGGCTATCTCTTGTT
	*CT270*_REV_T7	TAATACGACTCACTATAGGTCCCTACAGGAGGTTTAGAGGCTT

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
CTIG270 Probe -219 to -18	*CTIG270* FWD(-219)	TTTAGCCGGAATATTCTGAAGA AGTTGGTCTAAATG
	CTIG270 REVT7(-18)	TAATACGACTCACTATAGGAGA GTTACTTTTTTTGTGAAAGAGT AAAAAGACGAGTTC
CTIG270 Probe -106 to +67	*CTIG270*_FWD(-106)	GTCAAGCTTTGTCGACTAACTA CGTCATCTG
	*CTIG270*_REVT7(+67)	TAATACGACTCACTATAGGAGC GGGAGAGGAAACTAATAATAA AGA
CTIG270 Probe +83 to +264	*CTIG270*_FWD(+83)	CTGCACACTCATCCCAAAGCAA CA
	CTIG270 _REVT7(+264)	TAATACGACTCACTATAGGGAG AATAGCTTCTCGAGTGCTGCCT
CTIG270 Probe +275 to +494	*CTIG270*_FWD(+275)	AGCACACCGTCCTCCCATATAG
	*CTIG270*_REVT7(+494)	TAATACGACTCACTATAGGACT AAGGGATATTCTAGTGCAGGG
CTIG270 5’3’ SEQ	*CTIG270* FWD 5'3' SEQ	CACTCATCCCAAAGCAACAGAC CTCA
	CTIG270 REV 5'3'SEQ	AGTGTGCAGCCTGGCATAAGGA AA
CTIG270 5’RACE	*CTIG270*_GSP1	GAGGAATGGAATCGCAAATAG TGAG
	*CTIG270*_GSP2	TTGGGATGAGTGTGCAGCCTGG CATAA
ftsI-FLAG Cloning	*CT270*_NoATG FWD	AATCACCGTAGACAATTAACTC TGATCGTTGTTG
	CT270 FWD+0.5FLAG	GGACGACGATGACAAAAATCA CCGTAGACAATTAACTC
	CT270 FWD *FseI*/FLAG	ATTGGCCGGCCATGGACTACAA GGACGACGATGACAAAAATCA
	*CT270*_REV+Stop/*StuI*	ATTAGGCCTCTATTTGCGATTCC ATTCCTCATATAGCAGCT
	LacP/O/S.D.+*StuI*_FWD	ATTAGGCCTGCCGATTCATTAA TGCAGCTGG

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
ftsI-FLAG Cloning	LacP/O/S.D.+*FseI*_REV	AATCCGGCCGGTTTCCTGTGTG AAATTGTTA
CTIG270 Cloning	*CTIG270*_FWD/*HindIII*	ATGAAAGCTTAGAAATGTTTTG TTTGTGAGATGTAATTATGA
	*CTIG270*_REV/*HindIII*	ATTAAAGCTTTACTAAGGGATA TTCTAGTGCAGGG
CTIG327 Northern	*CTIG370*_FWD_T7	TAATACGACTCACTATAGGACA CCCGTCTCTAAGAAGTAGGCA
	*CTIG370*_REV_all gene	GTAACAAGGTGTTGTGTAATCA GAACT
	*CTIG370*_REV_PostTerm	TTCTGGTCTACAGAGGGCTTT
CTIG327 5’3’ SEQ	*CTIG370* FWD SEQ 3'5' RT	TTAAGTACTGCTCGAACCCGTA ACGC
	CTIG370 FWD 5'3' SEQ	CGGAGTCCTTCGAATTACACCA GGATCT
	CTIG370 REV SEQ3' 5'	GTCAAAGCAAGCCTTCTGGTCT ACAGAG
CTIG327 5’RACE	*CTIG370*_GSP1	CCTCTGTAGACCAGAAGGCTTG CTTTGACT
	*CTIG370*_GSP2	TCGGTCTCTATCCTTAAGTACTG CTCGAACCCGT
CTIG356 Northern	*CTIG356*_FWD_T7	TAATACGACTCACTATAGGTAA CAGCACGTTCCACTACTCGCA
	*CTIG356*_REV	GTTTATGTCTCAACTCAGCTAA ACGC
CTIG356 5’3’ SEQ	*CTIG356* FWD SEQ 3'5'	GGACCCAGCCGTCAGAACACTT T
	CTIG356 REV SEQ 3'5' RT	AGCTGCTTTGTTTGGGTGTTTGC C
CTIG360 Northern	*CTIG360*_FWD_T7	TAATACGACTCACTATAGGTAA GAACAAGGACCCGGCGATAAC
	*CTIG360*_REV	TCCTTTCATTTAGACTGGAGA
CTIG360 5’3’ SEQ	*CTIG360* FWD SEQ 3'5'	CCGGCCTTTAAACCAACTACAG ACATCC
	CTIG360 REV SEQ 3'5'	CCGGATACCTTAGTGTCGATTC GGTGAA

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
CTIG360 5'RACE	*CT359*_GSP1	CTGCTATCCCACACATAGGACTAC
	*CT359*_GSP2	AGAACAAGGACCCGGCGATAACCTTCACAA
CTIG498 Northern	*CTIG498* FWD	ACCTGGAACAGCAGAGTTGATTGTGGCT
	CTIG498 REV_T7	TAATACGACTCACTATAGGGGCTTTCACCAGGCGATGGGAAGAAG
CTIG498 5'3' SEQ	*CTIG498* REV SEQ 3'5' RT	CAGCATAATTTCGGAATCGAGCG
	CTIG498 REV SEQ 3'5'	CAGCATAATTTCGGAATCGAGCGAAGTCT
	CTIG498 FWD SEQ 3'5'	GGCTGTGAGTTCCCTGGATAAACAAGT
CTIG504 Northern	*CTIG504*_FWD_T7	TAATACGACTCACTATAGGACAAGAAGTTCTAGCACGCACGAC
	*CTIG504*_REV	CCGAAGAATCTCAAGAGCAGCCAA
CTIG504 5'3' SEQ	*CTIG504* REV SEQ3'5'	CTTCTTGTAAAGCTGTGGCGAGGACT
	CTIG504 FWD SEQ3'5' RT	AGTTCTAGCACGCACGACTGGAAA
CTIG643 Northern	*CTIG643*_FWD	GCAAATTCTTGCTTTAGCAGGCAGG
	*CTIG643*_REV_T7	TAATACGACTCACTATAGGTAAGCGGGATCTTAGTCTCTCGCA
CTIG643 5'3' SEQ	*CTIG643* FWD RT	AAAGAGCTTATATGTGAAAGACTTTGTGT
	CTIG643 FWD SEQ 3'5'	GCTCTTTTGAGAGCTTTCTTGGAAAGCTTGA
	CTIG643 REV SEQ 3'5'	CTCAAAAGAGCTTATATGTGAAAGACTTTGTGTCATA
	CTIG643 FWD SEQ 3'5'-2	AGCGCGCTTCTTCCGCTCTT
	CTIG643 REV SEQ 3'5'-2	AGCCTTGTCAGGGACTTTAAACGGG

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
CTIG675 Northern	*CTIG675*_FWD_T7	TAATACGACTCACTATAGGTTG CAAGTTGGTATTCTAACGCC
	*CTIG675*_REV	AAAGCCAAGAGAACCGGAGAT ACG
CTIG675 5’3’ SEQ	*CTIG675* FWD SEQ 3'5'	CCGTATCTCCGGTTCTCTTGGCT
	CTIG675 REV SEQ3'5' RT	GGCTAACTCCAGTCCATCTTGA CTTCC
CTIG675 5’RACE	*CTIG675*_GSP1	TACGGCTAACTCCAGTCCATCT TGACTTCC
	*CTIG675*_GSP2	CAAACACTAGAGTCAGAAGCTA TTCCATGGCGTT
CTIG675 Cloning	*CTIG675* FWD CLONING	ATTTTCATTAAGCCGAAAGTCA TCAGTAGCTTCGA
	CTIG675 REV CLONING	TACGAATGCGTTGCTCTACCAA CTGA
CTIG675 Cloning under tetracycline promoter	*CCTIG675* FWD/Tc1	GTGATAGAGATACTGAGCACAG TTGCAAGTTGGTATTCTAACGC C
	Add Tc1	AGAGATTGACATCCCTATCAGT GATAGAGATACTGAGCAC
	Add Tc2	TCGAGTCCCTATCAGTGATAGA GATTGACATCCCTATCAGTGAT
	Add Tc3	CGTAGGTACCTCGAGTCCCTAT CAGTGATAGAGATTGACA
CTIG684 Northern	*CTIG684*_FWD	GCAATACTCTCTCTGACTGTCT AGC
	*CTIG684*_REV_T7	TAATACGACTCACTATAGGAGA CAGCCTAGGGAAAGAGGATGT
CTIG684 5’3’ SEQ	*CTIG684* REV SEQ 3'5' RT	AGACATCCTCTTTCCCTAGGCT GTCT
	CTIG684 FWD SEQ 3'5'	GGATGTCTCTCCCTAGACCCAA GGAA
CTIG684 5’RACE	*CTIG684*_GSP1	ACAGCCTAGGGAAAGAGGATG TCT
	*CTIG684*_GSP2	AAGAGGATGTCTCTCCCTAGAC CCAA

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
Test SEQ Cloning	*CThctA* FWD Infusion	GATAGAGATACTAGGTAAAACT GAAAAAAATAGTTTAAAACAA CAACTAGAGGATATTTT
	Test SEQ Infusion REV_*CAT*	GATTTTTTTCTCAGGCATGAAA AAATATCCTCTAGTTGTTGT
	Test SEQ Infusion REV_tev	ATTTTCACCGCCAGGCATGAAA AAATATCCTCTAGTTGTTGT
	Test SEQ Infusion REV_*ccdB*	AACCTTAAACTGAGGCATGAAA AAATATCCTCTAGTTGTTGT
CAT/*ccdB*	*CcdB_CAT* FWD Infusion	TGCGATGAGTGGCAGCAGTTTA AGGTTTACACCTATAAAAGAGA
	ccdB REV Infusion	TCCCCCGGGCTGCAGGAATTTT ATATTCCCCAGAACATCAGGTT AATGGCGTTT
	CAT FWD Tc1+*StuI*	ATCAGTGATAGAGATACTAGGC CTGAGAAAAAAATCACTGGATA TACCACCGTTGATATA
	CAT FWD Tc2	TCCCTATCAGTGATAGAGATTG ACATCCCTATCAGTGATAGAGA TACTAGGCCTGAGAAA
	*CAT*_Tc FWD Infusion	TGGCAATTCCGACGTGACGTTC CCTATCAGTGATAGAGATTGAC ATCCCTATCAG
	CAT REV	CTGCCACTCATCGCAGTACTGT TGTAAT
ccdB/CAT	*CAT* FWD Infusion	GAGAAAAAAATCACTGGATAT ACCACCGTTG
	CAT Rev Infusion	CTGCCACTCATCGCAGTACTGT TGTAATTCATTAAGCATTCTGC CG
	Tc *ccdB* FWD1	CAGTGATAGAGATACTAGGCCT CAGTTTAAGGTTTACACCTATA AAAGAGAGAGCCGTTA
	Tc *ccdB* FWD2	TGATAGAGATTGACATCCCTAT CAGTGATAGAGATACTAGGCCT CAGTTTAAGGTTTACA

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
ccdB/CAT	Tc *ccdB* FWD3	TCCCTATCAGTGATAGAGATTGACATCCCTATCAGTGATAGAGATACTAG
	Tc FWD Infusion	CTGGCAATTCCGACGTTCCCTATCAGTGATAGAGATTGACATCCCTATCAG
	ccdB REV Infusion *CAT*	AGTGATTTTTTTCTCTATTCCCCAGAACATCAGGTTAATGGCGTTTTTGATGTCATTTTCGCGGT
Origin Change	SC101* FW Infusion	ATCCGCCGCCCTAGACCTAGG
	SC101* REV Infusion	GGGACTCTGGGGTTCGAGAGCTC
Sequencing	pZA24 Fwd	TGCGAAACGATCCTCATCCTGTCTCT
	pZSC101&15A REV SEQ	AGGAGAGCGTTCACCGACAAACAACA
pBAD Cloning	*TEV* FWD INFUSION LIBRARY	CCGCCCTAGACCTAGGACGTCTGTGTGGAATTGTGAGCGGATA
	pBAD REV+RBS	TATCCTTTCTCCTCTTTAATGAATTCTGTGTGAAATTGTTATCCG
Tig144 Cloning	*tig* FWD Infusion	AGAGGAGAAAGGATAATGCAAGTTTCAGTTGAAACCACTCAAGGC
	tig REV Infusion	TGTTGATGAACGCATCAGAGTATCCAGCATGCCGTCAACGTCAG
TEV for *Tig144* Cloning	*TEV* REV CLONING	AAACCCGTACCCTAGCATAGGCTTGGTTATGCCGGTACT
	TEV FWD CLONING2	ATGCGTTCATCAACAAGTTTGTACAAAAAAGCAGGCTCGGGAGAA
	TEV REV CLONING p15A	GGAATATATCCCTAGCATAGGCTTGGTTATGCCGGTACT
TEV for pBAD Cloning	*TEV* FWD CLONING for BAD	AGAGGAGAAAGGATAATGCGTTCATCAACAAGTTTG
	TEV REV CLONING	AAACCCGTACCCTAGCATAGGCTTGGTTATGCCGGTACT

Table 2-1 *continued.*

Application/ Use	Primer Name	Sequence
TEV for pBAD Cloning	*TEV* REV CLONING p15A	GGAATATATCCCTAGCATAGGC TTGGTTATGCCGGTACT
Tctev/*ccdB* _tev/*CAT* Cloning	Lib Tc infusion/lib FWD	CAGCTGGCAATTCCGACGTTCC CTATCAGTGATAGAGATTGACA T
	Lib Tc infusion/tev REV	AAAATACAGATTTTCACCGCCA GGCCTAGTATCTCTATCACTGAT
	Lib *ccdB* infusion/tev FWD	GAAAATCTGTATTTTCAGAGTC AGTTTAAGGTTTACACCTATAA A
	Lib *ccdB* infusion/tev REV	TCTGAAAATACAGATTTTCAGA ACCTATTCCCCAGAACATCAGG T
	Lib *CAT* infusion/tev FWD	ATCTGTATTTTCAGAGTGGCGA GAAAAAAATCACTGGATATACC A
	Lib *CAT* infusion/lib REV	CGGGCTGCAGGAATTTTACTGC CACTCATCGCAGTACTGTTGTAA
Tig144 StuI Removal	*tig144* REV *StuI*REMOVE	ACACGGCGGCCAAGGCCCTGAG TGGTTTCAACTGAAACTTGCATT
	tig144 FWD *StuI*REMOVE	ACCTATGAACTGTCGACTCGAG CATAGCATTTTTATCCATAAGAT
CcdB Add ATG	*ccdB* FWD1 ADDATG	TCAGGGCATGCAGTTTAAGGTT TACACCTATAAAAGAGAGAGCC
	ccdB FWD2 ADDATG	CCTGGCGGTGAAAATCTGTATTT TCAGGGCATGCAGTTTAAGGTT
	ccdB FWD infusion ADDATG	GATAGAGATACTAGGCCTGGCG GTGAAAATCTGTATTTTCAGGG C
	ccdB REV *BmgBI* infusion	TATCTGACAGCAGACGTGCACT GGCCAGGGGGATCA

2.6 ncRNA MAPPING USING 5' RACE

5' Rapid amplification of cDNA ends (5' RACE) is a technique used to determine the sequences at the beginning of RNA molecules; we used Super SMART^{TM}PCR cDNA Synthesis Kit (Clontech, CA) to determine the 5' ends of RNA. Briefly, RNA sample (1.5 µg) was mixed with gene specific primer(s)-1(GSP-1) (Table 2-1 under 5'RACE application) (84 pmoles total) and SmartOligo (84 pmoles), the mixture was heated to 65°C for 2 minutes. The reaction mixture was cooled to 42°C and mixed with reaction buffer (5x), DTT (2 mM, final concentration), dNTPs (2mM final concentration), RNase inhibitor (100 U, final concentration), and PowerScript Reverse Transcriptase (5 U, total) to generate the first-strand cDNA. First-strand cDNA was then column purified using NucleoSpin columns (Clontech, CA). cDNA was then amplified using Advantage II PCR kit (Clontech, CA) according to manufacturers instructions, we used SmartOligo primer and nested GSP-2 primers for amplification (Table 2-1 under 5'RACE application), the nested primer was designed downstream of GSP-1 to increase specificity. The reaction was titrated by removing 20 µl aliquots from the reaction starting at cycle 15 and from every third cycle following to a maximum of 33 cycles. Reaction products were resolved on TBE-PAGE gels (10%), stained with ethidium bromide, visualized and excised from the gel for, cloning and sequencing (similar to 2.5 above).

2.7 NORTHERN BLOTTING

Northern blotting was performed using 10% TBE-PAGE-Urea gels for RNA electrophoresis in 1 x TBE buffer. RNA was transferred by electro-blotting, using plate electrodes, for 1 h at 50 V, using 1 x TBE buffer as transfer medium. Zeta-probe GT nylon membranes (BIORAD, CA) were used for electro-blotting. Single-stranded biotin-labeled RNA Probes were prepared using a MAXIscript T7 Kit (Ambion, TX). Target sequences were amplified by PCR using chlamydial DNA as a template and a minimal T7 promoter sequence was added 5' to the anti-sense strand (Table 2-1 under Northern application). The resulting PCR fragment was fractionated on agarose gels, extracted, and used as a template for a larger scale PCR reaction. PCR product was purified and was used as a DNA template for an *in vitro* transcription reaction using MAXIscript T7 kit (Ambion, TX) and biotin-16-UTP (Roche, IN). The reaction was carried out for 1 h at 37°C, DNase treated for 15 min at 37°C, then the reaction was stopped by addition of 1 µl 0.5M EDTA. Buffer components and excess NTPs were removed by gel filtration using NucAway Spin Columns (Ambion, TX). Hybridization was carried out in UltraHyb (Ambion, TX) medium overnight at 65°C. Blots were washed using NorthernMax Low Stringency Buffer (Ambion, TX) at room temperature (2 x 5 min) and with NorthernMax High Stringency Buffer (Ambion, TX) at 65°C (2 x 60 min). Northern blots were developed using alkaline phosphatase and streptavidin according to manufacturer's protocols (BrightStar^{TM} BioDetect, Ambion).

2.8 GENERAL CLONING

For cloning we used two different systems, the first was the regular restriction/ligation/transformation and the second was recombination/transformation; recombination was more versatile in its use, as it was not generally connected to a specific restriction site, it was fast and reliable.

2.8.1 LIGATION

For ligation we used Fast-Link™ DNA Ligation Kit (Epicentre, WI) where target plasmid was linearized by single or double digestion using NEB restriction enzymes (NEB, MA), plasmid was either dephosphorylated using FastAP™ Thermosensitive Alkaline Phosphatase (Fermentase, MD) followed by ligation, or the plasmid was blunt ended using T4 DNA Polymerase (End-It™ DNA End-Repair Kit, Epicentre) followed by dephosphorylation for blunt end ligation. Insert was generally a PCR product, PCR products were gel purified, and either digested overnight by the appropriate restriction enzyme(s) for directional cloning or phosphorylated by T4 DNA Polynucleotide Kinase and ATP (End-It™ DNA End-Repair Kit, Epicentre) followed by ligation.

The End-It™ DNA End-Repair Kit (Epicentre, WI) was used to convert DNA with 5'-protruding and/or 3'-protruding ends to 5'-phosphorylated blunt-end DNA for blunt-end ligation. The conversion to blunt-end DNA was accomplished by exploiting the 5'→3' polymerase and the 3'→5' exonuclease activities of T4 DNA Polymerase, in addition, the enzyme mix also contained T4 DNA Polynucleotide Kinase and ATP carry out phosphorylation of the 5'-ends of PCR amplicons for subsequent ligation into a cloning vector. The protocol was carried out as follows, DNA (up to 5 µg) was resuspended (or eluted) in 34 µl TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA), 5 µl aliquots of 10X Buffer (330 mM Tris-acetate pH 7.8, 660 mM K acetate, 100 mM Mg acetate,5 mM DTT), ATP (10 mM), and dNTPs (2.5 mM each) were added to the DNA, followed by 1 µl of T4 DNA Polymerase/T4 Polynucleotide Kinase Enzyme Mix. The reaction mixture was kept at room temperature for 45 minutes followed by heat inactivation at 70°C for 10 min followed by column purification (MiniElute, Qiagen).

Ligation was carried out by Fast-Link Ligation kit (Epicentre, WI), where 1.5 µl of 10X Fast-Link Ligation Buffer (330 mM Tris-acetate pH 7.8, 660 mM K acetate, 100 mM Mg acetate,5 mM DTT), and 1.5 µl ATP (10 mM), were mixed with both vector and insert, the reaction volume was continued to 14 µl with water, then 1 µl Fast-Link DNA ligase (2U/µl) was added and the reaction kept at room temperature for 2 hour to over night. In case the reaction was a blunt end ligation, the amount of ATP was modified to 0.75 µl. The reaction was heated at 70°C for 15 min then transformed into chemically competent or electrocompetent *E. coli*. Calculation of insert and vector amounts used in ligation reaction were calculated as follows, the ratio of insert size to vector size was calculated and multiplied into the intended amount of vector to be used in nano grams then multiplied by 5, this equation calculated the amount in nano grams of insert used for ligation.

2.8.2 RECOMBINATION

Recombination was carried out using In-Fusion™ Advantage PCR Cloning Kit (Clontech, CA), recombination was used to join multiple pieces of DNA that have 15 bases of homology at their linear ends; recombination was typically used for fusing PCR products into vectors or joining different PCR products in one molecule with certain arrangement of DNA fragments dictated by the 15 base homology rule. Recombination offered a cloning technique that did not require the use of restriction enzymes, ligase or phosphatase.

The process first involved designing primers for insert amplification; the primers added 15 extra bases to each end of the amplicon, these extra bases were overlapping the ends of the linearized plasmid (Table 2-1 contains list of primers used for recombination under In Fusion application). In case of blunt cutters (*e.g.* StuI, BmgBI, ScaI) and 5' overhand cutters (AvrII, SalI, EcoRI) the 15 base overlap on the plasmid start at the cleaved nucleotide. In case of 3' overhand enzymes (AatII, SacI, KpnI), overlapping starts at the double stranded area of the plasmid ends, upstream of the single stranded staggered ends. Insert was amplified using Phusion™ High-Fidelity DNA Polymerase (Finnzymes Oy, MS), followed by gel/column purification (MiniElute clean up kit, Qiagen).

For recombining several PCR fragments together, all fragments were quantified and mixed to keep the molar ratio 1:1 as calculated by size ratio and concentration. For cloning an insert into a vector, the vector was digested for 3 hours to overnight with the appropriate restriction enzyme(s); restriction time depended on the stability of the enzyme in the reaction (NEB literature) and the units of enzyme used. Vector to insert molar ratio was 0.5, the amounts from each was calculated based on size ratios and concentrations. All recombining DNA fragments were mixed and volume was completed to 7 µl with water, then 2 µl of 5X buffer were mixed followed by 1 µl In-Fusion enzyme. The reaction mixture was then incubated at 37°C for 15 min followed by 15 min at 50°C. The reaction was stopped by adding 40 µl TE buffer then using 5 µl for transformation.

2.9 PROTEIN GEL ELECTROPHORESIS AND WESTERN BLOTTING

We used the NuPAGE® Bis-Tris Electrophoresis System (Invitrogen, CA). It is a neutral pH, discontinuous SDS-PAGE, pre-cast polyacrylamide mini-gel system. The neutral pH 7.0 environment during electrophoresis results in maximum stability of both proteins and gel matrix, providing sharp bands (Moos *et al*, 1988), samples were mixed with 2.5 µl of 4X NuPAGE® LDS Sample Buffer (106 mM Tris HCl, 141 mM Tris base, 2% (Lithium dodecyl sulfate [LDS], 10% Glycerol, 0.51 mM EDTA, 0.22 mM SERVA® Blue G250, 0.175 mM Phenol Red, pH 8.5). and 1 µl NuPAGE® Reducing Agent (500 mM DTT), water was added to complete the volume to 10 µl. Samples were heated at 70°C for 10 min and then loaded on the gel. Gels used were NuPAGE 4-12% Bis-Tris Gels, 1X running buffer was prepared by diluting 20X NuPAGE® MES SDS Running

Buffer (50 mM MES 50 mM Tris base 0.1% SDS 1 mM EDTA pH 7.3) and adding 500µl of NuPAGE® Reducing Agent (500 mM DTT) per 200 ml of 1X NuPAGE® MES SDS Running Buffer. The precast gel and running buffer were loaded into XCell *SureLock*. Mini-Cell according to manufacturer's instructions; gels loaded with preheated samples and protein Marker (BIORAD, CA) and were run at 200 volts for 1 hour. Gels were either stained or transferred to PVDF membranes for Western blotting. Staining was carried out as follows, NuPAGE® Gels were stained with SimplyBlue SafeStain (Invitrogen, CA), after electrophoresis, gels were placed in 100 ml of ultrapure water and microwaved on High (950 to 1100 watts) for 1 minute, gels were then put on an orbital shaker for 1 minute, the heating and shaking was repeated once more. Water was drained and 20 mls SimplyBlue SafeStain was added and the gel was microwaved again for 45 seconds to 1 minute, followed by shaking for 5 min, gels were washed with 100 ml of ultrapure water for 10 minutes on the shaker followed by 20 ml of 20% NaCl for 5 minutes where it was stored refrigerated.

Transfer was carried out using XCell II™ transfer cassette was assembled according the manufacturer's manual, transfer was carried out using 1X Transfer buffer, [20X NuPAGE® Transfer Buffer (25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA, pH 7.2) 50 ml, NuPAGE® Antioxidant 1 ml, Methanol 100 ml, Deionized Water 849 m)]. Electrotransfer was carried out at 30 volts for 1 hour. Membranes were blocked over night using Blocking Buffer (1X TBS, 0.1% Tween-20, 5% w/v nonfat dry milk), membranes were washed 4X15min using TBS-T (1X TBS, 0.1% Tween-20), followed by incubation with the primary antibody diluted in the blocking buffer over night. Membranes were then washed 4X15 min using TBS-T, then incubated with the HRP-secondary antibody diluted in blocking buffer for 3 hours, membranes were washed 4X15 min using TBS-T and developed by ECL (Abcam, MA), the result was visualized using Kodak BioMax XAR Film (Fisher Scientific, PA).

2.10 FUNCTIONAL ANALYSIS OF *CTIG270* IN A SURROGATE SYSTEM

To test the effect of expression of *CTIG270* on *ftsI*, a system was constructed where both *ftsI* and *CTIG270* were cloned in two compatible medium copy number plasmids and transformed into *E. coli*. *ftsI* was fused in frame to a 5' FLAG-tag sequence ($_N$-DYKDDDDK-$_C$) and cloned under the control of a lac promoter (from pCR2.1, Invitrogen) Figure 2-1.

The FLAG-tag was added using PCR (Table 2-1 *ftsI*-FLAG Cloning and *CTIG270* Cloning). Cloning was carried out in a promoterless pSMART-LCAmp (Lucigen, WI). PCR was used to add an *FseI* site to the 3' end and 5' end of the promoter and the FLAG-tagged *ftsI* respectively. Both amplicons were digested over night with *FseI* followed by ligation. The fusion product was then amplified by PCR and blunt end cloned into pSMART-LCAmp. *CTIG270* was cloned under the control of an arabinose responsive promoter (pBAD) in pRANGER-BTB vector (Lucigen, WI). *E. coli* transformed with both plasmids was induced with IPTG and/or arabinose. *E. coli* was harvested and analyzed using Northern blotting for both mRNA and ncRNA.

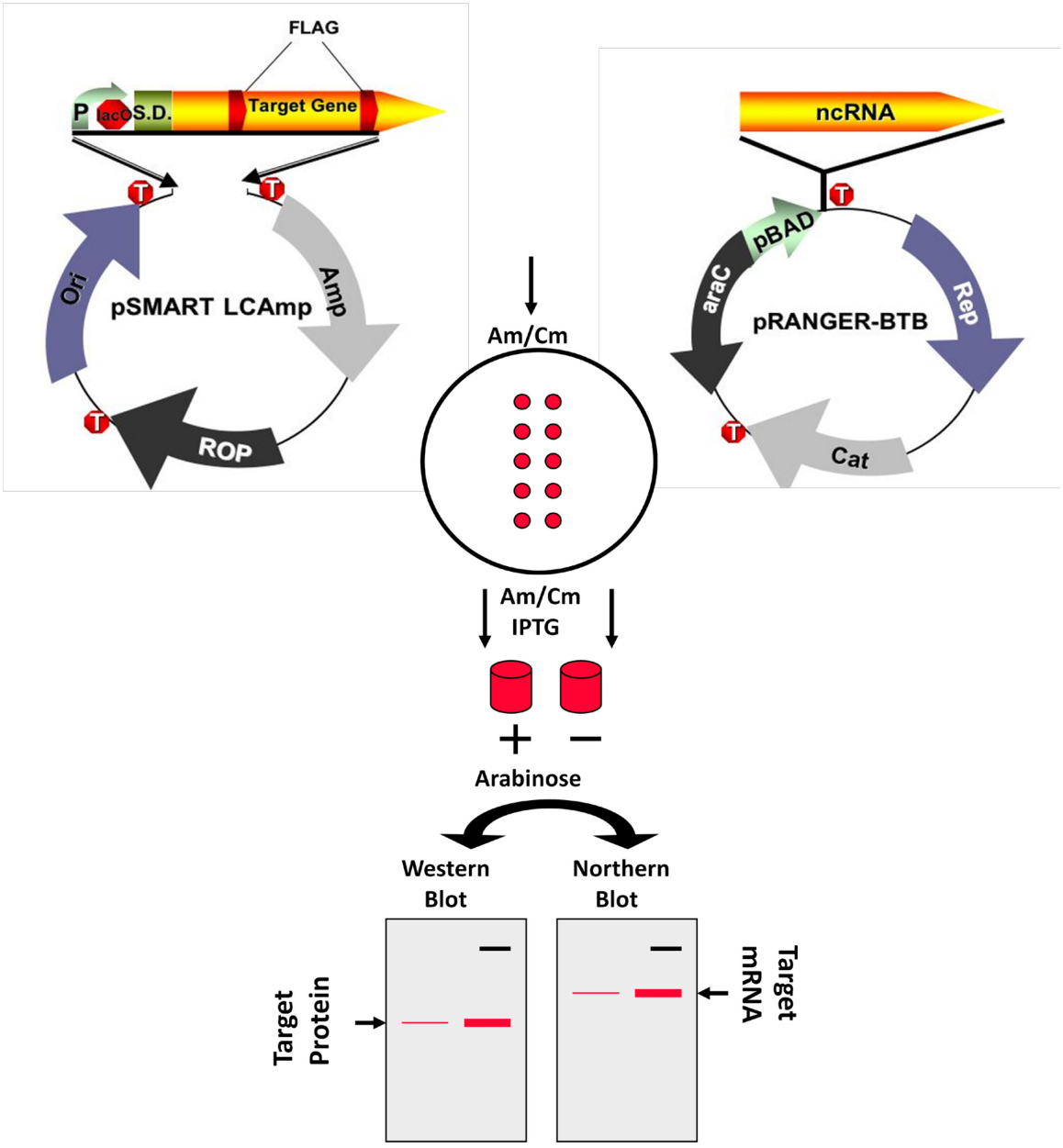

Figure 2-1 Scheme for testing the function of possible *cis* acting ncRNAs. The figure shows 2 plasmid system, one carries a FLAG-tagged target gene under pLAC promoter and the other carries the ncRNA under test controlled by pBAD promoter.

Northern blotting demonstrated transcription and stability of both RNAs. FtsI protein was also checked using Western blotting.

E. coli transformed with both plasmids was grown to an OD_{600} of 0.6. The culture was split into three tubes. One tube was induced with arabinose (0.2%), the second tube was induced with IPTG (1mM) and the third one was uninduced. All three tubes were incubated at 37°C on a shaker for 2.5 hours. A sample was collected from each test condition and cells were centrifuged. Bacterial cells were used to prepare RNA (MasterPure, Epicentre) for Northern Blotting and protein for Western blotting. Following sample collection each culture tube received the second inducer. The first tube was induced with IPTG (1mM), the second tube was induced with arabinose (0.2%) and the third tube was induced with both IPTG (1mM) and arabinose (0.2%). The culture tubes were incubated at 37°C on a shaker for 2.5 hours and a second sample was collected from each tube as previously mentioned. Culture tubes were incubated at 37°C on a shaker for 2.5 hours and the third sample was collected as previously mentioned. RNA and protein samples were resolved by gel electrophoreses and loading was normalized using OD_{600} measurements. Protein samples were resolved on a NuPAGE Bis Tris gel and transferred onto a PVDF membrane according to manufactures instructions (Invitrogen, CA). Western blotting was carried out using the ProteoQwest™ FLAG Chemiluminescent Western Blotting Kit (SIGMA-ALDRICH, MO). RNA samples used for *CTIG270* detection were resolved on a 10% Urea PAGE gels, while RNA samples for *ftsI* detection were resolved on a 1% formaldehyde agarose (NorthernMax™, Ambion). RNA was then transferred to a Zeta probe GT Nylon membranes according to manufacturer's instructions (BIORAD, CA). Northern blotting was carried out using biotinylated RNA probes specific to *CTIG270* and *ftsI*.

CHAPTER 3. DISCOVERY AND VALIDATION PHASE

3.1 CONSTRUCTION AND TESTING OF THE CHLAMYDIAL INTERGENIC MICROARRAY

3.1.1 CONSTRUCTION OF THE CHLAMYDIAL INTERGENIC MICROARRAY

Using intergenic microarrays for determining ncRNA expression represents an unbiased, hypothesis-neutral, and direct approach. We have designed and constructed a GeneChip Custom Affymetrix microarray to determine the expression levels of ORFs and untranslated elements in IGRs of *C. trachomatis* D (strain UW-3/CX) based on the genomic sequence (Stephens *et al*, 1998).

The IGRs of the *C. trachomatis* D and the *C. pneumoniae* AR-39 genomes that were 50 nucleotides or more were represented on the microarray. IGRs were represented as 25 nucleotides probes tiled head to tail on both strands. IGR probes present certain problems in the design process. The defined nature of the target sequences (*i.e.* tandem 25'mers) greatly restricted the selection process, and target probe acceptance values had to be lowered to a certain extent. The initial design submission returned 82% of the tiled probes as acceptable. To increase the flexibility of the algorithms selection process for the remaining 18% of the target regions were expanded by 5 nucleotides at the 5' and 3' ends (*i.e.* target sequence became 35 nucleotides) resulting in a coverage of 92% of the intergenic regions. This process was repeated by extending the remaining 8% region by 10 nucleotides in the 5' and 3' direction, resulting in 98% coverage of the intergenic region (6695 probes). Certain probes could not be synthesized (145) due to secondary structure or synthesis step problems. None of these were adjacent and therefore the transcriptional status of these regions can be conditionally estimated by comparison to flanking probes. The need to lower acceptance values for probe design in the IGRs made the synthesis of MM probes questionable and, at the suggestion of the Affymetrix Design Team; these probes were not included on the microarray. The expression results for the IGRs have therefore been interpreted in a contextual manner and require experimental controls.

An example of the microarray results are shown in Figure 3-1 in which the fluorescence intensities of the microarray have been converted to a genomic "Array Layout" display for ORFs and IGRs (Belland *et al*, 2003b). The region displayed in detail is the intergenic region containing *CT269* (*murE*) and *CT270* (*ftsI*) and a potential *cis*-acting ncRNA (*CTIG270*) that is expressed downstream of *ftsI* on the opposite strand. The display in Figure 3-1 shows each ORF as a single filled circle on the upper or lower strand depending on the orientation of the ORF (color represents the mean fluorescence intensity for the multiple ORF probes) while the intergenic region is shown as several double circles each representing a single 25 nucleotides probe on one of the two DNA strands (colors represent fluorescence intensity).

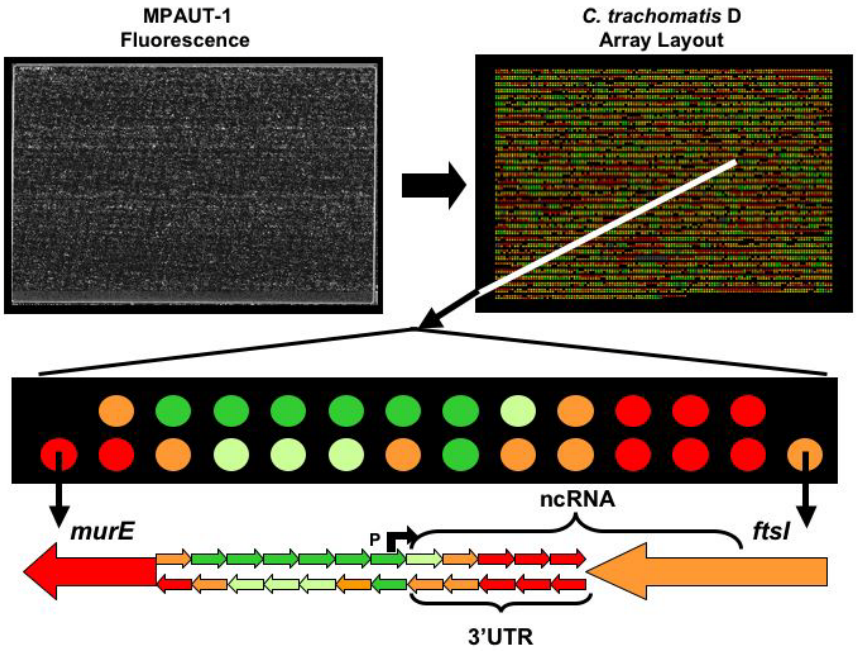

Figure 3-1 The MPAUT-1 microarray and a schematic display of experimental data. The figure shows the actual scanner-generated image of the array (the left hand side) and the array layout (right hand side): which is a computer generated image converting fluorescence intensities to colors of different shades (green being lowest expression and red being highest expression values). The figure also shows a magnified segment of the chromosome detailing *CT269*, *CT270* and a possible antisense RNA to *CT270*. Each ORF is represented by a circle, the direction of the ORF dictates whether the circle is located on the upper line (upper strand) or on the lower line (lower strand). The IGR is represented by a string of circles on both lines, each circle represent a single 25 mer probe.

3.1.2 TESTING THE MICROARRAY

Intergenic microarrays pose a challenge because of probe quality, the limited space available within the target sequence for maneuvering, and getting a probe with good hybridization parameters. Preliminary experiments were performed using *C. trachomatis* D chromosomal DNA. Chromosomal DNA offered a linear representation of all sequences on the array at a 1:1 ratio. We used labeled chromosomal DNA to check the intergenic microarray probe quality. DNA from *C. trachomatis* D was fragmented via partial DNase I digestion (Amersham, UK) and Biotin end-labeled with terminal transferase (Invitrogen, CA). The DNA array result summary is shown in Table 3-1.

ORFs had a mean fluorescence intensity of 2008 arbitrary fluorescence units and a standard deviation of ±244. This tight distribution is the consequence of each ORF being represented by 11-8 probe pairs (PM and MM) and the fluorescence represented the mean value according to the equation $\sum PM - MM/11$. On the other hand, probes representing IGRs had a mean fluorescence intensity of 2953 units and a standard deviation of ±1394. The apparent wide range of the IGR results was because the fluorescence of each probe was reported as an individual event. The DNA array analyses indicated that 399 out of 6695 (5.9%) probes in IGRs had a fluorescence value less than 500. The lower quality of these probes was taken into consideration when examining RNA microarray results.

3.2 RNA PREPARATION

RNA prepared from *Chlamydia*-infected cells is composed of a mixture of host RNAs (rRNA, mRNA, tRNA) and bacterial RNAs (rRNA, mRNA, tRNA) in which the former represents the vast majority. To enrich for bacterial mRNA we tested procedures that selectively remove host mRNA and bacterial and host structural RNAs (rRNA and tRNA) but retained small sized RNA species. Three enrichment procedures were carried out and compared using the MPAUT-1 intergenic microarray (summarized in Figure 3-2).

The first condition (labeled 1 in Figure 3-2) was a process the involved the removal of host mRNA by treatment with a $dC_{10}T_{30}$ oligo-nucleotide linked to polystyrene-latex particles (OligoTex^{TM}, Qiagen). The second condition involved

Table 3-1 *C. trachomatis* D-DNA microarray results summary.

Parameter	ORFs	IGRs
Mean fluorescence intensity	2008	2953
Median fluorescence intensity	2004	2892
Standard deviation	244	1394
Highest value	2981	8015
Lowest value	788	500

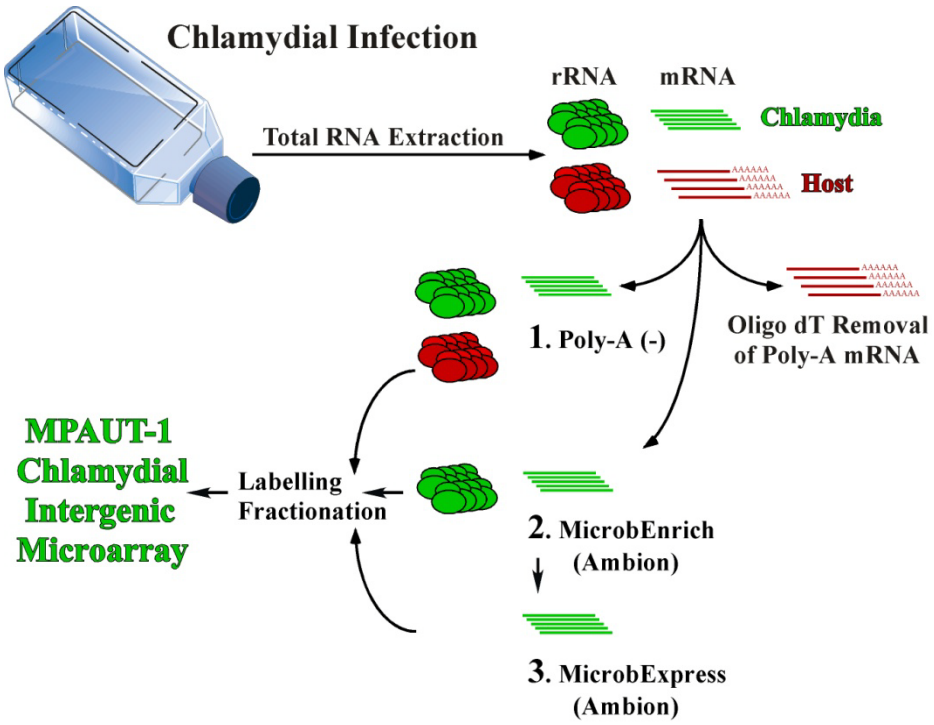

Figure 3-2 Enrichment procedures for chlamydial RNA for use with MPAUT-1.

the removal of host mRNA and structural RNAs (labeled 2 in Figure 3-2) using "capture" oligonucleotides coupled to magnetic beads that remove polyadenylated mRNA and 18S and 28S rRNAs (capture oligonucleotides are complementary to 18S and 28S sequences, MICROB*Enrich*™, Ambion). The third condition was an extension of the second procedure (labeled 3 In Figure 3-2) and involved further removal of bacterial 16S and 23S rRNAs using capture oligonucleotides (complementary to 16S and 23S rRNA sequences of *C. trachomatis*, MICROB*Express*™, Ambion).

The MPAUT-1 microarrays were used to compare the effect of these various enrichment methods on both sensitivity and specificity of the microarray. Equivalent quantities of RNA were used for random-primed cDNA synthesis, fragmentation with DNase I, and labeling with biotin using T4 Terminal transferase (according to Expression and Analysis Technical Manual, Prokaryotic Probe Processing, Affymetrix).

Comparison of different enrichment schemes indicated that all conditions showed acceptable levels of sensitivity. Sensitivity was estimated based on the number of ORFs flagged as present in each experiment (814, 760 and 864 for the three treatment protocols respectively). As shown in Figure 3-3, comparing the correlation coefficient of ORFs' fluorescence values for each treatment against the MICROB*Express*™ protocol showed high levels of reproducibility, (OligoTex™ to MICROB*Express*™ had an $R^2 = 0.8228$ and MICROB*Enrich*™ to MICROB*Express*™ had an $R^2 = 0.8764$).

The "scaling factor" calculation was used as an estimate of enrichment. Scaling is a mathematical technique used by GeneChip Operating Software (GCOS) to minimize differences in overall signal intensities between two or more arrays. This mathematical normalization allowed for a reliable detection of biologically relevant changes between different samples. GCOS calculates the overall intensity of an array by averaging the intensity values of a preset reference probe matrix (in this case we set chlamydial ORFs as the reference matrix). The average intensity of the array was then multiplied by the scaling factor to bring it to an arbitrary target intensity value (500). Thus, scaling allows a number of experiments to become normalized to one target intensity, and direct comparisons between any two experiments can be carried out. The scaling factor provided a measure of the brightness of the array. This measure reflected the availability of bacterial mRNA in matched test samples. By setting the reference probe matrix to chlamydial ORFs a high scaling factor can then be interpreted as a dim array, which might reflect lower abundance of chlamydial mRNA. On the other hand, low scaling factors reflect a bright array; an indication of abundant chlamydial mRNA on the array. The calculated scaling factors for every enrichment scheme reflected the relative enrichment of bacterial mRNA in each sample as follows; for OligoTex™ and MICROB*Enrich*™ it was 25.8 and 36.5 respectively, while for MICROB*Express*™ it was 2.5977. For this reason we chose the MICROB*Express*™ enrichment scheme for the following experiments.

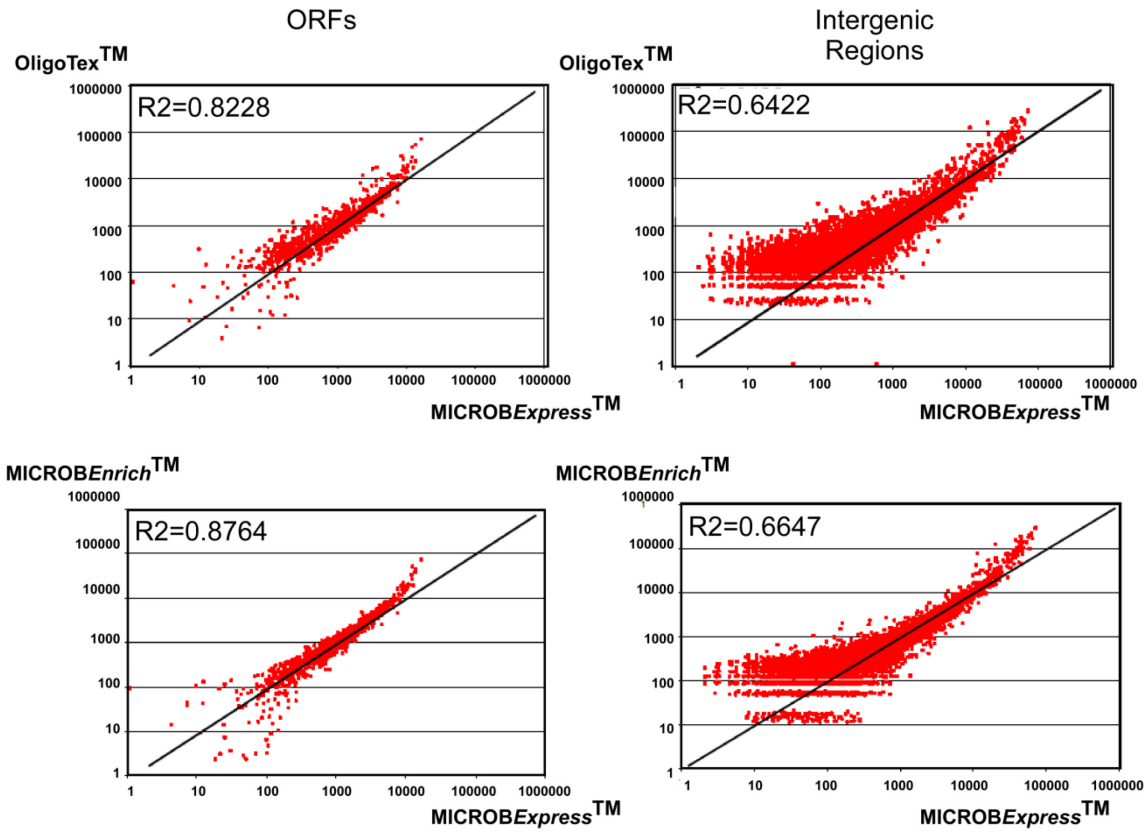

Figure 3-3 Comparison of RNA purification protocols for use with the MPAUT-1 microarray.

3.3 IDENTIFICATION OF *C. trachomatis* ncRNAs

3.3.1 GUIDELINES FOR IDENTIFICATION OF C. trachomatis ncRNAs

Potential ncRNAs were named using the IGR number corresponding to the downstream ORF using the CT numbers defined by Stephens *et al*, 1998 *e.g.* an ncRNA located upstream of *CT241* is referred to as *CTIG241*. *C. trachomatis* D RNA collected from 40 h post infection (PI) samples was processed and used to characterize the expression of ncRNAs. Identification of potential ncRNAs was based on a number of guidelines including: i) sequential probes showing simultaneous expression (3 or more), ii) high levels of fluorescence from at least one of the IGR probes (>1000), iii) relatively high G+C ratio of the expressed region (elevated G+C ratios have been shown to predict biological function (Zhang *et al*, 2004)), iv) presence of *rho*-independent terminators at the end of the ncRNA, v) location with respect to flanking ORFs, and vi) genomic conservation with other chlamydial species. An example of this type of analysis is shown in Figure 3-4 in which the expression of a potential ncRNA (*CTIG241*) is compared to expression of tRNA-Leu-2. The IGRs and flanking ORFs are shown with corresponding expression levels and G+C ratio.

3.3.2 RESULTS OF MICROARRAY SCREENING

As shown in Table 3-2, House-keeping ncRNAs were used as a validation tool for microarray results. We were able to detect all 37 tRNAs predicted in the genome of *C. trachomatis* D (including the tRNA-Leu-2 in Figure 3-4). In addition, a number of highly conserved ncRNAs found in other bacterial genera were also detected. These included: i) tmRNA (*CTIG019*), associated with the release of stalled ribosomes and the incorporation of a short peptide tag that targets the partial protein product for degradation (Withey and Friedman, 2003), ii) 4.5S RNA (*CTIG255*), which forms part of a ribonucleoprotein complex that plays a role in protein secretion (Poritz *et al*, 1990), iii) M1 RNA (*CTIG399*), which forms the catalytic subunit of RNase P, involved in 5' tRNA processing (Kleineidam *et al*, 1993) and iv) and both 5S rRNAs (*CTIG740* and *CTIG750*) which is involved in the ribosome assembly (reviewed in Gongadze *et al*, 2008).

Microarray analysis showed that *C. trachomatis* D expressed 34 other intergenic sequences that met the guidelines for ncRNA identification. These included ncRNAs that could be classified as *cis*-acting antisense molecules in that the expressed transcript overlapped the 3' end of a transcript from an annotated gene in the opposite direction (as shown for *CTIG270* ncRNA and *ftsI* gene in Figure 3-1, and for *CTIG153* ncRNA and *CT152* [*lolD*]). Potential intergenic ncRNAs were also found that fit into the category of *trans*-acting molecules. Table 3-2 summarizes the Northern blot results for all 34 potential ncRNAs (14 potential ncRNAs were detected by Northern blotting). An average G+C ratio of 0.41 was selected as a threshold ratio for ncRNAs. This ratio was selected because this was the average G+C ratio for *C. trachomatis* D chromosome. Out of the 14 transcripts detected by Northern blotting 3 ncRNAs had an average G+C

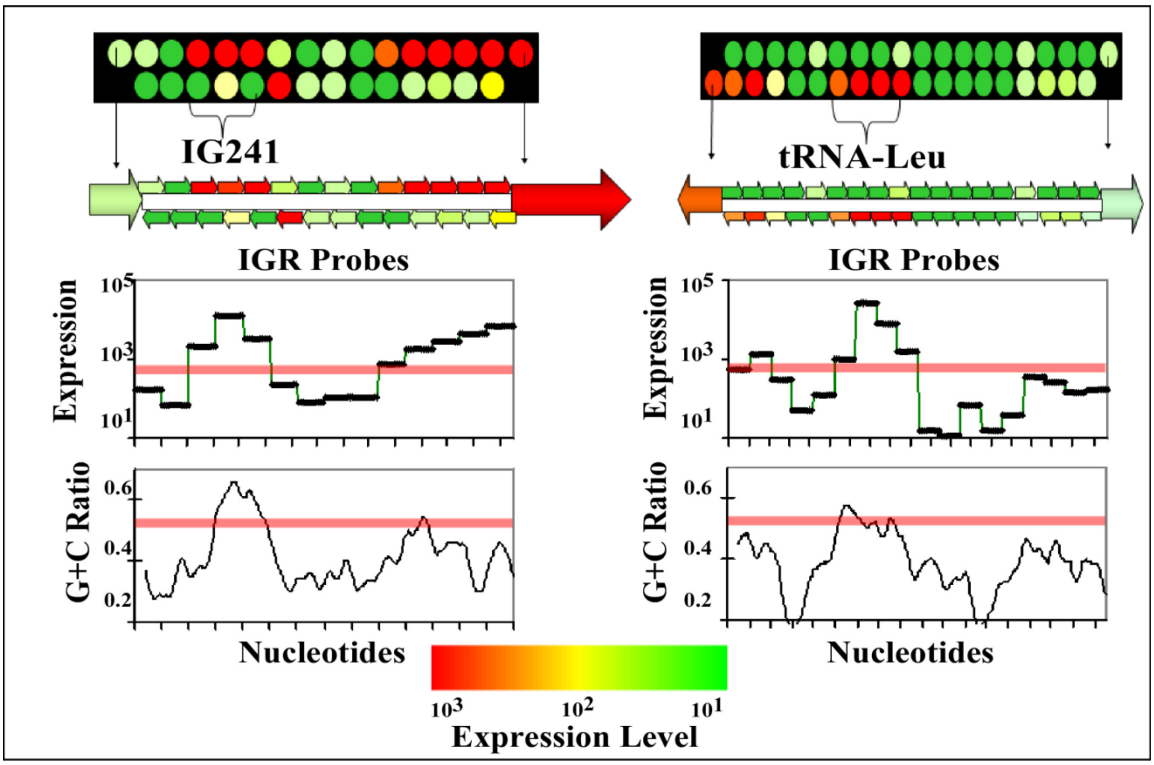

Figure 3-4 Guidelines for identification of the ncRNAs. Identification involved the expression levels of a contiguous intergenic regions, G+C ratio, and position with respect to flanking ORFs. Shown here is a comparison of the *CTIG241* potential ncRNA and the tRNA-Leu. Threshold levels for inclusion are shown as pink bars across the tiled intergenic regions.

Table 3-2 Results of the 40h PI microarray experiment showing all known ncRNAs (tRNA, tmRNA, 4.5SRNA, 5SRNA, M1RNA, and ihtA) in addition to potential ncRNAs.

Name	Approx. Size	Class[1]	A.F.U.[2]	Northern	*Rho* Indpt[3]	Average G+C Ratio
37tRNAs	75-100	tRNA	25927	NA	**NA**	0.51
tmRNA	525	PI	24598	NA	−	0.45
4.5S RNA	100	PI	15917	Positive	−	0.51
5S rRNA	125	Ribosomal	32001	NA	−	0.47
ihtA	150	*trans*/3'UTR	9938	Positive	√	0.46
M1 RNA	400	PI	20235	Positive	−	0.43
CTIG153	75	*cis*	884	Positive	−	0.52
CTIG241	75	*trans*	6073	Positive	−	0.53
CTIG270	100	*cis*	994	Positive	−	0.36
CTIG327	250	*trans*	1926	Positive	√	0.37
CTIG356	225	*trans*/UTR	7238	Positive		0.44
CTIG360	200	*trans*/Riboswitch	4533	Positive	√	0.43
CTIG498	100	*trans*/3'UTR	3976	Positive	√	0.46
CTIG504	75	*trans*/3'UTR	3959	Positive	√	0.42
CTIG582	200	*trans*	2135	Positive	√	0.39
CTIG643	650	*trans*	7279	Positive	√	0.41
CTIG684	100	*trans*	1420	Positive	√	0.47
CTIG805	200	*trans*	5677	Positive	−	0.37
CTIG809	125	*trans*	748	Positive	−	0.44
CTIG857	225	*trans*	772	Positive	−	0.45
CTIG001	125	*trans*	1396	Negative	−	0.35
CTIG059	50	*trans*	1281	Negative	−	0.57
CTIG072	175	*trans*	2780	Negative	−	0.39
CTIG181	75	*trans*	4001	Negative	−	0.36
CTIG195	200	*trans*	1852	Negative	−	0.4
CTIG237	325	*trans*	6378	Negative	−	0.4
CTIG256	150	*trans*	1941	Negative	−	0.38
CTIG268	175	*trans*/5'UTR	5765	Negative	−	0.43
CTIG323	225	*trans*	4479	Negative	−	0.41
CTIG433	75	*trans*/5'UTR	2723	Negative	−	0.44
CTIG442	200	*trans*/3'UTR	2339	Negative	−	0.4
CTIG444	125	*cis*	3620	Negative	−	0.37
CTIG449	125	*trans*	4248	Negative	√	0.49
CTIG592	150	*trans*	1335	Negative	−	0.36
CTIG660	125	*trans*/3'UTR	14347	Negative	−	0.43
CTIG663	250	*trans*	2784	Negative	−	0.44
CTIG775	175	*trans*	2877	Negative	−	0.39

Table 3-2 *continued*

Name	Approx. Size	Class[1]	A.F.U.[2]	Northern	*Rho* Indpt[3]	Average G+C Ratio
CTIG813	225	*trans*	1364	Negative	–	0.34
CTIG864	125	*trans*	19393	Negative	–	0.41

Approximate sizes are based on the array, the potential functional class, the average fluorescence of the probes representing the ncRNA, presence of *Rho* independent terminator, and the average G+C ratio across the region in question.
[1] Class of ncRNA refers to its presumed mode of action. PI is protein interaction.
[2] A.F.U. Average Fluorescence Units.
[3] *Rho* indpt is *Rho* Independent terminator as analyzed by www.softberry.com.

ratio bellow 0.41. All other tested ncRNAs and house-keeping ncRNAs had an average G+C ratio between 0.43 and 0.53.

Interestingly, 8 of these 14 ncRNAs had a potential *rho* independent terminator. ncRNAs that were not detected by Northern blotting generally had lower G+C ratio (11 out of 20 were bellow 0.41), and 1 had a potential *rho* independent terminator. Collectively these results pointed to the potential expression of new, previously uncharacterized, RNA molecules during the normal developmental cycle of *C. trachomatis* D.

3.3.3 VALIDATION OF ncRNA EXPRESSION

RNA samples were DNAsed (TurboDNAse, Ambion) and resolved on 10% UREA PAGE gels (Criterion, BioRad), transferred onto nylon membranes (GT-Zeta-Prob GT, BioRad), and probed with biotin-labeled single stranded RNA probes. The probes were prepared by adding a T7 minimal promoter sequence to PCR amplified IGRs (Northern Blotting, section 2.7 above). The microarray was used as guide to determine the orientation of the potential ncRNA (thus decide which strand to probe) and to narrow the target area, thus preventing interfering signals from nearby genes and/or tRNAs. T7 PCR products were then subjected to an *in vitro* transcription reaction containing 16-UTP-BIOTIN (Roche Applied Science, IN). The product was then DNAsed and stored at -80°C. By selecting the correct primer for the T7 promoter and by decreasing the size of the PCR to a minimal size (150-200 bases), we were able to probe every potential ncRNA. The Blots were then visualized by streptavidin bound alkaline phosphates and treated with the substrate, CDP-Star according to manufacturer's instructions (BrightStar, Ambion). Table 2-1 shows a list of primers used for probe preparation and Northern blotting.

A subset of the potential ncRNAs that were verified by Northern blotting (Figure 3-5) included *CTIG153, CTIG241, CTIG270, CTIG327, CTIG356, CTIG360, CTIG498, CTIG504, CTIG643,* and *CTIG684*. Figure 3-5 also showed the results of some known ncRNAs including tRNA leu, 4.5S RNA, M1 RNA, plasmid antisense transcripts (Fahr *et al*, 1992), and *ihtA* (Grieshaber *et al*, 2006). *ihtA* has been shown to translationally-silence expression of the chlamydial histone HctA in a heterologous *E. coli* co-expression system. A schematic of the chromosomal arrangement of the locus for each ncRNA is shown in Figure 3-6.

CTIG153 had two products ranging in size from 120 to 170 nucleotides (Figure 3-5). The chromosomal locus in Figure 3-6 showed that *CTIG153* is located on the lower strand running towards *CT152*. Upon further inspection, a *rho* independent terminator was detected 14 bases upstream from the *CT152* stop codon creating a possible overlap of 60 nucleotides between the ncRNA and *CT152* (14 bases upstream of the terminator plus 46 nucleotides that form the terminator stem loop). This architecture is strongly suggestive of an antisense mode of action. Moreover, the fact that the Northern blot showed two bands suggest that *CTIG153* is processed in the course of interaction with

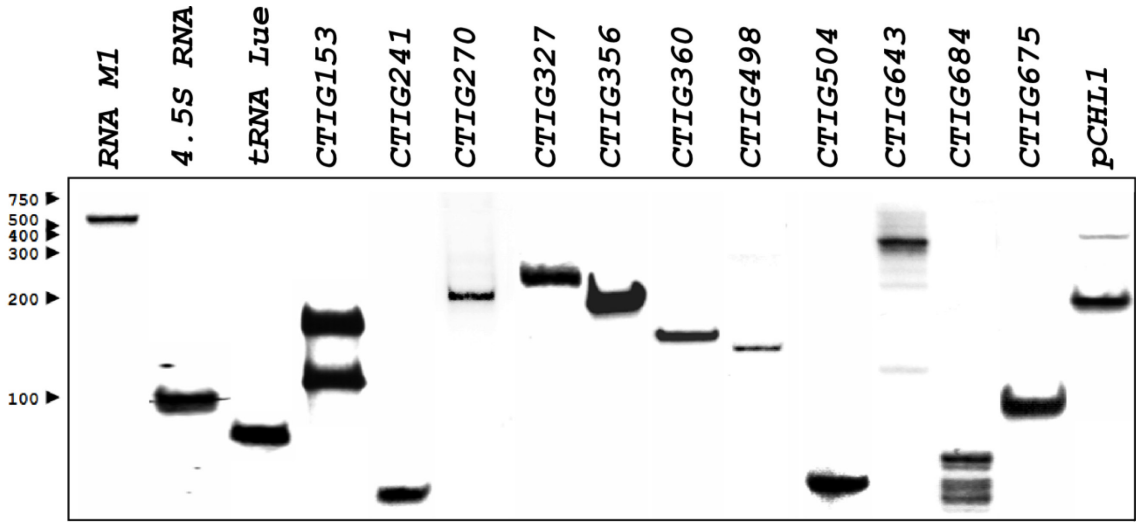

Figure 3-5 Northern blotting of potential ncRNAs from *C. trachomatis* at 40 h PI. House-keeping small RNAs (RNase P, 4.5S RNA, and tRNA-Leu-2) and several novel (*CTIG153, 241, 270, 327, 356, 360, 498, 504, 643,* and *684*) and previously described IhtA (Grieshaber *et al*, 2006) and plasmid antisense transcripts (Fahr *et al*, 1992) RNAs are shown. Chlamydial ncRNAs were probed with strand-specific, biotin-labeled probes using equivalent amounts of non-enriched RNAs.

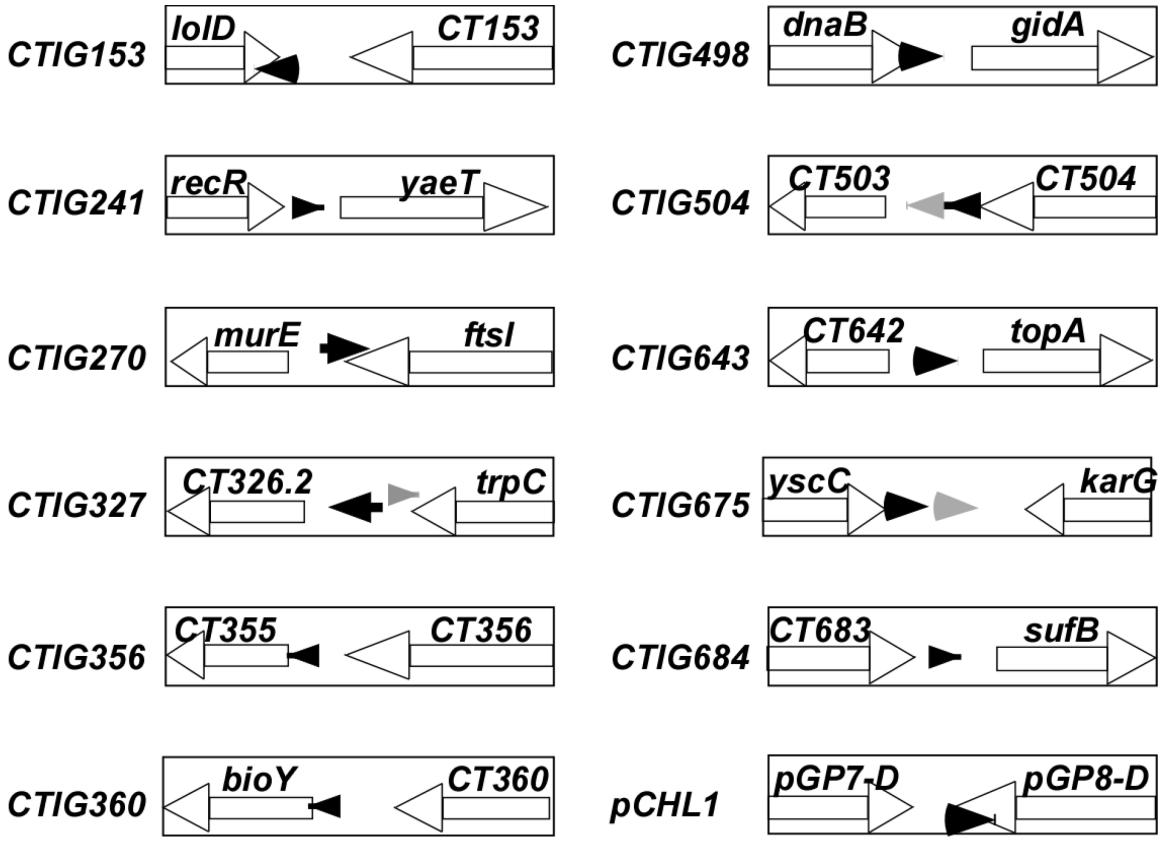

Figure 3-6 A schematic of the chromosomal locus for each ncRNA shown with the two flanking genes. Black arrowheads represent ncRNA and gray arrowheads represent tRNAs.

CT152. The small band might be a partially processed RNA that is missing its 3' end. In order to prove this hypothesis another set of probes were designed. One overlapped the 5' end of *CTIG153* and the other overlapped its 3' end. Northern blotting showed that only the 5' probe was able to identify both bands while the 3' probe identified the large product only (data not shown).

CTIG241 is located upstream of *CT241*. Generally ncRNAs located upstream of genes on the same strand might be 5'UTR for those genes. *CTIG241* is located about 50 nucleotides downstream of *CT240* and 250 nucleotides upstream of *CT241*. Although no obvious *rho* independent terminator was located in the vicinity of the ncRNA; a classical stem loop ending in a group of Us was located at the 3' end of the ncRNA. This potential terminator was not reported by the software (www.softberry.com) because it was destabilized by two bulges. The software might have assigned this terminator an inferior score. Analysis of the G+C ratio in this locus shows an unusual bias towards a high the G+C ratio in the ncRNA location (Figure 3-4). *CTIG241* does not overlap with flanking genes which makes it a potential *trans*-acting ncRNA.

CTIG270 deviated from the guidelines of selection of ncRNAs in several aspects. It had a low G+C ratio (0.36), and had no terminator structure within the IGR or in the downstream gene (Table 3-1). The sequence directly upstream of *CTIG270* contained several repeats suggesting a protein-DNA binding site(s). The signal detected on the Northern blot for *CTIG270* was considered the weakest in all the examined ncRNAs. The chromosomal arrangement showed that *CTIG270* ran on the upper strand and overlapped the 3' end of *CT270* (*ftsI*). The architecture suggests an anti sense mechanism of action that affects the stability of *ftsI*. On the contrary to *CTIG153*, a single major product was detected for *CTIG270*. This was possibly due to complete degradation of the ncRNA when it interacts with its target gene. Comparing the chromosomal locus in both *C. trachomatis* and *C. muridarum* revealed that the locus in *C. muridarum* contains an annotated ORF in place of *CTIG270*. The annotated ORF was located on the upper strand opposite to the flanking genes. Its first ATG overlapped *murE* coding sequence by 15 nucleotides (named *murG* in *C. muridarum*) and ends 7 nucleotides before the *ftsI* stop codon. The presence of this ORF led us to investigate if *CTIG270* as identified in *C. trachomatis* represented a pseudogene. In order to address this question, we synthesized a series of ~200 nucleotide probes spaced over the IGR from the start of *murE* to ~500 nucleotides inside *ftsI*, (Table 2-1 primers listed under Probe -219 to -18, Probe -106 to +67, Probe +83 to +264, and Probe +275 to +494, numbering starts at *CTIG270* TSS as +1) Northern blot analyses indicated that *CTIG270* was located 100-200 nucleotides upstream of *ftsI* stop codon, and about 100-200 nucleotides downstream of *murE* start codon. Probe -219 to -18 covered the area where the possible pseudogene would be located. This probe failed to detect *CTIG270*.

CTIG327 is located 200 nucleotides downstream of *CT326*.2 and 250 nucleotides upstream of *CT327*. All three genes were located on the lower strand and in the same orientation. There was a *rho* independent terminator located at the end of *CTIG327*. Upon examining the sequence within this intergenic region, a possible unannotated ORF of 204 nt was found overlapping *CTIG327*. The unannotated ORF ended 42 nt

downstream of the *rho* independent terminator. We investigated whether the product detected on the Northern blot represented *CTIG327* or represented the unannotated ORF. Two probes were designed. One overlapped the 5' end of the message and the other overlaps the region downstream of the terminator. Northern blotting result showed that the product could only be detected when using the probe overlapping the 5' end. This result showed that the product belongs to the ncRNA not the unannotated ORF (data not shown).

CTIG356 and ***CTIG360***; both had similar chromosomal arrangements. Their arrangements suggested that they might be 5' UTRs for there upstream genes (*CT355* and *CT359* respectively). Table 3-2 denoted *CTIG360* as a potential riboswitch. Riboswitches are RNA elements that undergo a shift in structure in response to binding of a small molecule. Such structural shifts affect stability and or translatability of the message. These elements are encoded within the transcript they regulate. They act in *cis* to control expression of the coding sequence within that transcript. Riboswitch RNAs control a broad range of genes in bacterial species, including those involved in metabolism, uptake of amino acids, cofactors, nucleotides, and metal ions (reviewed in Henkin, 2008). *CTIG360* was located directly upstream of *CT359*, which is a *bioY* gene. Orthologue of *bioY* in *Rhodobacter capsulatus* was able to confer biotin uptake phenotype to a mutant *E. coli* strain that was deficient in biotin uptake (Hebbeln *et al*, 2007). Latter results will add more to this discussion.

CTIG498 and ***CTIG504*** overlapped the 3'UTR of *CT497* and *CT504* respectively. Both genes were reminiscent of *ihtA* chromosomal arrangement. Both ended in a *rho* independent terminator. Moreover, *CTIG504* and *ihtA* have a tRNA directly downstream of the *rho* independent terminator.

CTIG643 and ***CTIG684*** were located in the middle of the IGR, on the upper strand and in the same direction as the downstream genes. Both ended in a *rho* independent terminator. *CTIG643* showed two products on the Northern blot. Two probes overlapping the sequences flanking the terminator were prepared. Northern blot showed that the small product represented the sequence upstream of the terminator, and the larger product contained both sequences upstream and downstream of the terminator. *CTIG684* On the other hand, was represented by several small products around 50-70 nucleotides. RNA probes upstream and downstream of the terminator showed that all products represented sequences upstream of the *rho* independent terminator.

3.4 DISCUSSION

3.4.1 USE OF THE TILING MICROARRAY PLATFORM

The primary hurdle facing ncRNA discovery is the absence of a clear signature element(s) in the primary genome sequences. Preset assumptions of possible signature sequences have been implemented to find candidate ncRNAs. The majority of confirmed

ncRNAs were identified primarily by comparative sequence analysis to identify conserved elements in IGRs. Additional sequence signatures including intergenic promoter elements, sequences predicted to form stable RNA secondary structures, and intergenic *rho*-independent terminator elements have also been used for the screening of candidate ncRNAs (Argaman *et al*, 2001; Rivas and Eddy, 2001a; Rivas *et al*, 2001b; Livny *et al*, 2005).

In such a new field there still not enough data that conclusively mandate using the aforementioned signature elements as the main method of screening for ncRNAs. In addition, the less biased shotgun cloning approach has identified highly expressed ncRNA candidates in *E. coli* that were missed by comparative analyses (Vogel *et al*, 2003).

We have used a different unbiased strategy to identify ncRNAs in *C. trachomatis*. In this strategy we used a custom tiled Affymetrix microarray combined with an experimental protocol and analysis scheme optimized for ncRNA detection. We first scored the quality of each individual intergenic probe by hybridizing chromosomal DNA to the microarray. We then tested RNA samples, and analyzed each IGR for simultaneous high levels of florescence of 3 or more consecutive probes. Each potential IGR was further inspected for *rho* independent terminators, localized high G+C ratio, conservation among various sequenced chlamydial species, and finally analyzed by Northern blotting using strand specific RNA probes.

The *E. coli* high-density microarrays were used to identify ncRNAs. Both total RNA and RNAs isolated by co-immunoprecipitation with the RNA binding protein, Hfq were used for screening (Tjaden *et al*, 2002; Zhang *et al*, 2003; Hu *et al*, 2006). High-density microarrays, combined with comparative genome analysis, had led to the identification of mostly growth phase dependent ncRNAs in *E. coli* and *Staphylococcus aureus* and of ncRNAs whose expression is under the control of sporulation in *Bacillus subtilis* (Wassarman *et al*, 2001; Pichon and Felden, 2005; Silvaggi *et al*, 2006). Although DNA microarrays are a valuable tool for both identification and transcription profiling of ncRNAs it, like all other approaches, bears some caveats. One of the most challenging aspects of this technique is the preparation and labeling of RNA samples. The small size and structure of ncRNAs make these transcripts poor substrates for amplification and labeling. In addition, shorter ncRNAs (< 50 nucleotides) might be harder to detect, especially if the probes were arranged head to tail with no overlap. Therefore, short ncRNAs and ncRNAs that are highly structured and/or modified were likely to be missed. This phenomenon was observed in case of *CTIG153*. Its average fluorescence on the array was 884 (Table 3-2) while the Northern blotting showed high level of transcription (exposure time for developing the Northern blot was seconds). A novel detection method employing antibodies specific for RNA–DNA hybrids circumvented the problem of labeling and was shown to improve sensitivity (Zhang *et al*, 2003; Hu *et al*, 2006), and probably would be worth checking for future microarray experiments, this method was also recommended by S. Gottesman (Personal communication, 2005).

3.4.2 PROBE QUALITY AND RESULT INTERPRETATION

DNA microarray result uncovered considerable variations in probe sensitivity. This observation was confounded by the absence of mismatch probes. It was recommended by the Affymetrix design team to disregard inclusion of the MM probes. It was reasoned that since IGR probes were tested individually, the statistical package that calculates the average fluorescence for each gene would not be fully functional and there was no need to include MM probes. Moreover, elimination of MM probes increased the available features on the microarray and increased the allowed synthesis steps by 4 rounds. Comparing the MPAUT-1 with the high density microarray for *E. coli* (Tjaden *et al*, 2002), we noticed the difference in probe selection protocol. MPAUT-1 probes were arranged head to tail with no intentional overlaps. On the other hand, the high density microarray probes were selected every 4 nucleotides. Whether this denser probe selection scheme offered more validity for the results is unknown, but the high density version offered a worthwhile alternative. Tjaden *et al*, 2002 preformed a similar set of quality control tests for the IGR probes. Initial analysis of the data across all experiments (total of 26 microarrays) showed a range of hybridization affinities for different probes. 2671 probes were eliminated from the analysis for which there was evidence of significant cross-hybridization or other non-specific hybridization. Bad probes were also determined by hybridizing *E. coli* genomic DNA labeled directly with terminal transferase to the microarray and eliminating the probes that failed to meet expected values.

Zhang *et al*, 2003 developed a rating system for analyzing microarray data. IGRs were rated 5 if the average probe florescence was ≥10,000, an average fluorescence intensity of ≥5000 was rated 4, an average probe florescence of ≥1000 was rated 3, and finally an average probe florescence of ≥200 was rated 2. Candidate IGRs were analyzed by Northern blotting.

The field of genome sequencing and genome wide transcriptome sequencing is getting faster, more efficient, and more cost effective by time. Lately a new technology has emerged where whole transcriptomes were sequenced in a bead based emulsion PCR. The cDNA was ligated to adapters, clonally selected on the beads and amplified via emulsion PCR. Several samples could be pooled (up to 20 samples) using bar coded adapters and all the beads were deposited on slides and sequenced in parallel by repeated rounds of hybridization and ligation of fluorescently labeled di-base probes followed by cleavage and detection (The SOLiD™ 3 System, Applied Biosystems). This system had the advantage of generating actual sequences as readouts rather than fluorescence. Based on the number of times a sequence appeared the relative abundance of the message was calculated. The SOLiD technology, being non sequence dependent, also had the advantage of getting sequences of previously un-identified transcripts. This sequencing based detection offered a solution to a problem that was seldom addressed using other techniques. It offered an easier way to detect antisense ncRNA that are totally buried within annotated genes. Future experiments are being planned involves using this platform.

3.4.3 POTENTIAL ncRNAs IN C. trachomatis D

Comparative genomics of ncRNAs in bacterial genomes found that on average, ncRNAs are found in approximately 30% of IGRs of each genome sequence. Of these, 25.7% are conserved among three or more organisms. Approximately 60% of the conserved ncRNAs are not located in orthologous IGRs, implying that ncRNAs may be shuffled in genomes (Luban and Kihara, 2007). The potential ncRNAs in *Chlamydia* were located 5'and 3' to annotated genes. Since no functions have been assigned to any of the ncRNAs (except *ihtA*), there remains the possibility that they represent processed 5' and 3' RNAs. Some established ncRNAs in other bacteria were located in the vicinity of flanking genes. 6S RNA, a major regulator of *E. coli* σ^{70}, was produced as a part of a longer message (Trotochaud and Wassarman, 2004; Wassarman, 2007). Another example is RNAIII, a bifunctional molecule that encodes the δ-hemolysin protein in its 5′ end while it also acts as an ncRNA (Janzon *et al*, 1989). Examining the chlamydial chromosome revealed several indications that several functional genetic elements are packed within a limited space. For example, *Chlamydia* has 6 tRNAs (out of 37) that were located just 24 to 37 nucleotides upstream of genes on the same strand. In order for the transcripts of these genes to be independent, a distance of at least 45 nt was required between them and the closest tRNA (35 nucleotides promoter and shine Dalgarno sequence). In addition 8 pairs of tRNAs belonged to bicistronic operons. Moreover, the only known chromosomally encoded ncRNA in *Chlamydia* (*ihtA*) was located 16 nucleotides downstream of yscC and directly upstream of a tRNA. These observations indicated that *Chlamydia* has limited space to fit several genetic elements within a confined spac.

3.4.4 CTIG270 AND ftsI

In *E. coli, ftsI* was involved in the biosynthesis and degradation of the murein sacculus and peptidoglycan. It was specifically localized to the division septum. Invagination of the cytoplasmic membrane was accompanied by the synthesis of peptidoglycan at the leading edge of the invagination. This septal specific peptidoglycan biosynthesis occurred in two stages (Woldringh *et al*, 1987; Wientjes and Nanninga, 1989; Nanninga, 1991). An early stage that is penicillin insensitive and a later stage that is sensitive to penicillin and requires PBP3, the product of the *ftsI* gene (Botta and Park, 1981; Spratt, 1977).

Identification of *CTIG270* as a potential antisense regulator of *ftsI* added more questions to an already puzzling situation (1.2.6 above). *dicF* an ncRNA inhibited the translation of *ftsZ* (Faubladier and Bouche, 1994; Tetart and Bouch, 1992). It was established as the smallest ncRNA that had an assigned function (53 nucleotides). Probably *Chlamydia* which lacks *ftsZ* was using another ncRNA to control another key enzyme in peptidoglycan synthesis. To our knowledge there was no precedence for an antisense RNA that regulates a penicillin binding protein. Whether *CTIG270* represents a key molecule in regulating cell division remains to be determined. Another question was

the involvement of other PBPs encoded by the chlamydial genome in this regulatory circuit.

Sequence inspection revealed more information about this ncRNA. *CTIG270* had a 15 nucleotides complementarity to a sequence 200 nucleotides downstream of *ftsI* start codon. Similar finding was observed in *C. muridarum CTIG270*/*ftsI* orthologues. The result of binding of *CTIG270* to *ftsI* could either be degradation or stabilization of *ftsI*. In *E. coli gadY* ncRNA was shown to overlap the 3' end of *gadX* gene. This overlap region was found to be necessary for the *gadY*-dependent accumulation of *gadX* mRNA. It was suggested that during the stationary phase *gadY* base paired with the 3'-untranslated region of *gadX* mRNA and conferred increased stability. This allowed *gadX* mRNA to accumulate and increased the expression of downstream acid resistance genes (Opdyke *et al*, 2004; Tramonti *et al*, 2008).

3.4.5 CTIG360 AND bioY

Biotin represents an important metabolite for both eukaryotes and prokaryotes. Biotin is a water-soluble vitamin that participates as a cofactor in gluconeogenesis, fatty acid synthesis and branched chain amino acid catabolism. It functions as the carboxyl carrier for biotin-dependent carboxylases (Gravel and Narang, 2005). Bacteria get biotin from two main sources, *de novo* synthesis (Pai and Lichstein, 1965; Gloeckler *et al*, 1990; Bower *et al*, 1996; Flint *et al*, 1997) and transporter based uptake (Pai, 1973; Guillen-Navarro *et al*, 2005; Hebbeln *et al*, 2007). In these systems the amount of biotin in the cell was controlled by transcription regulation. A bi-functional DNA binding protein BirA was responsible for this function (Streaker and Beckett, 1999). BirA was both a negative regulator of the biotin biosynthetic operon and also served as a biotin-activating enzyme (Buoncristiani *et al*, 1986; Rodionov *et al*, 2002). Separate domains were found to confer each function (Brennan *et al*, 1989). The *C. trachomatis* genome was missing key biotin biosynthesis pathway genes. Moreover, the DNA binding domain of the chlamydial *birA* gene was also missing. *C. trachomatis* expressed the *bioY* gene as the only known uptake mechanism to obtain biotin. The presence of a regulated 5'UTR might be a plausible alternative for the absence of a transcriptional regulator of biotin uptake. Further experiments will shed more light on this potential riboswitch.

CHAPTER 4. CHARACTERIZATION OF ncRNAs IN *C. trachomatis* D

4.1 MAPPING CHLAMYDIAL ncRNA

A subset of ncRNAs (10) that were detected during the discovery and validation phase of the study were chosen for further analysis. They were compared to the previously described *CTIG675* (*ihtA*) and the pCHL1.2 plasmid antisense transcripts. The selected ncRNAs included *cis*-acting ncRNAs (*CTIG153, 270*, and pCHL1.2), potential *trans*-acting ncRNAs (*CTIG241, 327, 356, 498, 504, 643,* and *684*), and a potential riboswitch-like ncRNA (*CTIG360*).

We determined the 5' and 3' ends of each molecule using an RNA circularization assay. The RNA circularization protocol was previously developed to assess the poly-A status of RNAs (Urban and Vogel, 2008; Vogel and Hess, 2001). A schematic representation of the procedure is shown in Figure 4-1.

Briefly, total RNA of a 24 h PI sample was treated with a tobacco acid pyrophosphatase (to convert primary triphosphate to monophosphate groups), and subjected to end circularization by T4 RNA ligase. Guided by the intergenic microarray results we designed primers for cDNA synthesis of circularized RNA (Table 2-1 for list of primers used, locate circularization primers under 5'3' SEQ). After cDNA synthesis, junction fragments were amplified by PCR (Figure 4-1), cloned, and sequenced (10 or more clones per reaction). This analysis revealed the transcription start sites (TSS in Table 4-1) for all 12 ncRNAs, including those with more than one species (*CTIG643.1/2* and *pCHL1.1/2*, each pair had the same 5' end but different 3' ends). The smaller fragment of *CTIG153* could not be determined due to the constraints on primer design presented by the small size of transcript and extensive secondary structures. The TSSs determined by the RNA circularization procedure were used to predict potential promoters for the ncRNAs (Table 4-1). The promoters identified were all of the σ^{66} family (the major sigma factor in *C. trachomatis*) but functional characterization of the promoters was not performed and these remain predicted promoter elements at this point.

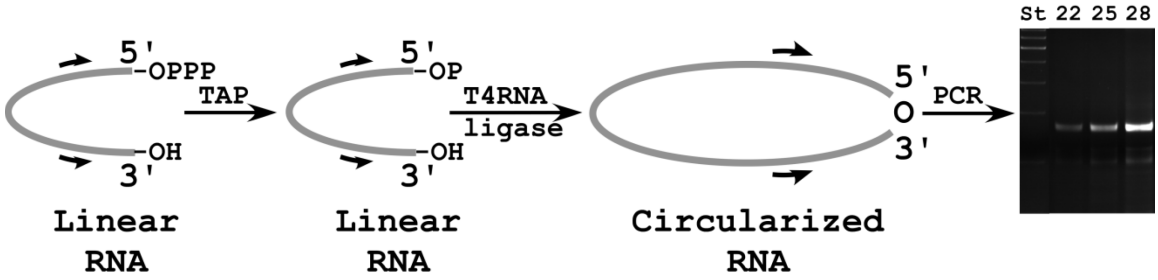

Figure 4-1 A schematic representation of the RNA circularization procedure. The procedure begins with the removal of the 5' pyrophosphate using tobacco acid pyrophosphatase (TAP) followed by circularization using T4 RNA ligase. Primers were then designed to amplify the 5'/3' junction.

Table 4-1 The 5' and 3' ends of the ncRNAs determined in this study.

Name	Promoter elements (P) -35	-10	Arm	TSS	3'end	Size
CTIG153	TTTACA	TGTGATAGC	N_7	TGTAGA	TTCAGC*	159
CTIG241	TTTTCA	TATATT	N_6	GGTTTA	TGCTTT*	80
CTIG270	TTTACT	TCTTAT	N_9	AAATGT	TCTAAT*	222
CTIG327	TTGACA	GATTAT	N_8	TTTAAG	TTTTAG*	308
CTIG356	TTTTTG	TATAAC	N_{12}	AAAGTG	GTAGTG	209
CTIG360	ATGCAT	TAGACT	N_6	AGTTTT	TCATAA	241
CTIG498	TTGTGG	TCCATT	N_3	TCTGTA	GTAACT	111
CTIG504	TTGCTT	TACAAT	N_6	TAGCTT	GTCGTT*	73
CTIG643.1	TTGACA	TATCAT	N_5	CCCGGG	CGCAGT*	126
CTIG643.2					GCACCT*	331
CTIG675	TGGAAA	TGTTATAAG	N_6	AAGTTG	TGGCTT*	106
CTIG684	TTGTAG	TACGAT	N_6	CCAACA	CTGTCT*	83
pCHL1.1	TTGCCA	TATATT	N_5	CATCTT	CCCCAC*	223
pCHL1.2					GCAAAC*	507

The 5' end is designated the TSS, the 3' end and overall size of the ncRNAs is listed. Non-coding RNAs that contained non-templated additions at the 3' end are indicated by an asterisk. Promoter predictions were made by examination of the areas immediately upstream of the TSS. All of the predicted promoters were of the σ^{66} type and two had an extended -10 sequence.

Two of the ncRNA promoters have sequences characterized as encoding "extended" -10 regions. All promoters had acceptable -35 sequences that ranged from 6 out of 6 matches to the -35 consensus [TTGACA] (*CTIG327*, and *CTIG643*), to 3 out of 6 match to consensus (*CTIG360, CTIG498, CTIG504,* and *CTIG684*). Moreover all potential promoters had acceptable -10 elements except for 2 cases, they ranged from 5 out of 6 ncRNAs had poor-10 elements, the match was only 2 out of 6 (*CTIG327* and *CTIG360*). match to the -10 consensus [TATAAT] (*CTIG365, CTIG504, CTIG643*, *CTIG675,* and *CHL1*), to 3 out of 6 match to consensus (*CTIG153,* and *CTIG498*). Collectively the combination of -10 and -35 promoter elements showed acceptable promoter strength. Two promoters were questionable as to whether they represent true promoters. These were *CTIG360* and *CTIG498;* both had either 3 or 2 matches to the consensus in both the -10 and -35 elements. These questionable promoter elements suggested that either these two ncRNAs were not the primary transcript or that they represent an unusual chlamydial promoter that might require chlamydial RNAP holoenzyme and/or transcription factors to be recognized.

The spacer between the -10 element and the TSS is ideally 5 to 9 nucleotides (Figure 1-2). 10 ncRNAs fall within this range and two were out of that range. For *CTIG356,* the -10 element was located 12 nucleotides upstream of the TSS. The promoter elements for *CTIG356* were strong (5 out of 6 matches to the consensus in both its -10 and -35 elements). The 12 nucleotide distance between the TSS and the -10 element can be interpreted as a post transcriptional processing from a longer RNA and the native TSS might be 3 to 7 nucleotides upstream. *CTIG498* had a short spacer arm of 3 nucleotides between the -10 element and the TSS. Combining this observation with its two moderate to weak promoter elements (3 out of 6 match to the consensus), and that the transcript is located at the 3'end of *CT497* suggests that *CTIG498* might not be a primary transcript. *CTIG498* might be a processed transcript from the *CT497* mRNA.

Two promoters had extended -10 elements, namely *CTIG153* and *CTIG675* (*ihtA*). The predicted promoter for *ihtA* differed from the previously published promoter (Grieshaber *et al*, 2006).

4.1.1 VERIFICATION OF TSS BY 5'RACE

We selected 7 ncRNA to verify their TSS using 5'RACE; namely *CTIG153, CTIG241, CTIG270, CTIG327, CTIG360, CTIG675,* and *CTIG684*. Primers used for these studies are listed in Table 2-1 under 5'RACE. RACE can be carried out in several ways. All depend on the addition of a known sequence to the 5' end of RNA (Adaptor ligation) or to the 3'end of the generated cDNA (template switching reverse transcriptase). We used template switching reverse transcriptase. It was a special RT enzyme in that, upon reaching the 5'end of the mRNA, it acquired a terminal transferase activity and added three Cs to the end of the newly formed cDNA. The newly added Cs formed a sticky end that allowed the end to base-pair with 3 Gs. The sticky end base paired with an adapter in the reaction mixture that had 3 Gs. The RT switched templates and resumed its RT activity using the adapter as its new template. The result was a

cDNA that had the complementary adapter sequence on its 3' end (5' end of the RNA). cDNA was amplified by PCR (gene specific primer and Universal adapter primer). The template switching technique was especially beneficial as it solved one of the problems associated with TSS analysis. RNA tended to acquire different secondary structures that required high temperatures to denature. The elevated temperatures were inhibitory to the RT enzyme. In the template switching technique the presence of secondary structure led to stalling of the RT, hindering the template switching and inhibiting amplification. When the RT enzyme reached a true 5' end of an RNA (be it a TSS or processed product) it acquired a terminal transferase activity and template switching took place. This technique selectively enriched for the full length cDNA rather than truncated products due to secondary structures. 24h PI RNA samples were DNased twice (TurboDNase, Ambion). cDNA was synthesized according to manufacturers instructions (Super SMART™ PCR cDNA synthesis, Clontech) using gene specific primers. The 3' end of cDNA was amplified, gel purified, cloned, and sequenced. Results confirmed the same TSS determined by RNA circularization (Table 4-1).

4.1.2 IhtA PROMOTER ANALYSIS

Grieshaber *et al*, 2006 determined the TSS for *ihtA* using a primer extension procedure. They determined the TSS to be 8 nucleotides downstream of the *CT674* stop codon. Surprisingly, when the TSS was determined using the RNA circularization protoccol it was determined to be 16 nucleotides downstream of the *CT674* stop codon (Figure 4-2B). Based on the differences in TSS, two different promoter elements were predicted. To address this discrepancy several experiments were carried out. First we determined the TSS using 5' RACE and the results confirmed the TSS of *ihtA* in Table 4-1 determined by RNA circularization. 2 clones for *ihtA* were constructed. *ihtA* was cloned under its own promoter (E/C in Figure 4-2A) in a promoterless vector pSMART-HCAmp (CloneSmart® Blunt Cloning, Lucigen). *ihtA* was also cloned under a tetracycline promoter such that its TSS would match the published TSS (E/T in Figure 4-2A) (for primers used see Table 2-1). Both clones were transformed into *E. coli* and allowed to grow to OD_{600} of 1. RNA was harvested (MASTERPURE, Epicentre) from both cultures and resolved on 10% UREA PAGE gels along with 40h PI chlamydial RNA. RNA was then electrotransferred onto nylon membranes. The blot was probed with an *ihtA* specific probe. Analyses indicated that both chlamydial RNA and *E. coli* RNA (E/C) contained the same size band of about 100 bases. The clone E/T showed a band migrating at a slightly larger size. The Northern blot showed that *ihtA* could be transcribed from its own chlamydial promoter in *E. coli*. The blot suggests that *ihtA* with the previously published TSS produced a larger product than native *ihtA*. In order to rule out the size difference between C, E/C, and E/T were due to differences in the 3' ends of the RNAs we carried out RNA circularization for both clones E/C and E/T. The results verified that the 5' end of *ihtA* in E/C was identical to the chlamydial *ihtA*. The results also verified that *ihtA* 5' end in E/T clone was identical to the published TSS. Finally the results confirmed that the 3' ends of all clones were identical, and the only difference between E/C and E/T was the difference in the predicted TSS.

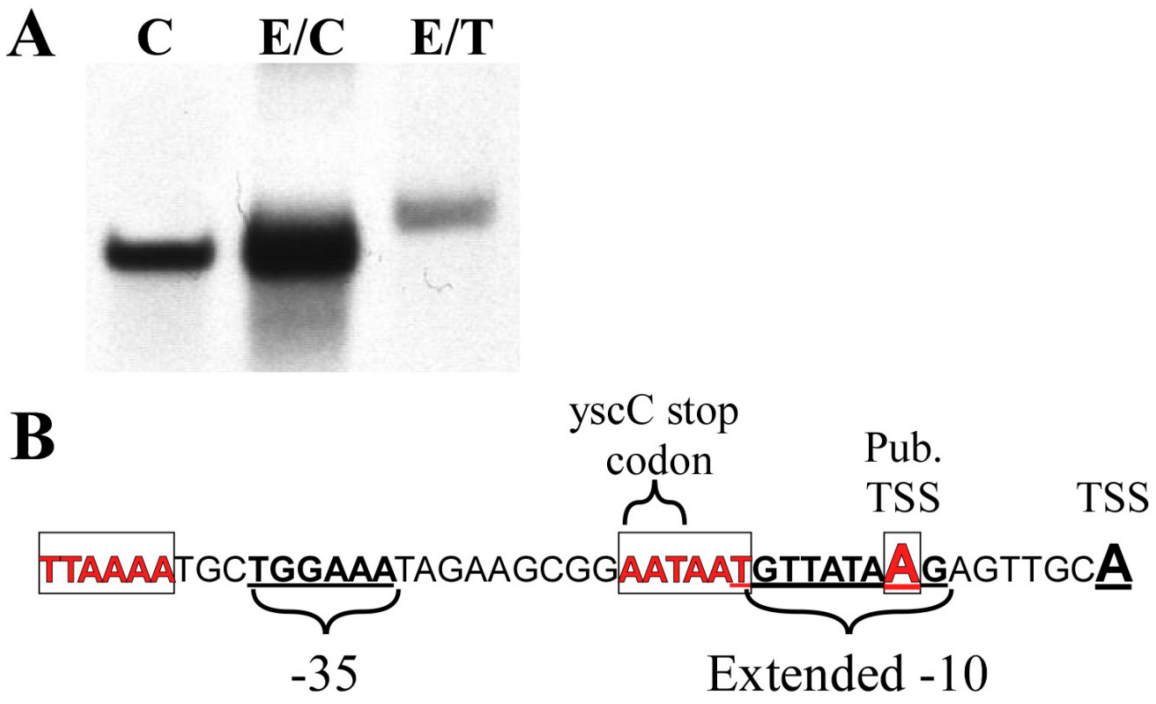

Figure 4-2 Resolving the *ihtA* promoter discrepancy. Panel A shows Northern blotting results. Lane C represents chlamydial RNA, lane E/C represents RNA from *E. coli* expressing *ihtA* from its native promoter, and lane E/T represents RNA from *E. coli* expressing *ihtA* from a tetracycline promoter such that the TSS matches the published TSS. Panel B represent a schematic of both promoter elements in which the red letters represent published promoter elements and the underlined letters represent the promoter as determined in this work. (Pub: published, TSS: transcription start site).

4.1.3 NON-TEMPLATED ADDITIONS TO ncRNAs

During mapping of the chlamydial ncRNAs, the 3' end of each transcript was determined. The 3' ends marked with an asterisk in Table 4-1 were determined to have non-templated additions of variable sizes. Table 4-2 shows a summary of the non-templated additions to ncRNAs' 3'ends. For simplification, the table only includes individual sequences that were composed of at least 4 nucleotides detected at the 3' end of RNAs. Generally, specific transcripts had between 3-5 added nucleotides (nt) but the largest found was 27 nt. Although the additions were predominantly poly-A, several non-templated additions contained G residues. Upon calculating the actual percentage of each nucleotide in the non-templated additions we found that A represented 74%, followed by G (16%), then T (8.2%) and finally C (1.7%).

We asked whether similar additions are present at the 3' ends of mRNAs encoding ORFs. We preformed RNA circularization assay for 5 chlamydial genes (*i.e. CT500, CT789, CT353, CT046,* and *CT099*). We were able to identify the TSS for each mRNA but none of these mRNA had non-templated additions at their 3' ends.

4.2 DEVELOPMENTAL EXPRESSION OF ncRNA IN *CHLAMYDIA*

To determine the expression patterns of ncRNAs, we chose to examine i) their expression during the normal developmental cycle, ii) their expression during IFN-γ-induced persistence and reactivation, and iii) their expression during carbenicillin-induced persistence and reactivation. DNA/RNA samples for the developmental cycle were taken at 0, 3, 8, 16, 24 and 48 h PI as described previously (Belland et al, 2003b) and shown in Figure 4-3. DNA/RNA was collected from IFN-γ-induced persistent samples at 24, 48, 72 h PI and 12, 24, 48, 72 h post reactivation (IFN and IFN R in Figure 4-3). IFN-γ was used to induce persistence (5ng/ml, as determined by a pilot titration experiment) and was added to the media 24 h prior to infection. Samples were reactivated at 24 h PI by changing media with IFN-γ-free media containing excess tryptophan (Belland et al, 2003a) and samples were taken at the times shown in Figure 4-3 (IFN and IFN R). Carbenicillin-induced persistence was induced by the addition of the antibiotic (2□g/ml, as determined by a pilot titration experiment) at the time of infection. DNA/RNA was collected from carbenicillin-treated samples at 24, 48 h PI and 24, 48, and 72 h post reactivation (Cb and Cb R in Figure 4-3). As shown in Figure 4-4, IFN-□ treatment resulted in a reduction of IFUs of ca. four logs at 48 h PI. Following reactivation for 48 h, the IFU values returned to ca. 98% of control cultures, indicating that RBs were present but were in a non-dividing or persistent state. Similarly, carbenicillin treatment resulted in a distinct decrease in IFUs at 48 h PI, but following removal of the antibiotic (Cb reactivation), IFU numbers returned to control levels (Figure 4-4). Genome copy numbers for each of the samples was determined by qPCR as previously described (Ouellette et al, 2006) and the samples were thoroughly DNased and RNA was then re-purified. Expression of ncRNAs was determined by Northern blotting of RNA samples from equivalent numbers of bacteria (as determined by genomic normalization) for the developmental cycle and carbenicillin persistence. Samples used

Table 4-2 Non-templated additions to chlamydial ncRNAs 3' ends.

Name	Sequence
CTIG153	AGAGAAA
	AAAAGTA
	AAAGTG
	AAGA
	AAAAAAA
	AAAAAAA
	AAAAA
	AAAAA
	AAAA
	AAAA
CTIG241	AAGAAATAAGAGAAGAG
	AGAAAATGAGAAAG
	AGGGAGAAGAAAG
	AAAAAAGGAAG
	AAAAA
CTIG270	AAAAAAA
	AAAAAA
CTIG327	GGGAGTGGAAG
	ATTTT
	TTTT
CTIG643-2	AAAAAAGC
CTIG675	GTAAAAAAAAAAACAACAAGAGACAAG
	AAATAAAAG
	GAAAGAGA
	AAAAAAA
	AAAAAAA
	AAAAAAA
	AAAAAAA
	AAAA
	AAAA
pCHL1.1	AAGGAGCAAAAAAAA

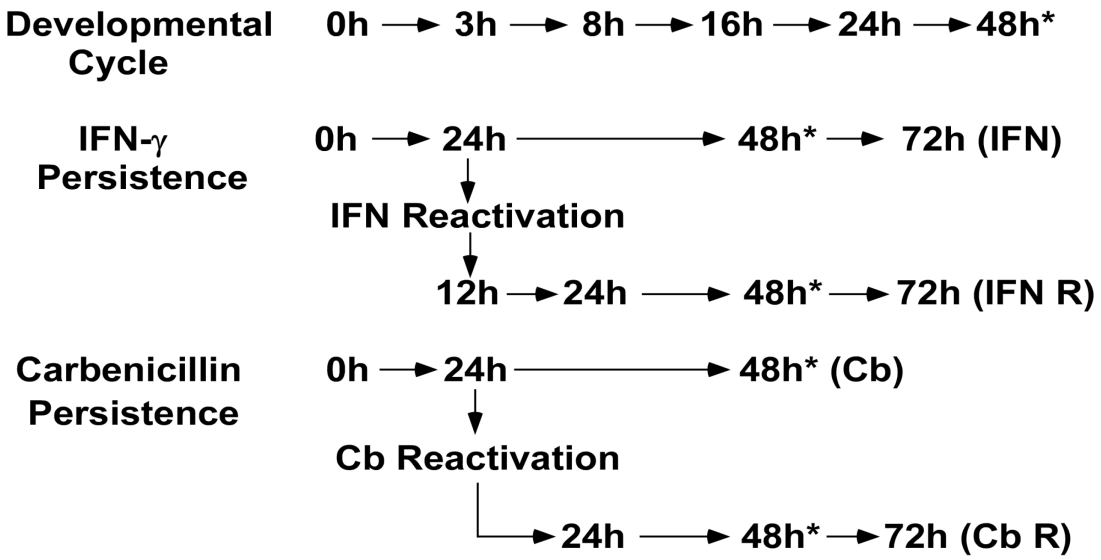

Figure 4-3 Scheme of nucleic acid sample collection during chlamydial developmental cycle, IFN-γ-induced and carbenicillin-induced persistence and reactivation. IFN: IFN-γ Persistence, Cb: Carbenicillin Persistence, R: Reactivation.

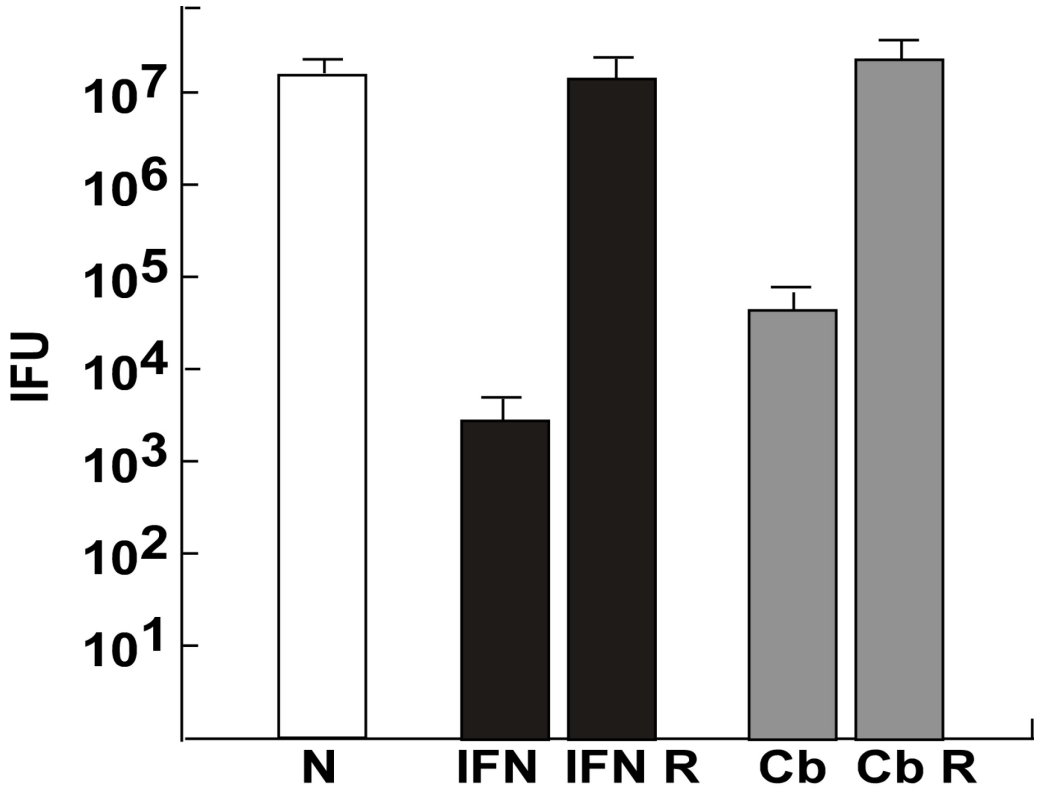

Figure 4-4 Titration results and establishment of persistence and reactivation. IFN: IFN-γ persistence, Cb: carbenicillin persistence, R: reactivation, IFU infectious units.

for IFN-γ persistence contained 20% of the genome numbers used for the other analyses due to the reduced chlamydial biomass associated with growth in the presence of IFN-γ. We chose to group results by ncRNAs rather than by condition, where all blots for each ncRNA are grouped in a single panel to allow for direct comparisons between normal and persistent growth conditions. The results in Figure 4-5 show ncRNA expression patterns for genes that predominantly express a single transcript while those in Figure 4-6 show the results for ncRNAs that are expressed as multiple species (both figures also show the genomic-locus of each ncRNA).

4.2.1 DEVELOPMENTAL CYCLE

Expression during the normal developmental cycle showed differences between the ncRNAs examined in the onset and point of maximal expression of different transcripts, indicating that ncRNA transcription is both tightly and temporally regulated. Four ncRNAs showed onset of expression at ca. 8 h PI *i.e. CTIG153.1/.2* (Figure 4-6A), *CTIG356* (Figure 4-5D), *CTIG360* (Figure 4-5E), and the plasmid antisense transcripts *pCHL1.1/.2*. Seven ncRNAs were first expressed at ca. 16 h PI *i.e. CTIG241* (Figure 4-5A), *CTIG327* (Figure 4-5C), *CTIG498* (Figure 4-5F), *CTIG504* (Figure 4-5G), *CTIG643.1/.2* (Figure 4-6B), *CTIG675/lhtA* (Figure 4-5H), and *CTIG684* (Figure 4-6C). The *cis*-acting *CTIG270,* that overlaps the *ftsI* gene, was the only ncRNA that was expressed late in the cycle at ca. 24 h PI.

4.2.2 IFN-γ-INDUCED PERSISTENCE

Six ncRNAs (*CTIG270, CTIG327, CTIG360, CTIG498, CTIG643.1/.2),* and *CTIG684*) were not detectable during IFN-γ persistence. *CTIG360* (Figure 4-5E) resumed expression as early as 12 h after reactivation and its expression peaked at 24h post-reactivation. *CTIG270* (Figure 4-5B), *CTIG327* (Figure 4-5C) and *CTIG684* (Figure 4-6C) were expressed at a much reduced intensity during the first 24 h post-reactivation and were undetectable at later time points while *CTIG498* (Figure 4-5F) and *CTIG643* (Figure 4-6B) did not appear to be expressed during reactivation.

Six ncRNAs (*CTIG153.1/.2*, *CTIG241, CTIG356, CTIG504, CTIG675,* and *pCHL1.1& (.2)*) were expressed to some degree during IFN-γ persistence. *CTIG153*.2 short-form (Figure 4-6A) was expressed at a high level during IFN-γ persistence, while expression of the longer form (*CTIG153.1*) declined rapidly. During reactivation, both forms declined at approximately the same rate and by 48 h PR they were no longer detectable. *CTIG241* (Figure 4-5A), *CTIG356* (Figure 4-5D), *CTIG675* (Figure 4-5H), and the plasmid antisense pCHL.1/.2 (Figure 4-6D) were expressed at low levels during IFN-γ persistence but showed increased expression levels following reactivation. *CTIG504* (Figure 4-5G) was expressed at ca. the same level throughout IFN-γ-induced persistence and reactivation.

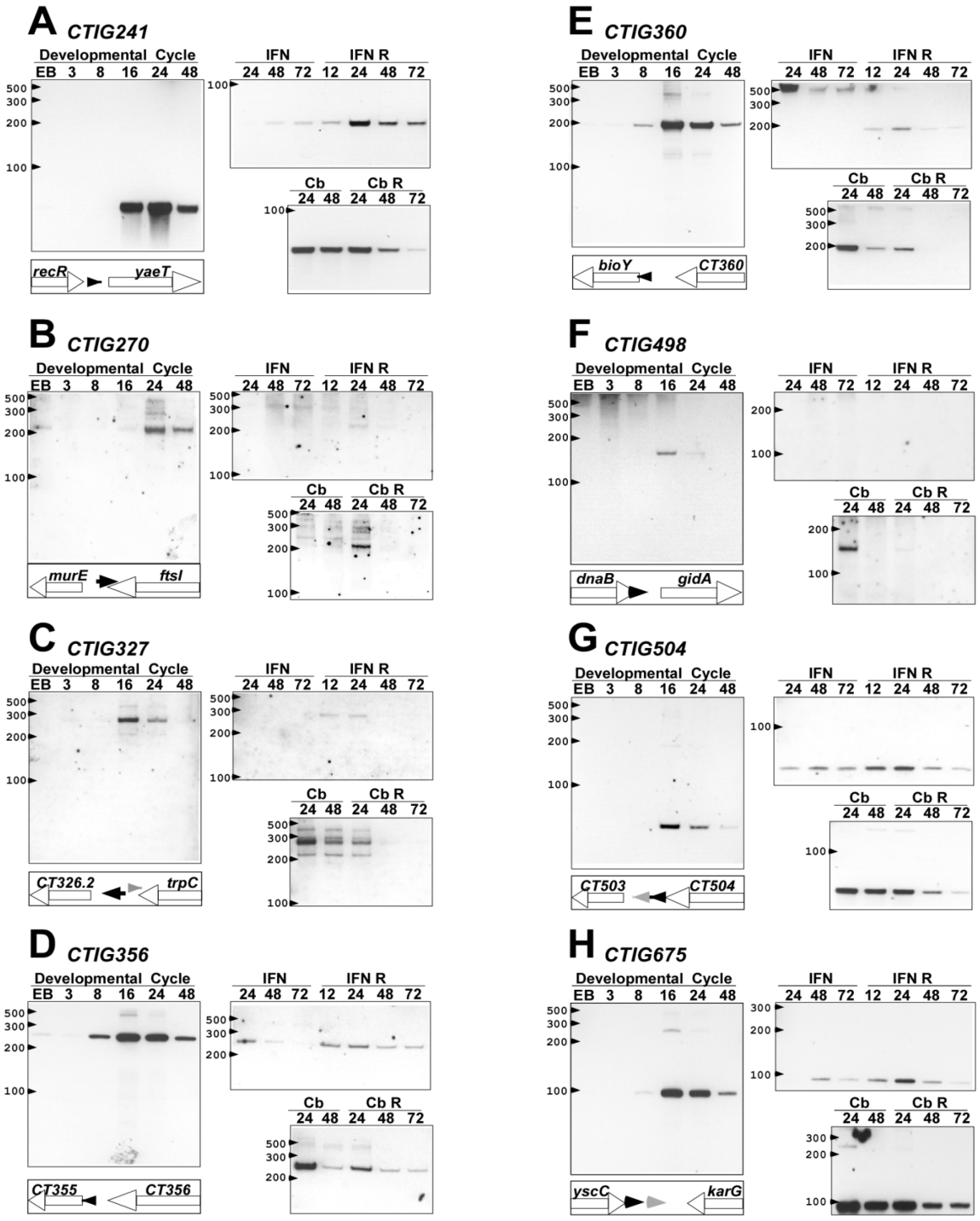

Figure 4-5 Developmental expression of ncRNAs in *C. trachomatis* during the developmental cycle, IFN-γ-induced and carbenicillin-induced persistence and reactivation. Each panel (A-H) shows the expression pattern of the particular ncRNA under the conditions described in Figure 4-3. In addition the genomic position of ncRNAs is shown in the schematic panel; black arrowheads represent ncRNAs, while gray arrowheads represent tRNA.

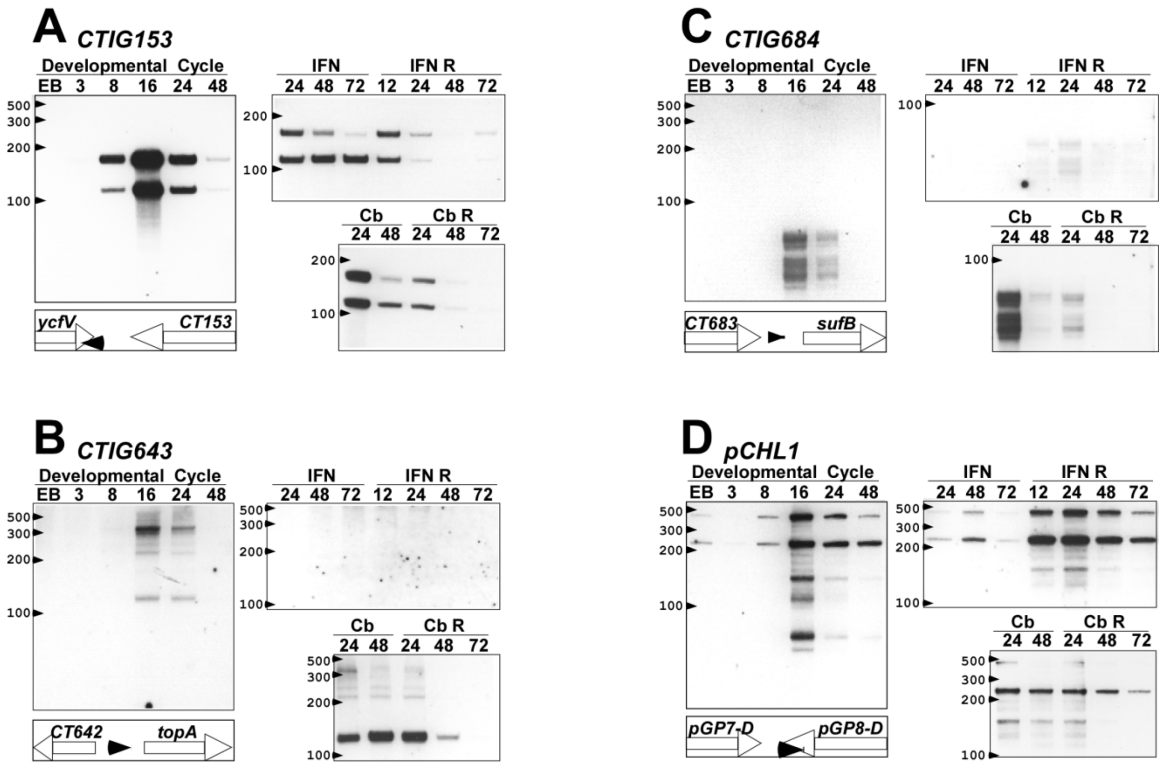

Figure 4-6 Developmental expression of the processed forms of ncRNAs in *C. trachomatis* during the developmental cycle, IFN-γ-induced and carbenicillin-induced persistence and reactivation. Each panel (A-D) shows the expression pattern of the particular ncRNA under the conditions described in Figure 4-3. In addition the genomic position and orientation of the ncRNAs are shown in the schematic panel under the developmental cycle panel.

4.2.3 CARBENICILLIN-INDUCED PERSISTENCE

In contrast to the situation with IFN-γ-induced persistence, virtually all tested ncRNAs showed significant expression during carbenicillin treatment. The exception was *CTIG270* (Figure 4-5B), which was undetectable during carbenicillin treatment and was transiently expressed 24 h after reactivation.

Nine ncRNAs (*CTIG153.1/.2), CTIG327, CTIG356, CTIG360, CTIG498, CTIG643.1, CTIG675, CTIG684, and pCHL1.1*) showed a decline in expression as the infection progressed from 24 h PI to 48 h PI. This decline in expression ranged from a slight reduction to complete disappearance of the transcripts, as evident with *CTIG498* (Figure 4-5F)*, CTIG684* (Figure 4-6C), and *pCHL1.1* (the longer plasmid antisense transcript, Figure 4-6D). In contrast, the only ncRNA to show a marked increase in transcript levels during carbenicillin treatment was *CTIG643.2* (the longer transcript, Figure 4-6B). Two ncRNAs showed little change in transcript levels during persistence (*CTIG241*; Figure 4-5A, and *pCHL1.2;* Figure 4-6D). Transcript levels for virtually all of the tested ncRNA were detectable 24h post-reactivation. Many of these declined to the point at which they could not be detected at 48h post-reactivation (*CTIG153.1& (.2), CTIG270, CTIG327, CTIG360, CTIG684*, and *pCHL1.1*) while the rest declined to some degree as the reactivation progressed from 24 to 72 hours. *CTIG498* (Figure 4-5F) was the exception at it was not detected during the course of reactivation from carbenicillin treatment.

4.3 DISCUSSION

4.3.1 RNA CIRCULARIZATION TECHNIQUE

We used an RNA circularization technique for mapping ncRNAs. The technique offered a reliable and cost effective method that determines both the 5' and 3' ends of RNA. Moreover, this technique was successful at identifying 3' non templated additions. RNA circularization was used to identify the ends of 11 mRNAs and rRNAs of the mitochondrial genome of *Drosophila melanogaster* (Stewart and Beckenbach, 2009). It was also applied in a protocol termed "ligation-mediated PCR" allowing the detection and mapping of cleavage products of specific nucleic acid molecules out of complex nucleic acid mixtures (Grange, 2008). Moreover, the RNA circularization technique was also used to monitor the editing processes that result in the structural retailoring of the aminoacyl acceptor stems of mitochondrial tRNAs (Lohan and Gray, 2007). From these examples and others, it is clear that RNA circularization is a reliable technique for mapping RNA ends.

4.3.2 POTENTIAL PROMOTER ELEMENTS FOR ncRNAs

In *E. coli* the consensus arrangement for the major sigma factor is shown in Figure 1-2 (reviewed in Ross and Gourse, 2009). Although there appear to be considerable levels of homology between *C. trachomatis* σ^{66} and *E. coli* σ^{70}, Mathews and Sriprakash reported that chlamydial RNA polymerase can tolerate considerably more variation at the -10 and -35 regions (Mathews and Sriprakash, 1994). In general, most published σ^{66} promoter elements have acceptable variation from the consensus (Engel and Ganem, 1987; Birkelund *et al*, 1989; Douglas and Hatch, 1995, 1996; Ochiai *et al*, 1999; Shen *et al*, 2000; Schaumburg and Tan, 2003; Grech *et al*, 2007). Therefore; except for *CTIG360* and *CTIG498*, the predicted promoters are σ^{66} promoters.

Two of the predicted promoters had extended -10 characteristics. Extended –10 promoter elements are present in ~20% of all *E. coli* genes (Burr *et al*, 2000). The primary σ factor (σ^{70}) has four regions of similarity. It is known that residues in region 2 recognize a −10 element (TATAAT) (Murakami *et al*, 2002), residues in region 3 recognize an extended TGn −10 element (positions −15 to −13) (Barne *et al*, 1997), and residues in region 4 recognize a −35 element (TTGACA) (Campbell *et al*, 2002). However, not all three of these promoter elements need to be present for promoter function. *E. coli* σ^{70}-dependent promoters have typically been characterized as either −10/−35 promoters, which have good matches to both the canonical −10 and −35 sequences and do not require the TGn motif (McClure *et al*, 1983), or as extended -10 promoters (TGn/-10 promoters) which have the TGn motif and an excellent match to the -10 consensus sequence and do not require a -35 element (Kumar *et al*, 1993; Lisser and Margalit, 1993; Browning and Busby, 2004). In *E. coli*, extended -10 promoters have several characteristics, Extended -10 promoters, tend to have longer spacer (distance between -10and -35 elements) lengths than promoters that do not. They also tend to show fewer matches to the canonical -35 elements (Mitchell *et al*, 2003). More importantly open promoter complexes can be formed on an extended -10 promoter at temperatures as low as 6°C a temperature at which complexes on most promoters are closed (Minakhin and Severinov, 2003).

The TSSs in Table 4-1 predict 2 potential extended -10 promoters, *CTIG153* and *ihtA*. Similar to σ^{70} in *E. coli,* chlamydial σ^{66} needs only regions 2 and 3 to interact with extended -10 promoters. A large and diverse family of proteins known as "anti-σ factors" regulates utilization of particular classes of bacterial promoters by targeting specific σ factors (Helmann, 1999; Hughes and Mathee, 1998). Typically, anti-σ factors interact with core binding determinants in their cognate σ factors, thereby preventing their association with the RNAP core enzyme (Campbell *et al*, 2008). The AsiA protein of bacteriophage T4 is one of the earliest anti-σ factors to be identified. It targets σ^{70} (Orsini *et al*, 1993; Severinova *et al*, 1996), however, unlike most other well-characterized anti-σ factors, AsiA binds to σ^{70} primarily through an interaction with σ^{70} conserved region 4.2 which has also been implicated in sequence-specific recognition of the -35 consensus promoter element (Severinova *et al*, 1998; Yuan *et al*, 2009). As a component of the σ^{70}-containing holoenzyme, AsiA inhibits transcription from the -10/-35 class of promoters, but does not inhibit transcription from extended-10 promoters (Severinova *et al*, 1998).

Interestingly, *CT663* interacts with the flap domain of the β subunit (beta-flap) of RNA polymerase core enzyme, and the conserved region 4 of the primary sigma subunit σ^{66} in *C. trachomatis*. *CT663* inhibits σ^{66}-dependent transcription *in vitro* (Rao *et al*, 2009). Thus there is a possibility that *CT663* may have a role in regulating transcription of TGn and nonTGn promoters.

4.3.3 NON-TEMPLATED ADDITIONS ON THE 3'ENDS OF ncRNA

It has been shown that prokaryotes posses two enzymes that can add nucleotides to the 3' ends of RNA without a template, namely Poly(A) polymerase I (PAPI) (Paschal *et al*, 2008) and Polynucleotide phosphorylase (PNPase) (Grunberg-Manago *et al*, 1955). Both enzymes are involved in modulating RNA stability (Carpousis *et al*, 1999); (Carpousis *et al*, 2008). PAPI, adds poly(A) extensions to the 3′ ends of mRNAs, as well as to tRNA and rRNA (Li *et al*, 1998). PNPase on the other hand, adds long, heteropolymeric tails *in vivo* (Mohanty and Kushner, 2000). Polyadenylation of RNAs by PAPI in *E. coli* plays a significant role in mRNA decay and general RNA quality control (Jasiecki and Wegrzyn, 2003; Mohanty and Kushner, 2006). Chlamydial genomes contain 2 paralogues of PAPI (*CT410* and *CT704*) and one PNPase (*CT842*). We believe that the observed non-templated tails are due to the activities of these enzymes. The tails composed of only As are probably due to the activity of PAPI, while the heteropolymeric tails are probably due to the activity of PNPase. In either case the modified RNA represent an intermediate for degradation. The heteropolymeric tails in *Chlamydia* were shown to contain the different nucleotides according to the following preference A > G > U > C. The same preference was also observed for *E. coli* PNPase (Mohanty *et al*, 2000).

Regulation of ncRNAs by polyadenylation was well characterized in the *glmS/glmZ/glmY* system (Kalamorz *et al*, 2007; Reichenbach *et al*, 2008). Translation of the *glmS* mRNA was normally weak since an internal hairpin structure containing the *glmS* Shine Dalgarno sequence limits ribosome access to the downstream *glmS* RBS and therefore GlmS synthesis (Kalamorz *et al*, 2007; Urban *et al*, 2008). *glmZ*, an ncRNA, functions as an anti-antisense and melts this hairpin, allowing *glmS* translation. *glmZ* is unstable and its stability is greatly increases by *glmY*, another ncRNA. Thus the levels of GlmS are affected by *glmZ*, which in turn is affected by *glmY* (Kalamorz *et al*, 2007). *glmY* levels are controlled by polyadenylation and it was found that about 50% of *glmY* is polyadenylated. This regulatory hierarchy signifies the effect of polyadenylation on the function of *glmS/glmZ/glmY* (Urban *et al*, 2008) reviewed in (Gorke and Vogel, 2008).

4.3.4 ncRNA EXPRESSION UNDER BOTH NORMAL AND STRESS CONDITIONS

The ncRNAs shown in this study may represent a subset of the total *C. trachomatis* ncRNAs since initial identifications were done using RNA from late stage infection (40h PI). For example, ncRNAs that were only expressed during the early

stages of the developmental cycle were not tested. The technical difficulties associated with obtaining sufficient RNA from early stages of the developmental cycle have hindered this approach. Most of the ncRNAs showed a decrease in expression during the late stages of the developmental cycle, a point in which the predominant developmental forms are the infectious, non-metabolizing EBs.

As is the case with many ncRNAs, their expression is tightly controlled. Their might be a significant portion of chlamydial ncRNAs that are not induced under normal developmental conditions but could be induced under various stress conditions. Thus it may be worthwhile to catalogue stress-induced ncRNAs, after establishing the appropriate system to screen their functions.

The predicted target of *CTIG270* is the *ftsI* transcript encoding FtsI. FtsI in other bacteria is associated with peptidoglycan modification during cell division. *CTIG270* was the only ncRNA to show late onset of transcription (24h PI). At this time chlamydial inclusions are easily visible under the microscope and some EBs can be seen inside inclusions. This timing of expression reinforces the hypothesis that this RNA controls *ftsI* mRNA, as one of the hallmarks of RB to EB differentiation is the inhibition of cell division. For further discussion concerning *CTIG270* and *ftsI*, please refer to section 5.2.1 below.

One of the interesting ncRNAs was *CTIG643*.1&.2 (long and short transcript respectively) as there is a significant difference in expression between the normal developmental cycle and carbenicillin treatment. The sequence of *CTIG643.2* shows a *rho* independent terminator marking its 3' end. During the developmental cycle the *rho* independent terminator was not functional and the major transcript detected was the longer transcript (*CTIG643.1*). During carbenicillin treatment, RNA polymerase terminated efficiently at the *rho* independent terminator producing primarily *CTIG643.2*. This interesting phenomenon can be interpreted in two ways. Either the RNA polymerase enzyme became increasingly sensitive to termination during carbenicillin treatment or there is a factor that bound the RNAP and prevented its termination during the normal developmental cycle. This factor might be unavailable during persistence. The *C. trachomatis* genome encodes three factors involved in termination/anti-termination; *NusA* (*CT097*), *NusG* (*CT320*), and *GreA* (*CT636*). *NusA* is an essential multifunctional transcription elongation factor that is universally conserved among eubacteria and archaea (Nudler and Gottesman, 2002). *NusA* may elicit opposite effects on transcription (Richardson, 1996). By itself, *NusA* stimulates pausing at some hairpins and *rho*-independent terminators. In complex with other Nus factors (*e.g. NusG*) it stimulates anti-termination at both *rho*-dependent and *rho*-independent terminators (Richardson, 1996; Nudler *et al*, 2002). *GreA*, functions when transcription elongation slows or stops completely at certain points of the template, resulting in formation of paused or arrested complexes. In these complexes, RNAP shifts along the DNA template in the direction opposite to that of transcription (Komissarova and Kashlev, 1997; Nudler *et al*, 1997). As a result, the 3' end of RNA disengages from the RNAP catalytic center making further elongation impossible. An arrested complex can resume transcript elongation only following endonucleolytic cleavage of the nascent RNA that generates a

new 3' end of the transcript in the RNAP catalytic center. The endonucleolytic reaction performed by the RNAP catalytic center is slow, but is greatly stimulated by transcript cleavage factors including *GreA* (Fish and Kane, 2002; Sergei *et al*, 2005; Stepanova *et al*, 2007). A recent study of iron induced persistence in *C. trachomatis* E showed that both *NusA* and *NusG* were regulated (Dill *et al*, 2009). This study is intriguing as it shows that Nus factors are regulated during persistence and might explain why *CTIG643* shows such discrepancy in termination under normal and stress conditions.

CTIG360 was described earlier as a potential riboswitch for the *bioY* gene. *CTIG360* was expressed early during the normal developmental cycle and was also detected during carbenicillin treatment. It was not detected during IFN-γ-induced persistence but its expression was readily detected during reactivation. IFN-γ induced persistence analyses indicated that, although *CTIG360* was not detected, a larger product (> 500 nt) was detected. The *bioY* gene is ca. 785 bp, including its 5'UTR. There is a possibility that this band represented the full length *bioY* mRNA. *Chlamydia* is a biotin auxotroph and might be sensing a need to transport biotin under IFN-γ treatment, thus allowing for more full length transporter gene to be transcribed. Under normal growth conditions there may be enough biotin that the transporter is down-regulated. This hypothesis needs more testing to verify such claims. We are currently undergoing detailed experiments using *E. coli* lacking *bioY* function to verify the function of the chlamydial *bioY* ortholog and to test the riboswitch hypothesis.

4.3.5 SIGNIFICANCE OF MULTIBLE BANDS FOR A SINGLE ncRNA

The expression of multiple species of ncRNAs has been previously described (*e.g.* for *gadY*, (Opdyke *et al*, 2004), *dsrA* (Repoila and Gottesman, 2001), and *glmY* (Reichenbach *et al*, 2008)) and is believed to involve post-transcriptional processing that is dependent on the balance between protection by RNA binding protein(s) and degradation by RNases. The functional significance has not been well-studied but, importantly, Davis and Waldor (Davis and Waldor, 2007) have shown that *MicX* of *Vibrio cholerae* can be detected as both an active precursor and a processed form and that the smaller, processed molecule is significantly more stable and abundant. Interestingly, processing occurs in an Hfq and RNase E-dependent manner but the regulatory effects of *MicX* on its target mRNA is not dependent on Hfq. *Chlamydia* spp. do not contain obvious orthologs of Hfq or RNase E, suggesting that regulation and processing may be very different in these organisms.

The presence of several bands representing a single ncRNA occurs because the ncRNA is processed (*CTIG153, pCHL1)* or the ncRNA might incompletely terminate at a *rho* independent terminator (*CTIG643*). The relative abundance of each band on the Northern blotting followed two different kinetic profiles (Figure 4-6). One profile shows identical relative intensities of both bands representing a single ncRNA. This group is represented by (*CTIG153, pCHL1, CTIG643,* and *CTIG684*) under normal developmental condition. The second profile shows different relative intensities, *e.g.* under carbenicillin treatment the small band was dominant for *CTIG153, CTIG643* and *pCHL1*. Previous

discussion of *CTIG643* indicated a likely mechanism, but for *CTIG153* and *pCHL1* there is no terminator in either RNA. Multiple bands in *CTIG153* and *pCHL1* might be due to processing of the 3' end as both ncRNAs are antisense transcripts and the short transcripts represent the 5' half of the molecule. Apparently the 3' end of the molecule was cleaved after forming double stranded RNA with the target gene. Processing of double stranded RNA may be carried out by chlamydial RNase III (*CT297*) and/or RNase G (*CT808*).

CHAPTER 5. FUNCTIONAL ANALYSIS OF ncRNA

5.1 *CIS*-ACTING *CTIG270*

5.1.1 VERIFICATION OF CTIG270 MAPPING

One of the *cis*-acting ncRNAs (*CTIG270*) (Figure 3-1) was selected for further analysis. Mapping of *CTIG270* transcript was carried out using the RNA circularization protocol (Table 4-1). *CTIG270* extends on the upper strand 142 and 80 nucleotides 5' and 3' to the *ftsI* stop codon respectively (Table 4-1). Mapping was verified using 5' RACE using *C. trachomatis* RNA from 24h PI (Super SMART^{TM} PCR cDNA synthesis, Clontech). The amplified 5'RACE product was TA cloned into pCR2.1 (Invitrogen, CA) and individual clones sequenced. Results identified a TSS for *CTIG270* similar to what was reported in (Table 4-1).

5.1.2 FUNCTIONAL ANALYSIS OF CTIG270 IN A SURROGATE SYSTEM

To test the effect of expression of *CTIG270* on *ftsI*, a system was constructed where both *ftsI* and *CTIG270* were cloned in two compatible medium copy number plasmids and transformed into *E. coli*. *ftsI* was fused in frame to a 5' FLAG-tag sequence ($_N$-DYKDDDDK-$_C$) and cloned under the control of a *lac* promoter (from pCR2.1, Invitrogen).

When *CTIG270* was induced (Figure 5-1 lane 1 panel C) neither the FLAG-*ftsI* protein nor its mRNA were detected (Figure 5-1 lane 1 panel A, B). When the second inducer was added (Figure 5-1 lanes 2 and 3), the FLAG-*ftsI* protein and its mRNA were undetectable. The amount of *CTIG270* detected was reduced by the end of the double induction period and none of the RNAs or protein was detectable (Figure 5-1 lane 3).

When the expression of the FLAG-tagged *ftsI* was induced by IPTG; both the mRNA and protein were detectable (Figure 5-1 lane 4). When the second inducer was added (Arabinose) *CTIG270* was detected and both FLAG-*ftsI* protein and its mRNA were transiently increased (Figure 5-1 lane 5). By the end of the double induction period, the FLAG-*ftsI* protein and its mRNA were barely detectable, and *CTIG270* was markedly reduced (Figure 5-1 lane 6).

The control culture (no induction) showed little to no expression (Figure 5-1 lane 7). When both inducers were added both RNAs and the protein were detected (Figure 5-1 lane 8). By the end of the co-induction experiment the FLAG-*ftsI* protein and its mRNA were not detectable and reduced amounts of *CTIG270* were detectable.

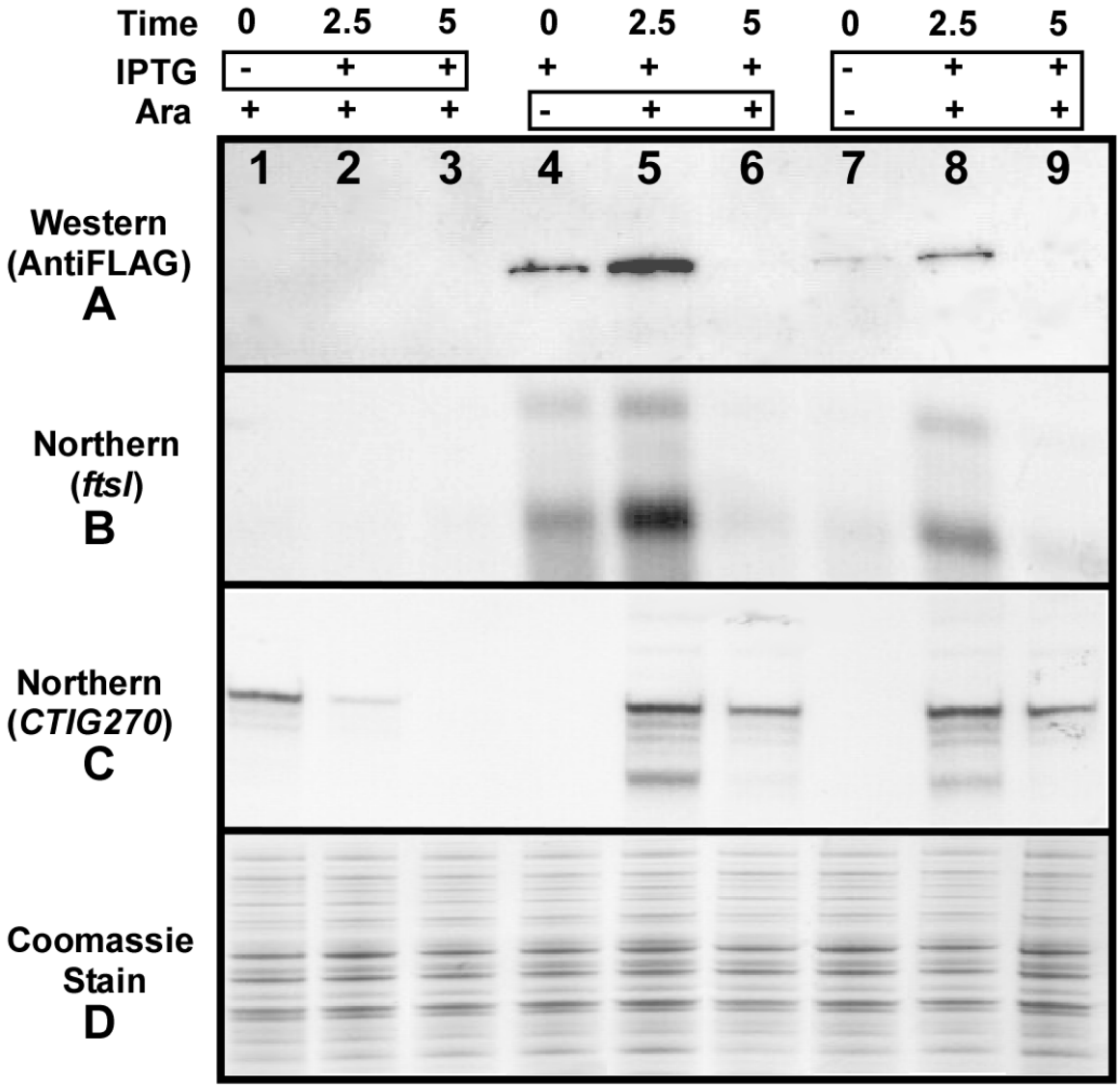

Figure 5-1 *CTIG270* analysis in an *E. coli* surrogate system. Lanes 1, 2, 3 represent induction by arabinose followed by IPTG, lanes 4, 5, 6 represent induction by IPTG followed by arabinose, lane 7 represents no induction, lanes 8, 9 represent double induction. Panel A showed Western blotting for a FLAG tagged *ftsI*, Panel B showed Northern blotting for *ftsI* mRNA, Panel C showed Northern blotting for *CTIG270*. Ara: arabinose.

5.1.3 IN-VIVO EXPRESSION OF CTIG270 AND ftsI

We analyzed the expression of both *CTIG270* and *ftsI* using matched Northern blotting and reverse transcription-PCR respectively. The expression of *ftsI* mRNA reached its maximum 16h PI. On the other hand, *CTIG270* was barely detectable at the same time point (Figure 5-2). By 24h PI, *CTIG270* expression reached its maximum while *ftsI* expression level was steadily decreasing. By 48h PI, both RNAs were barely detectable.

5.2 DISCUSSION

5.2.1 USING E. coli AS A SURROGATE SYSTEM

We used an *E. coli* surrogate expression system because we lack an established genetic system to manipulate *Chlamydia*. The results show possible mechanisms of action that might take place *in vivo* during an actual infection. The results must be understood in the contexts of an actual chlamydial infection (in a monolayer or in clinical samples). Knowing that both *E. coli* and *Chlamydia* are different and knowing that *E. coli* has the RNA chaperone Hfq, while *Chlamydia* lacks this protein, we examined the results.

The matched RT-PCR/Northern results for *ftsI*/*CTIG270* (Figure 5-2) show the levels of both species of RNA *in vivo*. When the expression of *ftsI* was at its maximum level, the expression of *CTIG270* was undetectable. On the contrary, when the expression of *CTIG270* reached its maximum level *ftsI* mRNA level was rapidly declining. This pattern is suggestive of an antisense regulation system, and it adds an indirect *in vivo* validation for results obtained from the surrogate system.

It has been shown that Arabinose can act as an inducer for *lac* derived promoters in *E. coli* (Narayanan *et al*, 2006). This is evident in Figure 5-1 lanes 5 which showed a transient increase in the FLAG-FtsI protein and RNA under arabinose induction. This is why we switched from using *lac* promoters to tetracycline promoters in later experiments (described later).

5.2.2 EXPRESSION OF CTIG270 DURING STRESS CONDITIONS

The results of the surrogate expression system showed an efficient tight regulation of the FLAG tagged *ftsI*. *CTIG270* controlled the expression of its target gene whether it was induced before or after target gene induction. In Figure 5-1 lanes 3, 6, and 9, both inducers were present and both RNAs were transcribed. Interestingly we could not detect either the FLAG-FtsI protein or its mRNA and very little *CTIG270* was detected. This result helped us to interpret the *CTIG270* expression pattern during IFN-γ persistence (Figure 4-5 panel B). The expression of *CTIG270* was greatly reduced during IFN-γ

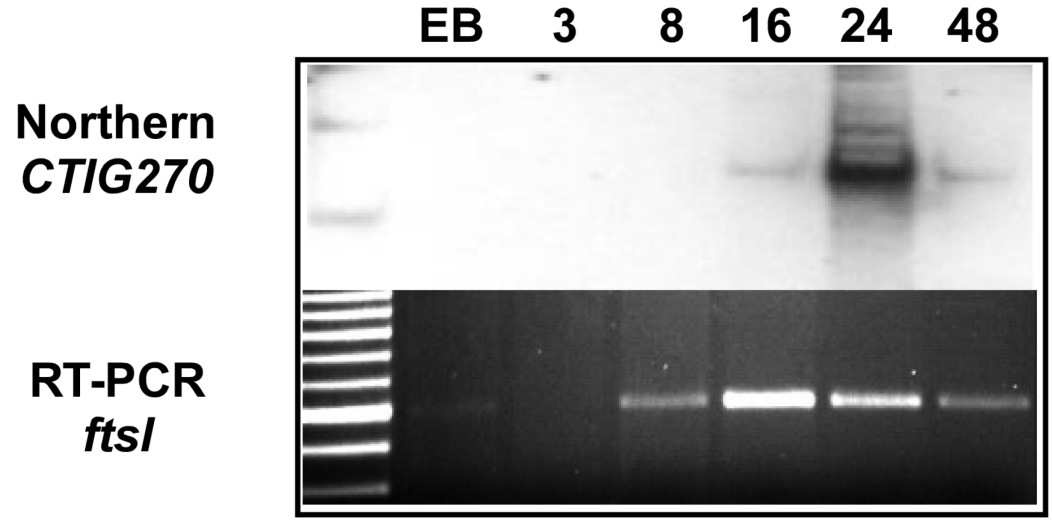

Figure 5-2 RT-PCR of *ftsI* mRNA and *CTIG270* Northern blotting during normal developmental cycle. EB: elementary bodies (time zero).

treatment. IFN-γ-induced persistence was associated with a block in RB cell division and the presence of large, multinucleate RBs within the inclusion (Beatty *et al*, 1993). The lack of *CTIG270* signal during IFN-γ-induced persistence implied that either the ncRNA was not transcribed or that the ncRNA is binding to its target and is being rapidly degraded by RNase. The idea appears radical, but if the function of *CTIG270* was to quantitatively degrade its target mRNA then the lack of signal on the Northern blotting might be because the ncRNA is highly active. *CTIG270* may be playing a role in the lack of bacterial cell division. *C. trachomatis*, although lacking Hfq and RNase E, encodes a homolog of RNase III which, among other functions, can degrade dsRNA (Viegas *et al*, 2007). To test this phenomenon we may need to genetically control the expression of *ftsI* inside *Chlamydia* and measure the effect on *CTIG270* transcript half life at various time points.

During carbenicillin treatment *CTIG270* was not detected and treatment with carbenicillin and other β-lactams results in very large, aberrant RBs that appear to have a distinct block in cell division. However, in this case the situation is somewhat different as FtsI (a penicillin binding protein) is also a target of carbenicillin. In other words, *FtsI* is already under tight functional control from the antibiotic. In summary, the delayed expression of *CTIG270* and then the complete disappearance of the ncRNA during persistence fits with our hypothesis that *CTIG270* is regulating *ftsI* mRNA.

5.3 DESIGN AND OPTIMIZATION OF A SCREENING TOOL FOR *trans*-ACTING ncRNA

5.3.1 RATIONAL

Screening the chlamydial genome for potential ncRNA targets was a novel goal. We proposed several criteria for designing a screening system. These criteria were pivotal so that the system would be functional and useful:

- Genomic DNA fragments will be cloned in a translational fusion with a reporter system (target plasmid). Upon inhibition of translation a phenotype can be visually detected.

- The phenotype elicited by translational inhibition of the reporter system should be easily distinguished. This phenotype should allow selection of positive clones from a larger number of negative background clones.

- During transformation, *E. coli* that takes up a self-ligated plasmid would not survive phase I selection, thus reducing plasmid load.

- Target plasmid with library inserts that are out of frame with the reporter system would not survive phase I selection. Phase I selection will reduce the genomic

fragments under test to include only genomic fragments in frame with the reporter system.

- Target plasmid should be a low or medium copy plasmid.
- Plasmid carrying the ncRNA (effector plasmid) should be a high copy plasmid compatible with the target plasmid.
- Expression of both the genomic DNA/reporter system and the ncRNA should be controllable.

We designed a screening system that would fulfill these criteria. The system underwent an optimization process until a target plasmid was constructed that adequately fulfilled the aforementioned criteria.

We used the pZ modular plasmid system designed by Lutz and Bujard (Lutz and Bujard, 1997). The pZ vector comprised of three main modules separated by unique cleavage sites. The modular organization of the plasmid made it easy to exchange promoters, origins of replication, and RBS. We used plasmid pZS*24MCS (Figure 5-3) as the parent plasmid, this plasmid had origin of replication SC101* (3-5 plasmid copy numbers/cell) (Manen *et al*, 1994) and had Kanamicin as a selectable marker. **Module I** contained the promoter/operator construct, and ribosomal binding site (RBS). This module was flanked by *AatII/EcoRI*. **Module II** harbored one of several compatible origins of replication (ColE1, p15A, pSC101*). This module was flanked by *AvrII/SacI*. **Module III** contained one of five antibiotic resistance markers for our purpose we used Kanamicin resistance marker.

5.3.2 SCREENING SYSTEM DESIGN

Target plasmid: in order to construct the target plasmid, we used life/death as a selection tool. Life/death selection was used for phase I selection which involved elimination of non-translational fusions and self ligations. Life/death selection was also used in phase II selection which involved selection of potential targets for ncRNA.

Life/death selection was established by using a fusion between a **positive selection gene** and a **negative selection gene**. The positive selection gene was an antibiotic marker, (chloramphenicol acetyltransferase (*CAT*)). The negative selection gene was a toxic gene (*ccdB*). *ccdB* is a member of the ccd poison/antidote system of the F plasmid (Critchlow *et al*, 1997; Maki *et al*, 1996). This system encodes a toxin targeting the essential *E. coli* DNA gyrase. CcdB inhibits DNA gyrase leading to unrepaired strand breaks and disintegration of the bacterial chromosome. CcdA, On the other hand, is the antidote, it interacts with CcdB protein to neutralize its toxicity (Bahassi *et al*, 1999). This toxin-antitoxin system was explored by other investigators (Gabant *et al*, 1997), and *ccdB* gene is in use in the Invitrogen Gateway system™. The reporter system was a translational fusion between *ccdB* and *CAT* (*ccdB*-*CAT*).

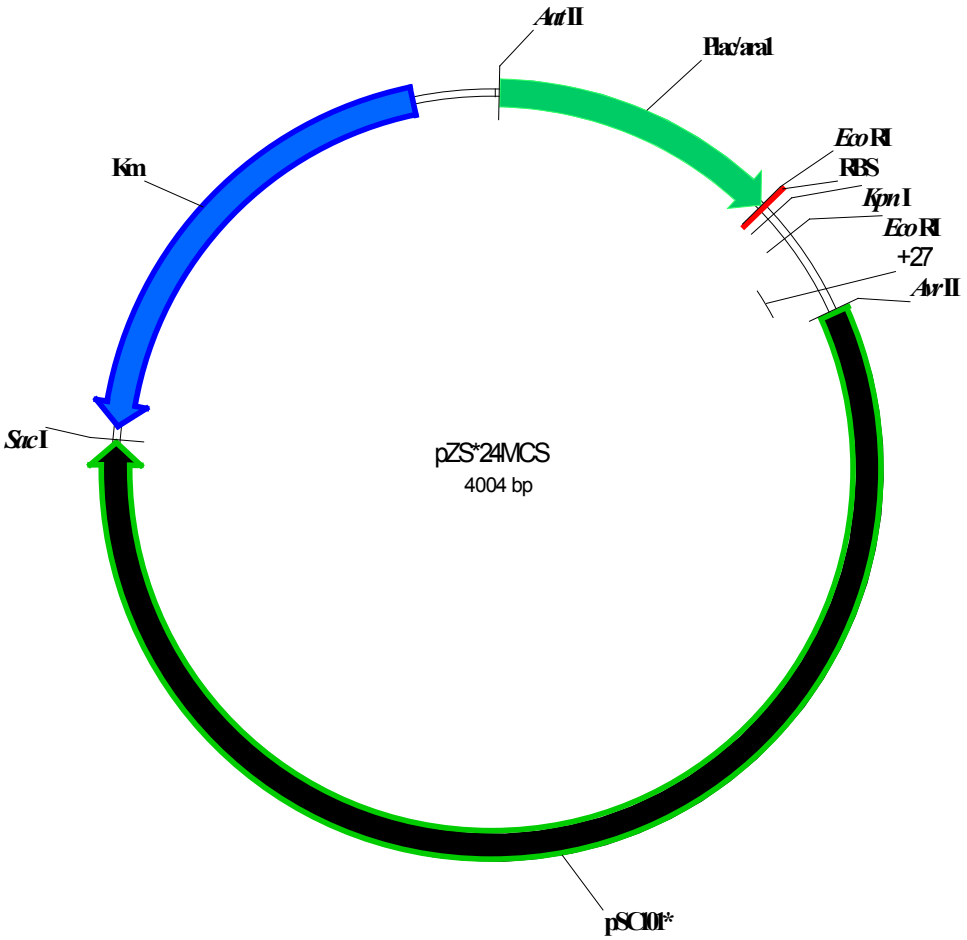

Figure 5-3 Vector pZS*24MCS. The vector had $P_{lac/ara}$ promoter, flanked by AatII/EcoRI, the map also shows pSC101* origin of replication flaked by AvrII/SacI.

The fusion gene was under a $P_{LtetO-1}$ promoter (Lutz *et al*, 1997). Figure 5-4 represents a schematic of the fusion gene. *ccdB* was cloned with no translational signal (RBS and the first ATG) at the 5' end. The absence of a translational signal upstream of the fusion genes prevented the ribosomes from binding to the newly formed mRNA. When the fusion gene was transcribed the cells maintained their sensitivity to chloramphenicol. The fusion gene was constructed as follows: briefly *ccdB* and *CAT* were amplified separately by PCR using the Gateway Vector Conversion System as a template (Invitrogen, CA). Sequential PCR reactions were used to add the $P_{LtetO-1}$ promoter sequence along with a *StuI* site directly upstream of *ccdB* second codon. PCR was also used to create a 15-base overlap between *ccdB* 3' end and the *CAT* 5' end. During primer design the first ATGs of both *ccdB* and *CAT* were eliminated along with *ccdB* stop codon. The 15-base overlaps between *ccdB* and *CAT* were designed to keep both genes in frame after recombination. PCR products were gel purified and recombined using PCR InFusion Cloning kit (Clontech, CA) according to manufacturer's instructions. The recombination product was selected and amplified using appropriate PCR primers. PCR was used to create 15 base overlaps between the fusion gene and the linearized target plasmid. Plasmid pZS*24MCS was digested with *AatII/EcoRI* (NEB, MA) to remove the promoter, RBS, and the multiple cloning site (MCS). The *ccdB-CAT* fusion PCR amplicon was gel purified and recombined with the linearized pZS*24MCS plasmid using PCR InFusion Cloning kit (Clontech, CA). The recombined plasmid was transformed into electrocompetent One Shot *ccdB* Survival Cells (Invitrogen, CA). The *ccdB* Survival cells were used because they were resistant to the cytotoxic effect of CcdB protein. These cells had a mutation in the *gyrA* gene (Arg462/Cys substitution) (Bernard and Couturier, 1992; Bernard *et al*, 1993). Transformants were selected on LB agar Kanamicin plates (25µg/ml). For list of primers see Table 2-1 under *ccdB/CAT* Cloning. The fusion gene lacked a translational start signal. For the fusion gene to be translated it required an in frame translational signal supplied by the chlamydial genomic library. *E. coli* where the fusion protein was produced were chloramphenicol resistant, while *E. coli* where the fusion protein was not produced were chloramphenicol sensitive.

5.3.3 FUNCTIONAL TESTING OF THE TARGET PLASMID

Testing the target plasmid was carried out by supplying a translational signal to *ccdB-CAT* fusion gene. This was followed by investigating whether the resulting plasmid confers chloramphenicol resistance to *ccdB* resistant *E. coli* (positive selection) and whether it subjects *ccdB* sensitive *E. coli* to *ccdB* mediated killing (negative selection).

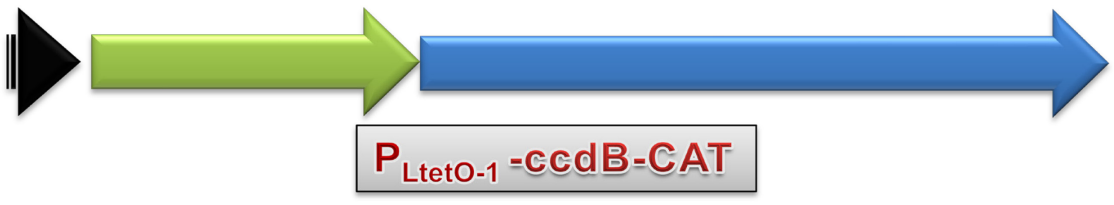

Figure 5-4 Initial design of the fusion reporter system.

To carry out this test, target plasmid was linearized by *StuI* digestion (NEB, MA). The translational signal was supplied by *hctA* leader/ATG. *hctA* leader sequence was amplified by PCR and the amplicon included all the 5'UTR of *hctA* plus the first ATG (Fahr *et al*, 1995) (Table 2-1 under Test SEQ Cloning). PCR was also used to add 15 base overlaps with both ends of the *StuI* linearized plasmid as shown in Figure 5-5. The resulting plasmid was transformed into electrocompetent One Shot *ccdB* Survival Cells and selected on Kanamicin (25µg/ml)/chloramphenicol (12.5µg/ml) LB Agar plates. Clones were screened by PCR and confirmed by sequencing. *ccdB* Survival cells containing translatable fusion gene were resistant to chloramphenicol. This implied that the fusion gene was translated and cells acquired chloramphenicol resistance phenotype. The Target plasmid containing translatable fusion gene was transformed into *ccdB* sensitive cells (Bl21A1 *E. coli*, Invitrogen) and grown on Kanamicin/chloramphenicol LB agar plates. Cells survived this selection. The lack of lethal effect of CcdB was not a result of lack of translation of the fusion gene but rather CcdB in the fused form might not be functional.

5.4 REDESIGN OF THE TARGET PLASMID

To address the loss of function of CcdB in the fusion form, we decided to insert two protease cleavage sites flanking *ccdB*. Upon translation of the fusion gene, a protease will regenerate the native CcdB.

Many proteases were used to cleave fusion proteins. For example, activated blood coagulation factor X (factor Xa), enteropeptidase (enterokinase) and α-thrombin have been used. There are many reports of fusion proteins that were cleaved by these

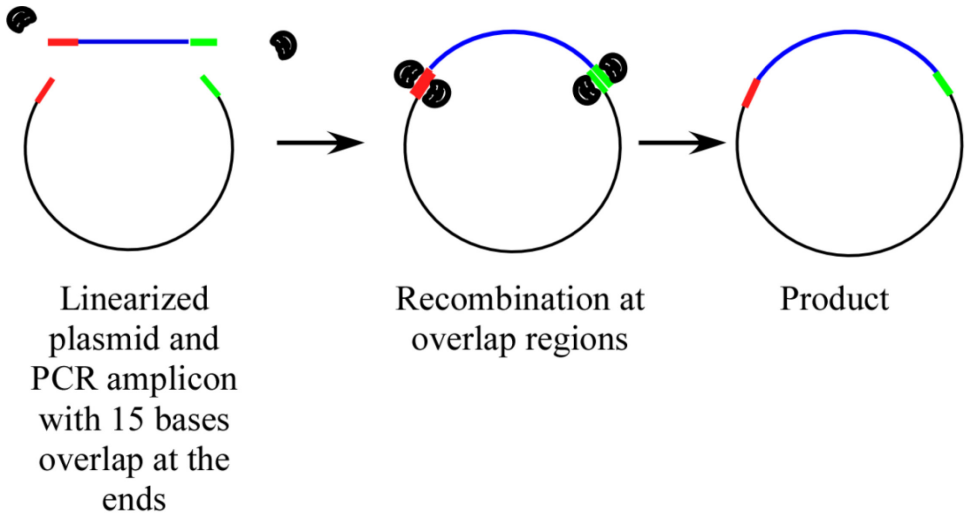

Figure 5-5 Schematic of the PCR infusion cloning.

proteases at locations other than the designed site (Forsberg *et al*, 1992; Forsberg *et al*, 1991; He *et al*, 1993; Stevens, 2000; Wagner *et al*, 1996), We selected tobacco etch virus protease (TEV protease) (Carrington and Dougherty, 1987, 1988) for use in the target plasmid, as the TEV protease had higher stringent sequence specificity than many other proteases (Dougherty *et al*, 1989; Kapust *et al*, 2002), it was already in use to cleave fusion proteins *in vivo* (in *E. coli*) (Shih *et al*, 2005) and *in vitro* for affinity purification (Babu *et al*, 2009; Song *et al*, 2009). TEV protease cleaves the amino acid sequence ENLYFQG/S between QG or QS with high specificity (Parks *et al*, 1994). Cleavage by TEV protease leaves one amino acid at the N-terminus of the protein of interest (Kapust *et al*, 2002). The use of the TEV protease presented two problems. The first problem was the low solubility of the protein in *E. coli* (Rachel and David, 1999). The second problem was the auto-digestion of the protease yielding fragments of very low activity (Lucast *et al*, 2001). We selected *TEV* protease developed by van den Berg *et al*, 2006. The group used TEV_{S219N} first created by Lucast *et al*, 2001. The single amino acid change, S219N inhibited the auto-proteolysis and subsequent loss of activity of the protease. The group then embarked on an accelerated mutagenesis scheme to develop a mutated version of the TEV protease that had more solubility than the wild-type protein. They developed a mutated *TEV* protease that was 5 times more soluble than the wild-type but retained wild-type levels of proteolytic activity. The mutant *TEVsh* had 3 amino acid changes from the wild-type (van den Berg *et al*, 2006).

Analysis of the crystal structure of both CcdB (Dao-Thi *et al*, 2004; Loris *et al*, 1999) and CAT (Andreeva *et al*, 2000) showed that a cleavage site between CcdB and CAT might be hidden from the protease. To address this potential problem we adopted a technique termed "Target-directed proteolysis at the ribosome" developed by Henrichs *et al*, 2005. Henrichs *et al.* developed a technique for SecA inactivation by proteolysis *in vivo*. Originally a technique called target-directed proteolysis allows complete proteolytic inactivation of target proteins *in vivo (Mondigler and Ehrmann, 1996)*. When investigating Target-directed proteolysis sites in SecA, it was found that lack of complete accessibility and protein folding caused the TEV cleavage site located downstream of residue 195 to be incompletely cleaved (Mondigler *et al*, 1996). Henrichs *et al.* devised a way to tether the TEV protease to the ribosome. The tethered TEV protease gained access to the nascent polypeptide as it exited the ribosome before folding occurs. The technique was successful at getting complete cleavage of the target protein (Henrichs *et al*, 2005). The tethering motive was cloned from an *E. coli* trigger factor (*tig*). Previous research showed that *E. coli tig* had a binding site on ribosomes (Hesterkamp *et al*, 1997). The domain that retained the ribosomal tethering activity was the N terminal 144 residues. Henrichs *et al*, 2005 fused this domain to the N terminus of TEV protease and the result was complete cleavage of SecA195. We adopted the same strategy and fused *tig144* to the *TEVsh* protease.

TEVsh protease gene was put under the control of a $P_{lac/ara\text{-}1}$ promoter (Lutz *et al*, 1997), which was a synthetic fusion of a *lac* promoter and an arabinose promoter. This synthetic promoter was regulated over a ~1800-fold range whereby derepression via IPTG caused a ~100-fold induction. Activation with arabinose caused a 15-20-fold

increase in promoter activity (thus both inducers lead to 1500-2000 fold activation) (Lutz *et al*, 1997).

We prepared four clones, all of them had the fused gene *ccdB/CAT* under the control of a $P_{LtetO-1}$ promoter and all four clones had two TEV protease cleavage sites (tev) flanking *ccdB* (The fusion gene is now *tev-ccdB-tev-CAT*). The differences between the constructs were the origin of replication (pSC101* of 3-5 copies per cell and p15A of 30-50 copies per cell), the *TEVsh* protease, and the N terminal fusion to *tig144* (Figure 5-6).

5.4.1 CONSTRUCTION OF THE REDESIGNED PLASMIDS

The constructs were built as separate pieces of DNA via PCR (Fusion gene, Origin of replication p15A/pSC101*, and *TEVsh/tig144TEVsh*). Different pieces of DNA were recombined sequentially into the destination plasmid pZS*24MCS to produce the four target clones (Figure 5-6).

5.4.2 CONSTRUCTION OF THE FUSION GENE ccdB/CAT

The fused gene was created as follows: briefly *ccdB* and *CAT* were amplified separately via Sequential PCR using Phusion™ High-Fidelity PCR (Finnzymes, MA). Sequential PCR was carried out as mentioned before, the PCR primers were used to add two tev sites (ENLFYQS), a $P_{LtetO-1}$ promoter sequence (Lutz *et al*, 1997) and *StuI* restriction site upstream of tev site 1 (Figure 5-7). The 15 base-overlaps the destination plasmid were added to the 5' end of the promoter. *CAT* was amplified via PCR with 15 base overlaps of the tev site 2 were added to the 5' end of *CAT* and 15 base overlaps in the destination plasmid were added to the 3' end of *CAT*. The two PCR fragments were gel purified and allowed to recombine using Infusion Advantage PCR cloning kit (Clontech, CA). The recombination product (promoter-tev1-*ccdB*-tev2-*CAT*) was amplified using appropriate primers and gel purified then used for recombination with the target plasmid. The target plasmid was digested by *AatII/EcoRI* (NEB, MA) to remove the plasmid's own promoter, RBS, and MCS. Promoter-*tev1-ccdB-tev2-CAT* fusion was recombined into the digested plasmid using Infusion Advantage PCR cloning (Clontech, CA). Recombination product was transformed into electrocompetent One Shot *ccdB* Survival Cells (Invitrogen, CA) and selected on kanamicin (25µg/ml) LB agar plates. Clones that carried the insert were verified by PCR and sequencing. Table 2-1 lists primers used for cloning under *ccdB/CAT*. Figure 5-8 shows the general scheme for constructing all four clones.

5.4.3 EXCHANGING pSC101 WITH p15A*

The pSC101* origin of replication was removed from the target plasmid via double digestion with *AvrII/SacII-HF* (NEB, MA). The p15A origin of replication was

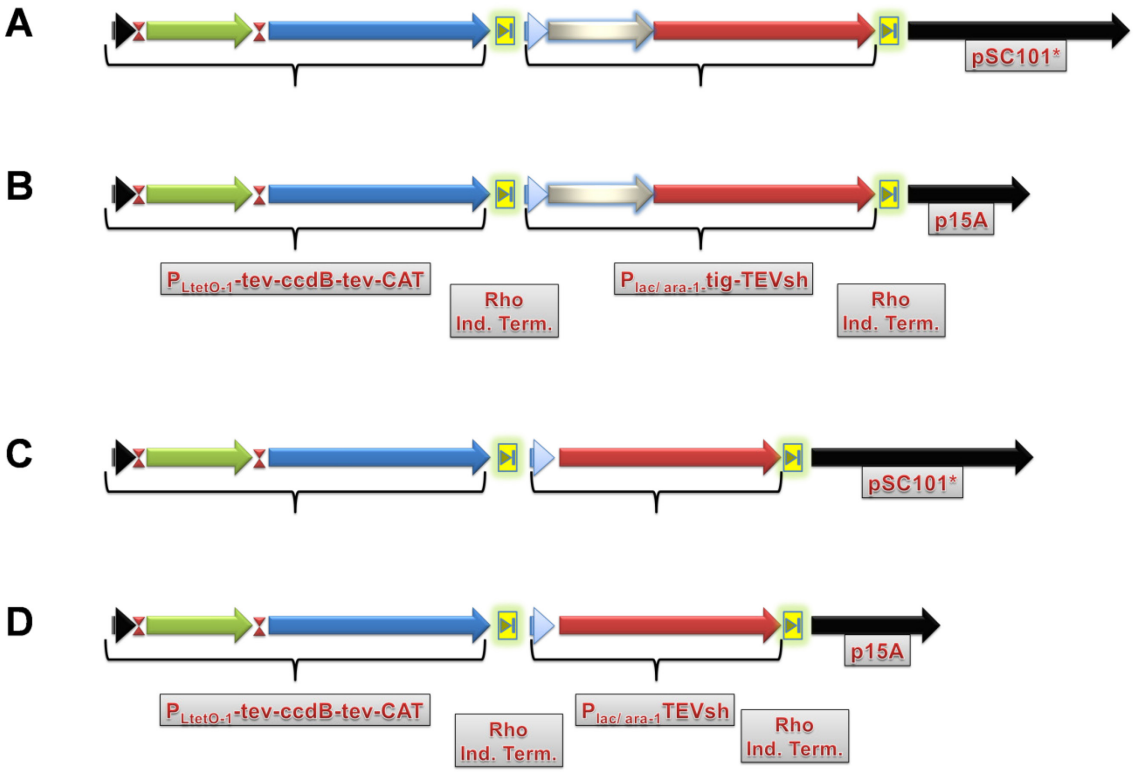

Figure 5-6 Schematic of the four clones for testing *ccdB* regeneration from target plasmid. Green ORF represent *ccdB*, blue ORF represent *CAT*, red ORF represent TEVsh, tig144 is represented by gray ORFs, tetracycline promoter are represented by black arrow heads and *ara/lac* promoters are represented by gray arrow heads, Origin of replication is represented by black arrows.

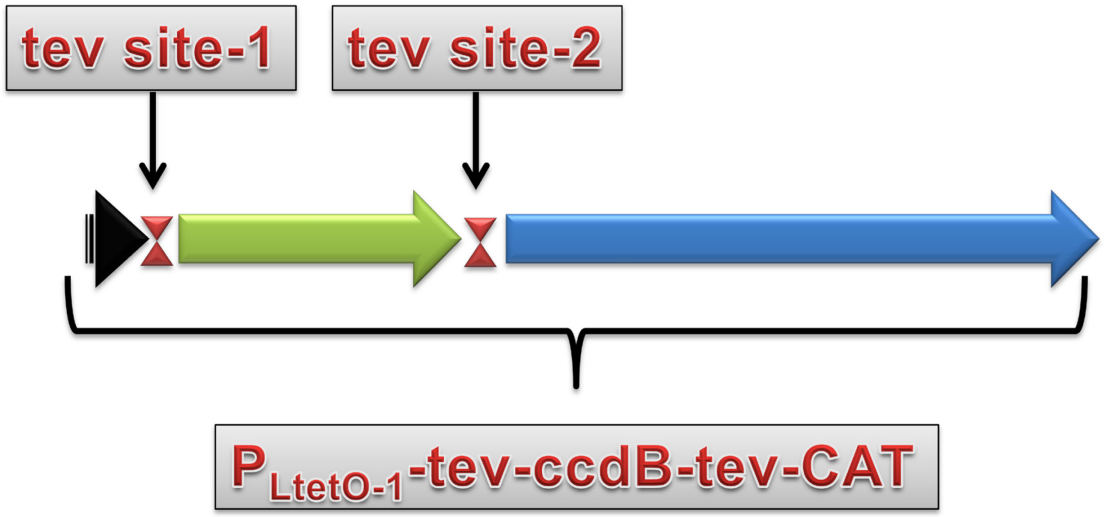

Figure 5-7 Modified fusion gene including tev sites.

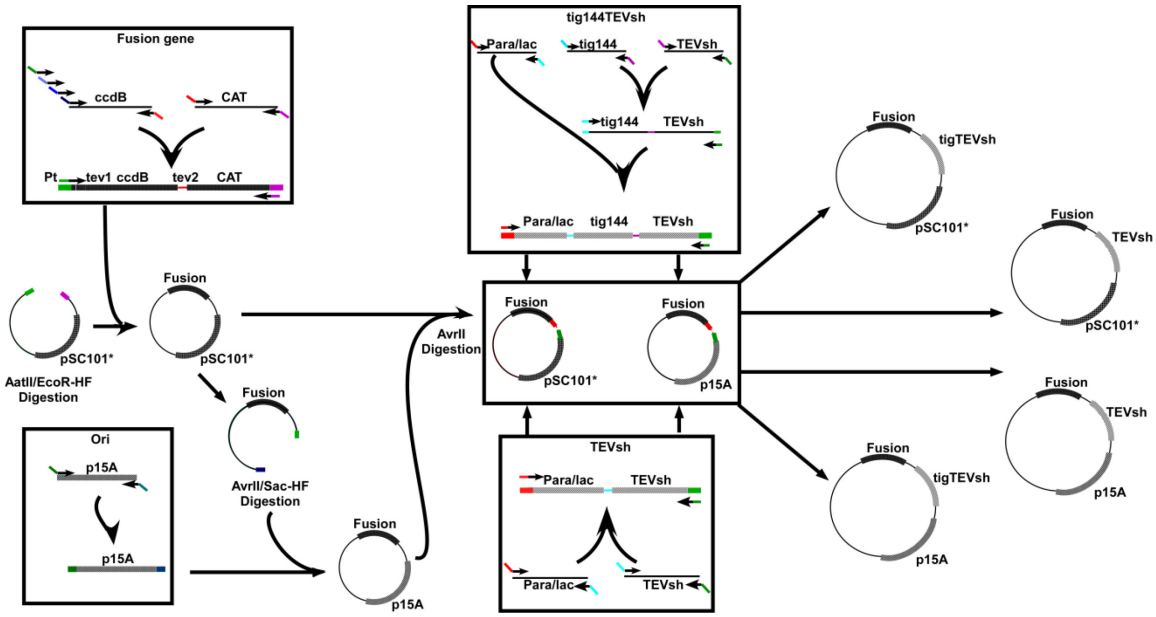

Figure 5-8 Scheme for the construction of different target plasmids.

amplified by PCR using pZA24MCS as a template. 15 base overlaps between the amplicon and the ends of the digested plasmid were created in the same PCR. Recombination, transformation and selection were carried out as mentioned before; clones were selected and verified by PCR and DNA sequencing. At this stage we have created two target plasmids (both have identical fusion genes) with a different origins of replication (pSC101* & p15A). Table 2-1 lists primers used for this reaction under "Origin Change".

5.4.4 INSERTION OF TEVsh AND tig144TEVsh INTO THE TARGET PLASMID

Two different TEV protease constructs were created. The first was a fusion between the *TEVsh* (van den Berg *et al*, 2006) and the promoter $P_{lac/ara\text{-}1}$ (Lutz *et al*, 1997). As mentioned before, PCR was used to amplify each fragment. *TEVsh* amplified from TH24 plasmid (gift from Dr Berglund) (van den Berg *et al*, 2006), and $P_{lac/ara\text{-}1}$ was amplified from pZS*24MCS. 15 base overlaps were created at the points of recombination between both fragments *i.e.* the TEVsh amplicon included the translation signals for the gene (RBS and first ATG). Recombination was carried out as described previously. The promoter/*TEVsh* fusion was PCR amplified and gel purified. The second TEV protease construct was composed of three pieces of DNA. One piece represented the promoter $P_{lac/ara\text{-}1}$ (Lutz *et al*, 1997), the second represented *tig144* (from the *E. coli* DH5α chromosome), and the third piece represented the *TEVsh* gene. PCR was used to amplify each fragment and to create 15 base overlaps between all fragments at the points where each fragment recombined with the other (Figure 5-8). An overlap was created between the 3'end of the promoter and the 5' end of *tig144*. Another overlap was created between the 3' end of *tig144* and the 5' end of *TEVsh*. Recombination was carried out between all three DNA fragments, and the tripartite product was amplified by appropriate PCR primers and the product was gel purified. Both *TEVsh* Constructs carried 15 bases-overlap with the destination plasmid at each end. Both target plasmids were linearized by *AvrII* digestion and then recombined with either of the *TEVsh* constructs. Recombination products (Figure 5-6) were subsequently transformed, selected and checked as before. (Table 2-1 has the list of primers under pBAD Cloning, *Tig144* Cloning, *TEV* for *Tig144* Cloning, *TEV* for pBAD Cloning).

5.5 OPTIMIZING THE TARGET PLASMIDS

The target plasmid construction went through an optimization process. Pieces of DNA were fused into the plasmid at different stages, as detailed in Figure 5-8. Optimizations led to the emergence of problems that were not foreseen in the original scheme and needed to be resolved in order for the target plasmid to be useful.

5.5.1 StuI RESTRICTION SITE OVERLAPPING dcm METHYLASE SITE

We planned to use *StuI* site for insertion of the chlamydial library. The *StuI* enzyme is sensitive to methylation. Methylation may take place because of an overlapping *dcm* methylation site. A *dcm* methylation site overlapping *StuI* site was created when tev site1 was added upstream of *ccdB* (AGGCCTGG the underlined sequence is *StuI* site, the 2 extra Gs after the *StuI* site created ***C^{m}***CTGG which is a *dcm* methylation site). To make the *StuI* site functional, methylation was prevented by transforming target plasmids into the *dcm* deficient strain Bl21A1 *E. coli* (Invitrogen, CA). Knowing that Bl21 was *recA*$^{+}$*/endA*$^{+}$ (Phue *et al*, 2008), we preformed an extra washing step during plasmid preps to remove the excess nucleases (as recommended by the manufacturer). BL21 was not used for long term storage of target plasmids.

5.5.2 REMOVAL OF StuI SITE

tig144 contained an *StuI* site at position Q9. In order to remove this site we introduced a silent mutation into Q9 as follows: target plasmids were grown in *dcm*$^{+}$ cells (TOP10, Invitrogen) and plasmids were double-digested with *StuI*/*SalI-HF* (NEB, MA). Under these conditions only a single *StuI* site (in *tig144*) was digested leaving the other methylated *StuI* site intact. The double digest removed a 171 nt DNA fragment overlapping the junction between the promoter and *tig144*. PCR was used to amplify this 171 nt fragment and to include 15 bases-overlap with digested plasmid ends. The reverse primer was used to introduce a silent mutation into *tig's* Q9. CA**A**ggcct was mutated to CA**G**ggcct; the capital letter represented the mutated codon (encodes for Q) and the underlined sequence represent the *StuI* site. The modified DNA fragment was recombined into the target plasmid. Each recombination reaction was subsequently transformed/selected and verified as before. Figure 5-9B shows 2 PCR amplicons digested with *StuI*. Lane one represented the modified *tig144*, lane two represented nonmodified *tig144*. The nonmodified *tig144* was cleaved by *StuI* (small band) while the modified *tig144* resisted *StuI* digestion. (Table 2-1 has list of primers under "Tig144: *StuI* Removal").

5.5.3 ADDITION OF THE FIRST CODON OF ccdB AND MODIFICATION OF tev SITE1

The translational signal of *ccdB* was removed during initial cloning of the fusion gene. A chlamydial library inserted upstream of *ccdB* was planned to provide the translational signal. After creating a tev cleavage site upstream of *ccdB* the first codon was no longer the first ATG of *ccdB* gene. In order to make *ccdB* similar to the wild-type gene we decided to insert the first codon of *ccdB*. The tev site1 in the fusion gene was ENLFYQS. Cleavage occurred between Q/S leaving an S on 5'end of CcdB. The polarity of Serine may have a detrimental effect on CcdB function. We decided to mutate S to G as both S and G are strong cleavage sites for the TEV protease. The target plasmid was digested with *StuI*/*BmgBI*. The double digestion removed 151 nt

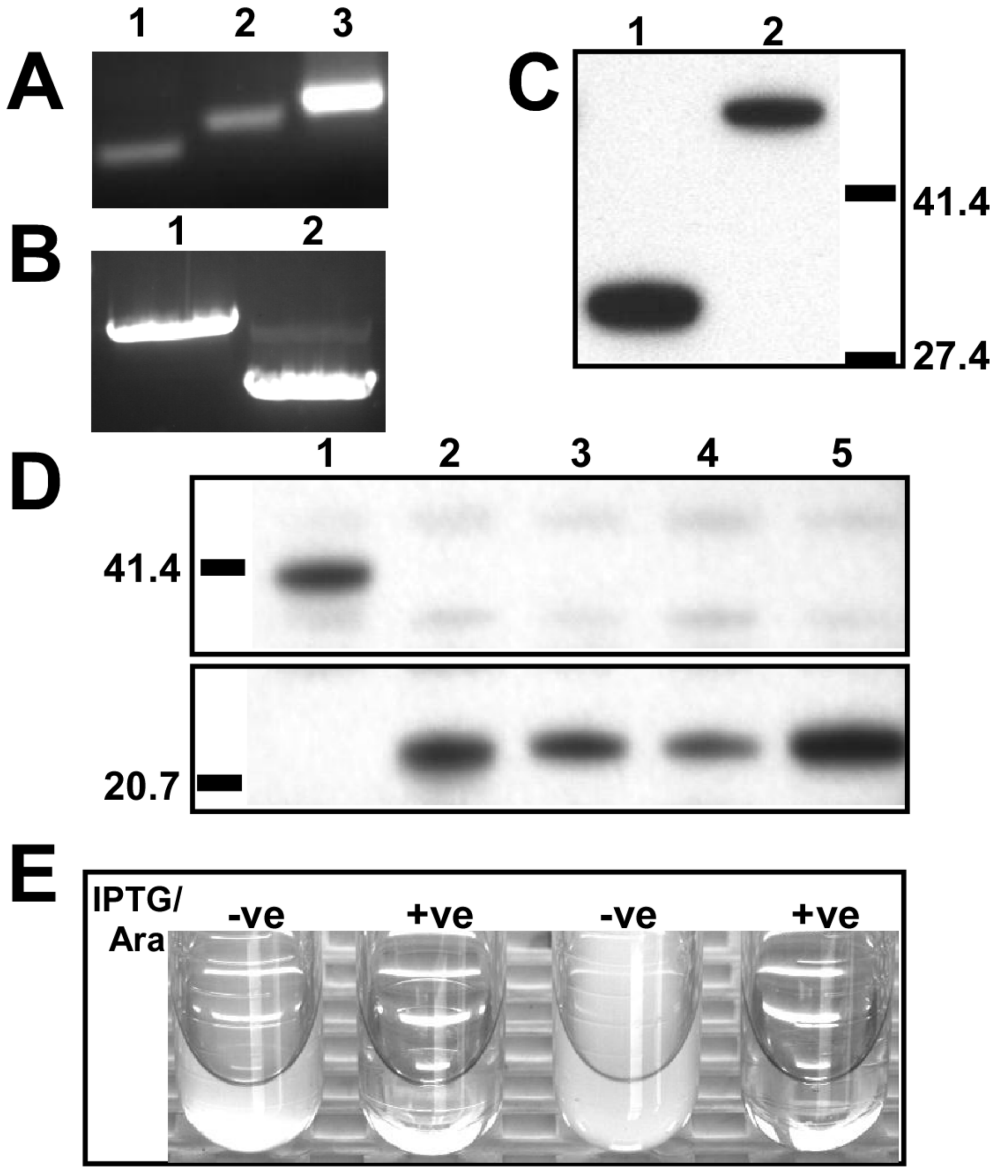

Figure 5-9 Functional analysis of target plasmids. Panel A shows sequencial nested PCR that added ATG and tev cleavage sites to *ccdB*. Panel B shows *StuI* digestion of wild-type and modified *tig144;* lane 1 shows modified *tig144* that is resistent to *StuI* digestion, lane 2 shows wild-type *tig144* sensitive to *StuI* digestion. Panel C shows Western blot results of TEVsh expression (the membrane was probed by anti-V5 antibody). Lane 1 shows TEVsh lane 2 shows tig144TEVsh fusion. Panel D shows western blotting results of reporter fusion protein expression (the fusion protein was probed by anti-CAT antibody): lane 1 shows fusion protein not digested by TEVsh, lanes 2 to 5 show digested fusion protein. The small size fragment represent CAT as seen in lanes 2 to 5, all four target plasmids exhibit processed fusion protein. Panel E shows the effect of expressing the fusion protein with TEVsh in ccdB sensitive *E. coli*. Turbid tubes represent clones where TEVsh expression was inhibited (no inducers) while clear tubes represent clones where TEVsh expression was induced.

piece of DNA. overlapping *ccdB* 5' half of the gene and tev site1. Sequential PCR was used to amplify the 151 nt and reconstruct the tev site1 and insert ATG as codon 1 in *ccdB* gene. The amplicon also had 15 bases-overlap with the digested plasmid. The amplicon was gel purified, recombined into the plasmids transformed/selected and verified as before. Figure 5-9A showed the results of the sequential PCR to create the modified 151 nt.

5.6 FUNCTIONAL ANALYSIS OF TARGET PLASMIDS

Target plasmids were checked for expression of TEVsh and tig144TEVsh proteases by Western blotting. *TEVsh* gene carried a C terminal V5 epitope tag and we used antiV5 antibodies for detection. Samples were grown overnight under inducing conditions (IPTG 1mM/Arabinose 0.2%). Cells were collected by centrifugation and solubilized for gel electrophoresis. Electrophoresis and Western blotting were carried out according to manufacture's instructions (NuPage, Invitrogen). The membrane was blocked overnight with nonfat dry milk and developed according to manufacturer's instructions. The primary antibody used was rabbit antiV5 antibody (Abgent, CA) and the secondary antibody was HRP conjugated goat anti rabbit (Abcam, MA). Bands were visualized by ECL reagent (Abcam, MA). Two bands were detected (Figure 5-9C). Lanes 1 and 2 represented TEVsh (33.44 kDa) and tig144TEVsh (49.25 kDa) respectively (the brackets represent the calculated molecular weights [MW]). The Western blot showed significant agreement between the calculated MWs and the observed ones.

To verify the expression and processing of the fusion protein, the four target plasmids were compared to a plasmid that contained the fusion gene but did not contain the *TEVsh* protease gene. The plasmids were digested with *StuI*. PCR amplified and gel purified *hctA* leader sequence was recombined into the linearized plasmids; (Section 5.3.3 above). Recombined plasmids were transformed into *ccdB* resistant cells and selected on Chloramphenicol (12.5µg/ml) LB agar plates. Positive clones were grown overnight under inducing conditions (IPTG 1mM, Arabinose 0.2%) in LB broth. Cells were collected by centrifugation and used for electrophoresis. Electrophoresis and Western blotting were carried out according to manufacture's instructions (NuPage, Invitrogen). The membrane was blocked overnight with nonfat dry milk and developed according to manufacture's instructions. The primary antibody used was rabbit anti-CAT (SIGMA, MS) and the secondary antibody was HRP conjugated goat anti-rabbit (Abcam, MA). Immunoreactive proteins were visualized using an ECL procedure (Abcam, MA). Target plasmids that did not express TEVsh protease showed the full length fusion protein. The apparent size of the fusion protein agreed with the calculated size (~41.4 vs 43.5kDa respectively, Figure 5-9D upper panel lane 1). The four target plasmids (Figure 5-6) showed complete processing of the fusion protein. The complete processing of the fusion protein was evident by the absence of the full length fusion-protein (upper panel Figure 5-9D, lanes 2-5). Free CAT protein accumulated after digestion (lower panel Figure 5-9D lanes 2-5) and its apparent size was in agreement with the calculated size (>20.7 vs 25.49kDa). Longer exposure time (data not shown) showed detectable

amounts of the unprocessed fusion protein in lanes 2 and 4 representing plasmids that carry *TEVsh*/pSC101*, *TEVsh*/p15A.

Plasmids (*TEVsh*/p15A, *tig144TEVsh*/p15A) were selected for further analysis. Both plasmids were transformed into BL21A1 cells (Invitrogen, CA). Transformed Bl21AI cells were grown overnight in LB under inducing (IPTG 1mM, Arabinose 0.2%) and non-inducing conditions. Figure 5-9E showed the results of a 24 hour culture. Induced BL21 cells did not grow while uninduced cells showed growth. These results confirmed that when TEVsh protein expression was induced the fusion protein was processed and free CcdB protein accumulated. Free CcdB formation led to inactivation of GyrA and *E. coli* growth was inhibited.

5.7 DISCUSSION

5.7.1 THE USE OF ccdB AS A REPORTER GENE

Initially we planned to use *LacZ*/*CAT* translational fusions to screen the chlamydial library for ncRNA targets. Several investigators have used *lacZ* translational fusions to analyze ncRNA/target interactions (Baker *et al*, 2007; De Lay and Gottesman, 2009; Mandin and Gottesman, 2009; Pulvermacher *et al*, 2009). We did a pilot experiment where we inserted test sequence carrying translational signals upstream of the *LacZ* fusion gene. The variability of the blue color (generated by X-gal) was significant (data not shown). The fact that the main phenotype to determine ncRNA/target interactions was variable for the same insert led us to switch to an alternative plan, LacZ would however be used for quantification of potential interactions (β-galactosidase assay). *ccdB* on the other hand, was used before in a limited number of applications. Insertional inactivation of the gene was used to select clones with DNA inserts in the gene (Gabant *et al*, 1997; Van Reeth *et al*, 1998). It was also used to screen for mutations or inhibitors that interfere with protein splicing mediated by the RecA intein of *Mycobacterium tuberculosis*. This screening procedure involved activation of the CcdB protein by protein splicing, such that host cells survive in the presence of inducer only when protein splicing was blocked (Lew and Paulus, 2002). CcdB is also used as a general selection marker in the gateway system (Invitrogen, CA). Clones that fail to undergo recombination and exchange the *ccdB* cassette with the insert do not survive CcdB mediated killing. CcdB produced a criterion phenotype that is clear cut *i.e.* cell death. A great deal is known about its function and it's crystal structure has been published (Dao-Thi *et al*, 2004; Dao-Thi *et al*, 1998; Loris *et al*, 1999). Moreover there is an *E. coli* strain that is resistant to CcdB mediated killing (Bernard *et al*, 1992).

5.7.2 CcdB LOSS OF FUNCTION AND REDSIGNING OF THE TARGET PLASMID

ccdB fusions are not commonly used in molecular biology. The single fusion that was reported was a *lacZ* N-terminal fusion to *ccdB* (Bernard, 1996; Gabant *et al*, 1997). Personal communications with Paulus H (Lew *et al*, 2002) (Personal communications, 2009) revealed that CcdB lost its lethal effect when 25 or more amino acid residues were fused to the N terminus. In this work we have presented the results of C terminal fusions. We have tried N terminal fusions (data not shown), but this construct was very toxic and we were not able to maintain an intact plasmid within the *ccdB* resistant cells. The plasmid instability forced us to use C terminal fusions. CcdB in C-terminal fusion lost its lethal effect. We shifted our plan to insert protease cleavage sites flanking *ccdB* so that we regenerate a CcdB protein molecule similar to wild-type.

The loss of a lethal effect with CcdB-CAT fusions led us to propose that the C terminal fusion to CcdB causes loss of activity. Bahassi *et al.* studied CcdB loss of function mutants and found seven mutations in CcdB 3' end that had no lethal effects. The mutations were all in the last 3 amino acids of the molecule (Bahassi *et al*, 1995). This finding suggested that the C terminal domain of the molecule was important for function and that C terminal fusions might not function properly. Dao-Thi *et al.* analyzed the crystal structure of CcdB bound to GyrA peptide (Dao-Thi *et al*, 2005) and showed that CcdB W99 (CcdB has 102 amino acids total) interacted with GyrA462R. Mutations in either of these two interacting amino acids led to loss of function for CcdB. Finally, *ccdB* resistant cells have an identical mutation in GyrA ($gyrA_{R462C}$).

In light of the previous argument, we reasoned that any deviation from wild-type CcdB might be detrimental to its function. Dao-Thi *et al.* shown that CcdB functions as a dimer. For dimerization to occur hydrophobic interactions between 6 methionines was required (Dao-Thi *et al*, 2005). Although the first methionine was not included in the interaction, we chose to re-insert M1. After TEV protease cleavage, the only residue that was left on the N terminus was S (or G). We chose to mutate S to G to eliminate a potential source of increased polarity in the region.

We designed four target plasmids. We reasoned that screening of chlamydial libraries will result in cloning many chlamydial translational signals with wide range of activities. The target plasmids therefore were designed with different copy number controlling regions to cover the spectrum of translational strength. Two different origins of replication were used. We also used two TEV protease constructs with different proteolytic efficacy that allowed us to control the amount of available CcdB protein within a cell. Although the Western blotting results does not show differences in efficacy of TEV cleavage (Figure 5-9D, lanes 2-5), over exposure showed traces of undigested fusion protein in clones carrying *TEVsh* (lane 2, 4). Total digestion of the fusion protein was evident for clones having *tig144TEVsh* (lane 3, 5). This result is in complete agreement with Henrichs *et al*, 2005.

CHAPTER 6. CONCLUDING REMARKS

C. trachomatis D/UW-3/Cx expresses numerous ncRNAs that were initially detected using an intergenic microarray and validated by Northern blotting analyses. These include the subset of potential ncRNAs in this study which are potential *cis*-acting (antisense), *trans*-acting, and riboswitches. We have shown that these ncRNAs have distinct temporal expression patterns during the normal developmental cycle and several appear to be processed. We have also shown that the expression of many ncRNAs was altered during growth conditions that induce persistent growth, particularly IFN-γ-induced persistence and carbenicillin-induced-persistence. We have mapped the ends of these ncRNAs, a prerequisite for cloning and functional analysis of these molecules. We have developed a system for testing *cis*-acting ncRNAs, and successfully used it for verification of the proposed function of *CTIG270*. We have also developed a screening system for *trans*-acting ncRNAs: a system based on positive/negative selection that is fully functional in *E. coli.* We believe that this screening tool can also be used to screen for ncRNA targets in other organisms.

Although these studies in *C. trachomatis* are primarily associated with the discovery and validation phase for ncRNAs, they provide a strong basis for mechanistic studies that may determine important regulatory roles for chlamydial ncRNAs in the development cycle and the changes that occur during persistent growth of *C. trachomatis*.

LIST OF REFERENCES

Akers JC, Tan M (2006) Molecular mechanism of tryptophan-dependent transcriptional regulation in *Chlamydia trachomatis*. *J Bacteriol* 188: 4236-4243.

Altuvia S, Weinstein-Fischer D, Zhang A, Postow L, Storz G (1997) A small stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* 90: 43.

Andreeva AE, Borissova BE, Mironova R, Glykos NM, Kotsifaki D, Ivanov I, Krysteva M, Kokkinidis M (2000) Crystallization of type I chloramphenicol acetyltransferase: an approach based on the concept of ionic strength reducers. *Acta Crystallogr D Biol Crystallogr* 56: 101-103.

Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EGH, Margalit H, Altuvia S (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Current Biology* 11: 941-950.

Babu M, Butl G, Pogoutse O, Li J, Greenblatt JF, Emili A (2009) Sequential peptide affinity purification system for the systematic isolation and identification of protein complexes from *Escherichia coli*. *Methods Mol Biol* 564: 373-400.

Bahassi EM, O'Dea MH, Allali N, Messens J, Gellert M, Couturier M (1999) Interactions of CcdB with DNA gyrase. Inactivation of Gyra, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. *J Biol Chem* 274: 10936-10944.

Bahassi EM, Salmon MA, Van Melderen L, Bernard P, Couturier M (1995) F plasmid CcdB killer protein: *ccdB* gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Mol Microbiol* 15: 1031-1037.

Baker CS, Eory LA, Yakhnin H, Mercante J, Romeo T, Babitzke P (2007) *CsrA* inhibits translation initiation of *Escherichia coli* hfq by binding to a single site overlapping the Shine-Dalgarno sequence. *J Bacteriol* 189: 5472-5481.

Balin BJ, Gerard HC, Arking EJ, Appelt DM, Branigan PJ, Abrams JT, Whittum-Hudson JA, Hudson AP (1998) Identification and localization of *Chlamydia pneumoniae* in the Alzheimer's brain. *Med Microbiol Immunol (Berl)* 187: 23-42.

Barbour AG, Amano K, Hackstadt T, Perry L, Caldwell HD (1982) *Chlamydia trachomatis* has penicillin-binding proteins but not detectable muramic acid. *J Bacteriol* 151: 420-428.

Barne KA, Bown JA, Busby SJ, Minchin SD (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J* 16: 4034-4040.

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.

Barry CE, 3rd, Hayes SF, Hackstadt T (1992) Nucleoid condensation in *Escherichia coli* that express a chlamydial histone homolog. *Science* 256: 377-379.

Beatty WL, Belanger TA, Desai AA, Morrison RP, Byrne GI (1994) Tryptophan depletion as a mechanism of gamma interferon-mediated chlamydial persistence. *Infect Immun* 62: 3705-3711.

Beatty WL, Byrne GI, Morrison RP (1993) Morphologic and antigenic characterization of interferon gamma-mediated persistent *Chlamydia trachomatis* infection *in vitro*. *Proc Natl Acad Sci U S A* 90: 3998-4002.

Belland RJ, Nelson DE, Virok D, Crane DD, Hogan D, Sturdevant D, Beatty WL, Caldwell HD (2003a) Transcriptome analysis of chlamydial growth during IFN-gamma-mediated persistence and reactivation. *Proc Natl Acad Sci U S A* 100: 15971-15976.

Belland RJ, Zhong G, Crane DD, Hogan D, Sturdevant D, Sharma J, Beatty WL, Caldwell HD (2003b) Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proc Natl Acad Sci U S A* 100: 8478-8483.

Bernard P (1996) Positive selection of recombinant DNA by CcdB. *Biotechniques* 21: 320-323.

Bernard P, Couturier M (1992) Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* 226: 735-745.

Bernard P, Kezdy KE, Van Melderen L, Steyaert J, Wyns L, Pato ML, Higgins PN, Couturier M (1993) The F plasmid CcdB protein induces efficient ATP-dependent DNA cleavage by gyrase. *J Mol Biol* 234: 534-541.

Birkelund S, Lundemose AG, Christiansen G (1989) Characterization of native and recombinant 75-kilodalton immunogens from *Chlamydia trachomatis* serovar L2. *Infect Immun* 57: 2683-2690.

Botta GA, Park JT (1981) Evidence for involvement of penicillin-binding protein 3 in murein synthesis during septation but not during cell elongation. *J Bacteriol* 145: 333-340.

Bower S, Perkins JB, Rogers Yocum R, Howitt CL, Rahaim P, Pero J (1996) Cloning, sequencing, and characterization of the *Bacillus subtilis* biotin biosynthetic operon. *J Bacteriology* 178: 4122-4130.

Brennan RG, Vasu S, Matthews BW, Otsuka AJ (1989) Crystallization of the bifunctional biotin operon repressor. *J Biol Chem* 264: 5.

Brickman TJ, Barry CE, 3rd, Hackstadt T (1993) Molecular cloning and expression of *hctB* encoding a strain-variant chlamydial histone-like protein with DNA-binding activity. *J Bacteriol* 175: 4274-4281.

Brown WJ, Rockey DD (2000) Identification of an antigen localized to an apparent septum within dividing *chlamydiae*. *Infect Immun* 68: 708-715.

Browning DF, Busby SJ (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* 2: 57-65.

Buoncristiani MR, Howard PK, Otsuka AJ (1986) DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (*BirA*) of *Escherichia coli*. *Gene* 44: 255-261.

Burgess RR, Travers AA, Dunn JJ, Bautz EK (1969) Factor stimulating transcription by RNA polymerase. *Nature* 221: 43-46.

Burr T, Mitchell J, Kolb A, Minchin S, Busby S (2000) DNA sequence elements located immediately upstream of the -10 hexamer in *Escherichia coli* promoters: a systematic study. *Nucl Acids Res* 28: 1864-1870.

Byrne GI, Lehmann LK, Landry GJ (1986) Induction of tryptophan catabolism is the mechanism for gamma-interferon-mediated inhibition of intracellular *Chlamydia psittaci* replication in T24 cells. *Infect Immun* 53: 347-351.

Campbell EA, Muzzin O, Chlenov M, Sun JL, Olson CA, Weinman O, Trester-Zedlitz ML, Darst SA (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol Cell* 9: 527-539.

Campbell EA, Westblade LF, Darst SA (2008) Regulation of bacterial RNA polymerase [sigma] factor activity: a structural perspective. *Curr Opin Microbiol* 11: 121-127.

Carabeo RA, Hackstadt T (2001) Isolation and characterization of a mutant Chinese hamster ovary cell line that is resistant to *Chlamydia trachomatis* infection at a novel step in the attachment process. *Infect Immun* 69: 5899-5904.

Carpousis AJ, Khemici V, Ait-Bara S, Poljak L (2008) Co-immunopurification of multiprotein complexes containing RNA-degrading enzymes. *Methods Enzymol* 447: 65-82.

Carpousis AJ, Vanzo NF, Raynal LC (1999) mRNA degradation: a tale of poly(A) and multiprotein machines. *Trends in Genetics* 15: 24-28.

Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. *Science* 301: 336.

Carrington JC, Dougherty WG (1987) Small nuclear inclusion protein encoded by a plant potyvirus genome is a protease. *J Virol* 61: 2540-2548.

Carrington JC, Dougherty WG (1988) A viral cleavage site cassette: identification of amino acid sequences required for tobacco etch virus polyprotein processing. *Proc Natl Acad Sci U S A* 85: 3391-3395.

Chopra I, Storey C, Falla TJ, Pearce JH (1998) Antibiotics, peptidoglycan synthesis and genomics: the chlamydial anomaly revisited. *Microbiology* 144 (Pt 10): 2673-2678.

Clifton DR, Fields KA, Grieshaber SS, Dooley CA, Fischer ER, Mead DJ, Carabeo RA, Hackstadt T (2004) A chlamydial type III translocated protein is tyrosine-phosphorylated at the site of entry and associated with recruitment of actin. *Proc Natl Acad Sci U S A* 101: 10166-10171.

Coles AM, Reynolds DJ, Harper A, Devitt A, Pearce JH (1993) Low-nutrient induction of abnormal chlamydial development: a novel component of chlamydial pathogenesis? *FEMS Microbiol Lett* 106: 193-200.

Crenshaw RW, Fahr MJ, Wichlan DG, Hatch TP (1990) Developmental cycle-specific host-free RNA synthesis in *Chlamydia* spp. *Infect Immun* 58: 3194-3201.

Critchlow SE, O'Dea MH, Howells AJ, Couturier M, Gellert M, Maxwell A (1997) The interaction of the F plasmid killer protein, CcdB, with DNA gyrase: induction of DNA cleavage and blocking of transcription. *J Mol Biol* 273: 826-839.

Dao-Thi MH, Van Melderen L, De Genst E, Afif H, Buts L, Wyns L, Loris R (2005) Molecular basis of gyrase poisoning by the addiction toxin CcdB. *J Mol Biol* 348: 1091-1102.

Dao-Thi MH, Van Melderen L, De Genst E, Buts L, Ranquin A, Wyns L, Loris R (2004) Crystallization of CcdB in complex with a GyrA fragment. *Acta Crystallogr D Biol Crystallogr* 60: 1132-1134.

Dao-Thi MH, Wyns L, Poortmans F, Bahassi EM, Couturier M, Loris R (1998) Crystallization of *ccdB*. *Acta Crystallogr D Biol Crystallogr* 54: 975-981.

Davis BM, Waldor MK (2007) *RNase E*-dependent processing stabilizes *MicX*, a *Vibrio cholerae* sRNA. *Mol Microbiol* 65: 373-385.

De Lay N, Gottesman S (2009) The Crp-activated small noncoding regulatory RNA *CyaR* (*RyeE*) links nutritional status to group behavior. *J Bacteriol* 191: 461-476.

Delihas N (1995) Regulation of gene expression by *trans*-encoded antisense RNAs. *Mol Microbiol* 15: 411.

Dill BD, Dessus-Babus S, Raulston JE (2009) Identification of iron-responsive proteins expressed by *Chlamydia trachomatis* reticulate bodies during intracellular growth. *Microbiology* 155: 210-219.

Dougherty WG, Cary SM, Parks TD (1989) Molecular genetic analysis of a plant virus polyprotein cleavage site: a model. *Virology* 171: 356-364.

Douglas AL, Hatch TP (1995) Functional analysis of the major outer membrane protein gene promoters of *Chlamydia trachomatis*. *J Bacteriol* 177: 6286-6289.

Douglas AL, Hatch TP (1996) Mutagenesis of the P2 promoter of the major outer membrane protein gene of *Chlamydia trachomatis*. *J Bacteriol* 178: 5573-5578.

Douglas AL, Hatch TP (2000) Expression of the transcripts of the sigma factors and putative sigma factor regulators of *Chlamydia trachomatis* L2. *Gene* 247: 209-214.

Dunman PM, Murphy E, Haney S, Palacios D, Tucker-Kellogg G, Wu S, Brown EL, Zagursky RJ, Shlaes D, Projan SJ (2001) Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *J Bacteriol* 183: 7341-7353.

Eb F, Orfila J, Lefebvre JF (1976) Ultrastructural study of the development of the agent of ewe's abortion. *J Ultrastruct Res* 56: 177-185.

Engel JN, Ganem D (1987) Chlamydial rRNA operons: gene organization and identification of putative tandem promoters. *J Bacteriol* 169: 5678-5685.

Everett KD, Bush RM, Andersen AA (1999) Emended description of the order Chlamydiales, proposal of *Parachlamydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new species, and standards for the identification of organisms. *Int J Syst Bacteriol* 49 Pt 2: 415-440.

Fahr MJ, Douglas AL, Xia W, Hatch TP (1995) Characterization of late gene promoters of *Chlamydia trachomatis*. *J Bacteriol* 177: 4252-4260.

Fahr MJ, Sriprakash KS, Hatch TP (1992) Convergent and overlapping transcripts of the *Chlamydia trachomatis* 7.5-kb plasmid. *Plasmid* 28: 247-257.

Faubladier M, Bouche JP (1994) Division inhibition gene *dicF* of *Escherichia coli* reveals a widespread group of prophage sequences in bacterial genomes. *J Bacteriol* 176: 1150-1156.

Fields KA, Hackstadt T (2000) Evidence for the secretion of *Chlamydia trachomatis CopN* by a type III secretion mechanism. *Mol Microbiol* 38: 1048-1060.

Fields KA, Mead DJ, Dooley CA, Hackstadt T (2003) *Chlamydia trachomatis* type III secretion: evidence for a functional apparatus during early-cycle development. *Mol Microbiol* 48: 671-683.

Fish RN, Kane CM (2002) Promoting elongation with transcript cleavage stimulatory factors. *Biochim Biophys Acta* 1577: 287-307.

Flint DH, Allen RM, Donald B. McCormick JWSaCW (1997) [37] Purification and characterization of biotin synthases. In *Methods in Enzymology*, Vol. Volume 279, pp 349-356: Academic Press.

Forsberg G, Baastrup B, Rondahl H, Holmgren E, Pohl G, Hartmanis M, Lake M (1992) An evaluation of different enzymatic cleavage methods for recombinant fusion proteins, applied on des(1-3)insulin-like growth factor I. *J Protein Chem* 11: 201-211.

Forsberg G, Brobjer M, Holmgren E, Bergdahl K, Persson P, Gautvik KM, Hartmanis M (1991) Thrombin and H64A subtilisin cleavage of fusion proteins for preparation of human recombinant parathyroid hormone. *J Protein Chem* 10: 517-526.

Gabant P, Dreze PL, Van Reeth T, Szpirer J, Szpirer C (1997) Bifunctional *lacZ* alpha-*ccdB* genes for selective cloning of PCR products. *Biotechniques* 23: 938-941.

Gerbase AC, Rowley JT, Mertens TE (1998) Global epidemiology of sexually transmitted diseases. *Lancet* 351 Suppl 3: 2-4.

Ghuysen JM, Goffin C (1999) Lack of cell wall peptidoglycan versus penicillin sensitivity: new insights into the chlamydial anomaly. *Antimicrob Agents Chemother* 43: 2339-2344.

Gloeckler R, Ohsawa I, Speck D, Ledoux C, Bernard S, Zinsius M, Villeval D, Kisou T, Kamogawa K, Lemoine Y (1990) Cloning and characterization of the *Bacillus sphaericus* genes controlling the bioconversion of pimelate into dethiobiotin. *Gene* 87: 63-70.

Gongadze GM, Korepanov AP, Korobeinikova AV, Garber MB (2008) Bacterial 5S rRNA-binding proteins of the CTC family. *Biochemistry (Mosc)* 73: 1405-1417.

Gorke B, Vogel J (2008) Noncoding RNA control of the making and breaking of sugars. *Genes Dev* 22: 2914-2925.

Gottesman S (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. *Trends Genet* 21: 399-404.

Gottesman S, Storz G, Rosenow C, Majdalani N, Repoila F, Wassarman KM (2001) Small RNA regulators of translation: mechanisms of action and approaches for identifying new small RNAs. *Cold Spring Harb Symp Quant Biol* 66: 353-362.

Gourse RL, Ross W, Gaal T (2000) UPs and downs in bacterial transcription initiation: the role of the alpha subunit of RNA polymerase in promoter recognition. *Mol Microbiol* 37: 687-695.

Grange T (2008) Sensitive detection of mRNA decay products by use of reverse-ligation-mediated PCR (RL-PCR). *Methods Enzymol* 448: 445-466.

Gravel RA, Narang MA (2005) Molecular genetics of biotin metabolism: old vitamin, new science. *J Nutr Biochem* 16: 428-431.

Grech B, Maetschke S, Mathews S, Timms P (2007) Genome-wide analysis of *chlamydiae* for promoters that phylogenetically footprint. *Res Microbiol* 158: 685-693.

Grieshaber NA, Fischer ER, Mead DJ, Dooley CA, Hackstadt T (2004) Chlamydial histone-DNA interactions are disrupted by a metabolite in the methylerythritol phosphate pathway of isoprenoid biosynthesis. *Proc Natl Acad Sci U S A* 101: 7451-7456.

Grieshaber NA, Grieshaber SS, Fischer ER, Hackstadt T (2006) A small RNA inhibits translation of the histone-like protein Hc1 in *Chlamydia trachomatis*. *Mol Microbiol* 59: 541-550.

Grunberg-Manago M, Oritz PJ, Ochoa S (1955) Enzymatic synthesis of nucleic acidlike polynucleotides. *Science* 122: 907-910.

Guillen-Navarro K, Araiza G, Garcia-de los Santos A, Mora Y, Dunn MF (2005) The *Rhizobium etli* bioMNY operon is involved in biotin transport. *FEMS Microbiol Lett* 250: 209-219.

Hahn DL, Dodge RW, Golubjatnikov R (1991) Association of *Chlamydia pneumoniae* (strain TWAR) infection with wheezing, asthmatic bronchitis, and adult-onset asthma. *JAMA* 266: 225-230.

Hatch TP (1999) Developmental biology. In *Chlamydia: Intracellular Biology, Pathogenesis, and Immunity*, Stephens RS (ed), pp 29–67. Washington, DC: ASM Press.

Haugen SP, Ross W, Manrique M, Gourse RL (2008) Fine structure of the promoter-sigma region 1.2 interaction. *Proc Natl Acad Sci U S A* 105: 3292-3297.

He M, Jin L, Austen B (1993) Specificity of factor Xa in the cleavage of fusion proteins. *J Protein Chem* 12: 1-5.

Hebbeln P, Rodionov DA, Alfandega A, Eitinger T (2007) Biotin uptake in prokaryotes by solute transporters with an optional ATP-binding cassette-containing module. *Proc Natl Acad Sci U S A* 104: 2909-2914.

Helmann JD (1999) Anti-sigma factors. *Curr Opin Microbiol* 2: 135-141.

Henkin TM (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. *Genes Dev* 22: 3383-3390.

Henrichs T, Mikhaleva N, Conz C, Deuerling E, Boyd D, Zelazny A, Bibi E, Ban N, Ehrmann M (2005) Target-directed proteolysis at the ribosome. *Proc Natl Acad Sci U S A* 102: 4246-4251.

Hesterkamp T, Deuerling E, Bukau B (1997) The amino-terminal 118 amino acids of *Escherichia coli* trigger factor constitute a domain that is necessary and sufficient for binding to ribosomes. *J Biol Chem* 272: 21865-21871.

Hogan RJ, Mathews SA, Mukhopadhyay S, Summersgill JT, Timms P (2004) Chlamydial persistence: beyond the biphasic paradigm. *Infect Immun* 72: 1843-1855.

Hsia R, Ohayon H, Gounon P, Dautry-Varsat A, Bavoil PM (2000) Phage infection of the obligate intracellular bacterium, *Chlamydia psittaci* strain guinea pig inclusion conjunctivitis. *Microbes Infect* 2: 761-772.

Hu Z, Zhang A, Storz G, Gottesman S, Leppla SH (2006) An antibody-based microarray assay for small RNA detection. *Nucl Acids Res* 34: e52.

Hughes KT, Mathee K (1998) The anti-sigma factors. *Annual Review of Microbiology* 52: 231-286.

Huttenhofer A, Vogel J (2006) Experimental approaches to identify non-coding RNAs. *Nucl Acids Res* 34: 635-646.

Ikemura T, Dahlberg JE (1973) Small ribonucleic acids of *Escherichia coli*. *J Biol Chem* 248: 5024.

Janzon L, Lofdahl S, Arvidson S (1989) Identification and nucleotide sequence of the delta-lysin gene, *hld*, adjacent to the accessory gene regulator (*agr*) of *Staphylococcus aureus*. *Mol Gen Genet* 219: 480-485.

Jasiecki J, Wegrzyn G (2003) Growth-rate dependent RNA polyadenylation in *Escherichia coli*. *EMBO Rep* 4: 172-177.

Kalamorz F, Reichenbach B, Marz W, Rak B, Gorke B (2007) Feedback control of glucosamine-6-phosphate synthase *GlmS* expression depends on the small RNA *GlmZ* and involves the novel protein *YhbJ* in *Escherichia coli*. *Mol Microbiol* 65: 1518-1533.

Kalman S, Mitchell W, Marathe R, Lammel C, Fan J, Hyman RW, Olinger L, Grimwood J, Davis RW, Stephens RS (1999) Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. *Nat Genet* 21: 385-389.

Kapust RB, Tozser J, Copeland TD, Waugh DS (2002) The P1' specificity of tobacco etch virus protease. *Biochem Biophys Res Commun* 294: 949-955.

Kawano M, Reynolds AA, Miranda-Rios J, Storz G (2005) Detection of 5'- and 3'-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. *Nucl Acids Res* 33: 1040-1050.

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

Kleineidam RG, Pitulle C, Sproat B, Krupp G (1993) Efficient cleavage of pre-tRNAs by *E. coli* RNAse P RNA requires the 2'-hydroxyl of the ribose at the cleavage site. *Nucl Acids Res* 21: 1097-1101.

Koehler JE, Burgess RR, Thompson NE, Stephens RS (1990) *Chlamydia trachomatis* RNA polymerase major sigma subunit. Sequence and structural comparison of conserved and unique regions with *Escherichia coli* sigma 70 and *Bacillus subtilis* sigma 43. *J Biol Chem* 265: 13206-13214.

Koehler L, Nettelnbreker E, Hudson AP, Ott N, Gerard HC, Branigan PJ, Schumacher HR, Drommer W, Zeidler H (1997) Ultrastructural and molecular analyses of the persistence of *Chlamydia trachomatis* (serovar K) in human monocytes. *Microb Pathog* 22: 133-142.

Komine Y, Kitabatake M, Yokogawa T, Nishikawa K, Inokuchi H (1994) A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc Natl Acad Sci USA* 91: 9223.

Komissarova N, Kashlev M (1997) Transcriptional arrest: *Escherichia coli* RNA polymerase translocates backward, leaving the 3' end of the RNA intact and extruded. *Proc Natl Acad Sci USA* 94: 1755-1760.

Koo IC, Stephens RS (2003) A developmentally regulated two-component signal transduction system in *Chlamydia*. *J Biol Chem* 278: 17314-17319.

Koo IC, Walthers D, Hefty PS, Kenney LJ, Stephens RS (2006) *ChxR* is a transcriptional activator in *Chlamydia*. *Proc Natl Acad Sci U S A* 103: 750-755.

Korzheva N, Mustaev A (2001) Transcription elongation complex: structure and function. *Curr Opin Microbiol* 4: 119-125.

Kumar A, Malloch RA, Fujita N, Smillie DA, Ishihama A, Hayward RS (1993) The minus 35-recognition region of *Escherichia coli* sigma 70 is inessential for initiation of transcription at an "extended minus 10" promoter. *J Mol Biol* 232: 406-418.

Kuo CC, Jackson LA, Campbell LA, Grayston JT (1995) *Chlamydia pneumoniae* (TWAR). *Clin Microbiol Rev* 8: 451-461.

Kuo CC, Shor A, Campbell LA, Fukushi H, Patton DL, Grayston JT (1993) Demonstration of *Chlamydia pneumoniae* in atherosclerotic lesions of coronary arteries. *J Infect Dis* 167: 841-849.

Kutlin A, Flegg C, Stenzel D, Reznik T, Roblin PM, Mathews S, Timms P, Hammerschlag MR (2001) Ultrastructural study of *Chlamydia pneumoniae* in a continuous-infection model. *J Clin Microbiol* 39: 3721-3723.

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.

Lew BM, Paulus H (2002) An in vivo screening system against protein splicing useful for the isolation of non-splicing mutants or inhibitors of the *RecA* intein of *Mycobacterium tuberculosis*. *Gene* 282: 169-177.

Li Z, Pandit S, Deutscher MP (1998) Polyadenylation of stable RNA precursors *in vivo*. *Proc Natl Acad Sci U S A* 95: 12158-12162.

Lisser S, Margalit H (1993) Compilation of *E. coli* mRNA promoter sequences. *Nucl Acids Res* 21: 1507-1516.

Livny J, Fogel MA, Davis BM, Waldor MK (2005) sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. *Nucl Acids Res* 33: 4096-4105.

Lohan AJ, Gray MW (2007) Analysis of 5'- or 3'-terminal tRNA editing: mitochondrial 5' tRNA editing in *Acanthamoeba castellanii* as the exemplar. *Methods Enzymol* 424: 223-242.

Loris R, Dao-Thi MH, Bahassi EM, Van Melderen L, Poortmans F, Liddington R, Couturier M, Wyns L (1999) Crystal structure of CcdB, a topoisomerase poison from *E. coli*. *J Mol Biol* 285: 1667-1677.

Luban S, Kihara D (2007) Comparative genomics of small RNAs in bacterial genomes. *Omics* 11: 58-73.

Lucast LJ, Batey RT, Doudna JA (2001) Large-scale purification of a stable form of recombinant tobacco etch virus protease. *BioTechniques* 30: 544-554.

Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl Acids Res* 25: 1203-1210.

Majdalani N, Cunning C, Sledjeski D, Elliott T, Gottesman S (1998) *DsrA* RNA regulates translation of *RpoS* message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci USA* 95: 12462.

Maki S, Takiguchi S, Horiuchi T, Sekimizu K, Miki T (1996) Partner switching mechanisms in inactivation and rejuvenation of *Escherichia coli* DNA gyrase by F plasmid proteins LetD (CcdB) and LetA (CcdA). *J Mol Biol* 256: 473-482.

Mandin P, Gottesman S (2009) A genetic approach for finding small RNAs regulators of genes of interest identifies *RybC* as regulating the *DpiA/DpiB* two-component system. *Mol Microbiol* 72: 551-565.

Manen D, Xia G, Caro L (1994) A locus involved in the regulation of replication in plasmid pSC101. *Mol Microbiol* 11: 875-884.

Mangold M, Siller M, Roppenser B, Vlaminckx BJ, Penfound TA, Klein R, Novak R, Novick RP, Charpentier E (2004) Synthesis of group A streptococcal virulence factors is controlled by a regulatory RNA molecule. *Mol Microbiol* 53: 1515-1527.

Masse E, Escorcia FE, Gottesman S (2003) Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. *Genes Dev* 17: 2374-2383.

Masse E, Vanderpool CK, Gottesman S (2005) Effect of *RyhB* small RNA on global iron use in *Escherichia coli*. *J Bacteriol* 187: 6962-6971.

Mathews SA, Sriprakash KS (1994) The RNA polymerase of *Chlamydia trachomatis* has a flexible sequence requirement at the -10 and -35 boxes of its promoters. *J Bacteriol* 176: 3785-3789.

Mathews SA, Stephens RS (1999a) DNA structure and novel amino and carboxyl termini of the *Chlamydia* sigma 70 analogue modulate promoter recognition. *Microbiology* 145 (Pt 7): 1671-1681.

Mathews SA, Timms P (2000) Identification and mapping of sigma-54 promoters in *Chlamydia trachomatis*. *J Bacteriol* 182: 6239-6242.

Mathews SA, Volp KM, Timms P (1999b) Development of a quantitative gene expression assay for *Chlamydia trachomatis* identified temporal expression of sigma factors. *FEBS Lett* 458: 354-358.

Matsumoto A (1973) Fine structures of cell envelopes of *Chlamydia* organisms as revealed by freeze-etching and negative staining techniques. *J Bacteriol* 116: 1355-1363.

Matsumoto A, Manire GP (1970) Electron microscopic observations on the effects of penicillin on the morphology of *Chlamydia psittaci*. *J Bacteriol* 101: 278-285.

Matsumoto DDRaA (2000) The Chlamydial developmental cycle. In *Prokaryotic Development*, Yves Brun LJS (ed), Y.V. Brun and L.J. Shimkets edn, pp 403–425. Washington, DC ASM Press.

McClure WR, Hawley DK, Youderian P, Susskind MM (1983) DNA determinants of promoter selectivity in *Escherichia coli*. *Cold Spring Harb Symp Quant Biol* 47 Pt 1: 477-481.

McCoy AJ, Sandlin RC, Maurelli AT (2003) *In vitro* and *in vivo* functional activity of *Chlamydia MurA*, a UDP-N-acetylglucosamine enolpyruvyl transferase involved in peptidoglycan synthesis and fosfomycin resistance. *J Bacteriol* 185: 1218-1228.

McManus MT, Sharp PA (2003) Gene silencing in mammals by small interfering RNAs. *Nat Rev Genet* 3: 737.

Meijer AaO, J.M. (2002) Description of wider diversity within the order Chlamydiales than currently classified. In *International Chlamydial Symposium*, J. Schachter J, G. Christiansen, I. N. Clarke, M. R. Hammerschlag, B. Kaltenboeck, C.-C. Kuo, R. G. Rank, G. L. Ridgway, P. Saikku, W. E. Stamm, R. S. Stephens, J. T. Summersgill, P. Timms, and P. B. Wyrick (ed) pp 13–16, San Francisco.

Minakhin L, Severinov K (2003) On the role of the *Escherichia coli* RNA polymerase sigma 70 region 4.2 and alpha-subunit C-terminal domains in promoter complex formation on the extended -10 *galP1* promoter. *J Biol Chem* 278: 29710-29718.

Mitchell JE, Zheng D, Busby SJW, Minchin SD (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucl Acids Res* 31: 4689-4695.

Mohanty BK, Kushner SR (2000) Polynucleotide phosphorylase functions both as a 3' right-arrow 5' exonuclease and a poly(A) polymerase in *Escherichia coli*. *Proc Natl Acad Sci U S A* 97: 11966-11971.

Mohanty BK, Kushner SR (2006) The majority of *Escherichia coli* mRNAs undergo post-transcriptional modification in exponentially growing cells. *Nucl Acids Res* 34: 5695-5704.

Moller T, Franch T, Udesen C, Gerdes K, Valentin-Hansen P (2002) Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. *Genes Dev* 16: 1696-1706.

Mondigler M, Ehrmann M (1996) Site-specific proteolysis of the *Escherichia coli* SecA protein *in vivo*. *J Bacteriol* 178: 2986-2988.

Moos M, Jr., Nguyen NY, Liu TY (1988) Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. *J Biol Chem* 263: 6005-6008.

Morfeldt E, Taylor D, von Gabain A, Arvidson S (1995) Activation of alpha-toxin translation in *Staphylococcus aureus* by the *trans*-encoded antisense RNA, RNAIII. *EMBO J* 14: 4569-4577.

Morita T, Maki K, Aiba H (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* 19: 2176-2186.

Moulder JW (1991) Interaction of *chlamydiae* and host cells *in vitro*. *Microbiol Rev* 55: 143-190.

Moulder JW (1993) Why is *Chlamydia* sensitive to penicillin in the absence of peptidoglycan? *Infect Agents Dis* 2: 87-99.

Murakami KS, Masuda S, Campbell EA, Muzzin O, Darst SA (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science* 296: 1285-1290.

Nanninga N (1991) Cell division and peptidoglycan assembly in *Escherichia coli*. *Mol Microbiol* 5: 791-795.

Narayan KS, Steele WJ, Busch H (1966) Evidence that the granular and fibrillar components of nucleoli contain 28 and 6S RNA, respectively. *Exp Cell Res* 43: 483-492.

Narayanan N, Hsieh MY, Xu Y, Chou CP (2006) Arabinose-induction of lac-derived promoter systems for penicillin acylase production in *Escherichia coli*. *Biotechnol Prog* 22: 617-625.

Nicholson TL, Olinger L, Chong K, Schoolnik G, Stephens RS (2003) Global stage-specific gene regulation during the developmental cycle of *Chlamydia trachomatis*. *J Bacteriol* 185: 3179-3189.

Novick RP, Iordanescu S, Projan SJ, Kornblum J, Edelman I (1989) pT181 plasmid replication is regulated by a countertranscript-driven transcriptional attenuator. *Cell* 59: 395-404.

Novick RP, Jiang D (2003) The staphylococcal saeRS system coordinates environmental signals with agr quorum sensing. *Microbiology* 149: 2709-2717.

Nudler E, Gottesman ME (2002) Transcription termination and anti-termination in *E. coli*. *Genes Cells* 7: 755-768.

Nudler E, Mustaev A, Goldfarb A, Lukhtanov E (1997) The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase. *Cell* 89: 33-41.

Ochiai Y, Fukushi H, Cai Y, Yamaguchi T, Hirai K (1999) Conservation of putative promoter sequences located upstream of chlamydial major sigma factor gene, *sigA* among *Chlamydia* spp. *Microbiol Immunol* 43: 419-424.

Opdyke JA, Kang JG, Storz G (2004) GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J Bacteriol* 186: 6698-6705.

Orsini G, Ouhammouch M, Le Caer JP, Brody EN (1993) The asiA gene of bacteriophage T4 codes for the anti-sigma 70 protein. *J Bacteriol* 175: 85-93.

Ouellette SP, Hatch TP, AbdelRahman YM, Rose LA, Belland RJ, Byrne GI (2006) Global transcriptional upregulation in the absence of increased translation in *Chlamydia* during IFNgamma-mediated host cell tryptophan starvation. *Mol Microbiol* 62: 1387-1401.

Paget MS, Helmann JD (2003) The sigma70 family of sigma factors. *Genome Biol* 4: 203.

Pai CH (1973) Biotin uptake in biotin regulatory mutant of *Escherichia coli*. *J Bacteriol* 116: 494-496.

Pai CH, Lichstein HC (1965) The biosynthesis of biotin in microorganisms. I. The physiology of biotin synthesis in *Escherichia coli*. *Biochim Biophys Acta* 100: 28-35.

Parks TD, Leuther KK, Howard ED, Johnston SA, Dougherty WG (1994) Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal Biochem* 216: 413-417.

Paschal BM, McReynolds LA, Noren CJ, Nichols NM (2008) RNA polymerases. *Curr Protoc Mol Biol* Chapter 3: Unit3 8.

Phue JN, Lee SJ, Trinh L, Shiloach J (2008) Modified *Escherichia coli* B (BL21), a superior producer of plasmid DNA compared with *Escherichia coli* K (DH5alpha). *Biotechnol Bioeng* 101: 831-836.

Pichon C, Felden B (2005) Small RNA genes expressed from *Staphylococcus aureus* genomic and pathogenicity islands with specific expression among pathogenic strains. *Proc Natl Acad Sci U S A* 102: 14249-14254.

Plaunt MR, Hatch TP (1988) Protein synthesis early in the developmental cycle of *Chlamydia psittaci*. *Infect Immun* 56: 3021-3025.

Poritz MA, Bernstein HD, Strub K, Zopf D, Wilhelm H, Walter P (1990) An *E. coli* ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. *Science* 250: 1111-1117.

Pulvermacher SC, Stauffer LT, Stauffer GV (2009) Role of the sRNA *GcvB* in regulation of *cycA* in *Escherichia coli*. *Microbiology* 155: 106-114.

Rachel BK, David SW (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science* 8: 1668-1674.

Rao X, Deighan P, Hua Z, Hu X, Wang J, Luo M, Wang J, Liang Y, Zhong G, Hochschild A, Shen L (2009) A regulator from *Chlamydia trachomatis* modulates the activity of RNA polymerase through direct interaction with the beta subunit and the primary sigma subunit. *Genes Dev* 23: 1818-1829.

Rasmussen AA, Eriksen M, Gilany K, Udesen C, Franch T, Petersen C, Valentin-Hansen P (2005) Regulation of ompA mRNA stability: the role of a small regulatory RNA in growth phase-dependent control. *Mol Microbiol* 58: 1421-1429.

Raulston JE (1997) Response of *Chlamydia trachomatis* serovar E to iron restriction in vitro and evidence for iron-regulated chlamydial proteins. *Infect Immun* 65: 4539-4547.

Read TD, Brunham RC, Shen C, Gill SR, Heidelberg JF, White O, Hickey EK, Peterson J, Utterback T, Berry K, Bass S, Linher K, Weidman J, Khouri H, Craven B, Bowman C, Dodson R, Gwinn M, Nelson W, DeBoy R, Kolonay J, McClarty G, Salzberg SL, Eisen J, Fraser CM (2000) Genome sequences of *Chlamydia trachomatis MoPn* and *Chlamydia pneumoniae* AR39. *Nucl Acids Res* 28: 1397-1406.

Read TD, Myers GS, Brunham RC, Nelson WC, Paulsen IT, Heidelberg J, Holtzapple E, Khouri H, Federova NB, Carty HA, Umayam LA, Haft DH, Peterson J, Beanan MJ, White O, Salzberg SL, Hsia RC, McClarty G, Rank RG, Bavoil PM, Fraser CM (2003) Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): examining the role of niche-specific genes in the evolution of the *Chlamydiaceae*. *Nucl Acids Res* 31: 2134-2147.

Reichenbach B, Maes A, Kalamorz F, Hajnsdorf E, Gorke B (2008) The small RNA *GlmY* acts upstream of the sRNA *GlmZ* in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucl Acids Res* 36: 2570-2580.

Repoila F, Gottesman S (2001) Signal transduction cascade for regulation of *RpoS*: temperature regulation of *DsrA*. *J Bacteriol* 183: 4012-4023.

Ribes V, Romisch K, Giner A, Dobberstein B, Tollervey D (1990) *E. coli* 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. *Cell* 63: 591-600.

Ricci S, Cevenini R, Cosco E, Comanducci M, Ratti G, Scarlato V (1993) Transcriptional analysis of the *Chlamydia trachomatis* plasmid pCT identifies temporally regulated transcripts, anti-sense RNA and sigma 70-selected promoters. *Mol Gen Genet* 237: 318-326.

Richardson J, and Greenblatt, J (1996) RNA chain elongation and termination. In *Escherichia coli and Salmonella : cellular and molecular biology*, Curtiss FCNR (ed), pp 822–848. Washington, DC: ASM Press.

Rivas E, Eddy SR (2001a) Noncoding RNA gene detection using comparative sequence analysis. *BMC Bioinformatics* 2: 8.

Rivas E, Klein RJ, Jones TA, Eddy SR (2001b) Computational identification of noncoding RNAs in *E. coli* by comparative genomics. *Curr Biol* 11: 1369-1373.

Robertson HD, Altman S, Smith JD (1972) Purification and properties of a specific *Escherichia coli* ribonuclease which cleaves a tyrosine transfer ribonucleic acid presursor. *J Biol Chem* 247: 5243-5251.

Rodionov DA, Mironov AA, Gelfand MS (2002) Conservation of the biotin regulon and the *BirA* regulatory signal in Eubacteria and Archaea. *Genome Res* 12: 1507-1516.

Ross W, Gourse RL (2009) Analysis of RNA polymerase-promoter complex formation. *Methods* 47: 13-24.

Saikku P, Leinonen M, Mattila K, Ekman MR, Nieminen MS, Makela PH, Huttunen JK, Valtonen V (1988) Serological evidence of an association of a novel *Chlamydia*, TWAR, with chronic coronary heart disease and acute myocardial infarction. *Lancet* 2: 983-986.

Schachter J (1978) Chlamydial infections (second of three parts). *N Engl J Med* 298: 490-495.

Schachter J, Stephens RS, Timms P, Kuo C, Bavoil PM, Birkelund S, Boman J, Caldwell H, Campbell LA, Chernesky M, Christiansen G, Clarke IN, Gaydos C, Grayston JT, Hackstadt T, Hsia R, Kaltenboeck B, Leinonnen M, Ojcius D, McClarty G, Orfila J, Peeling R, Puolakkainen M, Quinn TC, Rank RG, Raulston J, Ridgeway GL, Saikku P, Stamm WE, Taylor-Robinson DT, Wang SP, Wyrick PB (2001) Radical changes to chlamydial taxonomy are not necessary just yet. *Int J Syst Evol Microbiol* 51: 249; author reply 251-243.

Schaumburg CS, Tan M (2000) A positive *cis*-acting DNA element is required for high-level transcription in *Chlamydia*. *J Bacteriol* 182: 5167-5171.

Schaumburg CS, Tan M (2003) Mutational analysis of the *Chlamydia trachomatis* dnaK promoter defines the optimal -35 promoter element. *Nucl Acids Res* 31: 551-555.

Schaumburg CS, Tan M (2006) Arginine-dependent gene regulation via the *ArgR* repressor is species specific in *Chlamydia*. *J Bacteriol* 188: 919-927.

Scidmore MA, Rockey DD, Fischer ER, Heinzen RA, Hackstadt T (1996) Vesicular interactions of the *Chlamydia trachomatis* inclusion are determined by chlamydial early protein synthesis rather than route of entry. *Infect Immun* 64: 5366-5372.

Selinger DW, Cheung KJ, Mei R, Johansson EM, Richmond CS, Blattner FR, Lockhart DJ, Church GM (2000) RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nat Biotechnol* 18: 1262-1268.

Sergei B, Jookyung L, Oleg L (2005) Bacterial transcription elongation factors: new insights into molecular mechanism of action. *Molecular Microbiology* 55: 1315-1324.

Severinova E, Severinov K, Darst SA (1998) Inhibition of *Escherichia coli* RNA polymerase by bacteriophage T4 AsiA. *Journal of Molecular Biology* 279: 9-18.

Severinova E, Severinov K, Feny D, Marr M, Brody EN, Roberts JW, Chait BT, Darst SA (1996) Domain Organization of the *Escherichia coli* RNA Polymerase [sigma]70Subunit. *Journal of Molecular Biology* 263: 637-647.

Shaw EI, Dooley CA, Fischer ER, Scidmore MA, Fields KA, Hackstadt T (2000) Three temporal classes of gene expression during the *Chlamydia trachomatis* developmental cycle. *Mol Microbiol* 37: 913-925.

Shen L, Feng X, Yuan Y, Luo X, Hatch TP, Hughes KT, Liu JS, Zhang YX (2006) Selective promoter recognition by chlamydial sigma28 holoenzyme. *J Bacteriol* 188: 7364-7377.

Shen L, Shi Y, Douglas AL, Hatch TP, O'Connell CM, Chen JM, Zhang YX (2000) Identification and characterization of promoters regulating tuf expression in *Chlamydia trachomatis* serovar F. *Arch Biochem Biophys* 379: 46-56.

Shih YP, Wu HC, Hu SM, Wang TF, Wang AH (2005) Self-cleavage of fusion protein *in vivo* using TEV protease to yield native protein. *Protein Sci* 14: 936-941.

Silvaggi JM, Perkins JB, Losick R (2006) Genes for small, noncoding RNAs under sporulation control in *Bacillus subtilis*. *J Bacteriol* 188: 532-541.

Slepenkin A, Motin V, de la Maza LM, Peterson EM (2003) Temporal expression of type III secretion genes of *Chlamydia pneumoniae*. *Infect Immun* 71: 2555-2562.

Song J, Zhou C, Liu R, Wu X, Wu D, Hu X, Ding Y (2009) Expression and purification of recombinant arginine decarboxylase (*speA*) from *Escherichia coli*. *Mol Biol Rep*.

Spratt BG (1977) Temperature-sensitive cell division mutants of *Escherichia coli* with thermolabile penicillin-binding proteins. *J Bacteriol* 131: 293-305.

Stepanova E, Lee J, Ozerova M, Semenova E, Datsenko K, Wanner BL, Severinov K, Borukhov S (2007) Analysis of promoter targets for *Escherichia coli* transcription elongation factor *GreA in vivo* and *in vitro*. *J Bacteriol* 189: 8772-8785.

Stephens RS, Kalman S, Lammel C, Fan J, Marathe R, Aravind L, Mitchell W, Olinger L, Tatusov RL, Zhao Q, Koonin EV, Davis RW (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* 282: 754-759.

Stephens RS, Koshiyama K, Lewis E, Kubo A (2001) Heparin-binding outer membrane protein of *chlamydiae*. *Mol Microbiol* 40: 691-699.

Stevens RC (2000) Design of high-throughput methods of protein production for structural biology. *Structure* 8: R177-185.

Stewart JB, Beckenbach AT (2009) Characterization of mature mitochondrial transcripts in *Drosophila*, and the implications for the tRNA punctuation model in arthropods. *Gene* 445: 49-57.

Storz G, Opdyke JA, Zhang A (2004) Controlling mRNA stability and translation with small, noncoding RNAs. *Curr Opin Microbiol* 7: 140-144.

Streaker ED, Beckett D (1999) Ligand-linked structural changes in the *Escherichia coli* biotin repressor: the significance of surface loops for binding and allostery. *J Mol Biol* 292: 619-632.

Su H, Caldwell HD (1998) Sulfated polysaccharides and a synthetic sulfated polymer are potent inhibitors of *Chlamydia trachomatis* infectivity in vitro but lack protective efficacy in an *in vivo* murine model of chlamydial genital tract infection. *Infect Immun* 66: 1258-1260.

Su H, Raymond L, Rockey DD, Fischer E, Hackstadt T, Caldwell HD (1996) A recombinant *Chlamydia trachomatis* major outer membrane protein binds to heparan sulfate receptors on epithelial cells. *Proc Natl Acad Sci U S A* 93: 11143-11148.

Su H. YXZaRL (1985) Presence of muramic acid in *Chlamydia trachomatis* proved by liquid chromatography-mass spectrometry. *Ke xue tong bao* 30: 695–699.

Tamura A, Manire GP (1968) Effect of penicillin on the multiplication of meningopneumonitis organisms (*Chlamydia psittaci*). *J Bacteriol* 96: 875-880.

Tan M, Gaal T, Gourse RL, Engel JN (1998) Mutational analysis of the *Chlamydia trachomatis* rRNA P1 promoter defines four regions important for transcription *in vitro*. *J Bacteriol* 180: 2359-2366.

Tan M, Wong B, Engel JN (1996) Transcriptional organization and regulation of the *dnaK* and *groE* operons of *Chlamydia trachomatis*. *J Bacteriol* 178: 6983-6990.

Taraktchoglou M, Pacey AA, Turnbull JE, Eley A (2001) Infectivity of *Chlamydia trachomatis* serovar LGV but not E is dependent on host cell heparan sulfate. *Infect Immun* 69: 968-976.

Tetart F, Bouch JP (1992) Regulation of the expression of the cell-cycle gene *ftsZ* by *DicF* antisense RNA. *Mol Microbiol* 6: 615.

Tjaden B, Saxena RM, Stolyar S, Haynor DR, Kolker E, Rosenow C (2002) Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays. *Nucl Acids Res* 30: 3732-3738.

Tramonti A, De Canio M, De Biase D (2008) *GadX/GadW*-dependent regulation of the *Escherichia coli* acid fitness island: transcriptional control at the *gadY-gadW* divergent promoters and identification of four novel 42 bp *GadX/GadW*-specific binding sites. *Mol Microbiol* 70: 965-982.

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

Udekwu KI, Darfeuille F, Vogel J, Reimegard J, Holmqvist E, Wagner EG (2005) *Hfq*-dependent regulation of *OmpA* synthesis is mediated by an antisense RNA. *Genes Dev* 19: 2355-2366.

Urban JH, Vogel J (2008) Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol* 6: e64.

van den Berg S, Lofdahl PA, Hard T, Berglund H (2006) Improved solubility of TEV protease by directed evolution. *J Biotechnol* 121: 291-298.

Van Reeth T, Dreze PL, Szpirer J, Szpirer C, Gabant P (1998) Positive selection vectors to generate fused genes for the expression of his-tagged proteins. *Biotechniques* 25: 898-904.

Vanderpool CK, Gottesman S (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* 54: 1076-1089.

Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448: 157-162.

Viegas SC, Pfeiffer V, Sittka A, Silva IJ, Vogel J, Arraiano CM (2007) Characterization of the role of ribonucleases in *Salmonella* small RNA decay. *Nucl Acids Res* 35: 7651-7664.

Vogel J, Argaman L, Wagner EG, Altuvia S (2004) The small RNA *IstR* inhibits synthesis of an SOS-induced toxic peptide. *Curr Biol* 14: 2271-2276.

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. *Nucl Acids Res* 31: 6435-6443.

Vogel J, Hess WR (2001) Complete 5' and 3' end maturation of group II intron-containing tRNA precursors. *RNA* 7: 285-292.

Vogel J, Sharma CM (2005) How to find small non-coding RNAs in bacteria. *Biol Chem* 386: 1219-1238.

Wagner CR, Bergstrom CP, Koning KR, Hanna PE (1996) Arylamine N-acetyltransferases. Expression in *Escherichia coli*, purification, and substrate specificities of recombinant hamster monomorphic and polymorphic isozymes. *Drug Metab Dispos* 24: 245-253.

Wagner EG, Altuvia S, Romby P (2002) Antisense RNAs in bacteria and their genetic elements. *Adv Genet* 46: 361-398.

Wagner EGH, Brantl S (1998) Kissing and RNA stability in antisense control of plasmid replication. *Trends Biochem Sci* 23: 451.

Wassarman KM (2007) 6S RNA: a small RNA regulator of transcription. *Curr Opin Microbiol* 10: 164-168.

Wassarman KM, Repoila F, Rosenow C, Storz G, Gottesman S (2001) Identification of novel small RNAs using comparative genomics and microarrays. *Genes Dev* 15: 1637.

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

Weiss E (1950) The effect of antibiotics on agents of the psittacosis-lymphogranuloma group I. The effect of penicillin. *J Infect Dis* 87: 249-263.

Wichlan DG, Hatch TP (1993) Identification of an early-stage gene of *Chlamydia psittaci* 6BC. *J Bacteriol* 175: 2936-2942.

Wientjes FB, Nanninga N (1989) Rate and topography of peptidoglycan synthesis during cell division in *Escherichia coli*: concept of a leading edge. *J Bacteriol* 171: 3412-3419.

Willkomm DK, Hartmann RK (2005) 6S RNA - an ancient regulator of bacterial RNA polymerase rediscovered. *Biol Chem* 386: 1273-1277.

Wilson AC, Tan M (2002) Functional analysis of the heat shock regulator *HrcA* of *Chlamydia trachomatis*. *J Bacteriol* 184: 6566-6571.

Wilson AC, Wu CC, Yates JR, 3rd, Tan M (2005) Chlamydial *GroEL* autoregulates its own expression through direct interactions with the *HrcA* repressor protein. *J Bacteriol* 187: 7535-7542.

Withey JH, Friedman DI (2003) A salvage pathway for protein synthesis: tmRNA and *trans*-translation. *Annu Rev Microbiol* 57: 101.

Woldringh CL, Huls P, Pas E, Brakenhoff GJ, Nanninga N (1987) Topography of peptidoglycan synthesis during elongation and polar cap formation in a cell division mutant of *Escherichia coli* MC4100. *J Gen Microbiol* 133: 575-586.

Wood H, Fehlner-Gardner C, Berry J, Fischer E, Graham B, Hackstadt T, Roshick C, McClarty G (2003) Regulation of tryptophan synthase gene expression in *Chlamydia trachomatis*. *Mol Microbiol* 49: 1347-1359.

Wyllie S, Raulston JE (2001) Identifying regulators of transcription in an obligate intracellular pathogen: a metal-dependent repressor in *Chlamydia trachomatis*. *Mol Microbiol* 40: 1027-1036.

Yu HH, Di Russo EG, Rounds MA, Tan M (2006) Mutational analysis of the promoter recognized by *Chlamydia* and *Escherichia coli* sigma(28) RNA polymerase. *J Bacteriol* 188: 5524-5531.

Yu HH, Tan M (2003) Sigma28 RNA polymerase regulates *hctB*, a late developmental gene in *Chlamydia*. *Mol Microbiol* 50: 577-584.

Yuan AH, Nickels BE, Hochschild A (2009) The bacteriophage T4 *AsiA* protein contacts the beta-flap domain of RNA polymerase. *Proc Natl Acad Sci U S A* 106: 6597-6602.

Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S (2003) Global analysis of small RNA and mRNA targets of *Hfq*. *Mol Microbiol* 50: 1111.

Zhang L, Douglas AL, Hatch TP (1998) Characterization of a *Chlamydia psittaci* DNA binding protein (*EUO*) synthesized during the early and middle phases of the developmental cycle. *Infect Immun* 66: 1167-1173.

Zhang L, Kasif S, Cantor CR, Broude NE (2004) GC/AT-content spikes as genomic punctuation marks. *Proc Natl Acad Sci U S A* 101: 16855-16860.

Zhong J, Douglas AL, Hatch TP (2001) Characterization of integration host factor (*IHF*) binding upstream of the cysteine-rich protein operon (*omcAB*) promoter of *Chlamydia trachomatis* LGV serovar L2. *Mol Microbiol* 41: 451-462.

VITA

Yasser Mohammed Elsayed Metwally AbdelRahman, was born in Cairo, Egypt, on May 1. 1974. He graduated from the Tala' ElKamal School (The perfect pioneers), Heliopolis, Cairo, in 1992 and received his Bachelor's degree with honors in Pharmacy and Pharmaceutical Sciences from the College of Pharmacy, Cairo University in 1997. After graduation, he was appointed as a laboratory instructor in the department of microbiology in the same School. In 2001 he received his Masters degree titled "Development of Fast Serological Detection Systems for Human Fascioliasis". He enrolled in the Department of Molecular Sciences at the University of Tennessee Health Science Center, in Memphis, in August 2001. He worked as a research assistant in the department of Veteran Affairs in Memphis in August of 2001 under the supervision of Dr Malak Kotb. He joined the research group of Dr. Robert Belland in November 2003. During his doctoral study, he was invited to present his work as an oral presentation in the 2nd and 3rd *Chlamydia* Basic Research Society Meeting in April 2005 and March 2007.