THE RpoN REGULON OF *SALMONELLA* AND ITS COMPONENT RNA REPAIR SYSTEM

by

DAVID JOHN SAMUELS

(Under the Direction of Anna Karls)

ABSTRACT

RNA polymerase (RNAP) holoenzyme catalyzes transcription in eubacteria. The one variable subunit of this multi-subunit enzyme is σ, which mediates promoter recognition and promoter-holoenzyme interactions. In addition to a primary σ-factor, most cells have one or more alternative σ-factors that regulate different sets of genes in response to varying environmental conditions. One alternative σ-factor, σ^{54}, differs from other known σ-factors in structure, conserved promoter elements, and in its absolute requirement for an activator protein that, upon stimulation by a particular environmental condition, interacts with RNAP-σ^{54} and hydrolyzes ATP, generating the energy necessary for transcriptional initiation. This activator requirement can hinder global analysis of the σ^{54} regulon because σ^{54}-dependent promoters will be transcriptionally silent without proper environmental cues. To overcome this limitation, an engineered promiscuous, constitutively-active variant of the *Sinorhizobium meliloti* DctD activator was used to define the σ^{54} regulon in *Salmonella* Typhimurium. Using this engineered activator, microarray analysis was used to identify σ^{54}-dependent transcripts. This approach confirmed the regulation of 16 promoters previously predicted to be σ^{54}-dependent. Chromatin immunoprecipitation linked to microarray analysis revealed 70 sites throughout the genome interacting with either σ^{54} or RNAP-σ^{54}. Surprisingly, >50% of these sites were predicted to fall within coding sequences. Promoter fusion assays indicated that some of these intragenic sequences could function as σ^{54}-dependent promoters , raising the possibility of new regulatory roles for σ^{54}. One operon that was shown to be σ^{54}-dependent encodes a putative RNA repair system. The components of this system—an

RNA-binding ribonucleoprotein complex (Rsr-Y RNA), an RNA ligase (RtcB), and an RNA phosphate cyclase (RtcA)—are found throughout all domains of life. While functions for eukaryotic/archaeal homologs have been described, their role in bacteria remains enigmatic, largely because conditions that stimulate the activator protein, allowing transcription are unknown. We used quantitative, reverse transcriptase PCR to assess transcription after exposure to various stresses. Treatment with Mitomycin C, a nucleic acid alkylating agent, resulted in up-regulation of this operon. This finding supports our model in which these genes are expressed in response to nucleic acid damage, and their products interact with the damaged molecules to directing their repair or degradation.

INDEX WORDS: *Salmonella*, σ^{54}, RpoN, Transcription Regulation, RNA Repair

THE RpoN REGULON OF *SALMONELLA* AND ITS COMPONENT RNA REPAIR SYSTEM

by

DAVID JOHN SAMUELS

BS, Cornell University, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

THE RpoN REGULON OF *SALMONELLA* AND ITS COMPONENT RNA REPAIR SYSTEM

by

DAVID J. SAMUELS

Major Professor:	Anna Karls
Committee:	Timothy Hoover
	Lawrence Shimkets
	Duncan Krause
	Fred Quinn

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

ACKNOWLEDGEMENTS

First off, I owe a tremendous debt of gratitude to my mentor, Dr. Anna Karls. Without her guidance, none of this work would have been possible. While I could easily fill up 15 pages detailing innumerable ways that she's helped me, I'll limit myself to three salient points. 1) She has unconditionally supported me doing what is best for my life and has made sacrifices to facilitate that for me. 2) She has always demanded that I do good work. There were times when I tried to do work that was "good enough," but it would take more energy to convince her that "good enough" was an acceptable standard than it would to make the work "good." 3) The fact that I'm able to write this indicates that she has not murdered me, which would be an understandable act given the amount of antics, shenanigans, AND tomfoolery that I've engaged in over the years. How she managed to put up with this in (seemingly) good spirits is still beyond me.

At one point, Anna was reading her daughter's honor's thesis and it was dedicated "to the world's best mother." I thought it was strange that she would dedicate her thesis to my mother, but people do strange things all the time. Anyhow, this is an anecdotal lead in to me thanking my parents for all of their unconditional love and support, even though I'm not sure they knew exactly *what* it was that they were supporting. My sister, Ruth, has also helped mitigate some of my insanity. And my extended family (which provides a textbook example of how the confluence of nature and nurture can shape a person's behavior), has been tremendously supportive as well. It's difficult to articulate how much I appreciate how much I appreciate all of this.

Dr. Wendy Dustman has been a fantastic mentor for me as an educator. I've had the opportunity to teach under her guidance several times and she has always been supportive of my development as an educator, going so far as to give me an opportunity to sandbox a course idea in the 3500 Honors section. Additionally, she has always been willing to serve as a sounding board for any issues, teaching or

otherwise, that I may have. Additionally, Paul Quick and Denise Domizi at the CTL have provided fantastic educational support for me, particularly in the Future Faculty Program. Having this network of educationally-minded graduate students across the campus (and alumni across the country) has been a terrific resource.

In January 2010, I met this guy at a bar during the Micro-Ecology grad student happy hour. We talked about singing and a cappella music and he mentioned that he was starting a group. Given the amount of hare-brained ideas that I discuss in bars that never come to fruition, I didn't think too much of this. When I actually got a call from him the next day about a rehearsal, I had no idea how much being a part of the EcoTones would impact my life. Over the past 4 years, we've gone from singing in a living room to performing across the Southeast and between the singing, dancing, and laughing, these people have become a family to me. I don't know where I'd be without them. Blair, The Cheat, Ross, Spencer, Sander, Milton, Fish, Evan, Henrí, Kip, Forrest, Twerk, MoS, and everyone else #ToneLove.

I've had the opportunity to work with and learn from some fantastic lab-mates throughout my time here: John Buchner, Chandra Carpenter, Brian Higgins, Adrienne Cottrell, Ashley Bono, Jason Cloward, and Christy Hartman. John and Chandra especially helped me get my feet under me when I first started out. And discussions with Christie and Ashley in my later years were quite helpful.

My committee—Fred Quinn, Tim Hoover, Duncan Krause, and Larry Shimkets—has been incredibly helpful and supportive throughout this whole process. Throughout my first several years, rather than suggesting (based on my gross ineptitude) that I pursue other endeavors, they got excited about the work I was doing and suggested additional experiments. Toward the latter part of my graduate career, the ineptitude waned slightly; their enthusiasm and support did not.

Though none of the work they assisted on is discussed here, I would be remiss to not mention the help I received from Dr. Tom Krunkosky, Shelly Helms, Margie Lee, and Barbara Reaves at the UGA Vet School. My initial project involved working with a complicated tissue culture system, something I knew very little about. They spent a considerable amount of time teaching me the ins and outs of

mammalian cell culture and helped me set up my experiments. Though none of it ever worked, the fault was entirely my own.

Several professors/labs have been quite amenable to me mooching supplies and equipment. If they hadn't let me "borrow" things at the frequency I did, I'd probably be here for another couple of years. Eric Stabb, Diana Downs, Zach Lewis, Jorge Escalante (who was also instrumental in helping me land a postdoc), Tim Hoover, Rob Maier, Sidney Kushner, and Fred Quinn were my main targets, but others have certainly helped too.

The micro support staff has been fantastic. Dale Boyer has helped fix this computer more times than I can remember. Nancy Perkins, Christie Haynes, and Andrea Barnett have always stepped up and busted skulls for me whenever that became necessary. And Linda Sanders and Sophia Fleming have been responsive to all matter of trouble that I've managed to create.

I've had the opportunity to mentor several undergraduate students in the lab: Uchenna Ewonolu, Tanya Grancharova, Trevor Wright, Hollie Burrage, Janvi Wandiwala, Lauren McLeod, and Cameron Story. They've taught me many valuable lessons about how to be a good mentor. And perhaps more importantly, they've taught me many more about how not to be a good mentor. If I were smart, I'd apply these lessons moving forward.

Finally, I need to thank all of the wonderful friends I've made that have made my time here a pleasure. Whether it was discussing science or grabbing beers downtown (which often involved discussing science), biking, kayaking, playing hockey, bar trivia, tailgating before football games, going to Twilight, or any of the other things that make Athens a great place to be, they are largely responsible for my enjoyment. Honestly, if I try to name names, I'm going to have some embarrassing omissions. If we've done any of the things I've listed above together, please know you have helped me survive this ordeal.

TABLE OF CONTENTS

Page

CHAPTER 1:

INTRODUCTION AND LITERATURE REVIEW

Given the energy and resources required to express genes, as well as the potential harm that can result from genes being expressed at inappropriate times, it is not surprising that transcription (the conversion of DNA sequence information into an RNA molecule) is a highly regulated process in bacteria. One of the most common mechanisms of transcriptional regulation is via the σ factor, a variable subunit in the RNA polymerase (RNAP) holoenzyme. The σ factor is the subunit of the RNAP holoenzyme that recognizes and binds to the promoter region upstream of a gene or operon. While the primary σ factor, σ^{70} recognizes promoters for housekeeping genes, alternative σ factors recognize a variant promoter sequence upstream of a different subset of genes. By controlling the activity of the sigma factor, the cell can simultaneously regulate expression of all of these genes (collectively known as a regulon or sigmulon).

While most alternative σ factors are closely related to the housekeeping σ^{70}, σ^{54} is in a class by itself. One of the features that makes σ^{54} unique is its absolute requirement for a bacterial enhancer-binding protein (bEBP) to sense an environmental stimulus, become activated, and facilitate hydrolyze ATP to provide the energy for RNAP-σ^{54} to initiate transcription. The number of these bEBPs in a bacterial genome (13 in *Salmonella enterica* serovar Typhimurium) and the diversity of stimuli that they respond to present a challenge to the global definition of the σ^{54} regulon. This dissertation will describe the use of one of these bEBPs, which was engineered to be promiscuous and constitutively-active, to define the full extent of the σ^{54} regulon in *Salmonella enterica* subspecies *enterica* serovar Typhimurium (*S.* Typhimurium).

One of the operons regulated by σ^{54} in *Salmonella* is the *rsr-rtcBA* operon, which encodes a putative RNA repair system. Components of this system are conserved across all three domains of life

and have been characterized to varying extents in different organisms. The components of the bacterial operon are dispensable for growth under normal laboratory conditions and the activities of the RNA ligase and 2',3'-phosphate cyclase that are encoded by the operon have only been characterized *in vitro*. Accordingly, no physiological role for these gene products has been ascribed. The conservation of the RNA repair genes and the maintenance of their functionality suggest that the RNA repair system plays an important role under environmental conditions encountered by the bacteria. Research described in this dissertation determines the nature of the environmental stresses that activate transcription of the RNA repair genes in *S.* Typhimurium and addresses the role that the system plays in the response to these stresses.

In addition to the intrinsic value of a more profound knowledge of bacterial genetics and physiology, using *Salmonella* as a model organism has other benefits as well. *S.* Typhimurium is a causative agent of salmonellosis in humans, one of the most prevalent food-borne gastrointestinal diseases worldwide. In the United States alone, the Centers for Disease Control and Prevention estimates over 1 million incidences, 20,000 hospitalizations, and 400 deaths caused by salmonellosis each year [1]. While most cases of salmonellosis are self-limiting, ~5% of cases (especially those in children and immunocompromised patients) result in invasive bacteremia and require treatment with antimicrobials [2]. Given the rapid emergence of antimicrobial resistance in *Salmonella* [2] (and in other bacteria as well), it is imperative to continually find new targets for antimicrobial agents. Understanding the mechanisms that *Salmonella* uses to respond to changes in its environment may reveal novel targets for these agents.

Transcriptional Regulation by σ^{54} (RpoN)

Transcription in eubacteria is mediated by the RNA polymerase holoenzyme (Eσ), which has five constant subunits ($\alpha_2\beta\beta'\omega$) and a variable subunit (σ). The constant subunits constitute the RNA polymerase core (RNAP), which catalyzes polymerization activity. The σ subunit is responsible for

promoter recognition and binding as well as for the conversion of the transcriptionally inert closed complex to the active open complex, a process known as isomerization. In this way, the σ factor can modulate the specificity of the RNAP holoenzyme for different promoter sequences.

In bacteria, transcription of housekeeping genes (those that are generally constitutively on) is accomplished by the primary σ factor—σ^{70} (RpoD) in *E. coli* and *Salmonella.* In addition, most bacteria have a complement of alternative σ factors, which interact with a different sequence than σ^{70}, thereby directing transcription of a different set of genes. The number of alternative σ factors within a genome can vary from 0 (the streamlined *Mycoplasma genitalium* only possesses a primary σ factor) to 62 (in *Streptomyces coelicolor*) [3].

Most alternative σ factors control a set of genes that are generally involved with a common function. Among *S.* Typhimurium's five alternative σ factors, $\sigma^{S/38}$ controls genes that are critical to the cells' entry into stationary phase, $\sigma^{H/32}$ regulates expression of genes involved in the heat shock response, $\sigma^{E/24}$ guides transcription of genes in response to envelope stress, and $\sigma^{fliA/28}$ directs expression of flagellar biosynthesis genes [4]. The fifth alternative σ factor in *Salmonella*, σ^{54} (RpoN) does not control a regulon that responds to a single environmental stress or that has a coherent function in cell physiology. While σ^{54} was initially implicated in the transcription of genes involved in dealing with low nitrogen availability [5] (hence the N in RpoN), the repertoire of genes that it controls is much more diverse. Genes regulated by σ^{54} are involved in variety of cellular processes, including flagellar biogenesis, transport and metabolism of carbon substrates, and tolerance to heavy metals [6-9]. One recent analysis by Francke, et al. [10] used a comparative genomics approach in an attempt to elucidate a theme for this regulon. Using three different comparisons—*(i)* the taxonomic distribution of σ^{54}, *(ii)* the bEBP complement in bacteria encoding σ^{54}, and *(iii)* the genomic contexts of σ^{54}-dependent promoters—they concluded that σ^{54} was involved in determining the composition of the cell exterior under various growth conditions. In addition to controlling genes involved in flagellar biosynthesis, σ^{54} also regulates genes either directly or indirectly involved in the transport or synthesis of the precursors of extracellular polysaccharide,

lipopolysaccharide, peptidoglycan, phospholipids, and (lipo-) teichoic acid. This model is fairly comprehensive, yet it leaves out other known elements of the σ^{54} regulon, importantly, the Rtc RNA repair system (which received no mention in the paper [10]). While assessing the function of this operon, it may be useful to consider it within the framework of modulating the exterior of the cell.

Consistent with its role as an alternative σ factor, in most bacteria σ^{54} is not essential for growth under favorable conditions. However, in some bacteria, like *Myxococcus xanthus* [11] or *Geobacter sulfurreducens* [12], which are both members of the δ class of the Proteobacteria phylum, σ^{54} is essential for growth.

Prior to the work discussed in Chapter 2, there were six promoters in *S.* Typhimurium that were shown experimentally to be σ^{54}-dependent (Table 1.1)**.** There were also an additional 20 promoters that were predicted to be σ^{54}-dependent in *S.* Typhimurium (but not experimentally-verified). These predictions were based on *in silico* analyses indicating either homology to known σ^{54}-dependent operons in *E. coli* and other enteric bacteria or promoter sequence homology along with genetic proximity to predicted bEBP genes [6, 8, 13-25]. Chapter 2 describes the genomic and genetic analyses used to define the σ^{54} regulon in *S.* Typhimurium, confirming that 16 of these 20 predicted promoters are σ^{54}-dependent.

σ^{54} versus σ^{70}-type σ Factors

All characterized alternative σ factors are related to σ^{70} with the exception of σ^{54}, which is the sole representative of its σ factor family. There is extensive divergence in primary amino acid sequence between σ^{54} and the σ^{70} family leading to substantial differences in both the structure and function of these σ factors. The first major difference is in the promoter sequence that is recognized by these proteins. σ^{70}-type sigma factors recognize regions that are located -35 and -10 bp relative to the transcription start site (TSS). For any given σ factor, these consensus sequences are not highly conserved and the spacing is fairly flexible (i.e. insertion of deletion of 2-3 bp between the -35 and -10 site will not eliminate promoter function)[26]. In contrast, the promoter sequences recognized by σ^{54} lie at -24 and -12 bp upstream of the

TSS [27]. These promoters contain highly-conserved GG and GC dinucleotides at the -24 and -12 positions, respectively. Insertion or deletion of even 1bp between these two positions will abolish promoter activity, indicating that the orientation of these bases within the DNA helix is critical to σ^{54} recognition [27].

Another major difference between the σ^{70} family and σ^{54} lies in the mechanism of transcription initiation. Whereas $E\sigma^{70}$ can interact with promoter sequences and spontaneously isomerize to open complexes, the formation of the open complex (and, therefore, initiation of transcription) by $E\sigma^{54}$ is energy-dependent and requires a bEBP (Figure 1.1A-B). These proteins will be discussed in more detail below, but briefly: upon sensing an environmental stimulus, the bEBP becomes activated to form oligomers (generally hexamers), which can interact with a DNA sequence known as the upstream activator sequence (UAS) [28]. Compared to sequences that interact with σ^{70} regulatory proteins, the bEBP UAS is relatively far away (~100 bp) from the TSS [29]. A DNA looping event brings the bEBP into contact with $E\sigma^{54}$, bound to the promoter in a closed complex. ATP hydrolysis by the central domain of the bEBP provides the energy needed to transition to open complex. Between the reliance on an upstream enhancer sequence, the requirement for nucleotide hydrolysis, and the involvement of DNA looping, $E\sigma^{54}$ seems to share more similarities with eukaryotic polymerase II than other σ factors found in bacteria [30].

A final notable difference between the two sigma factor families is in the ability of the σ factor itself to bind DNA. Wild-type σ^{70} has only been shown to bind to promoter sequences when it is complexed with core RNAP as $E\sigma^{70}$ (variant forms of σ^{70} containing N-terminal deletions *were* shown to interact with DNA in the absence of RNAP core [31]). σ^{54}, on the other hand, is capable of interacting directly with some of its promoter sequences [32]. Though DNA binding by σ^{54} is much more efficient in the context of the holoenzyme, the ability of σ^{54} to bind directly to specific promoter sequences may play an important role in σ^{54} gene regulation. The binding of σ^{54} to DNA may facilitate the formation of RNAP holoenzyme *at* the promoter sequence, thereby compensating for the relatively (compared to σ^{70}) weak

affinity of σ^{54} for RNAP core in solution [33]. Alternatively, since different DNA sequences showed varying affinities for σ^{54} and $E\sigma^{54}$, binding of DNA by σ^{54} *per se* may be important. Understanding the physical differences between DNA sequences that interact with just σ^{54}, with $E\sigma^{54}$, and those that can actually function as promoters will provide valuable insight into σ^{54} regulation within the cell.

Bacterial Enhancer-Binding Proteins (bEBPs)

The requirement for an activator (and ATP hydrolysis) to stimulate $E\sigma^{54}$ open complex formation derives from the stability of the $E\sigma^{54}$ closed complex. Protein-DNA and protein-protein interactions contribute to the stability of the $E\sigma^{54}$ closed complex. In the stable closed complex, a transient DNA fork junction forms at the -12 region at which Region I of σ^{54} interacts with the open DNA, called the regulatory center. This nucleoprotein complex prevents the appropriate interactions between $E\sigma^{54}$ and the promoter sequence needed for extension of a transcription bubble to the +1 position [34-36]. Structural studies of $E\sigma^{54}$ and $E\sigma^{54}$ associated with a bEBP, including cryo-electron microscopy, show that σ^{54} Region I is positioned to block the active site of RNAP in the closed complex $E\sigma^{54}$ [37]. Interaction of bEBP with σ^{54} and the subsequent ATP hydrolysis cause structural changes in σ^{54} that release the inhibitory interactions at the regulatory center, moves Region I from the active site, and allows RNAP to form a stable open complex in which the transcription bubble reaches the +1 position [36, 37].

bEBPs are modular proteins that generally consist of three domains (Figure 1.1A). An N-terminal regulatory domain plays a role in signal detection. The central AAA+ ATPase domain mediates bEBP-σ^{54} contact and provides the energy necessary for closed complex isomerization. And the C-terminal domain often contains a helix-turn-helix DNA binding motif that enables enhancer site recognition. However, within this framework there is substantial variation with regard to the nature (and sometimes, the presence) of the C- and N-terminal domains among different bEBPs [10, 36].

Though the N- and C-termini may vary between different bEBPs, the central AAA+ ATPase domain is conserved among all bEBPs and is indispensible for σ^{54}-dependent transcription [36]. This

domain contains a characteristic GAFTGA motif, which mediates the interaction with σ^{54}. Two variant bEBPs that lack this motif—*Rhodobacter capsulatus* NtrC and *E. coli* TyrR—are unable to activate σ^{54}-dependent transcription and function instead with $E\sigma^{70}$ [36, 38, 39]. Experiments with truncated variants of the *Sinorhizobium meliloti* DctD dicarboxylic acid transport protein showed that this central ATPase domain is sufficient to stimulate transcription initiation by $E\sigma^{54}$ [40]. This truncated activator is used in the research detailed in Chapter 2 to define the σ^{54} regulon of *S.* Typhimurium.

Many bEBPs contain N-terminal response regulator domains that allow the cell to coordinate transcriptional activation with a specific environmental stimulus. This stimulus is generally perceived via either phosphorylation or ligand binding. Some bEBPs, such as NtrC, are members of two-component histidine kinase/response regulator systems. Under nitrogen-limiting conditions, NtrB phosphorylates NtrC, which stimulates oligomerization, thus activating it as a bEBP [41]. When nitrogen is abundant, the phosphatase activity of NtrB is engaged and non-phosphorylated NtrC does not stimulate transcription. In other cases, bEBPs respond to the presence of small molecules. XylR, which regulates genes involved in aromatic hydrocarbon catabolism in *Pseudomonas putida*, binds to molecules such as xylene and *o*- and *m*-nitrotoluene [42-44]. *E. coli* NorR, which regulates the nitric oxide (NO) detoxifying *norV* and *norW* genes, contains a non-heme iron center that allows the detection of NO in the cell [45].

Another class of bEBPs lack an N-terminal domain and instead relies on interactions with anti-activator proteins to regulate their activity [46]. This is the case for PspF, which is inactive when interacting with PspA [47]. It is believed that upon disruption of proton motive force, PspB and PspC bind to PspA, removing it from PspF. In the absence of PspA, PspF is able to activate transcription from the *psp* promoter. Similarly, experimental truncation of the N-terminal domain results in a constitutively active regulator (Figure 1.1C). These activators can stimulate transcription irrespective of environmental conditions [15, 48-50].

The method of bEBP activation will be important to consider when characterizing the function of the *S.* Typhimurium Rsr-RtcBA RNA repair system (Chapter 3), which is regulated by RtcR. Since RtcR

possesses an intact N-terminal regulatory region, stimulation is not likely to occur via removal of an antiactivator. Activation of RtcR, therefore, is likely dependent on phosphorylation or interaction with a small molecule. Potential activating stimuli for this bEBP are discussed later in this chapter.

The C-terminal domain of most bEBPs contains a helix-turn-helix DNA binding motif. This motif is responsible for recognizing and binding an enhancer sequence, also called an upstream activator sequence, thus mediating activator specificity for the proper promoter. There are usually one to three enhancer sequences ~80-150 bp upstream of σ^{54}-dependent promoters, but other configurations have been characterized. In *Pseudomonas aeruginosa*, FleQ activates σ^{54}-dependent promoters for multiple flagellar genes from enhancer sites between the transcription and translation start sites [51]; and the enhancer for *rocG* in *Bacillus subtilis* is 1.5 kb downstream from the promoter [52]. Engineered *glnA* enhancer sites in *E. coli* remain functional when located up to 3 kb upstream of the promoter [53], highlighting the importance of DNA looping in enhancer-promoter interactions. The fact that the enhancer does not have to be in a certain position with respect to the promoter can confound *in silico* predictions that look for promoter and enhancer consensus sequences within a prescribed range, as well as analysis of cloned promoter sequences. In either of these scenarios, failure to include enough DNA to account for a distant enhancer would result in a false negative.

bEBPs that lack a C-terminal domain (either naturally or via experimental manipulation) display promiscuous activation activity. That is, once activated via the N-terminal domain, these bEBPs can interact with $E\sigma^{54}$ at any σ^{54}-dependent promoter and stimulate transcription (Figure 1.1D)**.** Some bEBPs, such as *Chlamydia trachomatis* CtcC and *Helicobacter pylori* FlgR, naturally lack the C-terminal domain [54]. Since these are the only known bEBPs within these organisms, enhancer-mediated promoter specificity is unnecessary. In *Rhodobacter sphaeroides*, FleT is not the only bEBP, but it still lacks a DNA binding domain [54]. The specificity of FleT is maintained by interacting with only one of four different, non-interchangeable σ^{54} paralogs within the cell [55]. Constructed variants of PspF, NtrC, and DctD that lack the C-terminal domain are capable of stimulating transcription from solution (i.e., without

interacting with an enhancer) [56-58]. These previous characterizations of bEBP variants led to the use of a promiscuous, constitutively-active DctD variant (DctD250) to study cloned promoter sequences, as described in Chapter 2, thus avoiding the problem of not knowing how much of the surrounding sequence for an identified σ^{54}-dependent promoter is required to include the enhancer sequences.

Based on *in silico* analysis, *S.* Typhimurium is predicted to have 13 different bEBPs in its genome [13]. This is pertinent for testing the regulon definition methods described in Chapter 2. The presence of 13 bEBPs indicates that this regulon is capable of responding to many different conditions (as opposed to, say, *C. trachomatis* or *H. pylori*, which may be more limited). The moderate number of bEBPs in *S.* Typhimurium should yield a robust data set to analyze without being excessive.

Characterization of σ^{54} Regulons in Other Bacteria

Prior to the study described in Chapter 2, several attempts to characterize the global σ^{54} regulons in other bacteria were described. An *in silico* analysis was utilized to predict the σ^{54} regulon in *Pseudomonas putida* [59], which identified co-localized consensus σ^{54} promoter sequences and consensus enhancer sequences. Although some of the 46 potential σ^{54}-dependent promoters identified in this computational analysis have been shown *in vivo* to be part of the σ^{54} regulon, many of the sites have not been confirmed experimentally.

Since σ^{54} regulates genes involved in essential ammonia assimilation pathways in *Geobacter sulfurreducens*, a Δ*rpoN* mutant is not viable. To characterize this regulon, Leung, et al. [12] used microarray analysis to compare expression levels between wild type and RpoN-overproducing cells ($RpoN^+$). Using a relatively low cutoff (1.5-fold), they detected ~140 genes that were up-regulated and ~50 genes that were down-regulated in $RpoN^+$ cells compared to wild-type. These results facilitated the modeling of regulatory networks that might be useful when this bacterium is used in bioremediation or electricity generation.

Analysis of σ^{54} binding and transcription in *E. coli* was performed by chromatin immunoprecipitation microarray (ChIP-chip) and RNA microarrays comparing $RpoN^{+}$ cells to a Δ*rpoN* mutant [60]. This approach defined 40 targets of σ^{54}, including 22 that were previously identified. Using a lower cutoff (1.5-fold versus 2-fold) revealed an additional 30 putative promoters.

To examine the σ^{54} regulon in *Vibrio cholerae*, the Mekalanos lab utilized high throughput methods (i.e. ChIP-seq and RNA-seq) to compare differential σ^{54} binding and gene expression between wild type and Δ*rpoN* mutants [61]. This analysis identified 68 σ^{54} binding sites and 82 operons (with 144 genes) that are positively regulated by σ^{54}. These numbers likely include operons that are directly regulated by σ^{54} as well as those regulated by the products of other σ^{54}-dependent genes. Interestingly, this study revealed several σ^{54} binding sites that were not directly involved in transcription. In all, this work was able to show different regulation schemes for Type VI secretion systems between two pandemic strains of *V. cholerae.*

One critical limitation in these last three studies, which all of the authors acknowledge, is the absence of conditions necessary to activate bEBPs. Though their approaches (i.e. comparison of wild type cells to RpoN over-expression or mutant cells) revealed numerous promoters that showed differential expression, possibly due to a low level of spontaneous activation of bEBPs, it is likely that *bona fide* σ^{54}-dependent promoters were not identified. Performing these comparisons in the presence of a constitutively-active, promiscuous bEBP is likely to detect genes that these other approaches may have missed.

The *rsr-rtcBA* Operon of *Salmonella*

The *rsr-rtcBA* operon in *Salmonella* is a σ^{54}-dependent operon that encodes a putative RNA repair system. This operon contains three structural genes: *rsr*, a Ro-sixty related ribonucleoprotein; *rtcB,* an RNA ligase; and *rtcA,* an RNA phosphate cyclase (Figure 1.2)**.** Two small RNAs (the Y RNA partners

of Rsr), *yrlA* and *yrlB*, are encoded between *rsr* and *rtcB*. The operon is regulated by the bEBP RtcR, which is divergently-encoded on the opposite strand as the rest of the genes [15].

The prediction of Rsr-RtcBA as a putative RNA repair system comes primarily from two lines of evidence: (*i*) the characterization of Ro/Rsr orthologs and their small RNA partners (Y RNAs) from a wide variety of organisms (e.g. human, mouse, *Xenopus*, *Deinococcus*, etc.) and (*ii*) *in vitro* analysis of purified RtcB and RtcA from *E. coli*, which share 88% and 68% amino acid identity, respectively, with the *Salmonella* proteins (Figure 1.3)**.** Despite these predictions, a physiological role for this system in *Salmonella* remains elusive. This is chiefly because no conditions have been identified under which this operon is expressed. Because this operon is σ^{54}-dependent, it is presumably only expressed under conditions that lead to the activation of its bEBP, RtcR. To date, expression of these genes has only been seen in strains that either contain a truncated, constitutively-active bEBP [15, 62] or in mutant strains lacking the exoribonuclease polynucleotide phosphorylase (PNPase) [63]. In a comprehensive analysis of 22 "infection relevant" conditions, no conditions were identified which resulted in transcription of these genes above the limit of detection [64]. These results indicate that either the proper stimulus (or combination of stimuli) was not tested or that the operon is inactive in *Salmonella*. Substantial evidence, as described in detail below, supports the idea that the RNA repair operon is functional and the stimulus needed to activate RtcR has not been determined; the evidence includes the activity of the promoter when stimulated by bEBP variants [15, 62, 65], association of Rsr and Y RNA to form ribonucleoprotein (RNP) complexes [65], and RtcB functionality both *in vitro* [66, 67] and *in vivo* [68].

The "entire" RNA repair operon (i.e. *rtcR, rsr, rtcB, and rtcA*), as seen in *Salmonella*, is rare among bacteria. In fact, only ~5% of sequenced bacterial genomes encode an Rsr ortholog [69]. Analysis described in Chapter 3, reveals that this operon is present in many configurations in bacteria. Homologs with >40% identity to *Salmonella* RtcB are found widely throughout all three domains of life and are present in 19 different bacterial phyla. The co-localization of RtcR or RtcR and RtcA with RtcB is largely

limited to Proteobacteria. The occurrence of the entire operon is only observed in five genera: *Salmonella, Acidovorax, Variovorax, Pseudomonas, Sphingobium,* and *Sphingomonas.*

The following sections will focus on the *rsr-rtcBA* operon. First, I will discuss the each of the components of this operon. From there, I will move on to the functions of similar RNA repair systems that have been described from diverse organisms. With these functions in mind, I will then discuss the potential roles that this system may be playing in *Salmonella.*

Components of the *rsr-rtcBA* Operon

Ro-sixty related (Rsr) ribonucleoprotein

Ro/Rsr is a 60 kDa, ring-shaped protein with orthologs found in most animals and some prokaryotes. Ro was first identified in the 1960's-70's as a major antigen in two human rheumatic diseases, systemic lupus erythematosus and Sjögren's syndrome [70], but it wasn't until almost two decades later that biological roles began to emerge. Though these early characterizations revealed some functions of Ro, its full physiological role in the cell is still unclear.

One thing that is clear about Ro is that it interacts with RNA. Ro forms ribonucleoprotein (RNP) complexes with members of a class of small non-coding RNAs known as Y RNAs as well as other RNA molecules. The structure of Ro facilitates these interactions with RNA. X-ray crystallography shows Ro to be an elliptical torus comprised largely of HEAT-repeat domains connected by a von Willebrand factor A domain [71]. The central channel of the torus and a thin strip along the outer surface of the protein contain mostly basic residues. RNA binding occurs in these two regions. Recognition and binding of various RNAs are essential to Ro's predicted roles within the cell. There are currently three potential roles for Ro within the cell: quality control of non-coding RNA, response to RNA damage, and RNA metabolism during cell stress.

Non-coding (nc) RNAs are incorporated into a variety of cellular machines (e.g. ribosomes, the eukaryotic spliceosome). Since incorporating mutant forms of these ncRNAs into cellular machines can

have severe consequences [72, 73], it is likely that cells have developed ways to ensure that these variant ncRNAs are prevented from incorporating into these cellular machines. Ro is thought to be involved in this manner of ncRNA quality control. This hypothesis emerged after immunoprecipitation of Ro-RNP complexes from *Xenopus laevis* oocytes showed that Ro was interacting with both Y RNA and variant 5S rRNA molecules containing multiple point mutations [74]. In mouse embryonic stem cells, Ro was also shown to interact with variants of the U2 small nuclear RNA [75]. In both of these studies, Ro did not interact with wild type RNA. When looking at the 5S rRNA in *Caenorhabditis elegans* ribosomes, Δ*rop-1* incorporated >4-fold more variant RNA than wild type cells (8% vs. 1.6%) [76]. Examination of total cellular RNA showed that both strains produced the variant RNA at the same rate. Taken together, these findings support a ncRNA quality control role for Ro in which Ro binds to variant RNA molecules and prevents them from incorporating into cellular machines.

Two structural features of these variant RNA molecules are thought to be important in Ro binding. The first is having a single-stranded 3' end. The central channel of Ro—where these RNA molecules bind—is ~10-15 Å in diameter, large enough to accommodate single-, but not double-stranded RNA [71]. All of the variant 5S rRNAs bound to the *X. laevis* Ro had short (~5-8nt) 3' extensions that were single-stranded [74, 77]. EMSA analysis of multiple RNAs showed that those lacking 3' extensions had a decreased affinity for Ro [71, 78]. Another feature of RNA that influences Ro binding is secondary structure. All of the ncRNA variants that interacted with Ro were predicted to fold into alternative secondary structures [74, 75]. Elimination of these secondary structures led to decreased Ro binding [77]. These findings support a model where Ro recognizes mis-folded RNAs with single-stranded 3' extensions. These 3' ends are threaded through the central channel of Ro.

In addition to mutations that may occur as a result of a transcriptional error, RNA can be damaged by environmental factors, including UV light. UV light has been shown to cause RNA-RNA as well as RNA-protein cross links in bacteria [79]. Ro appears to play a role in response to UV stress. In *Deinococcus radiodurans*, both Rsr (Ro-sixty related, the bacterial Ro ortholog) and Y RNA were

upregulated in response to UV treatment [80]. Cells lacking Rsr were more susceptible to UV radiation than wild type cells. This trend was also seen in mouse embryonic cells, where $Ro^{-/-}$ cells showed increased susceptibility to UV treatment. Interestingly, a mouse $Ro^{-/-}$ strain was more photosensitive than a wild type, showing a 2-fold increase of apoptotic keratinocytes after UV irradiation [81]. Photosensitivity is a hallmark symptom of lupus, present in up to 90% of patients with anti-Ro antibodies [70].

Rsr may also be involved with ncRNA metabolism. In most Gram-negative cells, the three components of rRNA (16S, 23S, 5S) are transcribed as a polycistronic transcript [82]. This transcript folds into several double-stranded regions that can be cleaved by the dsRNA-specific RNaseIII, separating the three components. Further maturation results from a variety of endoribonucleases and 3'→ 5' exoribonucleases (an enzyme with 5' → 3' exoribonuclease activity has not been discovered in bacteria), often in the context of the assembling ribosome. Under normal growth conditions (growth at 30°), 23S rRNA maturation in *D. radiodurans* is inefficient, which results in precursors with 5' and 3' extensions [83]. At elevated temperatures (37°), maturation is more efficient; this increase in efficiency is dependent on Rsr and two 3'→5' exoribonucleases, RNase PH and RNase II. While these observations have not been fully explained, one theory is that interactions with Rsr may affect the secondary structure of the pre-23S rRNA, allowing access by the exoribonucleases. Mature 23S rRNA is not necessary for normal cellular growth. In *E. coli* cells lacking either RNase III or RNase T, pre-23S molecules were incorporated into ribosomes and cells were viable [84, 85]. Perhaps fully matured ribosomes are only critical under stress conditions?

In contrast to the role of Rsr in rRNA maturation during heat stress in *D. radiodurans,* Rsr has been implicated in rRNA degradation in response to starvation-induced stress in the same organism. Rsr is up-regulated in stationary phase and after 3 days growth, wild-type cells show a competitive advantage over Δ*rsr* cells. [86]. At this time, wild-type cells showed extensive (98-99%) rRNA degradation. Compared to the wild-type, cells lacking Rsr and PNPase showed rRNA levels roughly 5- and 15-fold

higher, respectively. This indicates that Rsr is responsible, but perhaps not directly, for rRNA degradation after starvation. *D. radiodurans* does not possess an ortholog of RNase E, the major component of the *Salmonella* degradosome, which is responsible for rRNA and mRNA degradation [87, 88].

In *D. radiodurans,* Rsr was up-regulated after several stress treatments, including heat [83], starvation [86], desiccation [89], and both UV [80] and γ-irradiation [89], indicating that Rsr may be important to the cells' stress response. The fact that PNPase sedimented with Rsr in both heat stress and starvation conditions supports a model in which these two proteins interact. Recent work from the Wolin lab examined Rsr-PNPase complexes and showed that these two proteins function as a molecular machine, using Y RNA as a tether to hold them together [65]. While Rsr bound to the Y RNA stem, the PNPase interacted with the loop region (see below for Y RNA structure). Since this RNP complex was better than PNPase alone at degrading structured RNA molecules, they proposed that the Rsr-RNP complex could sculpt PNPase for degrading structured RNAs. Alternatively, the Rsr-RNP could alter RNA structures to make them more accessible for PNPase degradation. This work also examined Rsr from *Salmonella* and showed that, while both YrlA and YrlB co-immunoprecipitated with Rsr when anti-Rsr antibodies were used, only YrlA co-immunoprecipitated with Rsr and $FLAG_3$-PNPase when anti-$FLAG_3$ antibody was utilized (the *rsr-rtcBA* operon was expressed using a constitutively-active form of RtcR [15]) [65].

Since the *Salmonella rsr-rtcBA* operon is not transcribed under standard laboratory growth conditions, it is unlikely that Rsr plays a role in general ncRNA quality control or maturation. Instead, it seems more likely that it likely plays a role in response to a particular stress condition.

Y RNA

Every genome that contains Ro/Rsr is predicted to contain at least one gene encoding a Y RNA. Y RNAs are short molecules, ~100 nt. Though there is little conservation among Y RNAs at the primary sequence level, structural elements are conserved [90]. Y RNAs contain a large internal pyrimidine-rich

loop and a long stem generated by base-pairing between the 5' and 3' ends. Within this stem, there is a conserved bulge caused by a single cytidine residue. This bulge is important for the interaction between Y RNA and Ro [91]. As cells that have the Ro gene deleted show a decrease in Y RNA levels, binding to Ro is believed to protect Y RNAs from degradation [76, 81]. Y RNA appears to have three functions within the cell: regulating access to the central channel of Ro by substrate RNAs, serving as a molecular tether between Ro and other proteins, and controlling the sub-cellular localization of Ro. These structural elements are visible in *Salmonella* YrlA (Figure 1.4).

Y RNAs sterically inhibit binding of substrate RNAs due to partially overlapping binding sites on the outer surface of Ro [71]. Mis-folded variants of 5S rRNA compete with Y RNA for binding to purified Ro protein [91]. A mutant form of Rsr that could not bind Y RNA was more effective at processing 23S rRNA than wild-type Rsr, suggesting that Y RNA hinders effective processing at 30° [83]. Binding of PNPase to the Y RNA loop appears to alter Y RNA-Rsr binding, exposing the central channel [65]. Perhaps the association of a partner protein, which is communicated by a shift in Y RNA binding site, is the signal needed to make the central channel of Rsr accessible to substrate RNAs.

Y RNA affects sub-cellular localization apparently by blocking a nuclear accumulation signal on Ro, thus localizing it to the cytoplasm. In cells lacking Y RNA or with Ro mutants that do not interact with Y RNA, Ro accumulates in the nucleus [81]. Nuclear accumulation of Ro may play an important role in the cellular response to UV irradiation and oxidative damage. After exposure to either of these conditions, Ro localized to the nucleus of mouse cells [75, 92]. Since bacterial cells lack nuclei, this is an unlikely function for Y RNA in bacteria. Indeed, Rsr from *D. radiodurans* did not localize to the nucleus when ectopically expressed in mouse ES cells [92].

Conditions for expression of *rsr* in eubacteria have only been determined in *D. radiodurans*, which does not encode σ^{54}. However, of the 165 bacterial species in the Integrated Microbial Genetics database with an Rsr ortholog, 131 also encoded RtcR adjacent to Rsr and also encoded σ^{54} in the genome, indicating that expression of many of these *rsr*-containing operons is likely σ^{54}-dependent.

Some important questions remain: What is the environmental stimulus needed to activate RtcR and thus turn on transcription of the *rsr-rtcBA* operon? Does Rsr interact with other proteins in *Salmonella* in addition to PNPase? Naturally, RtcB and RtcA would be likely candidates for forming a complex with Rsr, but complexes with other RNA modifying proteins are possible. What are the substrate RNAs that are repaired (or targeted for degradation) by Rsr-RNP complexes? Since the conditions that induce the operon have yet to be determined, the RNA substrates present in the cell at the same time as these proteins is still unknown.

RtcB

RtcB is an RNA ligase with orthologs found widely in Bacteria, Archaea, and metazoa [67, 93, 94], but not in plants or fungi. RtcB catalyzes the GTP- and divalent cation-dependent ligation of RNA molecules bearing 2',3'-cyclic phosphate (2',3'>P) and 5'-OH termini [66, 67, 95] to generate a 3'-5' phosphodiester linkage. This activity is in contrast to "classical" RNA ligases, which are ATP-dependent and function on substrates bearing 3'-OH and 5'-PO_4 termini [96, 97]. Accordingly, a bioinformatic analysis of RtcB revealed no homology to other known DNA or RNA ligases, including the conserved KXXG motif that defines a covalent nucleotidyltransferase superfamily [95, 98].

The unique reaction catalyzed by RtcB proceeds in four steps (Figure 1.5): *(i)* a conserved histidine residue on RtcB reacts with GTP to create an RtcB-GMP adduct [99]; *(ii)* an intrinsic cyclic phosphodiesterase (CPDase) activity hydrolyzes the 2',3'>P to a 2'-OH/3'-PO_4 terminus [99]; *(iii)* RtcB transfers the GMP onto the 3'-PO_4, generating an RNA(3')pp(5')G intermediate [99-101]; *(iv)* finally, an attack by the 5'-OH on this intermediate releases the GMP and creates the final 3'-5' phosphodiester linkage [99, 100]. Interestingly, the revelation that RtcB possessed this intrinsic CPDase showed that RNA molecules with 3'-PO_4 (in addition to 2',3'>P) could be ligated, increasing the potential range of substrates for these enzymes [66, 99].

The human ortholog of RtcB, HSPC117 (later dubbed RTCB), was recently shown to be an essential component of the tRNA splicing pathway [94]. RNAi silencing of RTCB abolished inter-strand tRNA splicing *in vivo* and *in vitro* experiments performed with HeLa cell extracts demonstrated that an RTCB variant altered in the conserved metal-binding cysteine residue was defective in tRNA splicing. Immunoaffinity chromatography revealed that RTCB forms a stable complex with four other proteins: DDX1, CGI-99, FAM98B, and ASW. When these other genes were silenced via RNAi, little (if any) inhibition of tRNA splicing was observed, indicating that RTCB was the only component of this complex that was essential for tRNA ligation. The molecular functions of CGI-99, FAM98B, and ASW have not been characterized, but DDX1 is in a family of DEAD-box RNA helicases and has been shown to be involved in mRNA processing as well as in recognition of double-strand DNA breaks [102]. It was therefore hypothesized to enhance the function of the RTCB complex. This idea was supported by the finding that formation of the RTCB-guanylate adduct was DDX1-dependent [103].

DDX1-dependent RTCB guanylylation was accelerated by another protein, Archease, a 16 kDa protein whose presence also expanded the NTP cofactor specificity of RTCB (from solely GTP to GTP/ATP/ITP/dGTP) [103, 104]. While archease was necessary for sustained function of human and archaeal RTCB enzymes, *E. coli* RtcB was capable of guanylate adduct formation on its own [66, 100]. This domain-specific requirement for archease is reflected in genomic analysis. Archease orthologs have been reported in "all three domains," but bacterial representation is limited to ten distantly-related genomes [105], supporting the notion that archease is not as crucial to RtcB function in bacteria as it is in Archaea and eukaryotes.

The salient point in these findings is that other proteins interact with the human and archaeal orthologs of RtcB, enhancing RNA ligation. These findings raise the possibility that RtcB forms complexes in bacteria as well. Neither the human nor archaeal RtcB complexes showed an interaction between RtcB and homologs of Ro or RtcA, the other two components found alongside RtcB in the

Salmonella chromosome. However, this does not necessarily exclude the possibility of Rsr or RtcA participating in potential bacterial complexes.

Several lines of evidence led to "RNA repair" becoming the dogma for RtcB function. The archaeal and mammalian forms of RtcB were identified from either *Methanopyrus kandleri* or HeLa cell extracts which were successively purified to find a fraction that could ligate 2',3'>P/5'-OH terminated RNA molecules [93, 94]. In bacteria, the RNA repair function of RtcB was predicted due to its presence in an operon with the RNA 2',3'-phosphate cyclase (see below), which was 500-fold more active on RNA than on DNA substrates [106]. This prediction led to the characterization of *E. coli* RtcB primarily with RNA substrates [67, 68]. Early challenges to the dogma that RtcB is involved exclusively in RNA repair came from an experiment showing that the enzyme was capable of ligating a RNA 2',3'>P molecule to a DNA 5'-OH molecule [68]. More robust evidence for RtcB activity on DNA came from a recent report that this enzyme could ligate DNA strands with "dirty" ends—3'-PO_4/5'-OH termini that cannot be repaired by classical DNA ligases [107]. RtcB can catalyze 3'-5' phosphodiester bond formation between single-stranded ends, but is only able to add a 3'-GMP "cap" (instead of fully repairing) nicked DNA substrates. Several stressors create DNA with "dirty" ends: nucleases (e.g., micrococcal nuclease and phosphodiesterase II), ionizing radiation, and certain chemotherapeutics (e.g., bleomycin and neocarzinostatin) [97]. The finding that RtcB can catalyze DNA repair in addition to RNA raises several new possibilities for its physiological function.

RtcA

RtcA is a terminal phosphate cyclase that can catalyze the conversion of an RNA 3'-PO_4 end into a 2',3'>P moiety. This enzyme is widely distributed throughout bacteria, archaea, and eukaryotes [106, 108]. Though RtcA is structurally unrelated to RtcB and has a dissimilar active site [101, 109, 110], the reaction it catalyzes is mechanistically similar in that it involves a three step reaction with a nucleotidyl transfer event between the enzyme and the RNA terminal prior to formation of the final product. First, the

enzyme becomes adenylylated on a conserved histidine residue, forming RtcA-AMP and releasing PP_i [110, 111]. The AMP group is then transferred to an RNA 3'-PO_4 to generate an RNA-N(3')pp(5')A intermediate [112]. An attack on this intermediate by the 2'-OH group forms the 2',3'>P with the concomitant release of AMP. In *E. coli*, this reaction requires divalent cations (Mn^{2+} is preferred, but Mg^{2+} also works) and ATP [15]. By far, the preferred substrate for this enzyme is RNA 3'-PO_4 (compared to RNA 3'-OH, and DNA 3'-PO_4 or 3'-OH); DNA 3'-PO_4 could become cyclized when present at concentrations 2-3 orders of magnitude higher than RNA 3-PO_4 [15].

When RtcA enzymes were first characterized, their function was obvious: they were required to convert a 3'-PO_4 into a 2',3'>P that could serve as a substrate for RtcB-type ligases. As more was learned about the enzymes that cleave RNA molecules (e.g., PrrC or RNA splicing endonuclease)—specifically, that these enzymes often generate the 2',3'>P end themselves—and about RtcB's ability to act directly on 3'-PO_4 groups, the role for RtcA became less obvious. This led to the search for potential new catalytic roles for RtcA.

One study investigated the activity of RtcA at polynucleotide (both RNA and DNA 5'-PO_4) ends and found that this enzyme can adenylylate a 5'-PO_4 terminal as efficiently as a 3'-PO_4, generating either A(5')pp(5')-DNA or-RNA products [109]. Similar to RtcB, RtcA can act at a nick in dsDNA and can catalyze the early steps in a traditional ligase pathway (i.e., activation of the 5' end), but it is unable to completely seal the nick. Several roles for this 5'-adenylyltransferase activity were proposed, including protection against exonucleolytic decay, analogous to a eukaryotic 5' mRNA cap; "marking" the end of the DNA/RNA for downstream functions (e.g., localization); and generation of App-5' ends that can serve as substrates for subsequent repair steps.

Another study looked at the potential of RtcA to act at RNA 2'-PO_4 ends [113]. Unlike RtcB, which was inactive on RNA ends bearing a 2'-PO_4 group [66], RtcA was able to use these ends as substrates and catalyze the cyclization of an RNA2'-PO_4 into an RNA 2',3'>P [113]. Though the reaction with a 2'-PO_4 was much slower than with a 3'-PO_4 (minutes vs. milliseconds), the enzyme was capable of

cyclizing a similar percentage of each substrate. This finding raises the possibility that a role for RtcA may be conversion of 2'-PO_4 groups, which are not substrates for RtcB, into 2',3'>P groups, which *can* be utilized by RtcB. This model would require a means to generate 2'-PO_4 RNA ends. One such method is through the action of LigT homologs [discussed further below]. Similar to RtcB, these enzymes convert 2',3'>P and 5'-OH RNA termini into a phosphodiester linkage using CPDase and ligase activities. In contrast, the linkage generated by LigT is an unusual 2'-5' phosphodiester bond, indicating a 2'- instead of a 3'-PO_4 CPDase intermediate. Since the ligase activity of LigT is much weaker than its CPDase activity, this could be a source of 2-PO_4 groups within the cell [114].

Based on its observed *in vivo* activities, there are several physiological roles posited for RtcA. Understanding the environmental cues which lead to its expression (and the expression of Rsr/RtcB) *in vivo* will help elucidate which of these roles are important within the cell.

Functions of Characterized RNA Repair Systems

Prior to the research reported in this dissertation, the conditions under which the Rsr-RtcBA system is expressed in *Salmonella* were unknown and the physiological role of the system could not be addressed. However, other RNA repair systems that are present throughout all domains of life and in viruses have been characterized, and the details of their activities in cellular processes provide insight into physiological role of the *Salmonella* Rsr-RtcBA system. As previously mentioned, "RNA ligase" is a broad designation which commonly refers to enzymes that catalyze phosphodiester bond formation between RNA 3'-OH and 5'-PO_4 termini [96]. While most RNA ligases catalyze analogous reactions, this discussion will be limited to RNA *repair* ligases, which are specific for RNA termini bearing 2',3'>P and 5'-OH moieties.

tRNA maturation

One of the best-studied functions of RNA repair systems is the removal of an intron from and subsequent re-ligation of pre-tRNA molecules to generate mature tRNAs. Introns in tRNAs are short intervening sequences that must be removed in the maturation process. They are found in all three domains of life. In Archaea, up to 70% of tRNA genes may harbor 16-44 nt introns inserted at various sites [102]. In contrast, eukaryotic tRNA introns are rarer—only ~20% of yeast and ~6% of human tRNAs are disrupted by introns [102]. The location of eukaryotic introns within the tRNA gene is less variable than in Archaea, with most known eukaryotic tRNA introns interrupting the anticodon loop at the 3' end of the anticodon [102, 115]. Intron removal has two basic steps: the tRNA is first cleaved at the cleavage sites flanking the intron by a splicing endonuclease. The two pre-tRNA halves are then joined together by an RNA ligase (see below for a more detailed description of this process).

Though some bacterial tRNA genes have been found which contain introns, those characterized to date are all Group I "self-splicing" introns [116, 117]. Acting as a ribozyme, they are able to perform two consecutive transesterification reactions to remove themselves from the RNA sequence without a protein catalyst; however, this reaction may not be completely protein-free. While true (protein-free) self-splicing has been observed *in vitro,* these reactions were performed under salt and temperature conditions that were not physiologically relevant and were still 10- to 50- fold slower than reactions observed *in vivo* [118]. The proteins that are involved in Group I intron splicing *in vivo* are hypothesized to act as molecular chaperones, helping the intron fold into the proper conformation to allow splicing [118, 119]. While Group I introns—like those found in some bacterial tRNA genes—may require protein factors to facilitate their removal, they do not appear to require an RNA ligase to splice the exon halves back together [116, 117].

Repair of tRNA Cleaved by Ribotoxins

While some RNA repair systems function to repair programmed breaks in tRNA under normal growth conditions, other systems play a role in repairing breaks generated by ribotoxins (ribonucleases with toxic effects). A well-characterized example of a ribotoxin that cleaves a specific tRNA and the RNA repair system that repairs the cleaved tRNA is the *E. coli* PrrC riboendonuclease and bacteriophage T4 Pnkp/Rnl1 RNA repair system. When bacteriophage T4 infects a host cell, it co-opts the bacterial transcription and translational machinery to replicate as quickly as possible and lyse the cell, so that the newly-produced virions can infect naive cells. In its arsenal to fight T4 infections, *E. coli* has an anti-codon nuclease, PrrC. Once a cell detects a viral infection it activates PrrC, which cleaves $tRNA^{Lys}$ in the anticodon loop [120]. Depletion of $tRNA^{Lys}$ inhibits translation of late T4 proteins. Though this altruistic method also prevents translation of bacterial proteins, the phage infection is thus contained. The RNA repair system of T4, consisting of polynucleotide kinase/phosphatase (Pnkp) and RNA ligase (Rnl 1), is capable of repairing these cleaved tRNAs, allowing infection to continue [121]; the biochemistry of this repair is described below. This is one example of how a RNA repair system has developed to combat ribotoxic stress.

Unfolded Protein Response

The unfolded protein response (UPR) is a stress response that occurs when unfolded proteins accumulate in the endoplasmic reticulum lumen. The UPR is conserved in all eukaryotes [122]. One element that is required for initiation of the full UPR is the transcription factor HAC1. Full-length HAC1 transcript is constitutively-produced but unstable. Under conditions that induce the UPR, an ER-bound kinase/endoribonuclease , Ire1, initiates cleavage at two sites within the HAC1 transcript, removing a 252nt intron [123]. Unlike most spliced mRNAs, ligation of the two exons was not mediated by the spliceosome, but rather by the tRNA ligase (Trl1) [122, 124]. After this alternative splicing, a new ORF is created and the protein can enter the nucleus and activate other genes involved in the UPR. This novel

function for an RNA ligase complex demonstrates another instance in which one of these systems is involved in a cellular stress response.

When expressed heterologously in yeast cells, *E. coli* RtcB is capable of catalyzing the alternative splicing of *HAC1* mRNA in the UPR [68]. In bacteria, unfolded proteins trigger the σ^E-mediated envelope stress response [125]. Since mRNA splicing has not been described as part of this response in bacteria, it seems unlikely that this represents a physiological role for the Rsr-RtcBA system.

Biochemical Mechanisms of RNA Repair

There are two basic pathways for repairing a "broken" RNA, i.e. an intact RNA molecule that has been separated into a 5'piece with a 2',3'>P end and a 3' piece with a 5'-OH terminus. One of these is the "heal and seal"-type pathway in which the 2',3'>P is removed and the 5'-OH is converted to a 5-PO_4 prior to formation of the 5'-3' phosphodiester. The other is "direct" pathway that joins two RNA halves using the phosphate present in the 2',3'>P moiety. "RNA ligase" is a broad designation that commonly refers to enzymes that catalyze phosphodiester bond formation between RNA 3'-OH and 5'-PO_4 termini [96]; however, this discussion will be limited to RNA *repair* ligases, which are capable of repairing RNA termini bearing 2',3'>P and 5'-OH moieties.

"Heal and Seal" pathways

There are two variations on the "heal and seal" RNA repair pathway: the phage system, which was first characterized in bacteriophage T4 [121], and the yeast/plant system, which has been characterized from *Arabidopsis thaliana* [126] and *Saccharomyces cerevisiae* [127, 128]. These two systems are mechanistically similar. The first step in either of these systems is hydrolyzing the 2',3'-cyclic phosphate. Next, a kinase adds a phosphate group onto the 5'-OH, generating a 5'-PO_4. The ligase is activated by an adenylylation of an active site lysine residue. This adenyl group is transferred onto the 5'-PO_4 and is released during the formation of the 5'-3' phosphodiester bond.

There are two major differences between the T4 and the yeast/plant systems. T4 uses one protein, Pnkp, to perform the phosphatase and kinase functions and another, Rnl1, to catalyze the ligation of the two RNA ends. In contrast, the *Sc*TRL1 (*S. cerevisiae*) and *At*RNL (*A. thaliana*) proteins are capable of catalyzing all three reactions necessary for ligation. The other major difference between these two pathways is the fate of the 2',3'-cyclic phosphate. The phosphatase activity of T4 Pnkp removes the phosphate group, yielding a 2',3'-*cis* diol end while the cyclic phosphodiesterase of the plant/yeast enzymes generates a 2'-PO_4. A separate phosphotransferase is used to remove this group and leave a 2'-OH [129].

One other important difference between the T4 and the plant/yeast systems is in the structural specificity of the substrate. The phage enzyme is limited to the specific structure of tRNAs [130]. This specificity likely arose due to the nature of the repaired ends. Since the ends recognized by the phage enzyme, a 2',3'-*cis* diol is common to many cellular RNAs, the phage system may have evolved the tRNA specificity to avoid spurious "recombination." In contrast, the 3'-OH/2'-PO_4 ends recognized by plant/yeast enzymes are fairly specific to RNA molecules that are being "healed." Therefore, structural specificity may not be as crucial for these enzymes.

Despite these differences, the phage and yeast/plant systems are complementary. When cloned into yeast cells, the T4Pnkp/Rnl1 system was able to carry out both tRNA and alternative HAC1 mRNA splicing and repair damaged tRNAs [130, 131]. Similarly, both *At*RNL and *E. coli* RtcB were able to repair damaged yeast tRNAs [68, 130]. Based on conserved domain architecture and residues between the proteins and conserved mechanistic features in the pathways—particularly, the fact that the phosphate present in the 2',3'>P is not used in the ultimate phosphodiester linkage—point to a common evolutionary origin [98, 129].

While the namesakes of these pathways are phages, yeast, and plants, "heal and seal" RNA repair enzymes have also been identified in bacteria (*Deinococcus radiodurans* and *Clostridium thermocellum)* and eukaryotic baculoviruses [132-134] and genome analysis indicates that a heal and seal pathway may

be present in additional bacterial genera [135, 136]. The bacterial proteins appear to emulate the phage proteins in that *Cth*Pnkp was able to generate 5'-PO_4 and 2',3'-*cis* diol termini but was unable to ligate pre-tRNA halves on its own [135]. Instead of a Rnl1-like enzyme, bacterial Pnkp orthologs interact with a Hen1 ortholog *in vitro* [137]. Hen1 is a 2'-O-methyltransferase that catalyzes the addition of a 2'-methyl residue to the 3' terminal of RNA [136]. This methylation was shown to activate the ligase activity of Pnkp [138]. A Hen1 ortholog is contained in a putative operon with Pnkp in all 40 bacterial genomes where the latter is found [136]. One effect of the 2'-O-methylation is resistance to ribotoxins. After *in vitro* ligation with a Pnkp/Hen1 complex, tRNA molecules were resistant to cleavage with both colicin D and E5 [137]. A recent crystallographic and structure/function analysis of the ligase domain of *Cth*Pnkp placed these bacterial enzymes in a novel family of RNA ligases [139]. Given the novel ligase biochemistry and the interaction with Hen1 which generates RNA with a 2'-OMe moiety at the splice junction, these bacterial Pnkp/Hen1 systems appear to form a third variation of "heal and seal" RNA repair pathway. Perhaps one explanation for the divergence between the phage and the bacterial "heal and seal" systems, which both seem to repair damage caused by anticodon nucleases, is the length of time the repaired tRNAs are needed. Since the phage may only require the repaired tRNA for a short time before host cell lysis, it may not be advantageous to repair tRNAs such that they cannot be re-cleaved.

"Direct" pathways

As with the "heal and seal" pathway, there are two variations on the "direct" RNA repair pathway as well. These variations are classified by the type of phosphodiester linkage generated after ligation. The first "direct" pathway is catalyzed by the RtcB family of RNA ligases and yields a 3'-5' phosphodiester linkage. [For a detailed description of RtcB and its orthologs, please see the previous section].

The other type of "direct" RNA repair pathway is catalyzed by LigT, and results in a 2'-5' phosphodiester linkage, using the phosphate present in the cyclic phosphate [127, 140]. LigT orthologs have been discovered and characterized from both bacteria and archaea but, with the exception of small

number of protist and fungal genomes, have not been identified in eukaryotic genomes [102, 114, 140]. Recent work with spinach chloroplasts demonstrated an enzyme capable of generating 2'-5' phosphodiester linkages from 2',3'>P and 5'-OH termini, but no LigT orthologs was observed in the genome [141].

In addition to the unique linkage that LigT ligases generate, they are distinct from other RNA ligases in that they require neither a metal cofactor nor additional NTPs for activity [127, 141]. Since *in vitro* reactions reached equilibrium between ligation products and substrates [140, 141], cleavage of 2'-5' phosphodiester linkages cannot be excluded as an *in vivo* function of LigT. However, pools of RNA oligoadenylates with 2'-5' linkages were observed in *E. coli* cells at concentrations of 50-300 nM [142]. Since no other enzymes are known to generate these linkages, ligation seems to be the more plausible *in vivo* function. The physiological role of these 2'-5' linked oligoadenylates and of LigT is unclear. Interestingly, the cellular pools of these small molecules increased significantly after infection of the bacteria with phage M13 [142]. This parallels a mammalian response to viral infection in which interferon induces expression of the enzyme 2'-5' oligoadenylate synthetase (OAS) which generates 2'-5' oligoadenylate molecules [143]. These molecules activate a latent RNase, RNase L, which aids in cellular defense against viruses. LigT and OAS bear little sequence homology and generate the 2'-5' linkages in different fashions (ligation of RNA molecules vs. synthesis from ATP) but the analogous generation of 2'-5' oligoadenylates in response to viral infection is an interesting phenomenon.

Annotations on the Integrated Microbial Genomics (IMG) database show that >200 sequenced *Salmonella* genomes contain a *ligT* ortholog. Interestingly, all of these genomes show that *ligT* is in the same gene neighborhood as *dksA*, which can affect transcription of rRNA genes as part of the stringent response [144], a glutamyl-tRNA synthetase, and *pcnB*, the poly-A polymerase, which has broad-reaching functions in RNA metabolism [145]. Though the physiological roles of LigT have not been fully characterized, the available evidence suggests that this ligase may play a role in re-directing RNA metabolism in response to phage infection.

Potential roles of *rsr-rtcBA* in *Salmonella*

Knowing the *in vitro* functions of *E. coli* RtcB and RtcA as well as the physiological functions of RNA repair systems in humans, plants, yeast, and Archaea, we should be able to draw some reasonable inferences about the physiological functions of the *rsr-rtcBA* operon in *Salmonella.*

tRNA Splicing

Although RtcB RNA repair systems splice programmed breaks to form mature tRNAs in humans and Archaea [93, 94], a role in routine tRNA maturation would necessitate constitutive expression of these genes. Additionally, no intron-containing tRNAs have been reported in *Salmonella.* Therefore, routine tRNA maturation is an unlikely physiological function for this system. However, it could be responsible for processing an as yet un-described intron-containing tRNA required to recognize rare codons under specific stress conditions. There is precedent for minor tRNAs being required to express certain characteristics under specific growth conditions. In uropathogenic *E. coli*, a rare $tRNA^{Leu}$, *leuX*, recognizes UUG codons present in the *fimB* recombinase transcript [146]. FimB mediates the inversion of the promoter for the Type I fimbriae, *fimA*, into the "on" position. In the absence of the $tRNA^{Leu}$, reduced Fim-mediated inversion results in reduced pathogenicity [146]. The *bldA* gene encodes a $tRNA^{Leu}$ that recognizes the rare AAU codon in the high G+C *Streptomyces coelicolor* [147]. This gene is dispensable in young cultures, but is required for production of antibiotics and formation of aerial hyphae in the late stages of growth. Insertional mutagenesis of a rare $tRNA^{Ser}$ gene in *Haemophilus influenza* resulted in a strain that grew normally under laboratory conditions but was much more sensitive oxidative damage and severely attenuated for virulence in an infant rat model [148]. Together, these results indicate that rare tRNAs, while non-essential during normal growth can play an important role in translating proteins in specific stress responses. Though none of these tRNAs contain introns, removal of an intron from a tRNA by a stress-induced RNA repair system may be a method that *Salmonella* uses to ensure that genes

containing codons recognized by these tRNAs are ONLY expressed efficiently under the appropriate circumstances.

Conveniently, a potential intron-containing tRNA in *S.* Typhimurium LT2 and 14028s lies in the intergenic region between *rsr* and *rtcB,* partially overlapping *yrlA.* This region is annotated on IMG as a "pseudo-tRNA" and has high sequence homology to $tRNA^{Asn}$ except for a ~40nt intervening sequence near the anticodon (several other *Salmonella* strains have a 114 nt sequence annotated as $tRNA^{Val}$). Since the annotation has not been updated to reflect the recent characterization of *yrlA* and *yrlB* [65], it is possible that the pseudo-tRNA is the result of spurious annotation. However, it is also possible that this pseudo-tRNA is important to the cell under the conditions when the *rsr-rtcBA* operon is expressed. If it were functioning as a mature tRNA in translation, the pseudo-tRNA would need to be processed from the full-length *rsr-rtcBA* transcript and undergo further (potentially Rsr/RtcBA-mediated) processing by an unknown riboendonuclease to remove the intron and then RtcB ligate the two pre-tRNA halves. Northern blotting with probes specific for the intron region, *yrlA,* and the predicted tRNA could determine whether this sequence is processed to a size that resembles a mature tRNA.

Alternative mRNA Splicing

One of the functions of the yeast TRL1 complex is alternative splicing of the HAC1 mRNA during the unfolded protein response. Since the *E. coli* RtcB is capable of complementing this system in yeast cells, it should be capable of recapitulating this function in bacteria, if the proper substrates are present. Given that the lack of a nuclear membrane in bacteria allows the coupling of translation with transcription, mRNA splicing is a rare phenomenon in bacteria. To this end, as with tRNAs, only a handful of introns have been detected in bacterial mRNAs. These introns are all either Group I or Group II introns [118]. Group II introns, like the Group I introns described above, possess the genetic information relevant to self-splice but may require a protein chaperone *in vivo* to facilitate the folding needed for catalysis [149, 150]. The fact that that none of the introns described in bacterial mRNAs are

spliced via a protein catalyst would argue against a role of RtcB in alternative mRNA splicing. However, until the discovery that HAC1 was activated by alternative splicing event, this too was an unexpected phenomenon [123]. The idea that an as yet unknown mRNA could undergo RtcB-mediated splicing under stress is, therefore, not completely implausible.

Recovery from Ribotoxin Damage

Colicin ribotoxins: In addition to PrrC, there are several other known forms of ribotoxic damage. One of these is colicin proteins. These proteins are produced and secreted by *E. coli* cells to inhibit the growth of nearby, but unrelated *E. coli* strains. A second, immunity (Im) protein is co-produced with the colicin and is not secreted, protecting any cell that produces the same colicin. Over 20 different colicins have been identified [151], but three--Colicins D, E3, and E5—are of particular interest here. These proteins cleave 16S rRNA (E3), $tRNA^{Arg}$ (D), and $tRNA^{Asp}$ (E5), yielding 2',3'>P and 5'-OH termini [151]. Though there are no reports of the interactions of these colicins and RNA repair systems *in vivo*, Colicins D and E5 have been used to generate and characterize tRNA molecules for *in vitro* analysis [137, 152]. Similar activity has been observed in eukaryotes in the γ-toxin of the dairy yeast *Kluyveromyces lactis* (zymocin) [153]. This toxin cleaves the wobble codon of $tRNA^{Glu}$, *inhibiting the growth of S. cerevisiae. In vivo* work with the *K. lactis* γ-toxin showed that exogenous expression of either T4 or plant RNA ligase could confer toxin resistance to *S. cerevisiae* whereas its native RNA ligase left it susceptible [130].

To the best of our knowledge, no work on how colicins and RNA repair systems interact has been done *in vivo*. Most studies of *Salmonella*/colicin interaction focus on human/animal health, looking at susceptibility of *Salmonella* to colicins produced by other enterics found in the intestines. That said, some *Salmonella* isolates have been shown to produce colicin-like proteins [154, 155] while others are susceptible to the colicins produced by *E. coli* strains [156]. Given that the products generated by secreted

ribotoxins are potential substrates for RtcBA enzymes, it is a tantalizing possibility that the two are somehow connected.

A genomic search for colicins in *Salmonella* on IMG returned scant results. After sifting through dubious annotations for strains 14028s and LT2 [157, 158], one gene, *cirA*, stands out. CirA is a receptor for the Col-Ib pore-forming colicin [159]. Col-Ib is annotated in 2 different strains in both the Typhimurium and Heidelberg serovars of *S. enterica.* A recent study showed that production of Col-Ib conferred a competitive advantage to *S.* Typhimurium SL1344 over *E. coli* in a mouse infection [159]. This advantage was dependent on gut inflammation, highlighting the complex interactions that occur between different strains of bacteria and the host environment. Though Col-Ib is not itself a ribotoxin, *Salmonella* may find itself in the vicinity of another bacterium that is secreting ribotoxins. In this case, Rsr-RtcBA may be counteracting their effects.

Ribotoxins of toxin-antitoxin Systems: Another potential source of ribotoxic stress that may necessitate RNA repair is from toxin-antitoxin systems (TA). These systems are widely distributed through bacteria and consist of a stable toxin protein and a labile antitoxin molecule. Once the antitoxin is depleted, the toxin is able to inhibit cell growth (or cause cell death). There are five classes of TA systems that are organized by the nature of the antitoxin (i.e. RNA vs. protein) and by the antitoxin mechanism of inhibition (e.g. anti-sense RNA binding to inhibit toxin translation, direct protein-protein/RNA-protein interactions, etc.). Most of the ribotoxin-containing TA systems are Type II systems, in which a protein antitoxin sequesters the toxin molecule [160]. A recent genomic analysis of *S.* Typhimurium revealed eleven TA loci, with nine found in the chromosome and two encoded by the pSLT virulence plasmid [161]. Two of these TA systems, VapBC and DinJ-YafQ, are of particular interest for their potential relationship Rsr-RtcBA RNA repair system.

VapBC is the most abundant type II TA locus described in eubacteria and there are two copies in *Salmonella*, one in the chromosome and one on the pSLT plasmid [161, 162]. Production of VapC leads to bacteriostasis via inhibition of translation [162, 163]. This is accomplished by cleaving the initiator

$tRNA^{f\text{-}Met}$ within the anticodon loop, yielding 2',3'>P and 5'-OH RNA termini [162]. The presence of two VapC loci in *Salmonella* and the fact that it generates RNA substrates that are similar to those repaired by other RNA repair systems makes VapC damage an attractive target for Rsr-RtcBA repair.

The genetic location of the *dinJ-yafQ* operon—108 bp downstream of *rtcA*—makes it an interesting candidate for investigation. YafQ is an endoribonuclease that associates with the 50S ribosomal subunit and cleaves incoming mRNA after a lysine (AAA) codon [164]. Despite the fact that it targets mRNA (as opposed to tRNA), molecular modeling revealed that YafQ possesses a structural fold that is similar to those found in the tRNA-specific Colicins, D and E5 [165]. While the genetic proximity may be coincidental, considering that the *rsr-rtcBA* operons in non-*Salmonella* strains do not co-localize with toxin-antitoxin genes, the chromosomal juxtaposition of the *Salmonella* RNA repair system with a ribotoxin warrants investigation.

Understanding the functions of TA systems in cells may help illuminate the role of Rsr-RtcBA in repairing toxin-caused damage. Though post-segregational killing of cells that don't receive plasmids—as occurs with the *hok/sok* and *kid/kis* systems—is a well-accepted rationale for plasmid-borne TA systems [160, 166], the explanation for TA systems on bacterial chromosomes is more contentious. One interesting hypothesis is that chromosomal TA systems are involved with the phenomenon of persistence, a high level of drug tolerance by a subset of a bacterial population that has stochastically converted to a slow-growing state [167]. Since deletion of 10 Type II TA loci in *E. coli* led to a marked decrease in persister cell formation, a model was proposed in which the activity of one or more of these toxins could result in dormancy [168]. The return from dormancy could require damaged RNAs to be re-ligated, thus employing the Rsr-RtcBA system.

Another explanation for chromosomal TA systems involves resistance to phage infection. Compared to a Δ*mazEF* mutant, wild-type *E. coli* cells showed significantly fewer PFU/ml after induction of a P1 prophage and after treatment with a non-lysogenic variant of P1 [169]. The explanation for this observation is that MazF, an mRNA interferase that cleaves in a codon-specific, ribosome-independent

manner [170], can lead to the programmed cell death of infected cells to prevent the spread of phage to the population, a role analogous to PrrC. If the role of the TA system in question is to effect cell death, then expression of an RNA repair system may be counterproductive.

Given the nature of the products generated by some toxins (e.g., VapC-cleaved tRNA with 2',3'>P/5'-OH termini) and the genetic proximity of *yafQ* to the *rsr-rtcBA* locus, Rsr-RtcBA mediated repair of TA-damaged RNA is an attractive theory to pursue. If Rsr-RtcBA were responsible for repairing this damage, it would add an interesting new facet to the debate over the biological role for these systems.

Recovery from Environmental Stresses

Mitomycin C: Mitomycin C (MMC) is a member of the mitomycin family of antibiotics, derived from multiple species of *Streptomyces*, including *S. caespitosus.* In humans, MMC displays broad-spectrum antitumor activity and is a commonly used chemotherapeutic agent in treatment of a variety of cancers. Once inside, MMC is enzymatically reduced to its active form, a bifunctional alkylating agent, which reacts with dG to form MMC-mono-dG adducts, intrastrand bi-adducts at -GpG-, and interstrand cross-links within the sequence 5'-CpG-3' in DNA [171].

The crosslink between complementary strands of DNA has long been cited as the basis for the cytoxicity of MMC [172]; however, this dogma has been challenged ever since it was first reported [173]. Reports from other labs showed that MMC would lead to degradation of RNA and the decomposition of ribosomes [174, 175]. A recent report acknowledged that MMC could bind DNA but questioned the biological relevance of this interaction [176]. They showed that upon addition of MMC, 18S rRNA transcript levels decreased rapidly, indicative of a drug interaction with the RNA itself, instead of the slow decline resulting from an inability to synthesize new transcripts. They also provided evidence of potential MMC-RNA adducts, including retardation in agarose gels and inhibition of ethidium bromide binding to treated RNA. Since their argument against DNA binding *in vivo* was predicated on the idea that the nuclear membrane would prevent the association of MMC, the required bioactivating enzymes,

and DNA, DNA may still be a target for MMC activity in bacteria. As demonstrated by the research presented in Chapter 3, treatment of *Salmonella* with MMC activates expression of *rsr-rtcBA.* The stimulus that activates RtcR may be MMC-directed damage to cellular RNA; or in light of the finding that RtcB can catalyze repair of "dirty" DNA breaks [107], the stimulus for RtcR activation may be MMC-mediated crosslinking of DNA.

Oxidative Damage: A poster at the 2013 ASM Conference on *Salmonella* reported preliminary data of global gene regulation in *Salmonella* in response to treatment with sub-lethal concentrations of cefotaxime (a cephalosporin antibiotic); these data indicated increased expression of the *rsr-rtcBA* operon following treatment with cefotaxine (personal communication, Ashley Bono and Katherine Miller, who attended the conference). Though these results have not yet been published, if true, they may lead to interesting discoveries about the effects of certain antibiotics on cells as well as the cells' responses to these antibiotics.

It is not initially clear why damage from cefotaxime, which inhibits cell wall formation by binding to penicillin binding proteins [177], would upregulate an RNA repair system but an explanation for this interaction could be found in a new hypothesis that is emerging about the action of bactericidal antibiotics. Traditionally, cell death has been thought to occur as a result of interactions between the antibiotic and its cellular target e.g., the cell wall (β-lactams), DNA gyrase (quinolones), or ribosomes (aminoglycosides). New evidence is emerging that a portion of the damage associated with these agents is due to oxidative damage. Treating cells with bactericidal antibiotics generates large amounts of hydroxyl radicals; the lethal effects of these antibiotics can be mitigated by inhibiting peroxide generation with the iron chelator 2,2'-dipyridyl or using a hydroxyl radical scavenger (thiourea) [178, 179]. These effects are exacerbated in mutants with defective oxidative stress responses (*kat, sod, ahp*). Oxidative damage is also observed when cells are exposed to sub-lethal concentrations of these antibiotics [180].

Since oxidative damage can be extremely toxic to nucleic acids, Rsr-RtcBA may help cells respond to this stress. In *D. radiodurans*, transcription of *rsr* increased after treatment with H_2O_2 [80].

Since a similar, brief H_2O_2 treatment did not upregulate these genes in *Salmonella* [64], these genes may only be turned on after a more sustained exposure to an oxidative agent or that they respond to a different source of oxidative damage.

Starvation: Because ribosomes comprise such a large portion of the cell mass, rRNA degradation can lead to the release of nutrients (particularly carbon, phosphate, and nitrogen) during periods of nutrient limitation [181, 182]. In *Salmonella*, the degradation of rRNA can be particularly extensive, with >90% of 23S and ~50% of 16S rRNA molecules degraded upon entry into stationary phase. Given that *rsr* in *D. radiodurans* was up-regulated when the cells reach stationary phase and that Rsr was shown to be important in rRNA degradation [86], the components of the *Salmonella rsr-rtcBA* operon may play a role in the degradation of rRNA that occurs in stationary phase. Though growth to "early" and "late" stationary phase (OD_{600}=2.0 and OD_{600}=2.0+6 hours, respectively) did not result in an up-regulation of these genes [64], the results seen in *Deinococcus* were observed after three days of incubation [86]. Therefore, Rsr/RtcBA-mediated rRNA degradation in *Salmonella* may be a response to more long-term starvation.

Ultraviolet Radiation: Exposure to ultraviolet radiation is well known to cause damage to DNA molecules within a cell. Despite the chemical similarity between DNA and RNA, RNA damage is often overlooked when discussing the cellular effects of UV radiation. However, UV damaged RNA may have important consequences for cells. Damage to small RNAs, particularly in dsRNA regions, was recently implicated as one of the stimuli that can lead to inflammation associated with sunburns [183]. In studies conducted in bacteria and *in vitro*, UV radiation was shown to cause similar damage to RNA as it does to DNA (particularly, lesions at pyrimidine dimers) [79, 184]. In mammalian cells, even low levels of UV exposure are able to damage nucleotides within the ribosome and elicit the ribotoxic stress response [185]. This stress response, which occurs when damage to the ribosome prevents translational elongation, is analogous to the stringent response or the cold/heat-shock responses in prokaryotes, which are mounted in response to amino acid starvation or treatment with ribosome targeting antibiotics, respectively [186].

Since UV radiation generally results in RNA lesions containing cyclobutane dimers between pyrimidine molecules and not backbone breaks with 2',3'>P and 5'-OH termini, this damage seems unlikely to trigger a response from Rsr-RtcBA. This is supported by work in *E. coli*, where an increase in transcription from P*rtcBA* after a 60-second UV treatment was not observed [187]. However, since a response by Rsr-RtcBA has not been ruled out in *Salmonella*, it is conceivable that this damage, either directly or through a downstream signal, could require RNA repair by this system.

SOS Response: Given that several of the conditions under which the Rsr-RtcBA system may be useful to *Salmonella* can result in genotoxic stress, an overview of the SOS response is *de rigueur*. The SOS response is a broad regulatory network of genes widely-conserved throughout bacteria that are up-regulated in response to DNA damage within a cell [188]. Activation of these genes results in repair of DNA. Some of this DNA repair is by inherently error-prone mechanisms and the resulting mutations may result in increased fitness.

The SOS response is initiated by the presence of single-stranded DNA [188]. This ssDNA is recognized by RecA, which can bind the ssDNA strand, generating a nucleoprotein filament. This filament can interact with DNA-bound LexA repressor protein and stimulate its auto-cleavage. De-repression of LexA results in expression of a cascade of genes that are involved in the SOS response. In *E. coli*, the SOS regulon encompasses ~40 different genes [187], but sizes vary between different bacterial species [188].

The ssDNA that stimulates RecA binding and precipitates the SOS response can come from many sources. In the absence of external damaging factors, events such as replication fork stalling, RNAP stalling, or collisions between the replication and transcription machinery can all lead to ssDNA. These events will either generate ssDNA directly or create double-stranded breaks, which are repaired via an ssDNA intermediate by the RecBCD complex. Additionally, a wide variety of external forces have been shown to induce the SOS response. Ultraviolet light [189, 190], γ radiation [190, 191], bile salts [192],

phage infection [193], and treatment with antibiotics like mitomycin C [194], and fluoroquinolones [195, 196], have all been shown to induce the SOS response.

A few studies have examined the SOS response in *Salmonella,* but have not implicated the *rsr-rtcBA* operon in this response. One examined the effects of sub-lethal concentrations of nine different fluoroquinolones [195]. Using a random promoter library, this study identified 26 promoters that were up-regulated in response to treatment. Though $P_{rsr\text{-}rtcBA}$ was not identified in this analysis, the library screened contained only ~3000 clones, many of which were duplicates (the 26 promoters identified came from 83 different clones). Another study used arbitrarily-primed PCR to probe the SOS response after MMC exposure. In response to treatment, 20 genes were up-regulated, 19 of which were dependent on RecA [194]. At the time this work was done, the *S.* Typhimurium LT2 genome sequence had not been completed [197] and 15 of these genes were described as "novel." Using a Basic Local Alignment Search Tool (BLAST) analysis, I was able to determine that most of these sequences were sequences for prophage genes (Fels, Gifsy1, Gifsy2). Given the relatedness of *E. coli and Salmonella,* the fact that this experiment failed to detect most of the genes involved in the *E. coli* SOS response [187] (as well as *rsr-rtcBA,* as I describe in Chapter 3) indicates that this method was not highly sensitive.

Conclusion

All characterizations of the components of the *rsr-rtcBA* operon point toward this operon playing a role in RNA metabolism: quality control, maturation, degradation, repair, or some combination thereof. However no biological function can be ascribed, since expression of this operon had not heretofore been observed in wild type bacterial cells and deletion mutations have no phenotype. The first step toward characterizing this system was examining gene expression under conditions (including, but not limited to, those discussed above) that are known to damage RNA—producing potential substrates for these enzymes. Interestingly, there is a good deal of overlap between these stresses. Many of the stresses known to damage RNA (e.g. UV light, MMC, oxidative damage, etc.) will activate the SOS response [187, 189]. The SOS response can, in turn lead to expression on TA systems [198, 199] or induce prophage replication [193, 200]. If these stresses are activating the Rsr-RtcBA repair system, they may be acting through a common element. Therefore, by comparing the stresses that do and do not activate this operon, as described in Chapter 3, we can glean an understanding of mechanisms resulting in its induction.

Dissertation Goals

The primary objective of my dissertation is to gain a better understanding of the σ^{54} regulon in S*almonella.* In the first part (Chapter 2), I examined the regulon as a whole. Given the challenges associated with global analysis of σ^{54} regulons, particularly missing genes due to an inactive bEBP, I utilize a system involving a constitutively-active, promiscuous bEBP to define the regulon. Using this approach, I was able to identify sequences in the genome that were capable of acting as σ^{54}-dependent promoters and experimentally confirm several other predicted promoters. I also identified numerous binding sites for σ^{54} or Eσ^{54} throughout the chromosome. Most of these sites appear to be within coding sequences. The physiological role of many of these sites has yet to be determined but—whether they are serving as promoters for alternate open reading frames or anti-sense transcripts or if binding, per se, is important—these findings reveal an additional means by which σ^{54} can regulate transcription.

Chapter 3 focuses on one of the operons that was confirmed to be σ^{54}-dependent by my early experiments. The *rsr-rtcBA* operon contains components that are involved in RNA repair and maturation in all three domains of life. As yet, no conditions have been described in which this operon is expressed in *Salmonella*. My goal is to characterize this operon, which includes determining the conditions necessary to activate RtcR and induce expression of the operon and also identifying the substrates and any potential interacting partners of these proteins. Knowing this information will prove a major step toward ascribing a physiological role to this operon.

References

1. CDC: ***Salmonella***. 2012.

2. Fabrega A, Vila J: ***Salmonella enterica*** **serovar Typhimurium skills to succeed in the host: virulence and regulation**. *Clinical microbiology reviews* 2013, **26**(2):308-341.

3. Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K *et al*: **Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor**. *Molecular microbiology* 2008, **68**(3):538-546.

4. Bang IS, Frye JG, McClelland M, Velayudhan J, Fang FC: **Alternative sigma factor interactions in *Salmonella*: sigma and sigma promote antioxidant defences by enhancing sigma levels**. *Molecular microbiology* 2005, **56**(3):811-823.

5. Ferro-Luzzi Ames G, Nikaido K: **Nitrogen regulation in *Salmonella typhimurium*. Identification of an *ntrC* protein-binding site and definition of a consensus binding sequence**. *The EMBO journal* 1985, **4**(2):539-547.

6. Leonhartsberger S, Huber A, Lottspeich F, Bock A: **The *hydH/G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system**. *Journal of molecular biology* 2001, **307**(1):93-105.

7. Palacios S, Escalante-Semerena JC: ***prpR, ntrA,*** **and *ihf* functions are required for expression of the *prpBCDE* operon, encoding enzymes that catabolize propionate in *Salmonella enterica* serovar typhimurium LT2**. *J Bacteriol* 2000, **182**(4):905-910.

8. Weiner L, Brissette JL, Model P: **Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms**. *Genes & development* 1991, **5**(10):1912-1923.

9. Niehus E, Gressmann H, Ye F, Schlapbach R, Dehio M, Dehio C, Stack A, Meyer TF, Suerbaum S, Josenhans C: **Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori***. *Molecular microbiology* 2004, **52**(4):947-961.

10. Francke C, Kormelink TG, Hagemeijer Y, Overmars L, Sluijter V, Moezelaar R, Siezen RJ: **Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior**. *BMC genomics* 2011, **12**.

11. Keseler IM, Kaiser D: **sigma54, a vital protein for *Myxococcus xanthus***. *Proceedings of the National Academy of Sciences of the United States of America* 1997, **94**(5):1979-1984.

12. Leang C, Krushkal J, Ueki T, Puljic M, Sun J, Juarez K, Nunez C, Reguera G, DiDonato R, Postier B *et al*: **Genome-wide analysis of the RpoN regulon in *Geobacter sulfurreducens***. *BMC genomics* 2009, **10**:331.

13. Studholme DJ: **Enhancer-dependent transcription in *Salmonella enterica* Typhimurium: new members of the sigmaN regulon inferred from protein sequence homology and predicted promoter sites**. *Journal of molecular microbiology and biotechnology* 2002, **4**(4):367-374.

14. van Heeswijk WC, Hoving S, Molenaar D, Stegeman B, Kahn D, Westerhoff HV: **An alternative PII protein in the regulation of glutamine synthetase in *Escherichia coli***. *Molecular microbiology* 1996, **21**(1):133-146.

15. Genschik P, Drabikowski K, Filipowicz W: **Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon**. *The Journal of biological chemistry* 1998, **273**(39):25516-25526.

16. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S: **Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**(26):14674-14679.

17. Kiupakis AK, Reitzer L: **ArgR-independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli***. *J Bacteriol* 2002, **184**(11):2940-2950.

18. Riordan JT, Tietjen JA, Walsh CW, Gustafson JE, Whittam TS: **Inactivation of alternative sigma factor 54 (RpoN) leads to increased acid resistance, and alters locus of enterocyte effacement (LEE) expression in *Escherichia coli* O157 : H7**. *Microbiology* 2010, **156**(Pt 3):719-730.

19. Gardner AM, Gessner CR, Gardner PR: **Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and sigma54 in the nitric oxide stress response**. *The Journal of biological chemistry* 2003, **278**(12):10081-10086.

20. Maier T, Binder U, Bock A: **Analysis of the *hydA* locus of *Escherichia coli*: two genes (*hydN* and *hypF*) involved in formate and hydrogen metabolism**. *Archives of microbiology* 1996, **165**(5):333-341.

21. Lutz S, Bohm R, Beier A, Bock A: **Characterization of divergent NtrA-dependent promoters in the anaerobically expressed gene cluster coding for hydrogenase 3 components of *Escherichia coli***. *Molecular microbiology* 1990, **4**(1):13-20.

22. Janaszak A, Majczak W, Nadratowska B, Szalewska-Palasz A, Konopa G, Taylor A: **A sigma54-dependent promoter in the regulatory region of the *Escherichia coli rpoH* gene**. *Microbiology* 2007, **153**(Pt 1):111-123.

23. Ramirez-Santos J, Collado-Vides J, Garcia-Varela M, Gomez-Eichelmann MC: **Conserved regulatory elements of the promoter sequence of the gene *rpoH* of enteric bacteria**. *Nucleic acids research* 2001, **29**(2):380-386.

24. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, Buck M: **Identification of a new member of the phage shock protein response in *Escherichia coli*, the phage shock protein G (PspG)**. *The Journal of biological chemistry* 2004, **279**(53):55707-55714.

25. Birkmann A, Sawers RG, Bock A: **Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli***. *Molecular & general genetics : MGG* 1987, **210**(3):535-542.

26. Merrick MJ: **In a class of its own--the RNA polymerase sigma factor sigma 54 (sigma N)**. *Molecular microbiology* 1993, **10**(5):903-909.

27. Barrios H, Valderrama B, Morett E: **Compilation and analysis of sigma(54)-dependent promoter sequences**. *Nucleic acids research* 1999, **27**(22):4305-4313.

28. Zhang X, Chaney M, Wigneshweraraj SR, Schumacher J, Bordes P, Cannon W, Buck M: **Mechanochemical ATPases and transcriptional activation**. *Molecular microbiology* 2002, **45**(4):895-903.

29. Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD: **The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor**. *J Bacteriol* 2000, **182**(15):4129-4136.

30. Kornberg RD: **Mechanism and regulation of yeast RNA polymerase II transcription**. *Cold Spring Harbor symposia on quantitative biology* 1998, **63**:229-232.

31. Dombroski AJ, Walter WA, Record MT, Jr., Siegele DA, Gross CA: **Polypeptides containing highly conserved regions of transcription initiation factor sigma 70 exhibit specificity of binding to promoter DNA**. *Cell* 1992, **70**(3):501-512.

32. Buck M, Cannon W: **Specific binding of the transcription factor sigma-54 to promoter DNA**. *Nature* 1992, **358**(6385):422-424.

33. Cannon W, Claverie-Martin F, Austin S, Buck M: **Core RNA polymerase assists binding of the transcription factor sigma 54 to promoter DNA**. *Molecular microbiology* 1993, **8**(2):287-298.

34. Cannon W, Gallegos MT, Casaz P, Buck M: **Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization**. *Genes & development* 1999, **13**(3):357-370.

35. Guo Y, Gralla JD: **Promoter opening via a DNA fork junction binding activity**. *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95**(20):11655-11660.

36. Bush M, Dixon R: **The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription**. *Microbiology and molecular biology reviews : MMBR* 2012, **76**(3):497-529.

37. Bose D, Pape T, Burrows PC, Rappas M, Wigneshweraraj SR, Buck M, Zhang X: **Organization of an activator-bound RNA polymerase holoenzyme**. *Molecular cell* 2008, **32**(3):337-346.

38. Morett E, Segovia L: **The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains**. *J Bacteriol* 1993, **175**(19):6067-6074.

39. Bowman WC, Kranz RG: **A bacterial ATP-dependent, enhancer binding protein that activates the housekeeping RNA polymerase**. *Genes & development* 1998, **12**(12):1884-1893.

40. Xu H, Gu B, Nixon BT, Hoover TR: **Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a sigma54-dependent transcriptional activator**. *J Bacteriol* 2004, **186**(11):3499-3507.

41. Dixon R, Kahn D: **Genetic regulation of biological nitrogen fixation**. *Nature reviews Microbiology* 2004, **2**(8):621-631.

42. Delgado A, Ramos JL: **Genetic evidence for activation of the positive transcriptional regulator Xy1R, a member of the NtrC family of regulators, by effector binding**. *The Journal of biological chemistry* 1994, **269**(11):8059-8062.

43. Delgado A, Salto R, Marques S, Ramos JL: **Single amino acids changes in the signal receptor domain of XylR resulted in mutants that stimulate transcription in the absence of effectors**. *The Journal of biological chemistry* 1995, **270**(10):5144-5150.

44. Perez-Martin J, De Lorenzo V: **The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor**. *Proceedings of the National Academy of Sciences of the United States of America* 1995, **92**(20):9392-9396.

45. D'Autreaux B, Tucker N, Spiro S, Dixon R: **Characterization of the nitric oxide-reactive transcriptional activator NorR**. *Methods in enzymology* 2008, **437**:235-251.

46. Studholme DJ, Dixon R: **Domain architectures of sigma54-dependent transcriptional activators**. *J Bacteriol* 2003, **185**(6):1757-1767.

47. Joly N, Engl C, Jovanovic G, Huvet M, Toni T, Sheng X, Stumpf MP, Buck M: **Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology**. *FEMS microbiology reviews* 2010, **34**(5):797-827.

48. Gu B, Lee JH, Hoover TR, Scholl D, Nixon BT: ***Rhizobium meliloti* DctD, a sigma 54-dependent transcriptional activator, may be negatively controlled by a subdomain in the C-terminal end of its two-component receiver module**. *Molecular microbiology* 1994, **13**(1):51-66.

49. Lee JH, Scholl D, Nixon BT, Hoover TR: **Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of *Rhizobium meliloti* DCTD, a sigma 54-dependent transcriptional activator**. *The Journal of biological chemistry* 1994, **269**(32):20401-20409.

50. Fernandez S, de Lorenzo V, Perez-Martin J: **Activation of the transcriptional regulator XylR of *Pseudomonas putida* by release of repression between functional domains**. *Molecular microbiology* 1995, **16**(2):205-213.

51. Jyot J, Dasgupta N, Ramphal R: **FleQ, the major flagellar gene regulator in *Pseudomonas aeruginosa*, binds to enhancer sites located either upstream or atypically downstream of the RpoN binding site**. *J Bacteriol* 2002, **184**(19):5251-5260.

52. Belitsky BR, Sonenshein AL: **An enhancer element located downstream of the major glutamate dehydrogenase gene of *Bacillus subtilis***. *Proceedings of the National Academy of Sciences of the United States of America* 1999, **96**(18):10290-10295.

53. Reitzer LJ, Magasanik B: **Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter**. *Cell* 1986, **45**(6):785-792.

54. Beck LL, Smith TG, Hoover TR: **Look, no hands! Unconventional transcriptional activators in bacteria**. *Trends in microbiology* 2007, **15**(12):530-537.

55. Poggio S, Osorio A, Dreyfus G, Camarena L: **The four different sigma(54) factors of *Rhodobacter sphaeroides* are not functionally interchangeable**. *Molecular microbiology* 2002, **46**(1):75-85.

56. North AK, Kustu S: **Mutant forms of the enhancer-binding protein NtrC can activate transcription from solution**. *Journal of molecular biology* 1997, **267**(1):17-36.

57. Jovanovic G, Rakonjac J, Model P: **In vivo and in vitro activities of the *Escherichia coli* sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFDeltaHTH**. *Journal of molecular biology* 1999, **285**(2):469-483.

58. Huala E, Stigter J, Ausubel FM: **The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription**. *J Bacteriol* 1992, **174**(4):1428-1431.

59. Cases I, Ussery DW, de Lorenzo V: **The sigma54 regulon (sigmulon) of *Pseudomonas putid*a**. *Environmental microbiology* 2003, **5**(12):1281-1293.

60. Zhao K, Liu M, Burgess RR: **Promoter and regulon analysis of nitrogen assimilation factor, sigma54, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis**. *Nucleic acids research* 2010, **38**(4):1273-1283.

61. Dong TG, Mekalanos JJ: **Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52**. *Nucleic acids research* 2012, **40**(16):7766-7775.

62. Samuels DJ, Frye JG, Porwollik S, McClelland M, Mrazek J, Hoover TR, Karls AC: **Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the sigma(5)(4) (RpoN) regulon of *Salmonella* Typhimurium LT2**. *BMC genomics* 2013, **14**:602.

63. Ygberg SE, Clements MO, Rytkonen A, Thompson A, Holden DW, Hinton JC, Rhen M: **Polynucleotide phosphorylase negatively controls spv virulence gene expression in *Salmonella enterica***. *Infection and immunity* 2006, **74**(2):1243-1254.

64. Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K *et al*: **An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium**. *Cell host & microbe* 2013, **14**(6):683-695.

65. Chen X, Taylor DW, Fowler CC, Galan JE, Wang HW, Wolin SL: **An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA**. *Cell* 2013, **153**(1):166-177.

66. Tanaka N, Chakravarty AK, Maughan B, Shuman S: **Novel mechanism of RNA repair by RtcB via sequential 2',3'-cyclic phosphodiesterase and 3'-Phosphate/5'-hydroxyl ligation reactions**. *The Journal of biological chemistry* 2011, **286**(50):43134-43143.

67. Tanaka N, Shuman S: **RtcB is the RNA ligase component of an *Escherichia coli* RNA repair operon**. *The Journal of biological chemistry* 2011, **286**(10):7727-7731.

68. Tanaka N, Meineke B, Shuman S: **RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing in vivo**. *The Journal of biological chemistry* 2011, **286**(35):30253-30257.

69. Sim S, Wolin SL: **Emerging roles for the Ro 60-kDa autoantigen in noncoding RNA metabolism**. *Wiley interdisciplinary reviews RNA* 2011, **2**(5):686-699.

70. Provost TT, Watson R, Simmons-O'Brien E: **Significance of the anti-Ro (SS-A) antibody in evaluation of patients with cutaneous manifestations of a connective tissue disease**. *Journal of the American Academy of Dermatology* 1996, **35**(2 Pt 1):147-169; quiz 170-142.

71. Stein AJ, Fuchs G, Fu C, Wolin SL, Reinisch KM: **Structural insights into RNA quality control: the Ro autoantigen binds misfolded RNAs via its central cavity**. *Cell* 2005, **121**(4):529-539.

72. Triman KL: **The 23S Ribosomal RNA Mutation Database (23SMDB)**. *Nucleic acids research* 1996, **24**(1):169-171.

73. Triman KL: **The 16S ribosomal RNA mutation database (16SMDB)**. *Nucleic acids research* 1996, **24**(1):166-168.

74. O'Brien CA, Wolin SL: **A possible role for the 60-kD Ro autoantigen in a discard pathway for defective 5S rRNA precursors**. *Genes & development* 1994, **8**(23):2891-2903.

75. Chen X, Smith JD, Shi H, Yang DD, Flavell RA, Wolin SL: **The Ro autoantigen binds misfolded U2 small nuclear RNAs and assists mammalian cell survival after UV irradiation**. *Current biology : CB* 2003, **13**(24):2206-2211.

76. Labbe JC, Hekimi S, Rokeach LA: **The levels of the RoRNP-associated Y RNA are dependent upon the presence of ROP-1, the *Caenorhabditis elegans* Ro60 protein**. *Genetics* 1999, **151**(1):143-150.

77. Shi H, O'Brien CA, Van Horn DJ, Wolin SL: **A misfolded form of 5S rRNA is complexed with the Ro and La autoantigens**. *Rna* 1996, **2**(8):769-784.

78. Fuchs G, Stein AJ, Fu C, Reinisch KM, Wolin SL: **Structural and biochemical basis for misfolded RNA recognition by the Ro autoantigen**. *Nature structural & molecular biology* 2006, **13**(11):1002-1009.

79. Doudney CO: **Ultraviolet light effects on the bacterial cell**. *Current topics in microbiology and immunology* 1968, **46**:116-175.

80. Chen X, Quinn AM, Wolin SL: **Ro ribonucleoproteins contribute to the resistance of Deinococcus radiodurans to ultraviolet irradiation**. *Genes & development* 2000, **14**(7):777-782.

81. Xue D, Shi H, Smith JD, Chen X, Noe DA, Cedervall T, Yang DD, Eynon E, Brash DE, Kashgarian M *et al*: **A lupus-like syndrome develops in mice lacking the Ro 60-kDa protein, a major lupus autoantigen**. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(13):7503-7508.

82. Deutscher MP: **Maturation and degradation of ribosomal RNA in bacteria**. *Progress in molecular biology and translational science* 2009, **85**:369-391.

83. Chen X, Wurtmann EJ, Van Batavia J, Zybailov B, Washburn MP, Wolin SL: **An ortholog of the Ro autoantigen functions in 23S rRNA maturation in *D. radiodurans***. *Genes & development* 2007, **21**(11):1328-1339.

84. King TC, Sirdeshmukh R, Schlessinger D: **RNase III cleavage is obligate for maturation but not for function of Escherichia coli pre-23S rRNA**. *Proceedings of the National Academy of Sciences of the United States of America* 1984, **81**(1):185-188.

85. Li Z, Pandit S, Deutscher MP: **Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T**. *Rna* 1999, **5**(1):139-146.

86. Wurtmann EJ, Wolin SL: **A role for a bacterial ortholog of the Ro autoantigen in starvation-induced rRNA degradation**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(9):4022-4027.

87. Carpousis AJ: **The RNA degradosome of *Escherichia coli*: an mRNA-degrading machine assembled on RNase E**. *Annual review of microbiology* 2007, **61**:71-87.

88. Hammarlof DL, Liljas L, Hughes D: **Temperature-sensitive mutants of RNase E in *Salmonella enterica***. *J Bacteriol* 2011, **193**(23):6639-6650.

89. Tanaka M, Earl AM, Howell HA, Park MJ, Eisen JA, Peterson SN, Battista JR: **Analysis of *Deinococcus radiodurans's* transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance**. *Genetics* 2004, **168**(1):21-33.

90. Teunissen SW, Kruithof MJ, Farris AD, Harley JB, Venrooij WJ, Pruijn GJ: **Conserved features of Y RNAs: a comparison of experimentally derived secondary structures**. *Nucleic acids research* 2000, **28**(2):610-619.

91. Green CD, Long KS, Shi H, Wolin SL: **Binding of the 60-kDa Ro autoantigen to Y RNAs: evidence for recognition in the major groove of a conserved helix**. *Rna* 1998, **4**(7):750-765.

92. Sim S, Weinberg DE, Fuchs G, Choi K, Chung J, Wolin SL: **The subcellular distribution of an RNA quality control protein, the Ro autoantigen, is regulated by noncoding Y RNA binding**. *Molecular biology of the cell* 2009, **20**(5):1555-1564.

93. Englert M, Sheppard K, Aslanian A, Yates JR, 3rd, Soll D: **Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(4):1290-1295.

94. Popow J, Englert M, Weitzer S, Schleiffer A, Mierzwa B, Mechtler K, Trowitzsch S, Will CL, Luhrmann R, Soll D *et al*: **HSPC117 is the essential subunit of a human tRNA splicing ligase complex**. *Science* 2011, **331**(6018):760-764.

95. Desai KK, Raines RT: **tRNA ligase catalyzes the GTP-dependent ligation of RNA with 3'-phosphate and 5'-hydroxyl termini**. *Biochemistry* 2012, **51**(7):1333-1335.

96. Pascal JM: **DNA and RNA ligases: structural variations and shared mechanisms**. *Current opinion in structural biology* 2008, **18**(1):96-105.

97. Huang SY, Pommier Y: **Cross-over of RNA 3'-phosphate ligase into the DNA world**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(51):20354-20355.

98. Shuman S, Lima CD: **The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases**. *Current opinion in structural biology* 2004, **14**(6):757-764.

99. Chakravarty AK, Shuman S: **The sequential 2',3'-cyclic phosphodiesterase and 3'-phosphate/5'-OH ligation steps of the RtcB RNA splicing pathway are GTP-dependent**. *Nucleic acids research* 2012, **40**(17):8558-8567.

100. Chakravarty AK, Subbotin R, Chait BT, Shuman S: **RNA ligase RtcB splices 3'-phosphate and 5'-OH ends via covalent RtcB-(histidinyl)-GMP and polynucleotide-(3')pp(5')G intermediates**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(16):6072-6077.

101. Englert M, Xia S, Okada C, Nakamura A, Tanavde V, Yao M, Eom SH, Konigsberg WH, Soll D, Wang J: **Structural and mechanistic insights into guanylylation of RNA-splicing ligase RtcB joining RNA between 3'-terminal phosphate and 5'-OH**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(38):15235-15240.

102. Popow J, Schleiffer A, Martinez J: **Diversity and roles of (t)RNA ligases**. *Cellular and molecular life sciences : CMLS* 2012, **69**(16):2657-2670.

103. Popow J, Jurkin J, Schleiffer A, Martinez J: **Analysis of orthologous groups reveals archease and DDX1 as tRNA splicing factors**. *Nature* 2014.

104. Desai KK, Cheng CL, Bingman CA, Phillips GN, Jr., Raines RT: **A tRNA splicing operon: Archease endows RtcB with dual GTP/ATP cofactor specificity and accelerates RNA ligation**. *Nucleic acids research* 2014, **42**(6):3931-3942.

105. Canaves JM: **Predicted role for the archease protein family based on structural and sequence analysis of TM1083 and MTH1598, two proteins structurally characterized through structural genomics efforts**. *Proteins* 2004, **56**(1):19-27.

106. Genschik P, Billy E, Swianiewicz M, Filipowicz W: **The human RNA 3'-terminal phosphate cyclase is a member of a new family of proteins conserved in Eucarya, Bacteria and Archaea**. *The EMBO journal* 1997, **16**(10):2955-2967.

107. Das U, Chakravarty AK, Remus BS, Shuman S: **Rewriting the rules for end joining via enzymatic splicing of DNA 3'-PO4 and 5'-OH ends**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(51):20437-20442.

108. Filipowicz W, Billy E, Drabikowski K, Genschik P: **Cyclases of the 3'-terminal phosphate in RNA: a new family of RNA processing enzymes conserved in eucarya, bacteria and archaea**. *Acta biochimica Polonica* 1998, **45**(4):895-906.

109. Chakravarty AK, Shuman S: **RNA 3'-phosphate cyclase (RtcA) catalyzes ligase-like adenylylation of DNA and RNA 5'-monophosphate ends**. *The Journal of biological chemistry* 2011, **286**(6):4117-4122.

110. Tanaka N, Smith P, Shuman S: **Structure of the RNA 3'-phosphate cyclase-adenylate intermediate illuminates nucleotide specificity and covalent nucleotidyl transfer**. *Structure* 2010, **18**(4):449-457.

111. Chakravarty AK, Smith P, Shuman S: **Structures of RNA 3'-phosphate cyclase bound to ATP reveal the mechanism of nucleotidyl transfer and metal-assisted catalysis**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(52):21034-21039.

112. Filipowicz W, Strugala K, Konarska M, Shatkin AJ: **Cyclization of RNA 3'-terminal phosphate by cyclase from HeLa cells proceeds via formation of N(3')pp(5')A activated intermediate**. *Proceedings of the National Academy of Sciences of the United States of America* 1985, **82**(5):1316-1320.

113. Das U, Shuman S: **2'-Phosphate cyclase activity of RtcA: a potential rationale for the operon organization of RtcA with an RNA repair ligase RtcB in *Escherichia coli* and other bacterial taxa**. *Rna* 2013, **19**(10):1355-1362.

114. Kanai A, Sato A, Fukuda Y, Okada K, Matsuda T, Sakamoto T, Muto Y, Yokoyama S, Kawai G, Tomita M: **Characterization of a heat-stable enzyme possessing GTP-dependent RNA ligase activity from a hyperthermophilic archaeon, *Pyrococcus furiosus***. *Rna* 2009, **15**(3):420-431.

115. Abelson J, Trotta CR, Li H: **tRNA splicing**. *The Journal of biological chemistry* 1998, **273**(21):12685-12688.

116. Hausner G, Hafez M, Edgell DR: **Bacterial group I introns: mobile RNA catalysts**. *Mobile DNA* 2014, **5**(1):8.

117. Reinhold-Hurek B, Shub DA: **Self-splicing introns in tRNA genes of widely divergent bacteria**. *Nature* 1992, **357**(6374):173-176.

118. Woodson SA: **Ironing out the kinks: splicing and translation in bacteria**. *Genes & development* 1998, **12**(9):1243-1247.

119. Lambowitz AM, Perlman PS: **Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing**. *Trends in biochemical sciences* 1990, **15**(11):440-444.

120. Kaufmann G: **Anticodon nucleases**. *Trends in biochemical sciences* 2000, **25**(2):70-74.

121. Amitsur M, Levitz R, Kaufmann G: **Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA**. *The EMBO journal* 1987, **6**(8):2499-2503.

122. Sidrauski C, Chapman R, Walter P: **The unfolded protein response: an intracellular signalling pathway with many surprising features**. *Trends in cell biology* 1998, **8**(6):245-249.

123. Cox JS, Walter P: **A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response**. *Cell* 1996, **87**(3):391-404.

124. Sidrauski C, Cox JS, Walter P: **tRNA ligase is required for regulated mRNA splicing in the unfolded protein response**. *Cell* 1996, **87**(3):405-413.

125. Truscott KN, Bezawork-Geleta A, Dougan DA: **Unfolded protein responses in bacteria and mitochondria: a central role for the ClpXP machine**. *IUBMB life* 2011, **63**(11):955-963.

126. Englert M, Beier H: **Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins**. *Nucleic acids research* 2005, **33**(1):388-399.

127. Greer CL, Javor B, Abelson J: **RNA ligase in bacteria: formation of a 2',5' linkage by an *E. coli* extract**. *Cell* 1983, **33**(3):899-906.

128. Phizicky EM, Schwartz RC, Abelson J: ***Saccharomyces cerevisiae* tRNA ligase. Purification of the protein and isolation of the structural gene**. *The Journal of biological chemistry* 1986, **261**(6):2978-2986.

129. McCraith SM, Phizicky EM: **A highly specific phosphatase from *Saccharomyces* cerevisiae implicated in tRNA splicing**. *Molecular and cellular biology* 1990, **10**(3):1049-1055.

130. Nandakumar J, Schwer B, Schaffrath R, Shuman S: **RNA repair: an antidote to cytotoxic eukaryal RNA damage**. *Molecular cell* 2008, **31**(2):278-286.

131. Schwer B, Sawaya R, Ho CK, Shuman S: **Portability and fidelity of RNA-repair systems**. *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**(9):2788-2793.

132. Martins A, Shuman S: **An end-healing enzyme from *Clostridium thermocellum* with 5' kinase, 2',3' phosphatase, and adenylyltransferase activities**. *Rna* 2005, **11**(8):1271-1280.

133. Martins A, Shuman S: **An RNA ligase from *Deinococcus radiodurans***. *The Journal of biological chemistry* 2004, **279**(49):50654-50661.

134. Martins A, Shuman S: **Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5'-kinase, and polynucleotide 3'-phosphatase activities**. *The Journal of biological chemistry* 2004, **279**(18):18220-18231.

135. Keppetipola N, Nandakumar J, Shuman S: **Reprogramming the tRNA-splicing activity of a bacterial RNA repair enzyme**. *Nucleic acids research* 2007, **35**(11):3624-3630.

136. Jain R, Shuman S: **Active site mapping and substrate specificity of bacterial Hen1, a manganese-dependent 3' terminal RNA ribose 2'O-methyltransferase**. *Rna* 2011, **17**(3):429-438.

137. Chan CM, Zhou C, Huang RH: **Reconstituting bacterial RNA repair and modification in vitro**. *Science* 2009, **326**(5950):247.

138. Wang P, Chan CM, Christensen D, Zhang C, Selvadurai K, Huang RH: **Molecular basis of bacterial protein Hen1 activating the ligase activity of bacterial protein Pnkp for RNA repair**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(33):13248-13253.

139. Smith P, Wang LK, Nair PA, Shuman S: **The adenylyltransferase domain of bacterial Pnkp defines a unique RNA ligase family**. *Proceedings of the National Academy of Sciences of the United States of America* 2012, **109**(7):2296-2301.

140. Arn EA, Abelson JN: **The 2'-5' RNA ligase of Escherichia coli. Purification, cloning, and genomic disruption**. *The Journal of biological chemistry* 1996, **271**(49):31145-31153.

141. Molina-Serrano D, Marques J, Nohales MA, Flores R, Daros JA: **A chloroplastic RNA ligase activity analogous to the bacterial and archaeal 2 -5' RNA ligase**. *RNA biology* 2012, **9**(3):326-333.

142. Trujillo MA, Roux D, Fueri JP, Samuel D, Cailla HL, Rickenberg HV: **The occurrence of 2'-5' oligoadenylates in *Escherichia coli***. *European journal of biochemistry / FEBS* 1987, **169**(1):167-173.

143. Chakrabarti A, Jha BK, Silverman RH: **New insights into the role of RNase L in innate immunity**. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* 2011, **31**(1):49-57.

144. Sharma UK, Chatterji D: **Transcriptional switching in *Escherichia coli* during stress and starvation by modulation of sigma activity**. *FEMS microbiology reviews* 2010, **34**(5):646-657.

145. Mohanty BK, Kushner SR: **Analysis of the function of *Escherichia coli* poly(A) polymerase I in RNA metabolism**. *Molecular microbiology* 1999, **34**(5):1094-1108.

146. Ritter A, Gally DL, Olsen PB, Dobrindt U, Friedrich A, Klemm P, Hacker J: **The Pai-associated *leuX* specific tRNA5(Leu) affects type 1 fimbriation in pathogenic *Escherichia coli* by control of FimB recombinase expression**. *Molecular microbiology* 1997, **25**(5):871-882.

147. Leskiw BK, Mah R, Lawlor EJ, Chater KF: **Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2)**. *J Bacteriol* 1993, **175**(7):1995-2005.

148. Gerlach G, Anthony M, Deadman M, Schoen C, Hood D, Reidl J: **Transposon insertion in a serine-specific minor tRNA coding sequence affects intraperitoneal survival of *Haemophilus influenzae* in the infant rat model**. *International journal of medical microbiology : IJMM* 2010, **300**(4):218-228.

149. Michel F, Ferat JL: **Structure and activities of group II introns**. *Annual review of biochemistry* 1995, **64**:435-461.

150. Bonen L, Vogel J: **The ins and outs of group II introns**. *Trends in genetics : TIG* 2001, **17**(6):322-331.

151. Papadakos G, Wojdyla JA, Kleanthous C: **Nuclease colicins and their immunity proteins**. *Quarterly reviews of biophysics* 2012, **45**(1):57-103.

152. Zhang C, Chan CM, Wang P, Huang RH: **Probing the substrate specificity of the bacterial Pnkp/Hen1 RNA repair system using synthetic RNAs**. *Rna* 2012, **18**(2):335-344.

153. Jablonowski D, Schaffrath R: **Zymocin, a composite chitinase and tRNase killer toxin from yeast**. *Biochemical Society transactions* 2007, **35**(Pt 6):1533-1537.

154. Fredericq P: **Colicins**. *Annual review of microbiology* 1957, **11**:7-22.

155. Simmons KW, Wooley RE, Brown J: **Comparison of virulence factors and R plasmids of *Salmonella* spp. isolated from healthy and ill swine**. *Applied and environmental microbiology* 1988, **54**(3):760-767.

156. Zihler A, Le Blay G, de Wouters T, Lacroix C, Braegger CP, Lehner A, Tischler P, Rattei T, Hachler H, Stephan R: **In vitro inhibition activity of different bacteriocin-producing *Escherichia coli* against *Salmonella* strains isolated from clinical cases**. *Letters in applied microbiology* 2009, **49**(1):31-38.

157. Fath MJ, Mahanty HK, Kolter R: **Characterization of a *purF* operon mutation which affects colicin V production**. *J Bacteriol* 1989, **171**(6):3158-3161.

158. Cariss SJ, Constantinidou C, Patel MD, Takebayashi Y, Hobman JL, Penn CW, Avison MB: **YieJ (CbrC) mediates CreBC-dependent colicin E2 tolerance in *Escherichia coli***. *J Bacteriol* 2010, **192**(13):3329-3336.

159. Nedialkova LP, Denzler R, Koeppel MB, Diehl M, Ring D, Wille T, Gerlach RG, Stecher B: **Inflammation fuels colicin Ib-dependent competition of *Salmonella* serovar *Typhimurium* and *E. coli* in enterobacterial blooms**. *PLoS pathogens* 2014, **10**(1):e1003844.

160. Yamaguchi Y, Park JH, Inouye M: **Toxin-antitoxin systems in bacteria and archaea**. *Annual review of genetics* 2011, **45**:61-79.

161. De la Cruz MA, Zhao W, Farenc C, Gimenez G, Raoult D, Cambillau C, Gorvel JP, Meresse S: **A toxin-antitoxin module of *Salmonella* promotes virulence in mice**. *PLoS pathogens* 2013, **9**(12):e1003827.

162. Winther KS, Gerdes K: **Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(18):7403-7407.

163. Winther KS, Gerdes K: **Ectopic production of VapCs from Enterobacteria inhibits translation and trans-activates YoeB mRNA interferase**. *Molecular microbiology* 2009, **72**(4):918-930.

164. Prysak MH, Mozdzierz CJ, Cook AM, Zhu L, Zhang Y, Inouye M, Woychik NA: **Bacterial toxin YafQ is an endoribonuclease that associates with the ribosome and blocks translation elongation through sequence-specific and frame-dependent mRNA cleavage**. *Molecular microbiology* 2009, **71**(5):1071-1087.

165. Armalyte J, Jurenaite M, Beinoraviciute G, Teiserskas J, Suziedeliene E: **Characterization of *Escherichia coli dinJ-yafQ* toxin-antitoxin system using insights from mutagenesis data**. *J Bacteriol* 2012, **194**(6):1523-1532.

166. Kamphuis MB, Monti MC, van den Heuvel RH, Lopez-Villarejo J, Diaz-Orejas R, Boelens R: **Structure and function of bacterial *kid-kis* and related toxin-antitoxin systems**. *Protein and peptide letters* 2007, **14**(2):113-124.

167. Gerdes K, Maisonneuve E: **Bacterial persistence and toxin-antitoxin loci**. *Annual review of microbiology* 2012, **66**:103-123.

168. Maisonneuve E, Shakespeare LJ, Jorgensen MG, Gerdes K: **Bacterial persistence by RNA endonucleases**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(32):13206-13211.

169. Hazan R, Engelberg-Kulka H: ***Escherichia coli* mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1**. *Molecular genetics and genomics : MGG* 2004, **272**(2):227-234.

170. Zhang Y, Zhang J, Hara H, Kato I, Inouye M: **Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase**. *The Journal of biological chemistry* 2005, **280**(5):3143-3150.

171. Bizanek R, McGuinness BF, Nakanishi K, Tomasz M: **Isolation and structure of an intrastrand cross-link adduct of mitomycin C and DNA**. *Biochemistry* 1992, **31**(12):3084-3091.

172. Tomasz M, Palom Y: **The mitomycin bioreductive antitumor agents: cross-linking and alkylation of DNA as the molecular basis of their activity**. *Pharmacology & therapeutics* 1997, **76**(1-3):73-87.

173. Iyer VN, Szybalski W: **A Molecular Mechanism of Mitomycin Action: Linking of Complementary DNA Strands**. *Proceedings of the National Academy of Sciences of the United States of America* 1963, **50**:355-362.

174. Suzuki H, Kilgore WW: **Decomposition of ribosomal particles in *Escherichia coli* treated with mitomycin C**. *J Bacteriol* 1967, **94**(3):666-676.

175. Suzuki H, Kilgore WW: **Effects of mitomycin C on macromolecular synthesis in *Escherichia coli***. *J Bacteriol* 1967, **93**(2):675-682.

176. Snodgrass RG, Collier AC, Coon AE, Pritsos CA: **Mitomycin C inhibits ribosomal RNA: a novel cytotoxic mechanism for bioreductive drugs**. *The Journal of biological chemistry* 2010, **285**(25):19068-19075.

177. **Cefotaxime: Drug Information**
[http://www.merckmanuals.com/professional/lexicomp/cefotaxime.html]

178. Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ: **A common mechanism of cellular death induced by bactericidal antibiotics**. *Cell* 2007, **130**(5):797-810.

179. Wang X, Zhao X: **Contribution of oxidative damage to antimicrobial lethality**. *Antimicrobial agents and chemotherapy* 2009, **53**(4):1395-1402.

180. Tkachenko AG, Akhova AV, Shumkov MS, Nesterova LY: **Polyamines reduce oxidative stress in *Escherichia coli* cells exposed to bactericidal antibiotics**. *Research in microbiology* 2012, **163**(2):83-91.

181. Davis BD, Luger SM, Tai PC: **Role of ribosome degradation in the death of starved *Escherichia coli* cells**. *J Bacteriol* 1986, **166**(2):439-445.

182. Zundel MA, Basturea GN, Deutscher MP: **Initiation of ribosome degradation during starvation in *Escherichia coli***. *Rna* 2009, **15**(5):977-983.

183. Bernard JJ, Cowing-Zitron C, Nakatsuji T, Muehleisen B, Muto J, Borkowski AW, Martinez L, Greidinger EL, Yu BD, Gallo RL: **Ultraviolet radiation damages self noncoding RNA and is detected by TLR3**. *Nature medicine* 2012, **18**(8):1286-1290.

184. Kladwang W, Hum J, Das R: **Ultraviolet shadowing of RNA can cause significant chemical damage in seconds**. *Scientific reports* 2012, **2**:517.

185. Iordanov MS, Pribnow D, Magun JL, Dinh TH, Pearson JA, Magun BE: **Ultraviolet radiation triggers the ribotoxic stress response in mammalian cells**. *The Journal of biological chemistry* 1998, **273**(25):15794-15803.

186. Iordanov MS, Pribnow D, Magun JL, Dinh TH, Pearson JA, Chen SL, Magun BE: **Ribotoxic stress response: activation of the stress-activated protein kinase JNK1 by inhibitors of the peptidyl transferase reaction and by sequence-specific RNA damage to the alpha-sarcin/ricin loop in the 28S rRNA**. *Molecular and cellular biology* 1997, **17**(6):3373-3381.

187. Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC: **Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli***. *Genetics* 2001, **158**(1):41-64.

188. Baharoglu Z, Mazel D: **SOS, the formidable strategy of bacteria against aggressions**. *FEMS microbiology reviews* 2014.

189. Janion C: **Inducible SOS response system of DNA repair and mutagenesis in *Escherichia coli***. *International journal of biological sciences* 2008, **4**(6):338-344.

190. Kuzminov A: **Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda**. *Microbiology and molecular biology reviews : MMBR* 1999, **63**(4):751-813, table of contents.

191. Wadhawan S, Gautam S, Sharma A: **A component of gamma-radiation-induced cell death in *E. coli* is programmed and interlinked with activation of caspase-3 and SOS response**. *Archives of microbiology* 2013, **195**(8):545-557.

192. Prieto AI, Ramos-Morales F, Casadesus J: **Bile-induced DNA damage in *Salmonella enterica***. *Genetics* 2004, **168**(4):1787-1794.

193. Campoy S, Hervas A, Busquets N, Erill I, Teixido L, Barbe J: **Induction of the SOS response by bacteriophage lytic development in *Salmonella enterica***. *Virology* 2006, **351**(2):360-367.

194. Benson NR, Wong RM, McClelland M: **Analysis of the SOS response in *Salmonella enterica* serovar typhimurium using RNA fingerprinting by arbitrarily primed PCR**. *J Bacteriol* 2000, **182**(12):3490-3497.

195. Yim G, McClure J, Surette MG, Davies JE: **Modulation of *Salmonella* gene expression by subinhibitory concentrations of quinolones**. *The Journal of antibiotics* 2011, **64**(1):73-78.

196. Ysern P, Clerch B, Castano M, Gibert I, Barbe J, Llagostera M: **Induction of SOS genes in *Escherichia coli* and mutagenesis in *Salmonella typhimurium* by fluoroquinolones**. *Mutagenesis* 1990, **5**(1):63-66.

197. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, Porwollik S, Ali J, Dante M, Du F *et al*: **Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2**. *Nature* 2001, **413**(6858):852-856.

198. Singletary LA, Gibson JL, Tanner EJ, McKenzie GJ, Lee PL, Gonzalez C, Rosenberg SM: **An SOS-regulated type 2 toxin-antitoxin system**. *J Bacteriol* 2009, **191**(24):7456-7465.

199. Dorr T, Vulic M, Lewis K: **Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli***. *PLoS biology* 2010, **8**(2):e1000317.

200. Waldor MK, Friedman DI: **Phage regulatory circuits and virulence gene expression**. *Current opinion in microbiology* 2005, **8**(4):459-465.

201. Klose KE, Mekalanos JJ: **Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates *Salmonella typhimurium* virulence**. *Infection and immunity* 1997, **65**(2):587-596.

202. Gopel Y, Luttmann D, Heroven AK, Reichenbach B, Dersch P, Gorke B: **Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae**. *Nucleic acids research* 2011, **39**(4):1294-1309.

203. Hirschman J, Wong PK, Sei K, Keener J, Kustu S: **Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a sigma factor**. *Proceedings of the National Academy of Sciences of the United States of America* 1985, **82**(22):7525-7529.

Table 1.1 K**nown, predicted, and novel σ^{54}-dependent operons and sRNA genes of *S.* Typhimurium**

Locus Tag[a]	Gene Symbol[b]	Function	bEBP[c]	Reference[d]
***Known σ^{54}-dependent operons and sRNA genes*:**				
STM0368-71	*prpBCDE*	Proprionate catabolism (putative)	PrpR	[6]
STM0830-28	*glnHPQ*	Glutamine high-affinity transporter	NtrC	[201]
STM2355	*argT*	Lysine/arginine/ornithine transport protein	NtrC	[5]
STM_R0152	*glmY*	GlmY sRNA	GlrR	[202]
STM_R0167	*glmZ*	GlmZ sRNA	GlrR	[202]
STM4007-05	*glnALG*	Glutamine synthetase	NtrC	[203]
***Predicted σ^{54}-dependent operons*:**				
STM0462-63	*glnK amtB*	hypothetical protein	NtrC	[14]
STM0577-72		PTS (putative)	STM0571	[13]
STM0649.S-53		Hydrolase (putative)	STM0652	[13]
STM0665-62	*gltIJKL*	Glutamate/aspartate transporter	NtrC	[13, 16]
STM1285-84	*yeaGH*	Serine protein kinase (putative)	NtrC	[13, 16]
STM1303-07	*astCABDE*	Arginine/ornithine/glutamine metabolism	NtrC	[17, 18]
STM1690-86	*pspABCDE*	Phage shock proteins	PspF	[8]
STM2360-56	------*ubiX*	Amino acid transport (putative)	STM2361	[13]
STM2840-41	*norV ygbD*	Nitric oxide reductase	NorR	[13, 19]
STM2843-42	*hydN hypF*	Hydrogenase maturation proteins	FhlA	[20]
STM2853-44	*hycABCDEFGHI-*	Hydrogenase 3	FhlA	[21]
STM2854-58	*hypABCDE*	Formate-hydrogen lyase system	FhlA	[21]
STM3521-18	*-rtcBA*	RNA repair system (putative)	RtcR	[15]
STM3568	*rpoH*	Heat shock sigma factor (σ^{32})		[22, 23]
STM3772-66		PTS (putative)	STM3773	[13]
STM4172	*zraP*	Zinc resistance-associated protein	ZraR	[6, 13]
STM4173-74	*hydHG*	Zinc resistance two-component system	ZraR	[6]
STM4244	*pspG*	Phage shock protein	PspF	[24]
STM4285	*fdhF*	Formate dehydrogenase	FhlA	[25]
STM4535-40.s		PTS (putative)	STM4534	[13]

[a]Locus tags for genes within operons or sRNA genes are grouped by those previously shown to be σ^{54}-dependent in *Salmonella*, previously predicted to be σ^{54}-dependent

[b]Genes that have not been assigned a gene symbol are represented by a dash (−).

[c]Known or predicted bacterial enhancer-binding protein (bEBP) that activates the σ^{54}-dependent operon.

[d]References for operons shown to be σ^{54}-dependent in *Salmonella* and for operons either determined to be σ^{54}-dependent in other bacterial genera or predicted to be regulated by σ^{54} in *Salmonella* are listed.

Figure 1.1: Activation of σ^{54}-dependent transcription. (A) The canonical bEBP modular domain structure. All bEBPs contain a central AAA+ family ATPase domain. Most of these proteins also contain N-terminal signal sensing and C-terminal DNA binding domains, but the presence of these domains is variable between different bEBPs [36, 46]. The schematic in grey boxes at the top of subsequent frames refers to the domain architecture shown here. **(B)** Standard, enhancer-dependent activation. σ^{54} (red subunit) directs binding of the RNA polymerase (dark blue subunit) holoenzyme ($E\sigma^{54}$) to the -12 and -24 promoter elements (light blue box). This closed complex is stable and cannot transition to open complex. In response to an environmental or cellular signal, the activator (bEBP; yellow dimers) oligomerizes. For most bEBPs, the oligomer binds to an enhancer (green box) 80 to 150 bp upstream of the promoter and DNA looping brings the activator in contact with σ^{54} in the $E\sigma^{54}$ closed complex. Hydrolysis of ATP by bEBP causes remodeling of $E\sigma^{54}$, which leads to open complex formation and transcription. **(C)** Some bacteria possess bEBPs that are missing the DNA binding domain; after oligomerization, these activators can bind to $E\sigma^{54}$ in closed complex with any promoter to stimulate open complex formation. **(D)** N-terminally truncated bEBP have lost their ability to respond to environmental stimuli. As a result, these proteins are constitutively-active and can interact with their specific enhancer sequence irrespective of growth conditions. **(E)** Combining both N- and C-terminal truncations leaves just the central AAA+ ATPase domain. This constitutively-active, promiscuous activator should be able to interact with any σ^{54}-dependent promoter in a cell under any growth condition.

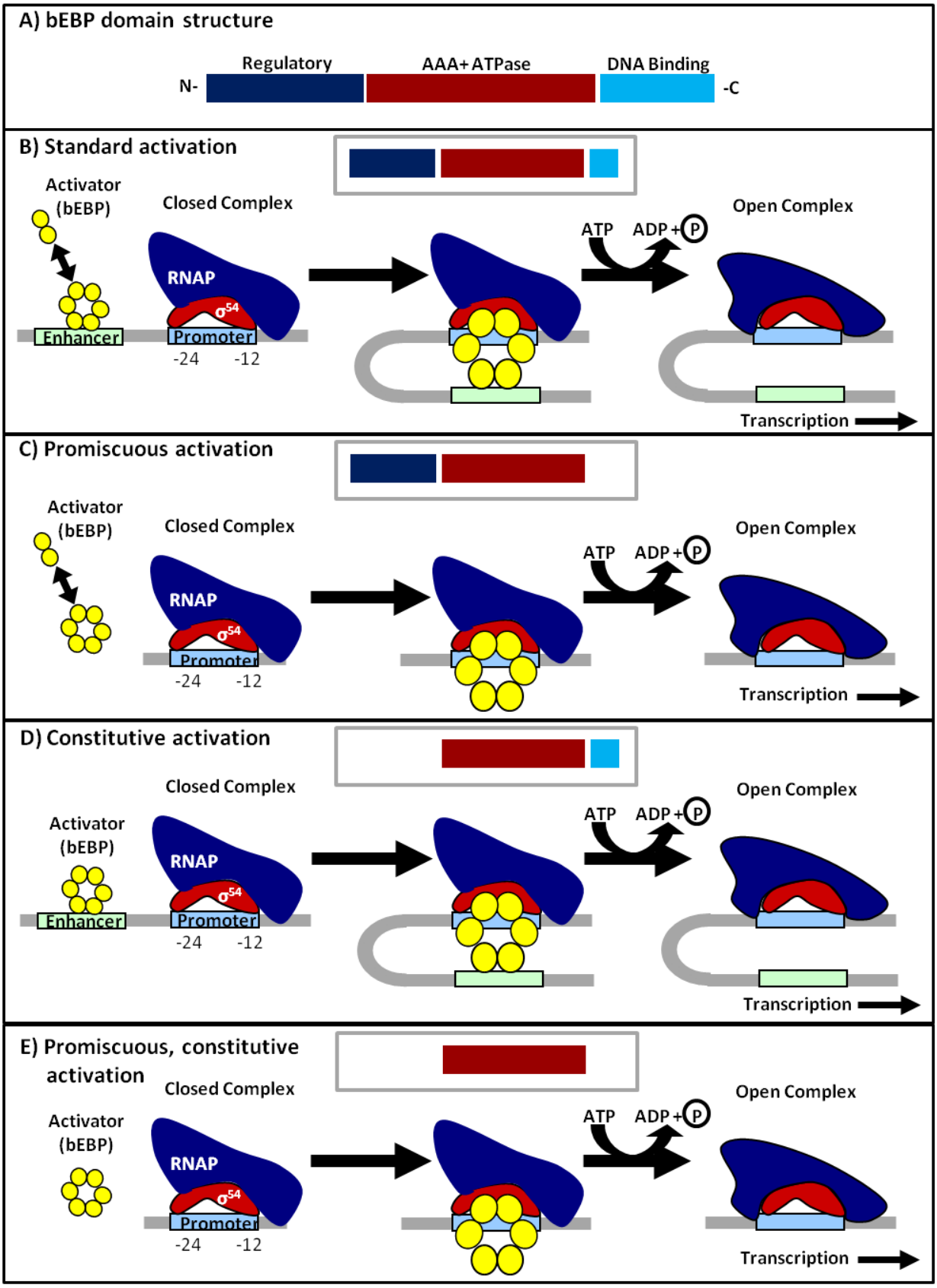

A) bEBP domain structure
Regulatory
AAA+ ATPase
DNA Binding
N-
-C
B) Standard activation
Activator (bEBP)
Closed Complex
RNAP
σ54
Enhancer
Promoter
-24
-12
ATP
ADP + P
Open Complex
Transcription
C) Promiscuous activation
Activator (bEBP)
Closed Complex
RNAP
σ54
Promoter
-24
-12
ATP
ADP + P
Open Complex
Transcription
D) Constitutive activation
Closed Complex
Activator (bEBP)
RNAP
σ54
Enhancer
Promoter
-24
-12
ATP
ADP + P
Open Complex
Transcription
E) Promiscuous, constitutive activation
Closed Complex
Activator (bEBP)
RNAP
σ54
Promoter
-24
-12
ATP
ADP + P
Open Complex
Transcription

Figure 1.2: Diagram of the *Salmonella rsr-rtcBA* operon. The σ^{54} enhancer binding protein, RtcR, and its σ^{70} promoter are shown in yellow and black, respectively. The other genes in the operon, including three structural genes—*rsr*, *rtcB*, and *rtcA*, shown in blue—and two small RNAs—*yrlB* and *yrlA*, shown in green—are expressed from a σ^{54}-dependent promoter, shown in red.

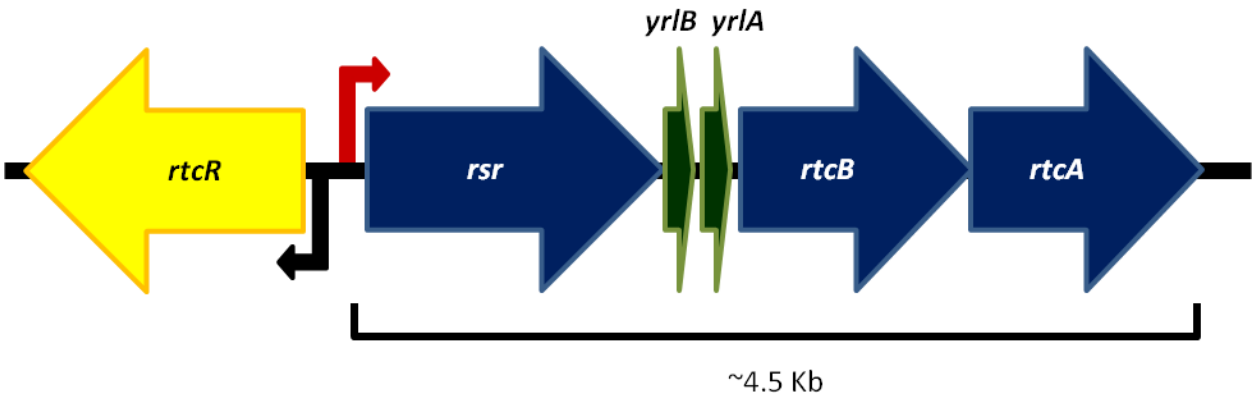

yrlB yrlA
rtcR
rsr
rtcB
rtcA
~4.5 Kb

Figure 1.3: RtcBA alignments. Pairwise alignment RtcB (A) and RtcA (B) from *S.* Typhimurium and the *E. coli* proteins characterized by Tanaka and Shuman [67] and Genschik, et al. [15]. Identical residues are indicated with dark gray shading, similar residues with light gray. The conserved RtcB Histidine (437) residue that becomes guanylylated is boxed in red.

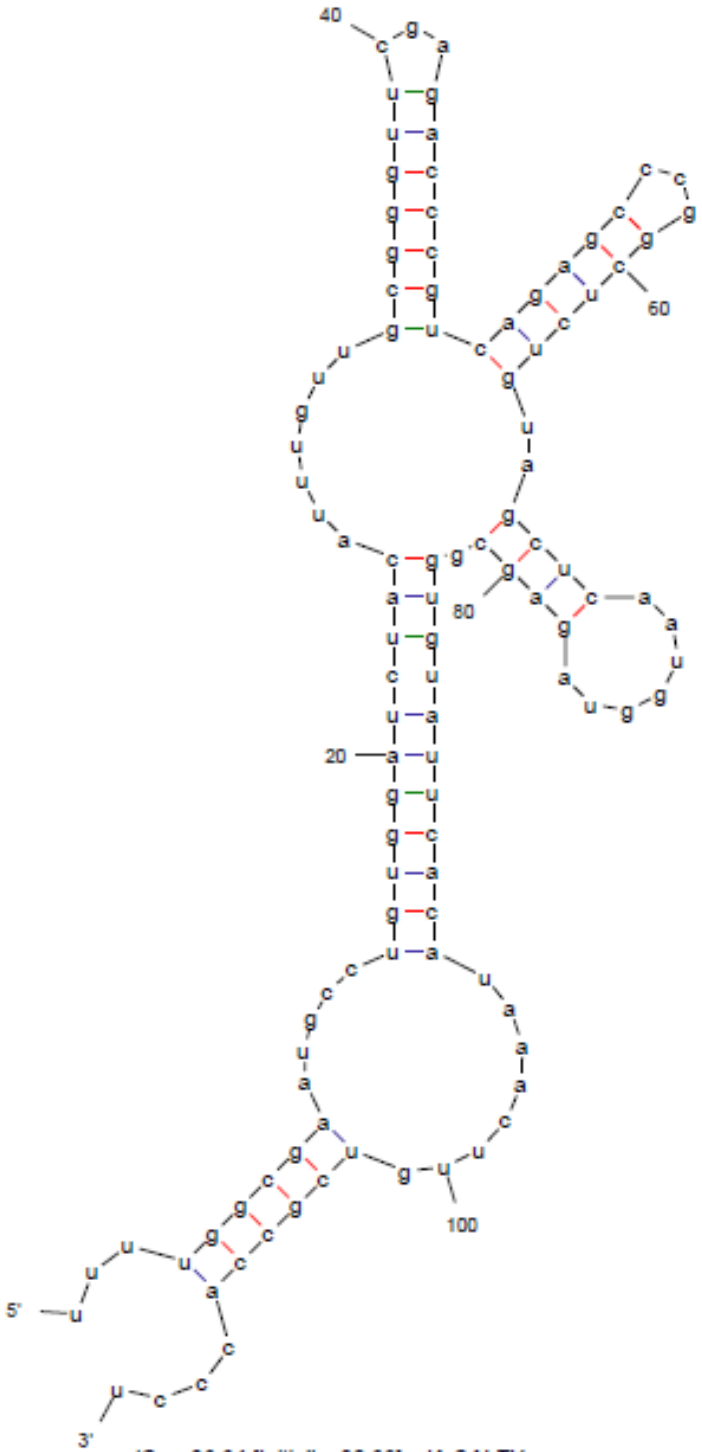
A) **RtcB** [88% identical; 93% similar]

```
1    MMNYELMTTQNAPVKMWTKGVPVEDDARQQLINTAKMPFIFKHIAVMPDVHLGKGSTIGS   60  Sty
1    -MNYELLTTENAPVKMWTKGVPVEADARQQLINTAKMPFIFKHIAVMPDVHLGKGSTIGS   59  Eco
      *****:**:************** ************************************

61   VIPTKGAIIPAAVGVDIGCGMNALRTSLTAADLPENLADLRSAIEAAVPHGRTTGRGHRD  120  Sty
60   VIPTKGAIIPAAVGVDIGCGMNALRTALTAEDLPENLAELRQAIETAVPHGRTTGRCKRD  119  Eco
     *************************:*** *******:**.***:********** :**

121  VGAWGNPPANVNEKWAQLEAGYQWLTQKYPRFLNTNNYKHLGTLGTGNHFIEICLDETDR  180  Sty
120  KGAWENPPVNVDAKWAELEAGYQWLTQKYPRFLNTNNYKHLGTLGTGNHFIEICLDESDQ  179  Eco
      *** ***.**: ***:****************************************:*:

181  VWIMLHSGSRGIGNAIGTYFIGLAQQEMQEQLETLPSRDLAYFNEGSEYFDDYLKAVHWA  240  Sty
180  VWIMLHSGSRGIGNAIGTYFIDLAQKEMQETLETLPSRDLAYFMEGTEYFDDYLKAVAWA  239  Eco
     ******************** ***:**** ************ **:********** **

241  QQFASLNREAMMENALAALQ----RCVEKPSALDMDEINCHHNYVQKEQHFGEEIYVTRK  296  Sty
240  QLFASLNRDAMMENVVTALQSITQKTVRQPQTLAMEEINCHHNYVQKEQHFGEEIYVTRK  299  Eco
     * ******:*****.::***    : *.:*.:* *:************************

297  GAVSARRGEFGIIPGSMGAKSFIVRGLGNEESFCSCSHGAGRVMSRTKAKKLFSVDDQIR  356  Sty
300  GAVSARAGQYGIIPGSMGAKSFIVRGLGNEESFCSCSHGAGRVMSRTKAKKLFSVEDQIR  359  Eco
     ****** *::*****************************************:****

357  ATAHVECRKDADVIDEIPMAYKDIDAVMAAQSDLVEIMYALRQVVCVKG  405  Sty
360  ATAHVECRKDAEVIDEIPMAYKDIDAVMAAQSDLVEVIYTLRQVVCVKG  408  Eco
     ***********:************************::*:*********
```

B) **RtcA**[68% identical; 76% similar]

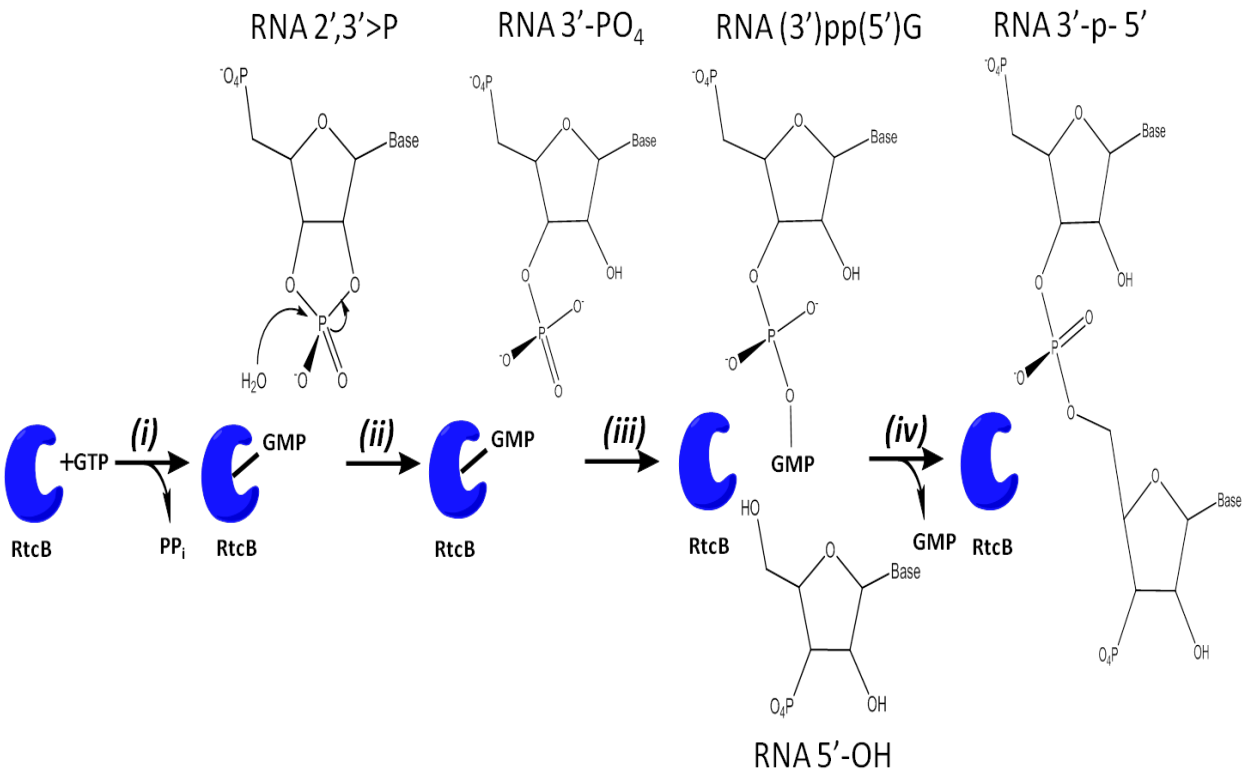
```
1    MARIIALDGAQGEGGGQILRSALSLSMITGQPFEMSGIRAGRAKPGLLRQHLTAVRAATE   60  Sty
1    MKRMIALDGAQGEGGGQILRSALSLSMITGQPFTITSIRAGRAKPGLLRQHLTAVKAATE   60  Eco
     * *:****************************** ::.******************:****

61   ICGAQVNGDELGSQQLRFTPGPIRGGEYRFAIGSAGSCMLVLQTVLPALWFADGSSRVEV  120  Sty
61   ICGATVEGAELGSQRLLFRPGTVRGGDYRFAIGSAGSCTLVLQTVLPALWFADGPSRVEV  120  Eco
     **** *:* *****:* * ** :***:********** *************** *****

121  HGGTHNQAAPSADFICRVWEPLLARMGISQRTTLIKHGFYPAGGGAAATVVEPAASLRGL  180  Sty
121  SGGTDNPSAPPADFIRRVLEPLLAKIGIHQQTTLLRHGFYPAGGGVVATEVSPVASFNTL  180  Eco
      ***.* :** **** ** *****::** *:***::*********..** *.*.**:. *

181  TLISRGETLRTTAEALLAAVPYHVGEREVATLEAHFPQAEKNVVALEGGCGPGNALSLMI  240  Sty
181  QLGERGNIVQMRGEVLLAGVPRHVAEREIATLAGSFSLHEQNIHNLPRDQGPGNTVSLEV  240  Eco
      * .**: ::  .*.***.** **.***:*** . *   *:*:  *    ****::** :

241  QSEQLTELFAAFGVKGTSAEAVANQVAHEARRYLASPAAVGEHLADQLILPLALAGEGAF  300  Sty
241  ESENITERFFVVGEKRVSAEVVAAQLVKEVKRYLASTAAVGEYLADQLVLPMALAGAGEF  300  Eco
     :**::** * ..* * .***.** *:.:*.:***** *****:*****:**:**** * *

301  TVARASAHLLTNIAVVERFLPVRFSCEATESGYLVRVSD  339  Sty
301  TVAHPSCHLLTNIAVVERFLPVRFSLIETDGVTRVSIE-  338  Eco
     ***: *.******************   *:.   * :.
```

Figure 1.4: Predicted YrlA structure. Mfold structure prediction for YrlA. The long stem and large internal loop characteristic of Y RNA molecules is present.



dG = -36.64 [Initially -38.60] yrlA-SALTY

Figure 1.5: RNA ligation by RtcB. The mechanisms of RtcB, as characterized by Tanaka, et al. [66, 67, 99, 100]. *(i*) RtcB is guanylylated at a conserved histidine residue. *(ii*) RtcB-GMP cyclic phosphodiesterase (CPDase) activity hydrolyzes an RNA 2',3'>P, yielding a 3'-PO_4. *(iii)* RtcB transfers the GMP group to the 3'-PO_4 of RNA. *(iv)* RtcB catalyzes the ligation of the RNA 3'-GMP to RNA 5'-OH, releasing GMP and forming a 3'-5' phosphodiester linkage. Alternatively, if the initial RNA substrate bears a 3'-PO_4 terminal, RtcB can proceed directly from *(i)* to *(iii)*.


RNA 2',3'>P
RNA 3'-PO4
RNA (3')pp(5')G
RNA 3'-p- 5'
Base
OH
H2O
+GTP
(i)
(ii)
(iii)
(iv)
GMP
PPi
RtcB
HO
RNA 5'-OH

CHAPTER 2:

USE OF A PROMISCUOUS, CONSTITUTIVELY-ACTIVE BACTERIAL ENHANCER-BINDING PROTEIN TO DEFINE THE σ^{54} (RpoN) REGULON OF *SALMONELLA* TYPHIMURIUM LT2[1]

[1] Samuels, D.J., Frye, J.G., Porwollik, S., McClelland, M., Mrázek, J., Hoover, T.R., and A.C. Karls. 2013. *BMC Genomics*. **14**:602. Reprinted here with permission of the publisher.

Abstract

Background

Sigma54, or RpoN, is an alternative σ factor found widely in eubacteria. A significant complication in analysis of the global σ^{54} regulon in a bacterium is that the σ^{54} RNA polymerase holoenzyme requires interaction with an active bacterial enhancer-binding protein (bEBP) to initiate transcription at a σ^{54}-dependent promoter. Many bacteria possess multiple bEBPs, which are activated by diverse environmental stimuli. In this work, we assess the ability of a promiscuous, constitutively-active bEBP—the AAA+ ATPase domain of DctD from *Sinorhizobium meliloti*—to activate transcription from all σ^{54}-dependent promoters for the characterization of the σ^{54} regulon of *Salmonella* Typhimurium LT2.

Results

The AAA+ ATPase domain of DctD was able to drive transcription from nearly all previously characterized or predicted σ^{54}-dependent promoters in *Salmonella* under a single condition. These promoters are controlled by a variety of native activators and, under the condition tested, are not transcribed in the absence of the DctD AAA+ ATPase domain. We also identified a novel σ^{54}-dependent promoter upstream of STM2939, a homolog of the *cas1* component of a CRISPR system. ChIP-chip analysis revealed at least 70 σ^{54} binding sites in the chromosome, of which 58% are located within coding sequences. Promoter-*lacZ* fusions with selected intragenic σ^{54} binding sites suggest that many of these sites are capable of functioning as σ^{54}-dependent promoters.

Conclusion

Since the DctD AAA+ ATPase domain proved effective in activating transcription from the diverse σ^{54}-dependent promoters of the *S.* Typhimurium LT2 σ^{54} regulon under a single growth condition, this approach is likely to be valuable for examining σ^{54} regulons in other bacterial species. The *S.*

Typhimurium σ^{54} regulon included a high number of intragenic σ^{54} binding sites/promoters, suggesting that σ^{54} may have multiple regulatory roles beyond the initiation of transcription at the start of an operon.

Keywords

Sigma54, RpoN, Bacterial enhancer-binding protein, Regulon, Sigma factor, *Salmonella*

Background

Transcription in eubacteria is mediated by the RNA polymerase holoenzyme (Eσ), which has five constant subunits ($\alpha_2\beta\beta'\omega$) and a variable subunit (σ). The constant subunits constitute the RNA polymerase core (RNAP), which has the polymerization activity; the σ subunit determines promoter recognition and functions in the Eσ-promoter transition from closed complex to open complex (isomerization). The primary σ factor in a bacterium, such as σ^{70} in *Escherichia coli*, controls transcription of most housekeeping genes in the cell; alternative sigma factors have specialized regulons that function in the response to environmental stressors or morphological changes, or in developmental systems (for review see [1]). In many bacteria the alternative σ factor σ^{54} (also called RpoN or NtrA) has unusually diverse regulons, with genes that function in a variety of cellular processes, including flagellar biogenesis, response to nitrogen starvation, transport and metabolism of carbon substrates, and tolerance to heavy metals [2-6].

Multiple features, including protein structure, promoter consensus sequence, and mode of activation, distinguish σ^{54} from all other primary and secondary sigma factors, which constitute the σ^{70} family (reviewed in [1,7]). Although both σ^{54}- and σ^{70}-type sigma factors associate with the β and β' subunits of RNAP and mediate the binding of Eσ to specific promoter sequences, σ^{54} differs extensively from σ^{70}-type sigma factors in primary amino acid sequence and domain organization (reviewed in [8]). The essential promoter features for $E\sigma^{54}$ recognition and binding center around conserved GG and TGC elements at -24 and -12, respectively, relative to the transcription start site (TSS) [9], while holoenzymes with the various σ^{70}-type sigma factors generally recognize and bind promoter elements at -35 and -10 with the consensus sequences TTGACA and TATAAT, respectively (reviewed in [1]). Perhaps the most important feature of $E\sigma^{54}$ that differs from $E\sigma^{70}$ is the isomerization process (Figure 1A). For $E\sigma^{70}$ the transition from closed complex to open complex is usually spontaneous and rapid, so regulation of transcription initiation frequently occurs at the level of closed complex formation. Initiation of transcription by $E\sigma^{54}$ more closely resembles eukaryotic Pol II systems in that $E\sigma^{54}$ forms a stable closed

complex that requires a bacterial enhancer-binding protein (bEBP) and ATP hydrolysis for isomerization to open complex (reviewed in [10]). The bEBPs add an additional level of complexity to the σ^{54} regulon.

bEBPs have a modular structure that is generally conserved: an N-terminal regulatory domain, a central AAA+ ATPase domain, and a C-terminal DNA binding domain (Figure 1B; reviewed in [8]). These proteins activate transcription from σ^{54}-dependent promoters in three basic steps (Figure 1A). First, the bEBP receives an environmental stimulus through phosphorylation, ligand binding, or protein-protein interactions with the N-terminal regulatory domain that stimulates the bEBP to multimerize through the AAA+ ATPase domain and bind to an upstream activator sequence (UAS or enhancer) via the C-terminal DNA binding domain. The bEBP-UAS complex is then brought into contact with the $E\sigma^{54}$-promoter closed complex via a DNA looping event and interactions between highly conserved regions of the AAA+ ATPase domain of bEBP and σ^{54}. Finally, ATP hydrolysis drives isomerization, allowing the initiation of transcription.

The requirement for bEBP-mediated activation of σ^{54}-dependent transcription presents two problems for global analysis of a σ^{54} regulon. The first is the need for the proper environmental stimulus to activate bEBPs. Since the $E\sigma^{54}$ closed complex requires an activated bEBP, σ^{54}-dependent promoters are usually transcriptionally silent in the absence of the specific stimulus for the bEBP [8]. Analysis of transcription from σ^{54}-dependent promoters under any single growth condition would miss operons whose bEBPs are not activated under the condition tested. Secondly, the requirement for the UAS or enhancer by most bEBPs presents a challenge for predicting whether a $E\sigma^{54}$ binding site is functioning as a promoter or not. There is no common consensus sequence for the enhancer and their position relative to the promoter can be quite variable. For many σ^{54}-dependent promoters the UAS sequence lies ~70-150 bp upstream of the promoter, but other configurations have been characterized, such as enhancers located 1.5 kb downstream of the *rocG* promoter in *Bacillus subtilis* [11] and up to 3 kb upstream of the promoter in artificial constructs of the *glnA* operon from *E. coli* [12]. If a σ^{54} binding site is examined for promoter

activity in isolation, such as in a promoter-reporter vector, it is difficult to discern whether a site is inactive because it is not a promoter or because the enhancer was not included in the cloned sequence.

Previous studies to define the σ^{54} regulons of *Escherichia coli* [13], *Vibrio cholerae* [14] and *Geobacter sulfurreducens* [15] have recognized the limitations presented by the requirement for activated bEBPs in the characterization of the full σ^{54} regulon, even when σ^{54} is overexpressed from a heterologous promoter. Our approach to overcoming these problems in the global characterization of σ^{54} regulons in bacteria is the utilization of a constitutively-active, promiscuous bEBP, the AAA+ ATPase domain of *Sinorhizobium meliloti* DctD [16,17]. We chose to assess the efficacy of this approach in *Salmonella enterica subsp*. *enterica* serovar Typhimurium LT2 (hereafter referred to as *S*. Typhimurium LT2), a widely-used laboratory strain, because it has a moderately-sized σ^{54} regulon with 13 known or predicted bEBPs [18], providing sufficient diversity in bEBPs to test our hypothesis.

We report here that use of this constitutively-active, promiscuous bEBP in DNA microarrays and promoter function assays permitted detection of nearly all known and predicted σ^{54}-dependent operons. These studies also revealed a new σ^{54}-dependent promoter expressing a putative *cas1* gene in *S*. Typhimurium LT2 (STM2938). In addition, chromatin immunoprecipitation-microarray (ChIP-chip) analysis combined with bioinformatics identified 70 $E\sigma^{54}$ or σ^{54} binding sites, of which 41 appear to be within open reading frames (ORFs). This surprising number of intragenic sites suggests regulatory roles for σ^{54} or $E\sigma^{54}$ that may involve repression, transcriptional interference, or expression of cis- or trans-acting small RNA (sRNA) [19,20].

Results and Discussion

Utility of a promiscuous, constitutive bEBP in characterizing the σ^{54} regulon

Since all known σ^{54}-dependent promoters require an activated bEBP for transcription initiation, it is a challenge to find a condition under which all promoters can be detected within the σ^{54} regulon of a bacterium. In the recent mapping of the *S*. Typhimurium SL1344 transcriptome using early stationary

phase cultures in rich media (Lennox broth), only one of the known or predicted σ^{54}-dependent gene transcripts was detected, *pspA* [21]. The currently favored approach is over-expression of σ^{54} to facilitate detection of σ^{54}-dependent promoters, which assumes a reasonable basal level of activation of the bEBPs. Using relatively low cutoffs for the fold-change (1.5- to 2-fold) in transcript levels between the σ^{54}-overexpression strain and wild type or Δ*rpoN* strains, a considerable portion of the σ^{54}-dependent transcriptome was defined in *Escherichia coli* [13], *Vibrio cholerae* [14] and *Geobacter sulfurreducens* [15]. However, not all previously-identified σ^{54}-dependent operons were detected for *E. coli* and *G. sulfurreducens*, and evidence from the *V. cholera* and *G. sulfurreducens* studies suggests that overexpression of σ^{54} may repress expression from some σ^{54}-dependent promoters and alter expression of σ^{54}-independent promoters [13-15]. We hypothesize that a promiscuous and constitutive variant of the bEBP DctD from *S. meliloti* can activate transcription from all σ^{54}-dependent promoters in *S.* Typhimurium LT2 at wild-type levels of σ^{54} under a single growth condition, thereby facilitating global characterization of the σ^{54} regulon without overexpression of σ^{54}. This promiscuous and constitutive DctD variant is missing the N-terminal response regulator and C-terminal DNA binding domains, leaving only the central AAA+ ATPase domain, residues 141 to 390 of DctD and referred to hereafter as DctD250 [17]. Previous work showed that DctD250 was able to interact with $E\sigma^{54}$ in *E. coli* to drive transcription from the chromosomal *glnA* promoter and from the *S. meliloti dctA* promoter in the absence of native DctD and without an enhancer sequence [16,17].

The σ^{54}-dependent promoters of *S.* Typhimurium LT2 are normally responsive to one or more of thirteen known and predicted bEBPs under various growth conditions [18], so to initially assess DctD250 activation of transcription from σ^{54}-dependent promoters that respond to different bEBPs in *Salmonella*, the σ^{54}-dependent promoters for the *glnKamtB* (STM0462) and *rtcBA* (STM3521) operons were introduced upstream of a promoter-less *lacZ* gene and the reporter plasmids were transformed into a derivative of *S.* Typhimurium LT2 (wild-type; WT) and WT containing the DctD250 expression plasmid (WT + DctD250) to perform β-galactosidase assays. The *glnKamtB* and *rtcBA* promoters were chosen

because neither has predicted σ^{70}-dependent promoters within the cloned promoter region and each is responsive to a different bEBP: NtrC for *glnKamtB* [22] and RtcR for *rtcBA* [23]. In the WT strain, the *glnKamtB* and *rtcBA* operon promoters expressed *lacZ* at very low levels; but in the presence of DctD250, *lacZ* was expressed at 150- and 16-fold higher levels, respectively (Table 1). To compare the level of expression stimulated by DctD250 to the level that is seen under physiological conditions that activate the promoter-associated bEBP, *lacZ* expression from the *glnKamtB* promoter was assayed in the WT strain in nitrogen-limiting medium, which activates NtrC. Under nitrogen-starvation conditions NtrC multimerizes, binds the enhancer in the cloned promoter region, and hydrolyzes ATP to stimulate transcription by $E\sigma^{54}$ at the *glnKamtB* promoter (see Figure 1A). In the presence of activated NtrC, the *glnKamtB* promoter expresses *lacZ* at a nearly 10-fold higher level than in the presence of DctD250. This reduced level of activation by DctD250 relative to the cognate bEBP under activation conditions is consistent with previous studies comparing the activity of truncated versions of bEBPs, which must interact with $E\sigma^{54}$ from solution, to that of the wild type bEBPs, which are directed to the target σ^{54} promoter via binding to the enhancer sequence [17,24]. The control reporter plasmids pDV6, which has the σ^{70}-dependent, circle junction promoter from IS*492* [25], and the promoter-less pDS12 expressed *lacZ* at approximately the same level in WT as WT + DctD250 (Table 1). Based on these results, DctD250 activates transcription from σ^{54}-dependent promoters that are normally responsive to different bEBPs under different growth conditions. Therefore, we performed DNA microarray and promoter-reporter analyses in the presence of the promiscuous, constitutive activator DctD250 to assess the efficacy of this approach in defining the σ^{54} regulon of *S*. Typhimurium LT2.

Microarray analysis of σ^{54}-dependent transcripts in *Salmonella* expressing DctD250

To determine the genes whose transcription is controlled by σ^{54} in *S*. Typhimurium LT2 we performed a microarray analysis comparing WT+DctD250 to an isogenic strain with a deletion of *rpoN* (Δ*rpoN*+DctD250). RNA collected during mid-log phase growth in nutrient medium was reverse

transcribed and cDNAs from each strain were differentially labeled and applied to a complete ORF array containing all annotated open reading frames for *S*. Typhimurium LT2 [26]. Open reading frames that were transcribed in WT at a level > 3-fold higher than in the Δ*rpoN* strain, with a p value <0.02, were considered up-regulated and, for the purpose of the initial categorization of these results, an operon was considered up-regulated if at least one gene met these criteria. In three biological replicates, the same 33 operons were up-regulated in the presence of σ^{54}. The microarray results for *S*. Typhimurium LT2 genes within operons that meet the criteria for up-regulation, or that are known or predicted to be σ^{54}-dependent, are shown in Table 2 and Additional file 1. Only 4 genes, STM2722, STM2724, STM2729, and STM2730, which are part of 2 operons in the Fels-2 prophage, were down-regulated >3-fold with a p-value <0.02 in the WT strain as compared to the Δ*rpoN* strain.

Known σ^{54}-dependent operons and sRNA

If our hypothesis is correct, then in the presence of DctD250 we should observe up-regulation of operons (one or more structural genes) and sRNA genes that are known to have σ^{54}-dependent promoters, even though they are normally activated by different bEBPs. Previously, four *Salmonella* operons have been experimentally shown to be regulated by σ^{54}: *prpBCDE* [4], *glnHPQ* [27], *argT* [2], and *glnALG* [29]. Additionally, two sRNA genes, *glmY* and *glmZ*, have also been shown to have σ^{54}-dependent promoters [28]. Table 2 summarizes the genes, functions, bEBPs, and microarray results for the known σ^{54}-dependent operons and sRNA genes of *Salmonella*.

The DNA microarrays showed up-regulation of all four known σ^{54}-dependent operons in *Salmonella*, *prpBCDE*, *glnHPQ*, *argT*, and *glnALG* (Table 2). The two sRNA genes with known σ^{54}-dependent promoters did not appear up-regulated by σ^{54}. This result was not surprising since in *S*. Typhimurium both *glmY* and *glmZ* possess σ^{70}-dependent promoters that fully overlap the σ^{54}-dependent promoters, such that the $E\sigma^{70}$ and $E\sigma^{54}$ compete for binding to their respective promoters [28]. Gopel et al. [28] demonstrated that the level of *glmY* transcription was similar in wild type and Δ*rpoN* cells and that

transcription of *glmZ* actually increased in the *rpoN* mutant, reflecting that the σ^{70}-dependent promoter for *glmZ* is stronger than the σ^{70}-dependent promoter for *glmY*. The presence of a σ^{70} promoter does not necessarily preclude detection of a σ^{54}-dependent promoter controlling expression of a gene or operon in these microarray assays, though; the promoter region of *glnA* has non-overlapping σ^{70}- and σ^{54}-dependent promoters [29], yet was up-regulated 48-fold. Taken together, these results for the known σ^{54}-dependent promoters are consistent with our hypothesis that DctD250 can promiscuously and constitutively activate σ^{54}-holoenzyme at a variety of σ^{54}-dependent promoters.

Confirmation of predicted σ^{54}-dependent operons

There are 20 operons that we define as 'predicted' σ^{54}-dependent operons in *Salmonella*. These predictions are based on *in silico* analyses indicating either homology to known σ^{54}-dependent operons in *E. coli* and other enteric bacteria or promoter sequence homology along with genetic proximity to predicted bEBP genes [3,5,18,22,23,30-39]. However, σ^{54}-dependent transcription of these operons has not previously been experimentally demonstrated in *Salmonella*. In the DNA microarrays, 16 of the 20 operons that have been predicted to have σ^{54}-dependent promoters in *Salmonella* were up-regulated in WT+DctD250 as compared to Δ*rpoN*+DctD250 (Table 2), providing experimental evidence that these genes are, in fact, regulated by σ^{54} in *S.* Typhimurium LT2.

For these 16 up-regulated σ^{54}-dependent operons there are 11 different bEBPs that either are known or predicted to activate expression from their σ^{54}-dependent promoters (Table 2). Five of the up-regulated operons, STM0577-0572, STM0649.s-0653, STM2360-2356, STM3772-3766, and STM4535-4540.s, were predicted to be σ^{54}-dependent based on linkage to a predicted bEBP and an upstream sequence with the essential -12 and -24 elements of a σ^{54}-dependent promoter [18]. There are no orthologs in *E. coli* for the predicted bEBPs associated with these operons; three of these predicted bEBPs, STM0571, STM3773 and STM4534, are similar to the LevR-type EBPs found in Gram-positive bacteria [18]. In addition to the microarray evidence presented here for σ^{54} regulation of these operons,

we know that STM3773 is the bEBP controlling expression of STM3772-3776 and that this operon encodes the components of a phosphotransferase system permease for D-glucosaminic acid and enzymes required for catabolism of this acid sugar [40]. These results show that DctD250 can activate expression at σ^{54}-dependent promoters that are normally regulated by the LevR-type bEBPs.

Of the four predicted σ^{54}-dependent operons that did not fulfill our criteria for upregulation in the microarray, at least two have additional σ^{54}-independent promoters, which may have masked the effect of σ^{54} on transcription levels. The heat shock sigma factor gene *rpoH* has been shown to be under the control of additional promoters and other regulatory proteins in *E. coli* [36]. The conservation of this promoter region for *rpoH* in *S.* Typhimurium LT2 suggests that a similar complex regulatory scheme may be involved [37], thereby reducing the effects of the Δ*rpoN* mutation. The *yeaGH* operon, which was just below the 3-fold cutoff for up-regulation in the microarray analysis, has previously been shown to be under control of σ^{S} in *Salmonella* [41]; however, our assays utilized *S.* Typhimurium LT2, which has a defective *rpoS* gene due to a transversion mutation in the start codon [42]. The promoter-reporter assay with the *yeaGH* promoter region, described below, suggests there is a σ^{54}- and σ^{S}-independent promoter expressing the *yeaGH* operon in both the WT+DctD250 and Δ*rpoN*+DctD250 strains.

The frequency of alternate promoters seen for the σ^{54}-dependent operons in *Salmonella* (at least 15% for the known and predicted promoters in our analyses) is not unique. Zhao et al. [13] estimate that 14% of σ^{54}-dependent genes in *E. coli* are transcribed by σ^{70}-associated RNA polymerase and suggest that expression of σ^{54}–dependent genes from alternate promoters allows for differential expression under various environmental conditions.

New potential σ^{54}-dependent genes

In addition to the σ^{54}-dependent expression of known or predicted genes and operons, the DNA microarray analysis revealed up-regulation of a gene, STM2938, which has not previously been reported or predicted to be σ^{54}-dependent. STM2938 is the penultimate gene in a nine-gene operon that is

annotated as a group of CRISPR-associated (*cas*) genes. Although none of the other genes in this operon seem to be controlled by σ^{54}, further evidence is presented below that supports the presence of a σ^{54}-dependent promoter within the gene upstream of STM2938. CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and *cas* genes constitute an adaptive immune system in bacteria and archaea that protects against invading mobile DNA, such as phage and plasmids [43]. The response to phage infection, which is referred to as phage shock, is regulated by σ^{54} and the bEBP PspF in *E. coli* [5]; thus, it would not be surprising for essential components of the bacterial immune response in phage infection to be regulated similarly. The potential σ^{54}-dependent gene STM2938 is a homologue of the *cas1* gene, which is an endonuclease that is associated with all CRISPR loci and is most likely involved in the adaptation phase of the CRISPR-*cas* immune system [44]. The regulation of this *cas1*-like gene by PspF in *Salmonella* is currently under investigation.

There were 12 additional ORFs that met the 3-fold cutoff for up-regulation by σ^{54} in the microarray assay, including genes for pilin biosynthesis (*hofB*), histidine ammonia lyase (*hutH*), bEBPs (*ygaA*, *fhlA*), propanediol utilization (*pduG*), siderophore production (*iroD*), and cell invasion (*invG*). The whole genome chromatin immunoprecipitation assays described below did not reveal σ^{54} binding sites associated with these ORFs; thus the expression of these genes may be indirectly affected by the absence of σ^{54} in the Δ*rpoN* mutant, or constitute false positives (Additional file 1).

ChIP-chip analysis of genome-wide σ^{54} binding sites in *Salmonella*

In the characterization of the σ^{54} regulon of *Salmonella*, determination of the genomic binding sites for the $E\sigma^{54}$ allows confirmation of primary transcripts indicated by microarray analysis and recognition of potential σ^{54}-regulated genes that might not have been detected due to instability of the transcripts. To assess the binding of $E\sigma^{54}$ in the *S.* Typhimurium LT2 genome, we isolated σ^{54}-DNA complexes from WT and Δ*rpoN* strains that did not contain the DctD250 expression plasmid. Since bEBPs do not activate transcription by recruiting $E\sigma^{54}$ to promoter sequences [8], inclusion of DctD250

should not be necessary to detect binding of holoenzyme to promoter sequences in the ChIP-chip assay. Protein-DNA complexes containing either $E\sigma^{54}$ or σ^{54} are pulled down in the ChIP with α-σ^{54}. The σ^{54} subunit is most likely to interact with the genome in the context of the RNA polymerase holoenzyme; however, σ^{54} has been shown to specifically bind in the absence of the core RNA polymerase at σ^{54}-dependent promoters that have a T-tract upstream of the GC in the -12 promoter element [45]. DNA fragments from the α-σ^{54} ChIP were labeled and applied to the same complete open reading frame arrays as used in the microarray analysis.

Since the use of the ORF arrays did not allow direct mapping of the binding sites, we combined the ChIP-chip data with *in silico* analysis to determine the potential σ^{54} DNA binding sites. A Position-Specific Score Matrix (PSSM) was created using 27 known or previously predicted σ^{54}-dependent promoters from *S*. Typhimurium LT2 (Additional file 2); the extent of each promoter sequence used for the PSSM (18 bp) was based on the consensus sequence for σ^{54}-dependent promoters defined by Barrios et al. [9] and comparison analysis of the known *Salmonella* σ^{54}-dependent promoters. This PSSM was applied with the Motif Locator program [46] to the enriched ORF sequence and 1000 bp of flanking sequence on both sides of the ORF to identify potential σ^{54} DNA binding sites. The size range of DNA fragments that were pulled down via ChIP and amplified by ligation-mediated PCR was 200–1000 bp long, as determined by agarose gel electrophoresis, suggesting that intergenic binding sites up to 1000 bp from the enriched ORF might be detected in the ChIP-chip assays.

σ^{54} binding to promoters for known, predicted, and novel σ^{54}-dependent operons

In the ChIP-chip assays with the WT and Δ*rpoN* strains, the promoter-proximal gene for all the 24 known and predicted σ^{54}-dependent operons and the 2 sRNA genes (Table 2) were enriched, as defined by a stringent cut-off, i.e. signal ratio ≥ 3 and p-value <0.02 (Table 3). The associated promoter sequences, as determined by the *in silico* analysis, had PSSM scores ranging from 10.9 to 23.6 and were within 27 to 154 bp of the enriched ORF. In the DNA microarrays, six of the known or predicted σ^{54}-dependent

operons did not appear up-regulated; but in the ChIP-chip assays the promoter regions for all six operons gave signal ratios ranging from 3.2- to 39-fold greater in WT than in Δ*rpoN* cells. The detection of the σ^{54}-dependent promoters for all the other known and predicted σ^{54}-dependent operons supports the efficacy of our approach to mapping potential σ^{54}-binding sites.

The ChIP-chip analyses also showed that only one (STM2938) of the 13 newly-identified, potential σ^{54}-dependent operons from the DNA microarray assays has a σ^{54} DNA binding site associated with it (Table 3), suggesting that the other 12 operons may be indirectly regulated by σ^{54}. The σ^{54} DNA binding site associated with STM2938, the *cas1*-like gene, is within the upstream gene, STM2939 (539 bp from the start of STM2938). Further characterization of this potential σ^{54} promoter is described below in the promoter-reporter analysis.

σ^{54} binding to newly identified potential promoter and regulatory sites

In total, 70 ORFs were each found to be enriched in 3 replicate samples for the WT cells as compared to the Δ*rpoN* cells in the ChIP-chip assays (Table 3). The potential σ^{54} binding site with the highest PSSM score within each enriched ORF or up to 1000 bp of flanking intergenic sequence was identified by Motif Locator and is reported in Table 3. For the 70 enriched ORFs, 29 of the associated binding sites mapped to intergenic regions and 41 of the potential σ^{54} binding sites were located within the enriched ORF (Figure 2). In determining the most likely binding site for an enriched ORF, sequence within an adjacent non-enriched ORF was not considered for potential σ^{54} binding sites, since the ORF containing the binding site should be enriched; therefore, even if a site with a higher PSSM score was located in an immediately adjacent non-enriched ORF, the next highest scoring site found in either the enriched ORF or adjacent intergenic sequence was reported as the potential binding site in Table 3. This reflects a limitation of the *in silico* prediction of σ^{54} binding sites based on a PSSM that was created with known and predicted intergenic promoter sequences; the sequences for intragenic promoters or for σ^{54}

binding sites that are regulatory sites, but not promoters, may differ enough to appreciably affect PSSM scores.

Consensus sequences were generated using WebLogo [47] for the intergenic and intragenic potential σ^{54} binding sequences and for the promoter sequences used to generate the PSSM (Figure 3). Noteworthy differences in the consensus sequence for the intragenic σ^{54} binding sites, as compared to the consensus sequences for the intergenic σ^{54} binding sites and PSSM promoters, are at the −23 and −11 positions, which each contribute in different ways to σ^{54}-promoter DNA interactions. The −23 A-T base pair is important in promoter recognition by σ^{54}; the winged helix-turn-helix DNA binding motif of σ^{54} makes base-specific contacts with the top strand GG at positions −26 and −25 and with the bottom strand T at position −23 [48]. The base pairs immediately adjacent to the conserved GC element in the -12 region of the promoter are involved in Eσ^{54} binding to form the stable closed complex; the bases on the bottom strand of the promoter at the -12 and −11 positions interact with σ^{54} in a short region of 'early melting' that stabilizes closed complex until the bEBP binds σ^{54} and activates the holoenzyme to transition to open complex [49]. The reduced conservation of nucleotide sequence at the −23 and −11 positions for the potential intragenic σ^{54} binding sites may reflect varied functionality of these intragenic sites, or a level of inaccuracy inherent to *in silico* prediction of the binding sites associated with enriched ORFs in the ChIP-chip assays.

The position and orientation for each potential σ^{54} binding site are indicated in Table 3 and summarized for all the binding sites in Figure 2. This information is useful in considering possible functions for the binding sites. For example, the 16 intragenic σ^{54} binding sites oriented in the opposing direction of the gene might regulate by transcription interference and/or anti-sense RNA [19]. Four intragenic σ^{54} binding sites are within 250 bp of the 5' end of a downstream gene, or a large intergenic region (>100 bp), and oriented in the direction such that they might act as promoters for the downstream gene or a sRNA [20]. Binding sites located near a functional σ^{54} promoter may serve to accelerate the search for the promoter by Eσ^{54} sliding from the secondary sites [50]; while binding sites adjacent to, or

overlapping, a σ^{70}- or σ^{54}-promoter may bind Eσ^{54} or σ^{54} and repress or activate transcription from the other promoter [51]. The possible functions of the σ^{54} binding sites are quite varied and many are dependent on whether the binding site can function as a promoter.

It is likely that our initial approach to defining the global binding sites of σ^{54} in *S.* Typhimurium LT2 resulted in an underestimation of the number of binding sites. Multiple sites within ~2,000 bp of an ORF would enrich one or two adjacent ORFs and, in our analysis, would have been counted as one site. In addition, since Eσ^{54}-promoter closed complexes are reversible [52], some complexes might have not been detected due to high disassociation rates; detection of these sites may be improved in the presence of DctD250, which stimulates conversion of closed complex to the more stable open complex, and rifampicin, which prevents extension of RNA past the second or third nucleotide [53], thus improving the chances of cross-linking Eσ^{54} at the promoter sequence [54].

Promoter-reporter analysis to determine activity for predicted promoter sequences

To assess the functionality of σ^{54} binding sites defined by the ChIP-chip and PSSM analyses, promoter-*lacZ* fusion assays were performed using several of these sequences. We had two goals in performing these assays. First, we wanted to further confirm the σ^{54}-dependent promoter activity for some of the predicted σ^{54}-dependent promoters that were up-regulated in the DNA microarrays and enriched in the ChIP-chip assays. Secondly, we wanted to test the σ^{54} binding site predictions from the ChIP-chip combined with PSSM analyses; i.e. does a potential σ^{54} binding site equate to a σ^{54}-dependent promoter? This promoter function assay is the initial exploration of the roles for σ^{54} binding sites located within intragenic regions.

Potential promoters were introduced upstream of a promoter-less *lacZ* gene in a reporter vector, either pDS11 or pDS12 (which differ only in their MCS sequence). These promoter-reporter plasmids were co-transformed along with the DctD250 expression plasmid into either WT or Δr*poN* cells. After induction of DctD250 expression, standard β-galactosidase assays were performed. The results from WT

were compared to those from the Δ*rpoN* mutant to determine whether activity seen was σ^{54}-dependent (Figure 4). For intergenic sequences that were known or predicted to be σ^{54}-dependent promoters, the results matched those observed in the DNA microarray assays (Figure 4A, Table 2). The *glnA*, *glnK*, and STM3521 (*rtcBA* operon) promoters showed strong σ^{54}-dependent activity. For the *glmY*, *glmZ*, *rpoH*, and *yeaG* promoters, transcription in the Δ*rpoN* mutant was either as high as or higher than in wild type cells. This is likely due to the presence of σ^{70}-type promoters in the cloned sequence. In addition to the σ^{54} dependent promoters, other promoters have been reported upstream of *glmY*, *glmZ*, *rpoH*, and *yeaG* [28,36,41].

A total of eight intragenic sites identified in the ChIP-chip assay were selected for functional analysis (Figure 4B). All of these predicted sites had PSSM scores >10. As shown in Figure 4C, the sites chosen represent a variety of configurations with regard to their position and orientation within the ORF as well as the position and orientation of downstream ORFs. Given the possible functions for an intragenic promoter sequence (e.g. promoter for a downstream gene or sRNA, generation of antisense RNA, etc.), results of our analysis allow us to determine which, if any, of these roles may be attributable to any of these promoters.

Comparing the levels of *lacZ* expression in wild type cells to those in Δ*rpoN* mutants, we found that four of the eight intragenic sites were able to function as σ^{54}-dependent promoters. For these sites, the difference observed between WT and Δ*rpoN* cells varied from 4.6-fold for the sequence located in STM2957 to 8.9-fold for the sequence within STM2939. Overall, the level of transcription from these promoters was relatively low, with Miller units ranging from ~30-100. The low activity levels may indicate that $E\sigma^{54}$ has a low affinity for these sequences or that the DctD250 is inefficient in productively engaging closed complexes formed at these sites. A subset of the promoter-reporter plasmids with intragenic sites that exhibited σ^{54}-dependent transcription were also assayed in WT cells versus WT-DctD250 to determine the dependence of transcription on the promiscuous, constitutive bEBP (Table 1). All three intragenic promoters assayed, STM0699, STM2430, and STM2939, gave low levels of β-

galactosidase activity in the absence of DctD250 and from 4.7- to 34-fold higher levels of β-galactosidase activity in the presence of DctD250. The σ^{54}-dependent transcription from intragenic binding sites suggests previously unrecognized regulatory functions for σ^{54} in *Salmonella*; however, it will be critical to characterize transcription from their chromosomal loci before biological functions can be ascribed.

Some of the potential promoter sequences that were assayed failed to show any transcriptional activity. There are a number of possible reasons for the lack of promoter activity for these sites. Two likely explanations are: 1) the wrong sequence was chosen as the binding site based on the PSSM score and proximity to the enriched ORF in the ChIP-chip assays, i.e. a lower scoring sequence near the enriched ORF was the actual σ^{54} binding site; or 2) the σ^{54} binding site does not function as a promoter but serves another regulatory role, such as an operator site for regulating promoter activity, a site for transient binding in facilitated diffusion, or a site for sequestering $E\sigma^{54}$ in order to increase local concentration (since σ^{70} has a higher affinity for RNAP [55]).

Summary of *S*. Typhimurium LT2 σ^{54} regulon and comparison to σ^{54} regulons of other bacteria

Figure 5 summarizes the results from the DNA microarray and promoter-fusion assays performed in the presence of DctD250 and ChIP-chip in the absence of DctD250 to characterize the σ^{54} regulon of *S.* Typhimurium LT2. Based on DNA microarray, there are 33 up-regulated operons (76 genes; Additional file 1); global ChIP-chip combined with *in silico* analysis revealed at least 70 σ^{54} binding sites (Table 3), of which 21 were associated with up-regulated operons from the DNA microarrays. The promoter-*lacZ* fusions with seven of the 29 intergenic σ^{54} binding sites and eight of the 41 intragenic σ^{54} binding sites showed DctD250- and σ^{54}-dependent expression for three intergenic sites (associated with up-regulated operons) and four intragenic σ^{54} binding sites (Table 1, Figure 4). The cellular functions impacted by genes in the σ^{54} regulon of *S.* Typhimurium LT2 are quite diverse, ranging from carbon-source and amino acid metabolism to response to stressors, such as nitric oxide and toxic levels of zinc (Table 2). Our results suggest that a new cellular process may be added to this extensive list—cell immunity through the

CRISPR system; the role of σ^{54} in regulating a *cas1*-related gene within an operon of CRISPR-associated genes is presently being investigated.

The σ^{54} global regulon of *S.* Typhimurium LT2 may differ from that of virulent *S.* Typhimurium isolates due to accumulated mutations in this extensively-used, laboratory strain, particularly the *rpoS* mutation that contributes to attenuation of the LT2 strain [42]. Changes in the level of expression of one sigma factor can alter the expression of genes that are expressed by different sigma factors [56]; for example, it has been shown that deletion of *rpoN* alters expression of σ^{S}-dependent promoters in *E. coli* [57]. We are currently characterizing the σ^{54} global regulon of the virulent strain *S.* Typhimurium 14028s.

The σ^{54} regulons in other δ/γ-proteobacteria have been characterized experimentally to varying extents [13-15,58-60]. Only in *Vibrio cholera* 037 strain V52 have both global transcripts and binding sites been characterized experimentally [14]. In *E. coli* MG1655 and *Geobacter sulfurreducens*, the global σ^{54} transcriptomes were determined and local σ^{54} binding sites associated with up-regulated genes were assessed by computational analysis and selected promoters were assessed experimentally [13,15]. The number and diversity of the operons that are directly controlled by σ^{54}-promoters in these δ/γ-proteobacteria are comparable to that of *S.* Typhimurim LT2. The greatest variability in the σ^{54} regulons of the γ-proteobacteria appears to be the location of σ^{54} binding sites. Zhao et al. [13] estimated 70 σ^{54} promoters in *E. coli* MG1655, of which 13 (18%) were intragenic or located between convergently transcribed genes. In *V. cholera*, Dong and Mekalanos [14] identified a total 68 σ^{54} binding sites, of which 35 (51%) were intragenic and, similarly, we found 70 potential σ^{54} binding sites of which 41 (58%) appear to be located in intragenic regions.

Does the success with DctD250 in characterizing the *S.* Typhimurium σ^{54} regulon predict utility of this constitutive, promiscuous activator in defining σ^{54} global regulons in bacteria from other classes in the Proteobacteria phylum, or from other phyla? The key to activation of $E\sigma^{54}$ by DctD250 in diverse bacteria is the ability of the activator to make the appropriate interactions with σ^{54} in the context of the $E\sigma^{54}$-promoter closed complex; thus, comparison of interacting regions of σ^{54} and bEBPs between *S.*

Typhimurium and phylogenetically diverse bacteria is a good predictor of success. Extensive characterization of bEBP activation of $E\sigma^{54}$ in closed complex has shown that the GAFTGA motif of the AAA+ ATPase domain plays a primary and essential role for productive interactions with $E\sigma^{54}$, which lead to transcriptional activation (reviewed in [8]); the GAFTGA motif is very highly conserved among bEBPs in all bacteria that encode σ^{54}, which includes bacteria from a majority of the eubacterial phyla [61]. It has not yet been determined which specific residues of σ^{54} are contacted by Loop 1 of the bEBP AAA+ ATPase domain, but it has been clearly demonstrated that multiple residues within the amino-terminal 50 amino acids of σ^{54} (Region I) are key determinants for activator interaction [62] and there is extensive conservation of amino acid sequence in Region I for σ^{54} from phylogenetically diverse bacteria [63]. Thus, the comparison of interacting regions of σ^{54} and the AAA+ ATPase domain among diverse bacteria predicts that DctD250 will be a valuable tool in characterizing the σ^{54} regulons in many bacteria.

Conclusions

The results of DNA microarray and promoter-*lacZ* fusion analyses of the σ^{54} regulon of S. Typhimurium LT2 in the presence of DctD250 support our initial hypothesis: the AAA+ ATPase activation domain of DctD can stimulate transcription from σ^{54}-dependent promoters in a constitutive and promiscuous manner, thereby facilitating the global characterization of σ^{54} regulons. Sixteen previously predicted σ^{54}-dependent operons were confirmed, and a new σ^{54}-dependent gene, *cas1*, was identified by the DNA microarray and ChIP-chip analyses. In addition, the ChIP-chip analyses indicate an excess of σ^{54} binding sites compared to the number of σ^{54}-dependent transcripts and a high percentage of intragenic binding sites, suggesting that $E\sigma^{54}$ and σ^{54} may have more regulatory functions than transcription initiation at the start of an operon or sRNA. The number of functional promoters located inside genes suggests a need to consider such promoters in bioinformatic analyses of transcription factor binding sites.

Materials and Methods

Bacterial strains, media, and enzymes

The parental strain, designated wild-type, in these experiments was *Salmonella enterica subspecies enterica serovar* Typhimurium LT2 derivative MS1868 [*leuA414*(Am) *hsdSB*(r^-m^+) $Fels^-$] [64]. An isogenic derivative, TRH134, has a deletion in *rpoN* (*ntrA*) from codons 8 through 455, rendering it auxotrophic for glutamine [65]. *S*. Typhimurium strains were cultured in either nutrient broth (NB; Difco Laboratories), MOPS minimal media [66], or nitrogen-limiting MOPS [67]. Media supplement concentrations were 5 mM L-glutamine (Gln), 40 μg/ml L-Leucine (Leu), and 10 mM L-glutamate (Glu). Cloning procedures were performed in *E. coli* DH5α cultured in Luria-Bertani medium (LB; Fisher Scientific). All strains were grown at 37°C. Antibiotics (Sigma-Aldrich) were used at the following concentrations (μg/ml) for *E. coli*/*S.* Typhimurium (NB)/ *S.* Typhimurium (MOPS), respectively: ampicillin (Amp) 80/120/50; spectinomycin (Spc) 50/125/50; streptomycin (Str) 25/75/0. All enzymes were purchased from New England Biolabs, unless otherwise indicated, and were used according to manufacturer's recommendations.

Plasmids

Plasmid pPBHP92 is a derivative of the expression vector pTrcHisC (Invitrogen) that expresses the *Sinorhizbium meliloti* DctD AAA+ ATPase domain (E141-S390, designated DctD250) with an N-terminal 6x-His tag. This plasmid was constructed by digestion of pHX182 [17] with NdeI, filling in the 5'-overhang with T4 DNA polymerase and subsequent digestion with XhoI. The blunt-XhoI fragment containing the truncated *dctD* was cloned into pTrcHisC, which had been cut with NheI, blunt-ended, and cut with XhoI. The truncated *dctD* is under control of P_{trc} and subject to repression by the vector-encoded LacI. The reporter plasmids used in these studies, pDS11 and pDS12, are both derivatives of pDV6 [25] that contain a promoter-less copy of *lacZ* downstream of a MCS region. The MCS region was generated by annealing two oligonucleotide primers (Additional file 3) which were then ligated into a pDV6

backbone that had been digested with BamHI and HindIII. pDS11 and pDS12 differ only in MCS sequence. Potential promoter sequences were amplified from *S.* Typhimurium LT2 genomic DNA using Taq polymerase and the primers in (Additional file 3) and cloned into pCR2.1 (Invitrogen). Sequencing analysis to determine accuracy and orientation was performed for all plasmids by Genewiz, Inc. (South Plainfield, NJ). Depending on their orientation in pCR2.1 potential promoter sequences were sub-cloned into pDS11/12 using XbaI and either KpnI or HindIII. Plasmid pTG4, which encodes the DctD AAA+ ATPase domain under control of P_{tac}/*lacI*^{q}, was created by amplifying the corresponding region of pPBHP92 using primers DctD-F/R (Additional file 3), digesting the product with BamHI and HindIII, followed by ligation into the similarly digested pKH66 [68].

Transcriptional profiling by microarrays

S. Typhimurium strains MS1868 and TRH134, each bearing plasmid pPBHP92 (WT+DctD250 and Δ*rpoN*+DctD250, respectively), were grown overnight at 37°C in NB-Amp. Cultures were sub-cultured in fresh medium and grown to mid-log phase (OD_{600} ≈ 0.8). Since the basal level of DctD250 expression from pPBHP92 was shown to optimally activate transcription from a σ^{54}-dependent *dctA*'-'*lacZ* reporter [17], IPTG induction was not used for these cultures. RNA isolated using the RNAeasy kit (Qiagen) was used to generate differentially labeled cDNA using reverse transcriptase as previously described [69]. Labeled cDNA was hybridized to DNA microarrays containing complete open reading frames (ORFs) from *S.* Typhimurium LT2 printed in triplicate [70]. Microarrays were scanned with a ScanArray Lite laser scanner (Packard BioChip Technologies, Billerica, MA) using ScanArray Express 1.1 software. Signal intensities were quantified using QuantArray 3.0 (Packard). The ratio of WT+DctD250 signal to Δ*rpoN*+DctD250 signal was determined for each of the triplicate spots and the median value for each ORF was used in the statistical analysis [70]. Data shown is the result of three biological replicates with statistical analysis performed using the WebArrayDB program [71,72]. The intensity values for the three biological replicates of WT+DctD250 and of Δ*rpoN*+DctD250 were

compared for the calculation of the p-values, where the null hypothesis was that the intensities for WT+DctD250 and Δ*rpoN*+DctD250 would be equivalent. Genes that displayed a WT+DctD250/Δ*rpoN*+DctD250 signal ratio of >3-fold with a p-value of <0.02 were considered to be up-regulated.

Chromatin immunoprecipitation (ChIP)

ChIP was carried out using the ChIP Assay kit (USB Corporation) essentially as described by the manufacturer's instructions. Briefly, 100 ml cultures of *S.* Typhimurium strains MS1868 (WT) and TRH134 (Δ*rpoN*) were grown overnight in NB at 37°C and sub-cultured in fresh medium the next day. Once cultures reached mid-log phase ($OD_{600} \approx 0.7$), cells were treated with formaldehyde (3 ml of a 37% solution per 100 ml of culture) for 10 min. at room temperature to cross-link proteins to DNA. Cross-linking was quenched by the addition of glycine (10 ml of 1.33 M solution per 100 ml of culture) and incubation at 4°C for 30 min. Cells were harvested, washed and lysed in accordance with kit instructions. Cells were lysed in two passages through a French pressure cell at 10,000 psi. Cell extracts were clarified and pre-cleared with the provided protein A-Sepharose bead slurry per the kit instructions. 0.6 ml of the resulting extracts were mixed with 2 µl of rabbit anti-serum against *S.* Typhimurium σ^{54} [73] and incubated with gentle shaking overnight at 4°C. The next day, 50 µl of protein A-Sepharose bead slurry was added to each sample, incubated 1 hr at room temperature and collected by centrifugation. The beads were washed, and protein-DNA complexes were eluted from the beads and disrupted per the supplier's instructions. DNA was purified from each sample using the Qiagen PCR purification kit.

ChIP-chip assays

Purified ChIP DNA was amplified by ligation-mediated PCR, adapting the procedure found at [http://www.flychip.org.uk/protocols/archive_protocols/lm_pcr.php]. Linkers consisting of complementary oligonucleotides (LM-PCR; Additional file 3) were ligated to the ends of purified DNA

repaired with T4 DNA polymerase. Ligated was purified using the Qiagen PCR purification kit and the DNA was amplified with Taq polymerase (Fermentas; Burlington, ON) using LM-PCR-R as the PCR primer and the following cycling conditions: 55°C—2 min (1×); 72°C—5 min (1×); 94°C—5 min (1×); 94°C—1 min, 55°C—1 min, 72°C—1 min (24×); 72°C—5 min (1×); 4°C—hold. The resulting amplicons, most of which were 300-800 bp, were purified using the Qiagen PCR purification kit and to prepare dye-labeled DNA (Cy3 or Cy5) for hybridization to the *S.* Typhimurium complete ORF microarray. Microarrays were scanned and analyzed as above. ChIP-chip was performed on three biological replicates for the WT and Δ*rpoN* strains; the statistical analysis of the data was performed as described for the microarray data.

Identifying candidate σ^{54} binding sites in the *S.* Typhimurium genome

The Motif Locator program [http://www.cmbl.uga.edu/software.html] was used to identify candidate σ^{54} binding sites. The program applies the standard position-specific score matrix (PSSM) described in [46]. We used a PSSM derived from the alignment of 27 high-confidence sites supported by experimental evidence in either *Salmonella* or *E. coli* (Additional file 2). Background nucleotide frequencies were assigned in accordance with the genomic G+C content. Pseudo-counts equal to the background frequencies were used in PSSM construction. For ORFs discovered in the ChIP-chip assay, this matrix was used to determine the most likely binding site either within the ORF itself or in the region ±200, 500, or 1000 bp surrounding the gene.

β-Galactosidase assays

The DctD250 expression plasmid pTG4 was introduced into *S.* Typhimurium MS1868 and TRH134 by electroporation using a GenePulser 2 system (BioRad; Hercules, CA) and the resulting transformants were electroporated with pDS11 or pDS12 reporter constructs containing potential σ^{54}-dependent promoter sequences. Overnight cultures grown in MOPS-LeuGln, or nitrogen-limiting MOPS-

Glu, with the appropriate antibiotics were sub-cultured into fresh medium, grown to $OD_{600} \approx 0.2$, and induced with 50 μM IPTG (empirically determined IPTG concentration for optimal expression of DctD250 from pTG4 to activate known σ^{54}-dependent promoters on the reporter plasmids). Cultures were induced for 6 hours and β-galactosidase activity was measured as described previously [74] with the following changes: 1) assays were performed at 37°C, and 2) after stopping reaction, samples were centrifuged and OD_{420} of the supernatant was measured, eliminating the OD_{550} correction for cell debris. Activity was calculated as Miller units: $[(OD_{420} \times 1000)/(OD_{600} \times \text{Time (min)} \times \text{volume (ml)})]$ [74]. Ratios of activity in wild type/Δ*rpoN* cells were compared and analyzed using a 2-tailed Student's T-test. Data shown for each promoter construct represents ≥3 biological replicates.

Accession number for microarray and ChIP-chip data

The DNA microarray and ChIP-chip data were deposited in NCBI GEO under accession number GSE25849.

Back Matter

Abbreviations

Amp, Ampicillin; bEBP, Bacterial enhancer-binding protein; ChIP, Chromatin immunoprecipitation; Eσ, RNA polymerase holoenzyme; LB, Luria-Bertani media; LM-PCR, Ligation-mediated PCR; NB, Nutrient broth; RNAP, RNA polymerase; Spc, Spectinomycin; Str, Streptomycin; UAS, Upstream activation sequence

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

DS harvested RNA, carried out promoter-reporter assays, and drafted and revised and prepared the manuscript. JF prepared cDNA and preformed DNA microarrays assays and applied DNA from the ChIP to the microarrays. SP and MM produced the microarrays and protocols and performed the statistical analysis of the microarray therein. JM performed the bioinformatic analyses. TH conceived, designed, and coordinated the study, harvested RNA, and performed ChIP pulldowns. AK conceived, designed, and coordinated the study, harvested RNA, and drafted and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank the following people for their assistance on this project: Ashley Bono, Uchenna Ewulonu, and Trevor Wright generated reporter constructs. Tanya Grancharova sub-cloned DctD250. Sonya Chelliah performed ChIP pull-downs. Jennifer Turpin prepared and applied cDNA to the microarrays. This work was supported by National Science Foundation Grant MCB-1051175 (to T.R.H. and A.C.K.). MM and SP were supported in part by NIH Contract No. HHSN272200900040C from PATRIC and grants AI039557 AI052237, AI073971, AI075093, AI077645 AI083646, USDA grants 2009-03579-30127 and 2011-67017-30127, the Binational Agricultural Research and Development Fund. JM was supported in part by the National Science Foundation grant DBI-0950266.

References

1. Gruber TM, Gross CA: **Multiple sigma subunits and the partitioning of bacterial transcription space.** *Annu Rev Microbiol* 2003, **57:**441–466.

2. Ferro-Luzzi Ames G, Nikaido K: **Nitrogen regulation in *Salmonella typhimurium*. Identification of an NtrC protein-binding site and definition of a consensus binding sequence.** *EMBO J* 1985, **4**(2)**:**539–547.

3. Leonhartsberger S, Huber A, Lottspeich F, Bock A: **The *hydH*/*G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system.** *J Mol Biol* 2001, **307**(1)**:**93–105.

4. Palacios S, Escalante-Semerena JC: ***prpR, ntrA,* and *ihf* functions are required for expression of the *prpBCDE* operon, encoding enzymes that catabolize propionate in *Salmonella enterica* serovar Typhimurium LT2.** *J Bacteriol* 2000, **182**(4)**:**905–910.

5. Weiner L, Brissette JL, Model P: **Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma^{54} and modulated by positive and negative feedback mechanisms.** *Genes Dev* 1991, **5**(10)**:**1912–1923.

6. Niehus E, Gressmann H, Ye F, Schlapbach R, Dehio M, Dehio C, Stack A, Meyer TF, Suerbaum S, Josenhans C: **Genome-wide analysis of transcriptional hierarchy and feedback regulation in the flagellar system of *Helicobacter pylori*.** *Mol Microbiol* 2004, **52**(4)**:**947–961.

7. Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, *et al*: **Modus operandi of the bacterial RNA polymerase containing the σ^{54} promoter-specificity factor.** *Mol Microbiol* 2008, **68**(3)**:**538–546.

8. Bush M, Dixon R: **The role of bacterial enhancer binding proteins as specialized activators of sigma^{54}-dependent transcription.** *Microbiol Mol Biol Rev* 2012, **76**(3)**:**497–529.

9. Barrios H, Valderrama B, Morett E: **Compilation and analysis of sigma(54)-dependent promoter sequences.** *Nucleic Acids Res* 1999, **27**(22)**:**4305–4313.

10. Ghosh T, Bose D, Zhang X: **Mechanisms for activating bacterial RNA polymerase.** *FEMS Microbiol Rev* 2010, **34**(5)**:**611–627.

11. Belitsky BR, Sonenshein AL: **An enhancer element located downstream of the major glutamate dehydrogenase gene of *Bacillus subtilis*.** *Proc Natl Acad Sci U S A* 1999, **96**(18)**:**10290–10295.

12. Reitzer LJ, Magasanik B: **Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter.** *Cell* 1986, **45**(6)**:**785–792.

13. Zhao K, Liu M, Burgess RR: **Promoter and regulon analysis of nitrogen assimilation factor, sigma^{54}, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis.** *Nucleic Acids Res* 2010, **38**(4)**:**1273–1283.

14. Dong TG, Mekalanos JJ: **Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52.** *Nucleic Acids Res* 2012, **40**(16)**:**7766–7775.

15. Leang C, Krushkal J, Ueki T, Puljic M, Sun J, Juarez K, Nunez C, Reguera G, DiDonato R, Postier B, *et al*: **Genome-wide analysis of the RpoN regulon in *Geobacter sulfurreducens*.** *BMC Genomics* 2009, **10:**331.

16. Brahmachary P, Dashti MG, Olson JW, Hoover TR: ***Helicobacter pylori* FlgR is an enhancer-independent activator of σ^{54}-RNA polymerase holoenzyme.** *J Bacteriol* 2004, **186**(14)**:**4535–4542.

17. Xu H, Gu B, Nixon BT, Hoover TR: **Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a sigma^{54}-dependent transcriptional activator.** *J Bacteriol* 2004, **186**(11)**:**3499–3507.

18. Studholme DJ: **Enhancer-dependent transcription in *Salmonella enterica* Typhimurium: new members of the sigma^{N} regulon inferred from protein sequence homology and predicted promoter sites.** *J Mol Microbiol Biotechnol* 2002, **4**(4)**:**367–374.

19. Georg J, Hess WR: **cis-antisense RNA, another level of gene regulation in bacteria.** *Microbiol Mol Biol Rev* 2011, **75**(2)**:**286–300.

20. Gottesman S, Storz G: **Bacterial small RNA regulators: versatile roles and rapidly evolving variations.** *Cold Spring Harbor perspectives in biology* 2011, **3:**12.

21. Kroger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, *et al*: **The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium.** *Proc Natl Acad Sci U S A* 2012, **109**(20)**:**E1277–1286.

22. van Heeswijk WC, Hoving S, Molenaar D, Stegeman B, Kahn D, Westerhoff HV: **An alternative PII protein in the regulation of glutamine synthetase in *Escherichia coli*.** *Mol Microbiol* 1996, **21**(1)**:**133–146.

23. Genschik P, Drabikowski K, Filipowicz W: **Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma^{54}-regulated operon.** *J Biol Chem* 1998, **273**(39)**:**25516–25526.

24. Huala E, Stigter J, Ausubel FM: **The central domain of *Rhizobium leguminosarum* DctD functions independently to activate transcription.** *J Bacteriol* 1992, **174**(4)**:**1428–1431.

25. Perkins-Balding D, Duval-Valentin G, Glasgow AC: **Excision of IS*492* requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*.** *J Bacteriol* 1999, **181**(16)**:**4937–4948.

26. Porwollik S, Frye J, Florea LD, Blackmer F, McClelland M: **A non-redundant microarray of genes for two related bacteria.** *Nucleic Acids Res* 2003, **31**(7)**:**1869–1876.

27. Klose KE, Mekalanos JJ: **Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates *Salmonella typhimurium* virulence.** *Infect Immun* 1997, **65**(2)**:**587–596.

28. Gopel Y, Luttmann D, Heroven AK, Reichenbach B, Dersch P, Gorke B: **Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae.** *Nucleic Acids Res* 2011, **39**(4)**:**1294–1309.

29. Hirschman J, Wong PK, Sei K, Keener J, Kustu S: **Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription in vitro: evidence that the *ntrA* product is a sigma factor.** *Proc Natl Acad Sci U S A* 1985, **82**(22)**:**7525–7529.

30. Zimmer DP, Soupene E, Lee HL, Wendisch VF, Khodursky AB, Peter BJ, Bender RA, Kustu S: **Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation.** *Proc Natl Acad Sci U S A* 2000, **97**(26)**:**14674–14679.

31. Kiupakis AK, Reitzer L: **ArgR-independent induction and ArgR-dependent superinduction of the *astCADBE* operon in *Escherichia coli*.** *J Bacteriol* 2002, **184**(11)**:**2940–2950.

32. Riordan JT, Tietjen JA, Walsh CW, Gustafson JE, Whittam TS: **Inactivation of alternative sigma factor 54 (RpoN) leads to increased acid resistance, and alters locus of enterocyte effacement (LEE) expression in Escherichia coli O157: H7.** *Microbiology* 2010, **156**(Pt 3)**:**719–730.

33. Gardner AM, Gessner CR, Gardner PR: **Regulation of the nitric oxide reduction operon (*norRVW*) in *Escherichia coli*. Role of NorR and σ^{54} in the nitric oxide stress response.** *J Biol Chem* 2003, **278**(12)**:**10081–10086.

34. Maier T, Binder U, Bock A: **Analysis of the *hydA* locus of *Escherichia coli*: two genes (*hydN* and *hypF*) involved in formate and hydrogen metabolism.** *Arch Microbiol* 1996, **165**(5)**:**333–341.

35. Lutz S, Bohm R, Beier A, Bock A: **Characterization of divergent NtrA-dependent promoters in the anaerobically expressed gene cluster coding for hydrogenase 3 components of *Escherichia coli*.** *Mol Microbiol* 1990, **4**(1)**:**13–20.

36. Janaszak A, Majczak W, Nadratowska B, Szalewska-Palasz A, Konopa G, Taylor A: **A sigma^{54}-dependent promoter in the regulatory region of the *Escherichia coli rpoH* gene.** *Microbiology* 2007, **153**(Pt 1)**:**111–123.

37. Ramirez-Santos J, Collado-Vides J, Garcia-Varela M, Gomez-Eichelmann MC: **Conserved regulatory elements of the promoter sequence of the gene *rpoH* of enteric bacteria.** *Nucleic Acids Res* 2001, **29**(2)**:**380–386.

38. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, Buck M: **Identification of a new member of the phage shock protein response in *Escherichia coli*, the phage shock protein G (PspG).** *J Biol Chem* 2004, **279**(53)**:**55707–55714.

39. Birkmann A, Sawers RG, Bock A: **Involvement of the *ntrA* gene product in the anaerobic metabolism of *Escherichia coli*.** *Mol Gen Genet* 1987, **210**(3)**:**535–542.

40. Miller KA, Phillips RS, Mrazek J, Hoover TR: ***Salmonella*** **utilizes D-glucosaminate via a mannose family phosphotransferase system permease and associated enzymes**. *J Bacteriol* 2013, **195** (18): 4057-4066.

41. Ibanez-Ruiz M, Robbe-Saule V, Hermant D, Labrude S, Norel F: **Identification of RpoS (sigma^{S})-regulated genes in *Salmonella enterica* serovar typhimurium.** *J Bacteriol* 2000, **182**(20)**:**5749–5756.

42. Swords WE, Cannon BM, Benjamin WH Jr: **Avirulence of LT2 strains of *Salmonella typhimurium* results from a defective *rpoS* gene.** *Infect Immun* 1997, **65**(6)**:**2451–2453.

43. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P: **CRISPR provides acquired resistance against viruses in prokaryotes.** *Science* 2007, **315**(5819)**:**1709–1712.

44. Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA: **Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense.** *Structure* 2009, **17**(6)**:**904–912.

45. Buck M, Cannon W: **Specific binding of the transcription factor σ^{54} to promoter DNA.** *Nature* 1992, **358**(6385)**:**422–424.

46. Mrazek J: **Finding sequence motifs in prokaryotic genomes–a brief practical guide for a microbiologist.** *Brief Bioinform* 2009, **10**(5)**:**525–536.

47. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: a sequence logo generator.** *Genome Res* 2004, **14**(6)**:**1188–1190.

48. Doucleff M, Pelton JG, Lee PS, Nixon BT, Wemmer DE: **Structural basis of DNA recognition by the alternative sigma-factor, σ^{54}.** *J Mol Biol* 2007, **369**(4)**:**1070–1078.

49. Cannon W, Wigneshweraraj SR, Buck M: **Interactions of regulated and deregulated forms of the sigma^{54} holoenzyme with heteroduplex promoter DNA.** *Nucleic Acids Res* 2002, **30**(4)**:**886–893.

50. Li GW, Elf J: **Single molecule approaches to transcription factor kinetics in living cells.** *FEBS Lett* 2009, **583**(24)**:**3979–3983.

51. Johansson LU, Solera D, Bernardo LM, Moscoso JA, Shingler V: **sigma^{54}-RNA polymerase controls sigma^{70}-dependent transcription from a non-overlapping divergent promoter.** *Mol Microbiol* 2008, **70**(3)**:**709–723.

52. Friedman LJ, Gelles J: **Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation.** *Cell* 2012, **148**(4)**:**679–689.

53. Feklistov A, Mekler V, Jiang Q, Westblade LF, Irschik H, Jansen R, Mustaev A, Darst SA, Ebright RH: **Rifamycins do not function by allosteric modulation of binding of Mg^{2+} to the RNA polymerase active center.** *Proc Natl Acad Sci U S A* 2008, **105**(39)**:**14820–14825.

54. Raffaelle M, Kanin EI, Vogt J, Burgess RR, Ansari AZ: **Holoenzyme switching and stochastic release of sigma factors from RNA polymerase in vivo.** *Mol Cell* 2005, **20**(3)**:**357–366.

55. Maeda H, Fujita N, Ishihama A: **Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase.** *Nucleic Acids Res* 2000, **28**(18)**:**3497–3503.

56. Grigorova IL, Phleger NJ, Mutalik VK, Gross CA: **Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA.** *Proc Natl Acad Sci U S A* 2006, **103**(14)**:**5332–5337.

57. Dong T, Yu R, Schellhorn H: **Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*.** *Mol Microbiol* 2011, **79**(2)**:**375–386.

58. Chaudhuri RR, Yu L, Kanji A, Perkins TT, Gardner PP, Choudhary J, Maskell DJ, Grant AJ: **Quantitative RNA-seq analysis of the *Campylobacter jejuni* transcriptome.** *Microbiology* 2011, **157**(Pt 10)**:**2922–2932.

59. Damron FH, Owings JP, Okkotsu Y, Varga JJ, Schurr JR, Goldberg JB, Schurr MJ, Yu HD: **Analysis of the *Pseudomonas aeruginosa* regulon controlled by the sensor kinase KinB and sigma factor RpoN.** *J Bacteriol* 2012, **194**(6)**:**1317–1330.

60. Dasgupta N, Wolfgang MC, Goodman AL, Arora SK, Jyot J, Lory S, Ramphal R: **A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in *Pseudomonas aeruginosa*.** *Mol Microbiol* 2003, **50**(3)**:**809–824.

61. Francke C, Groot Kormelink T, Hagemeijer Y, Overmars L, Sluijter V, Moezelaar R, Siezen RJ: **Comparative analyses imply that the enigmatic Sigma factor 54 is a central controller of the bacterial exterior.** *BMC Genomics* 2011, **12:**385.

62. Xiao Y, Wigneshweraraj SR, Weinzierl R, Wang YP, Buck M: **Construction and functional analyses of a comprehensive sigma54 site-directed mutant library using alanine-cysteine mutagenesis.** *Nucleic Acids Res* 2009, **37**(13)**:**4482–4497.

63. Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD: **The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor.** *J Bacteriol* 2000, **182**(15)**:**4129–4136.

64. Grana D, Youderian P, Susskind MM: **Mutations that improve the *ant* promoter of *Salmonella* phage P22.** *Genetics* 1985, **110**(1)**:**1–16.

65. Kelly MT, Hoover TR: **Mutant forms of *Salmonella typhimurium* sigma^{54} defective in transcription initiation but not promoter binding activity.** *J Bacteriol* 1999, **181**(11)**:**3351–3357.

66. Maloy S, Stewart VJ, Taylor RK: *Genetic Analysis of Pathogenic Bacteria: A Laboratory Manual.* Plainview, NY: Cold Spring Harbor Laboratory Press; 1996.

67. Kustu SG, McFarland NC, Hui SP, Esmon B, Ames GF: **Nitrogen control of *Salmonella typhimurium*: co-regulation of synthesis of glutamine synthetase and amino acid transport systems.** *J Bacteriol* 1979, **138**(1)**:**218–234.

68. Hughes KT, Youderian P, Simon MI: **Phase variation in *Salmonella*: analysis of Hin recombinase and hix recombination site interaction in vivo.** *Genes Dev* 1988, **2**(8)**:**937–948.

69. Wang Q, Frye JG, McClelland M, Harshey RM: **Gene expression patterns during swarming in *Salmonella typhimurium*: genes specific to surface growth and putative new motility and pathogenicity genes.** *Mol Microbiol* 2004, **52**(1)**:**169–187.

70. Porwollik S, McClelland M: **Lateral gene transfer in *Salmonella*.** *Microbes Infect* 2003, **5**(11)**:**977–989.

71. Xia XQ, McClelland M, Porwollik S, Song W, Cong X, Wang Y: **WebArrayDB: cross-platform microarray data analysis and public data repository.** *Bioinformatics* 2009, **25**(18)**:**2425–2429.

72. Xia X, McClelland M, Wang Y: **WebArray: an online platform for microarray data analysis.** *BMC Bioinforma* 2005, **6:**306.

73. Kelly MT, Hoover TR: **The amino terminus of *Salmonella enterica* serovar Typhimurium sigma^{54} is required for interactions with an enhancer-binding protein and binding to fork junction DNA.** *J Bacteriol* 2000, **182**(2)**:**513–517.

74. Miller JH: *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria.* Plainview, NY: Cold Spring Harbor Press; 1992.

Table 2.1 DctD250-dependent activity of predicted and potential σ^{54}-dependent promoters

	Miller Units[b]		
Promoter[a]	**WT**	**WT + DctD250**	**WT-N**
Intergenic:			
STM0462 (*glnKamtB*)	1.1 ±0.1	180 ±22	1778 ±28
STM3521 (*rtcBA*)	14 ±6.8	220 ±55	N.D.
Intragenic:			
STM0699	5.3 ±1.6	56 ±6.2	N.D.
STM2430	19 ±3.8	90 ±5.5	N.D.
STM2939	2.2 ±1.4	76 ±17	N.D.
Controls:			
IS*492*-CJ (pDV6)	6400 ±2200	9000 ±4100	N.D.
Empty vector (pDS12)	2.0 ±0.5	8.4 ±3.5	4 ±0

[a]Promoters assayed in *lacZ*-reporter plasmids pDS11 and pDS12: intergenic promoters are predicted promoters for the *glnALG*, *glnKamtB* and *rtcBA* operons; intragenic promoters were identified in the ChIP-chip assay (see Results). Controls are the σ^{70}-dependent, circle-junction promoter from IS*492* and the empty vector, pDS12.

[b]β-galactosidase assays were performed in MOPS minimal medium (WT and WT+DctD250) or in nitrogen-limiting MOPS (WT-N).

Table 2.2 Microarray results for known, predicted, and novel σ^{54}-dependent operons and sRNA genes of *S*. Typhimurium

Locus Tag[a]	Gene Symbol[b]	Function	bEBP[c]	WT/Δ*rpoN*[d]	Ref.[e]
***Known σ^{54}-dependent operons and sRNA genes*:**					
STM0368-71	*prpBCDE*	Proprionate catabolism (putative)	PrpR	**45**	[4]
STM0830-28	*glnHPQ*	Glutamine high-affinity transporter	NtrC	**7.1**	[27]
STM2355	*argT*	Lysine/arginine/ornithine transport protein	NtrC	**3.5**	[2]
STM_R0152	*glmY*	GlmY sRNA	GlrR	0.9	[28]
STM_R0167	*glmZ*	GlmZ sRNA	GlrR	1.1	[28]
STM4007-05	*glnALG*	Glutamine synthetase	NtrC	**48**	[29]
***Predicted σ^{54}-dependent operons*:**					
STM0462-63	*glnK amtB*	hypothetical protein	NtrC	**3.6**	[22]
STM0577-72		PTS (putative)	STM0571	**67**	[18]
STM0649.S-53		Hydrolase (putative)	STM0652	**11**	[18]
STM0665-62	*gltIJKL*	Glutamate/aspartate transporter	NtrC	1.8	[18,30]
STM1285-84	*yeaGH*	Serine protein kinase (putative)	NtrC	2.5	[18,30]
STM1303-07	*astCABDE*	Arginine/ornithine/glutamine metabolism	NtrC	**2.4**[f]	[31,32]
STM1690-86	*pspABCDE*	Phage shock proteins	PspF	**17**	[5]
STM2360-56	------*ubiX*	Amino acid transport (putative)	STM2361	**100**	[18]
STM2840-41	*norV ygbD*	Nitric oxide reductase	NorR	**16**	[18,33]
STM2843-42	*hydN hypF*	Hydrogenase maturation proteins	FhlA	**13**	[34]
STM2853-44	*hycABCDEFGHI-*	Hydrogenase 3	FhlA	**26**	[35]
STM2854-58	*hypABCDE*	Formate-hydrogen lyase system	FhlA	**5.6**	[35]
STM3521-18	*-rtcBA*	RNA repair system (putative)	RtcR	**71**	[23]
STM3568	*rpoH*	Heat shock sigma factor (σ^{32})		1.7	[36,37]
STM3772-66		PTS (putative)	STM3773	**39**	[18]
STM4172	*zraP*	Zinc resistance-associated protein	ZraR	**16**	[3,18]
STM4173-74	*hydHG*	Zinc resistance two-component system	ZraR	**3.7**	[3]
STM4244	*pspG*	Phage shock protein	PspF	1.4	[38]

STM4285	*fdhF*	Formate dehydrogenase	FhlA	**29**	[39]
STM4535-40.s		PTS (putative)	STM4534	**16**	[18]
***Novel σ^{54}-dependent operon*:**					
STM2944-2937		CRISPR-associated genes		**1.6**[f]	--

[a]Locus tags for genes within operons or sRNA genes are grouped by those previously shown to be σ^{54}-dependent in *Salmonella*, previously predicted to be σ^{54}-dependent, or identified in this study as encoded in a novel σ^{54}-dependent transcript. Locus tags for operons that are up-regulated are in bold type.
[b]Genes that have not been assigned a gene symbol are represented by a dash (−).
[c]Known or predicted bacterial enhancer-binding protein (bEBP) that activates the σ^{54}-dependent operon.
[d]Signal ratio for the first gene in the operon in WT and Δ*rpoN* strains expressing DctD250 from pPHBP92. Operons with at least one gene with a signal ratio >3 and p-value <0.02 are considered up-regulated by RpoN; signal ratios above the 3-fold cut off are in bold type. Data for all genes in these operons can be found in Additional File1.
[e]References for operons shown to be σ^{54}-dependent in *Salmonella* and for operons either determined to be σ^{54}-dependent in other bacterial genera or predicted to be regulated by σ^{54} in *Salmonella* are listed.
[f]The first gene in the operon was <3-fold up-regulated, but other genes in the operon were >3-fold up-regulated.

Table 2.3 ChIP-chip signal ratios, PSSM scores, and predicted binding sites for ORFs enriched in the presence of σ^{54}

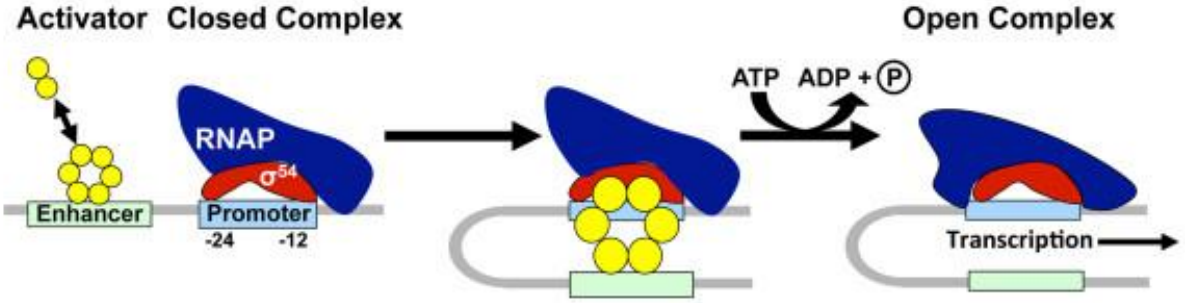
Locus Tag[a]	Gene Name	Signal Ratio[b]	Orientation[c]	PSSM[d]	Start	End	Sequence
Sites located within intergenic regions:							
STM0368	*prpB*	11	+	20.7	417914	417931	TGGCATAGCCTTTGCTTT
STM0448	*clpP*	4.6	+	14.6	503028	503045	TGTCACGTATTTTGCATG
STM0462	*glnK*	3.2	+	20.0	520445	520462	TGGCACATCCTTTGCAAT
STM0577		8.3	+	17.9	636883	636866	TGGCACGCCGTTTGCCAT
STM0649.S		6.9	+	18.7	711945	711962	TGGCACGCCTTTTGATTA
STM0665	*gltI*	3.2	+	22.1	730107	730090	TGGCACGTCTATTGCTTT
STM0830	*glnH*	16	+	20.8	897079	897062	TGGCATGATTTTTTCATT
STM1285	*yeaG*	4.1	+	21.5	1363884	1363867	TGGCATGAGAGTTGCTTT
STM1303	*astC*	4.0	+	21.7	1382105	1382122	TGGCACGAATGCTGCAAT
STM1690	*pspA*	23	+	20.3	1782486	1782469	TGGCACGCAAATTGTATT
STM2355	*argT*	3.5	+	16.2	2466359	2466376	TGGCATAAGACCTGCATG
STM2360		4.8	+	23.6	2472731	2472714	TGGCATGCCTTTTGCTTT
STM_R0152	*glmY*	31	+	20.6	2707874	2707857	TGGCACAATTACTGCATA
STM2809	*proV*	9.5	-	14.6	2955839	2955822	TGGCATGAATATTGCGAG
STM2840		6.5	+	20.1	2985009	2985026	TGGCACACTAGCTGCAAT
STM2843	*hydN*	23	+	17.1	2990721	2990704	TGGCACGATTCGTGTATA
STM2853	*hycA*	31	+	17.9	2999639	2999622	TGGCATGGAAAATGCTTA
STM2854	*hypA*	71	+	22.4	2999753	2999770	TGGCATAAATATTGCTTT
STM3521		15	+	21.2	3684734	3684717	TGGCACGCTGGTTGCAAT
STM3568	*rpoH*	22	+	18.9	3736836	3736819	TGGCACGGTTGTTGCTCG
STM3772		3.6	+	20.3	3972484	3972467	TGGCACAACCTTTGCTCT
STM_R0167	*glmZ*	15	+	19.5	4141620	4141637	TGGCACGTTATGTGCAAT
STM4007	*glnA*	4.2	+	19.2	4217110	4217093	TGGCACAGATTTCGCTTT
STM4172	*zraP*	29	+	17.4	4388217	4388234	TGGCACGGAAGATGCAAG
STM4173	*hydH*	4.8	+	20.1	4388385	4388402	TGGCATGATCTCTGCTTA
STM4244	*pspG*	39	+	19.4	4465042	4465059	TGGCATGATTTTTGTAAG
STM4285	*fdhF*	10	+	18.2	4527564	4527547	TGGCATAAAACATGCATA
STM4367	*yjeB*	3.8	+	14.1	4610407	4610424	TGGCAGATATTTTGCTTG
STM4535		12	+	18.3	4794881	4794898	TGGCACGCCGCTTGCTCT
Sites located within the enriched ORF:							
STM0131	*ftsQ*	7.6	+	6.2	153598	153615	TGGAACGCGTCTTGCAGG
STM0155		4.1	+	9.5	182767	182784	CGGCATGGCATTTGCCAG
STM0322	*proA*	7.8	-	11.3	368058	368041	CGGCACAGTTTATGCAAG
STM0332		3.0	-	8.1	376286	376269	TGGCCAGAAATATGCTTA
STM0526	*ylbA*	4.3	+	9.1	588233	588216	TGGCATTAATGCTGCATC
STM0699		14	+	13.7	761691	761674	TGGCATCGATATTGCAAA
STM0879^	*potH*	5.2	+	12.1	951550	951567	TGGCAGGAGTTTTTCAAT
STM0884	*ulaA*	5.2	+	10.0	955545	955562	CGGCACGATTTTTTCCAT
STM0901		3.9	+	11.7	971761	971778	TGGCATGAAACTTGTCAC
STM0940		9.5	+	13.7	1018097	1018080	TGGCCTGAATCTTGCTAA
STM0961		7.6	-	17.7	1041686	1041669	TGGCATGAAAGCTGCTCA

STM1361	*ydiM*	3.6	+	11.6	1443903	1443886	TGGCATTCTTTATGCTCA
STM1390	*orf242*	8.9	-	12.9	1475563	1475546	TGGCATCATTATTGCCTA
STM1409	*ssaJ*	5.0	+	6.5	1490273	1490290	TGGCATGAAGGTTCATCG
STM1586		6.6	-	13.5	1672845	1672862	TGGCAAGAATATTGCCAT
STM1594	*srfB*	4.8	+	13.4	1681565	1681582	TGGCACACGTTTTGCGCT
STM1665		4.2	-	11.7	1759185	1759168	TGGCATCATTTTTTCAAG
STM1904^	*yecN*	3.6	+	5.9	1998988	1999005	TGGCAAACCTGTGGTATA
STM1928	*otsA*	5.3	+	8.1	2023398	2023381	TGGCAGGAGCGTTTTATT
STM1990	*yedA*	5.6	+	14.4	2072998	2073015	TGGCGCGCTTTTTGCCTT
STM2033	*cbiC*	4.4	-	2.2	2111221	2111238	CGGTATAAATAATGCACG
STM2115	*wcaA*	4.3	-	9.5	2198775	2198792	TGGCATATAAATTGAGAT
STM2181	*yohJ*	15	+	4.9	2277993	2278010	AGGCATTTTTCTTGCATC
STM2430	*cysK*	4.8	-	11.7	2544207	2544190	TGGCATCACTGTTGCAGT
STM2475		9.0	-	1.0	2585621	2585638	TGGCACATCAGGCAAAAG
STM2476	*ypfG*	3.1	+	12.7	2586874	2586857	TGGCAGGTCACCTGCAAT
STM2517	*sinH*	4.6	-	11.1	2650462	2650479	TGGTACGGATCTTGCCAT
STM2563	*yfhG*	4.7	-	6.8	2705786	2705803	CGGCGTAATTTTTGCATC
STM2939	*ygcH*	10	+	10.9	3080061	3080044	CGGCACAGCTCTTGCATC
STM2957	*rumA*	5.5	+	14.5	3105809	3105792	TGGAACGCTTTTCGCATT
STM3072		4.1	-	6.9	3234181	3234164	TGGCCCATTGAATGCATC
STM3302	*yhbE*	5.5	+	12.6	3472042	3472025	TGGCATGATGGTCGCCAG
STM3535	*glgA*	8.0	+	11.6	3702315	3702298	AGGCATGTTTTATGCAAA
STM3721	*rfaP*	13.5	+	8.3	3916283	3916300	TGGTACGTAAAATGCACG
STM3863^		7.5	+	11.1	4072959	4072942	TGGCGCGATTATTGCCAG
STM3919	*wzzE*	4.2	+	11.0	4128295	4128312	TGGCCTGCTATTTGCCCT
STM3924	*wecD*	22	+	11.8	4133232	4133249	TGGCGCGGAAATTGCACA
STM4013.S		3.6	-	13.6	4222708	4222725	TGGCATAAAACCTGAAAA
STM4226	*yjbA*	6.4	-	4.2	4446318	4446301	AGGCGCGAATAATGCATC
STM4290	*proP*	13	+	10.1	4532022	4532039	TGGCCTGATTTTTGCAGG
STM4572	*stjB*	8.3	-	8.2	4826908	4826925	TGGCGTGGCGATTTCAAT

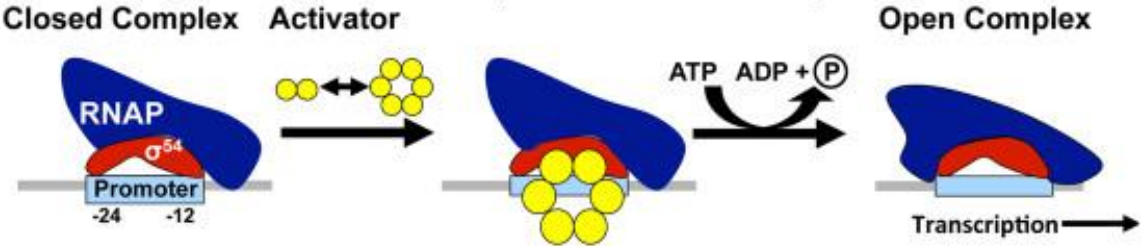
[a]Loci listed in bold are known or predicted σ^{54}-dependent promoters.
[b]Ratio of signals between WT and Δ*rpoN* cells.
[c]Orientation of the predicted binding site with respect to the listed ORF. (+) binding site is in *same* direction as ORF; (−) binding site is in *opposite* direction as ORF.
[d]PSSM score for the best predicted binding site within 1 kb of enriched ORF using the position-specific scoring matrix derived from the sequences in Additional File 2. As a reference to interpret the PSSM scores, the *S*. Typhimurium LT2 chromosome contains 3 sites with PSSM scores ≥22.0, 21 sites with PSSM scores ≥18.0, 61 sites with PSSM scores ≥14.0, and 401 sites with PSSM scores ≥10.0.
^The predicted binding site is a potential promoter for a neighboring gene based on its orientation and location within 250 bp of the 5' end of a neighboring gene or a long intergenic region (>100 bp) that may encode a sRNA.

Figure 2.1 Activation of σ^{54}-dependent transcription and activator structure. A) σ^{54} (red subunit) directs binding of the RNA polymerase (dark blue subunit) holoenzyme ($E\sigma^{54}$) to the -12, -24 promoter elements (light blue box). This closed complex is stable and cannot transition to open complex. In response to an environmental or cellular signal, the activator (bEBP; yellow dimers) oligomerizes. For most bEBPs, the oligomer binds to an enhancer (green box) 80 to 150 bp upstream of the promoter and DNA looping brings the activator in contact with σ^{54} in the $E\sigma^{54}$ closed complex. Hydrolysis of ATP by bEBP causes remodeling of $E\sigma^{54}$, which leads to open complex formation and transcription. There are a few bacteria with bEBPs that are missing the DNA binding domain; after oligomerization, these activators can bind to $E\sigma^{54}$ in closed complex with any promoter to stimulate open complex formation (promiscuous activation). **B)** The domain structure for the *Sinorhizobium meliloti* bEBP, DctD, is typical of most bEBPs. The amino-terminal regulatory domain (dark blue box) inhibits assembly of the bEBP oligomer until it interacts with an activation signal; the AAA+ ATPase domain (red box) mediates ATP binding and hydrolysis, as well as the protein-protein interactions between bEBPs (oligomerization) and between bEBP and σ^{54}; the carboxyl-terminal DNA binding domain (aqua box) contains a helix-turn-helix motif for binding the enhancer. The truncated DctD variant, DctD250, is missing the regulatory and DNA binding domains, so that it is constitutively active and promiscuous in stimulating transcription from σ^{54}-dependent promoters.

A. Enhancer-dependent activation of transcription by Eσ^{54}:



Enhancer-independent (promiscuous) activation of transcription by Eσ^{54}:



B. Domain structure of DctD from *S. meliloti*:

Constitutive, promiscuous variant of DctD, DctD250:

Figure 2.2 Binding sites predicted by ChIP-chip analysis. A) Location of the predicted binding sites for the 70 ORFs enriched by α-σ^{54} pulldown. Outer bars represent further breakdown by location and orientation of the binding site relative to the enriched ORF, as diagrammed in **(B)**. A (+) indicates that the binding site is in the same orientation as the ORF while (−) indicates that the binding site is in the opposite orientation as the enriched ORF.

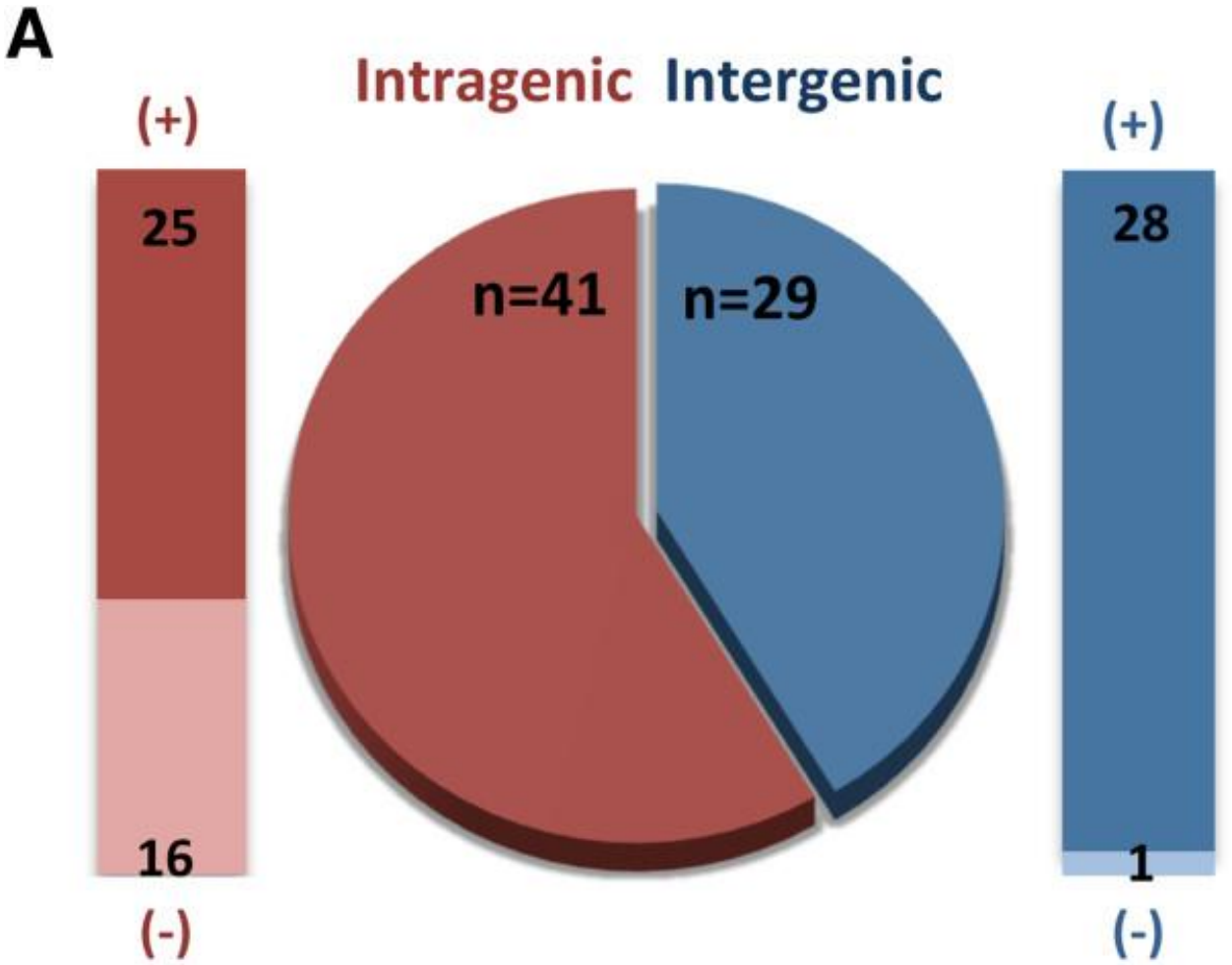

A
Intragenic Intergenic
(+)
25
16
(-)
n=41
n=29
(+)
28
1
(-)
B
(+)
Intragenic
(-)
(+)
Intergenic
(-)

Figure 2.3 Alignment of σ^{54} binding sites. Weblogos show the consensus sequence for **A)** 27 known/predicted promoter sequences used to generate the position-specific scoring matrix **B)** 29 predicted intergenic binding sites for ORFs enriched in ChIP-chip analysis or **C)** 41 predicted intragenic binding sites from within ORFs enriched in ChIP-chip analysis. Weblogos were generated using the online program available at http://weblogo.berkeley.edu/.

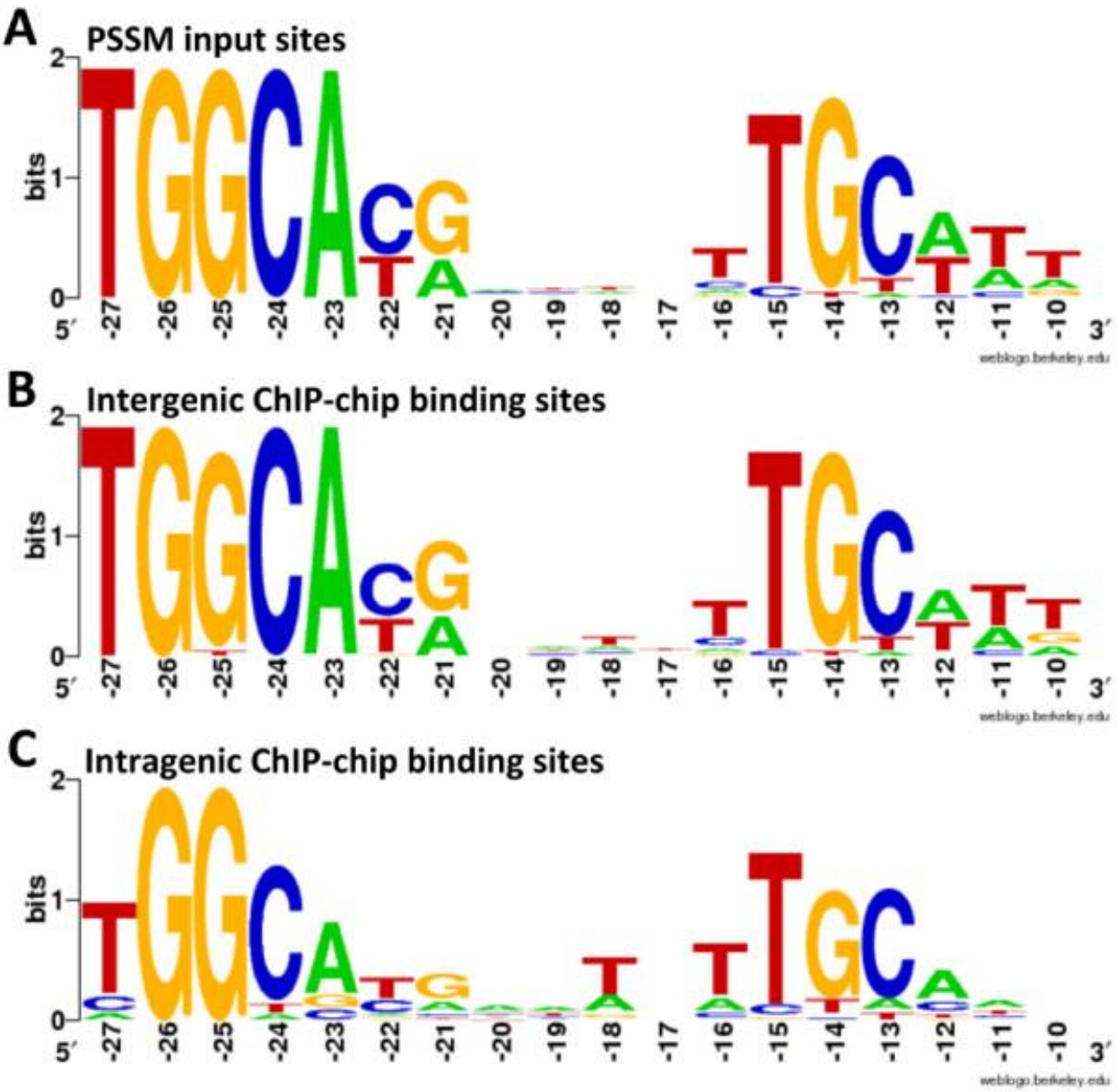

A
PSSM input sites
B
Intergenic ChIP-chip binding sites
C
Intragenic ChIP-chip binding sites
bits
5′
3′
-27 -26 -25 -24 -23 -22 -21 -20 -19 -18 -17 -16 -15 -14 -13 -12 -11 -10
weblogo.berkeley.edu

Figure 2.4 Promoter location, orientation, and activity for selected σ^{54} binding sites. The ratio of β-galactosidase activity (Miller Units) in WT+DctD250 vs. Δ*rpoN*+DctD250 cells is shown for **(A)** known (light blue bars) and predicted (dark blue bars) σ^{54}-dependent promoters, and **(B)** potential intragenic promoter sequences (red bars) in the promoter reporter vectors, pDS11 or pDS12 (black bars). Double asterisks denote significant increase in β-galactosidase activity in WT+DctD250 versus Δ*rpoN*+DctD250 (p-value <0.02). Circled numbers below locus tags indicate orientation of the potential promoter sequence, as illustrated in **(C)**. Orientation of potential intragenic promoter sequence is: **1)** same as ORF and >300 bp from 3' end; **2)** same as ORF and <300 bp from 3' end of a convergent downstream gene; **3)** opposite of ORF and >300 bp from 5' end; **4)** opposite of ORF and <300 bp from the 3' end of an upstream gene; and **5)** opposite of ORF and <300 bp from 5' end of gene, but >300 bp from the 3' end of an upstream gene.

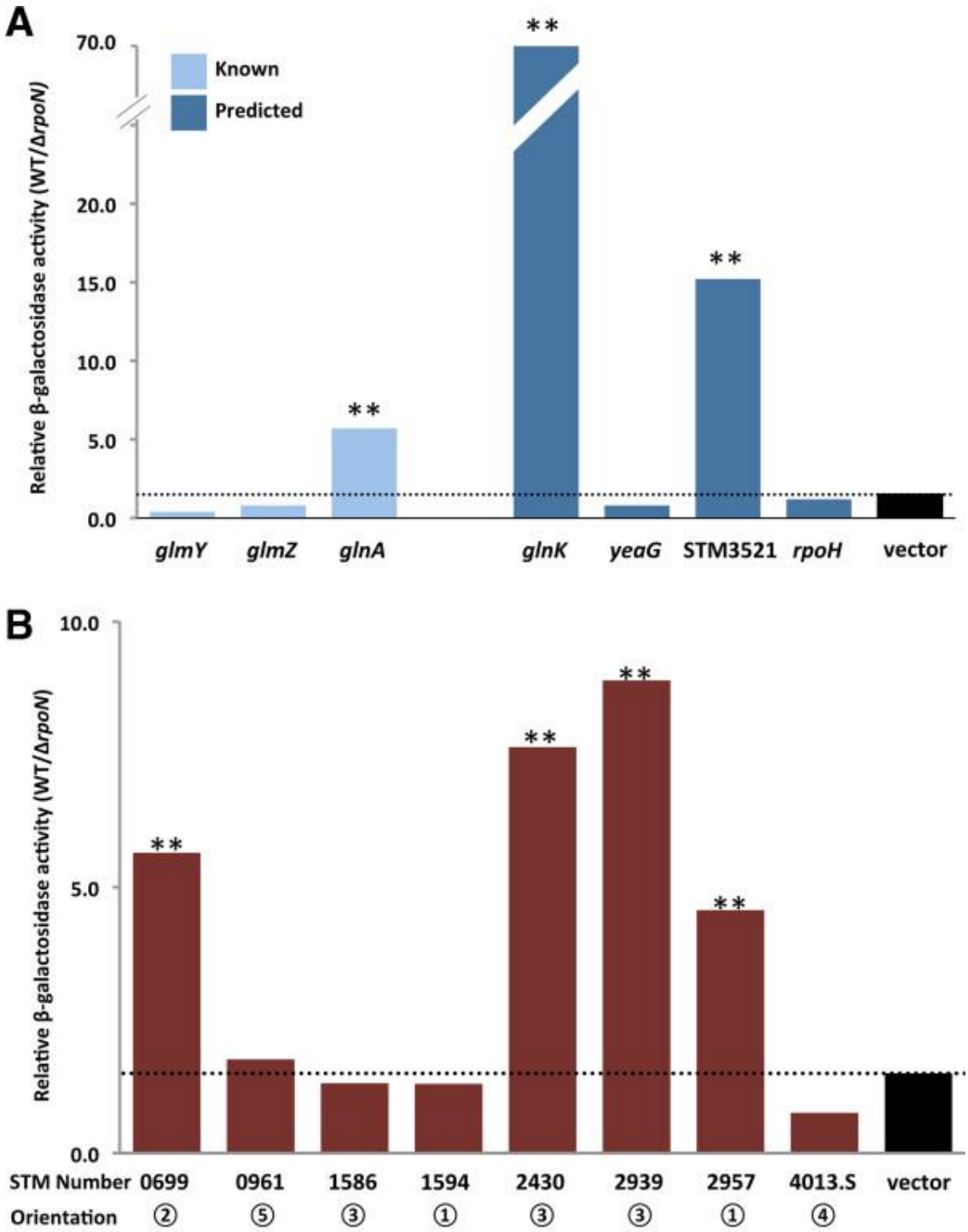
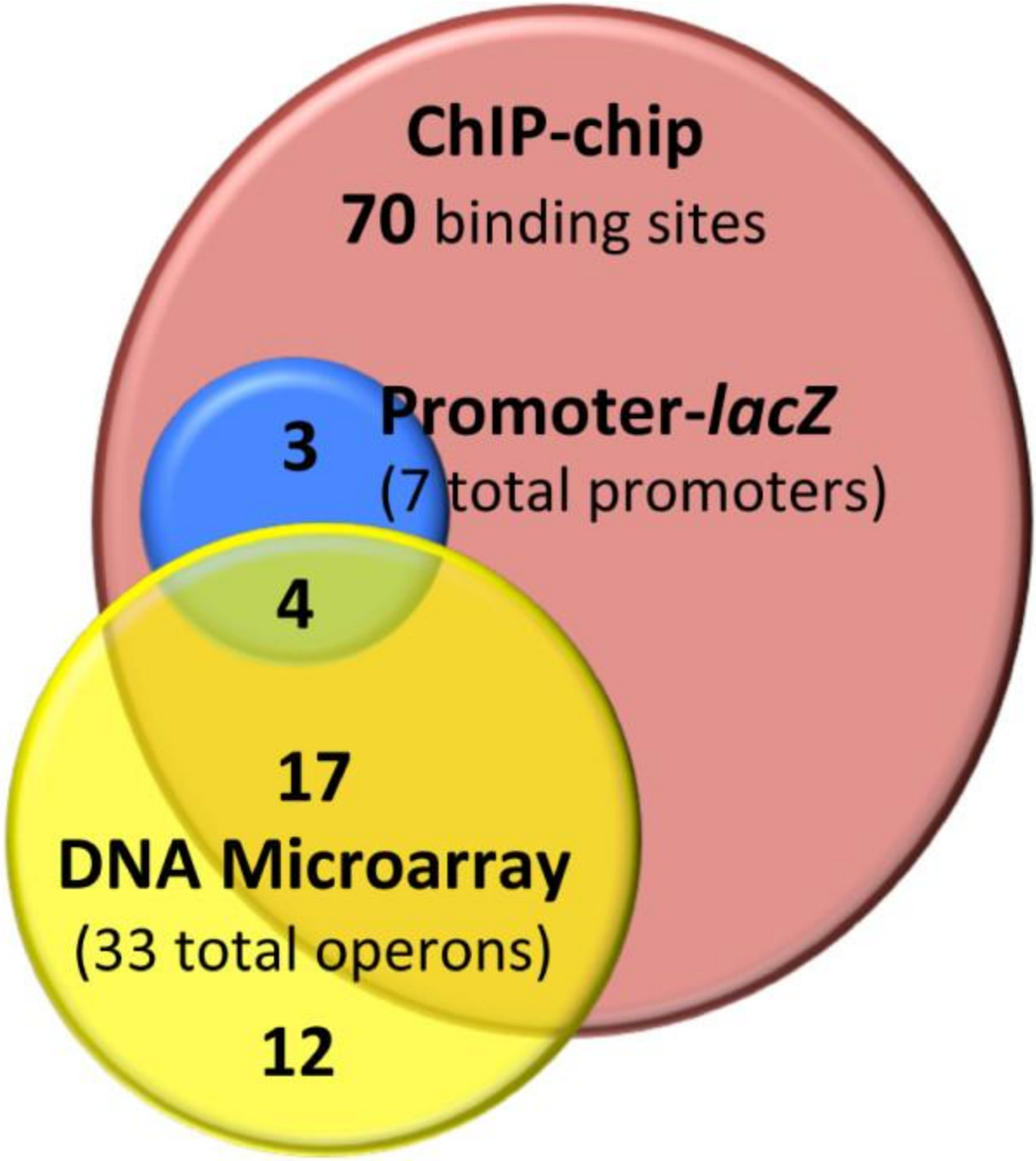

A
70.0
20.0
15.0
10.0
5.0
0.0
Known
Predicted
Relative β-galactosidase activity (WT/ΔrpoN)
**
**
**
glmY
glmZ
glnA
glnK
yeaG
STM3521
rpoH
vector
B
10.0
5.0
0.0
Relative β-galactosidase activity (WT/ΔrpoN)
**
**
**
**
STM Number 0699 0961 1586 1594 2430 2939 2957 4013.S vector
Orientation ② ⑤ ③ ① ③ ③ ① ④
C
①
②
③
④
⑤

Figure 2.5 Comparison of positive results from characterization of the σ^{54} regulon for *S.* Typhimurium LT2. Positive results are promoter sequences that were up-regulated >3-fold, or displayed a significant increase in β-galactosidase activity in WT+DctD250 compared to Δ*rpoN*+DctD250 in DNA microarray, and promoter-*lacZ* fusion assays, respectively or enriched >3-fold in WT compared to Δ*rpoN* cells in ChIP-chip assays. Regions of overlap indicate promoters that were positive in multiple experiments.


ChIP-chip
70 binding sites
Promoter-lacZ
(7 total promoters)
3
4
17
DNA Microarray
(33 total operons)
12

Additional File 2.1: Microarray results for WT and Δ*rpoN S.* Typhimurium LT2 strains expressing DctD250

Locus Tag[a]	Gene Symbol	Function	WT/Δ*rpoN*[b]
Known σ^{54}-dependent operons and sRNA genes:			
STM0368	*prpB*	Proprionate catabolism (putative)	**45**
STM0369	*prpC*		**25**
STM0370	*prpD*		**72**
STM0371	*prpE*		**21**[c]
STM0830	*glnH*	Glutamine high-affinity transporter	**7.1**
STM0829	*glnP*		**2.4**
STM0828	*glnQ*		**2.9**
STM2355	*argT*	Lysine/arginine/ornithine transport	**3.5**
STM_R0152	*tke1*	GlmY sRNA	**0.9**
STM_R0167	*sraJ*	GlmZ sRNA	**1.1**
STM4007	*glnA*	Glutamine synthetase	**48**
STM4006	*glnL*		**3.5**
STM4005	*glnG*		**3.6**
Predicted σ^{54}-dependent operons:			
STM0462	*glnK*	Nitrogen regulatory protein pII	**3.6**
STM0463	*amtB*	Ammonium transport (putative)	**190**
STM0577		PTS system (putative)	**67**
STM0576			**3.8**
STM0575			**3.5**
STM0574			**59**
STM0573			**18**
STM0572			**23**
STM0649.S		Hydrolase (putative)	**11**
STM0650			**42**
STM0651			**8.8**
STM0652			**1.2**
STM0653	*ybeL*		**1.2**
STM0665	*gltI*	Glutamate/aspartate transporter	**1.8**
STM_R0126	*sroC*		**0.81**
STM0664	*gltJ*		**1.3**
STM0663	*gltK*		**1**
STM0662	*gltL*		**1.9**

Locus Tag[a]	Gene Symbol	Function	WT/Δ*rpoN*[b]
STM1285	*yeaG*	Serine protein kinase (putative)	**2.5**
STM1284	*yeaH*		**2.5**
STM1303	*astC*	Arginine/ornithine/glutamine metabolism	**2.4**
STM1304	*astA*		**2.4**
STM1305	*astD*		**2.6**
STM1306	*astB*		**3.3**
STM1307	*astE*		**2.4**
STM1690	*pspA*	Phage shock proteins	**17**
STM1689	*pspB*		**0.9**
STM1688	*pspC*		**4**
STM1687	*pspD*		**1.3**
STM1686	*pspE*		**1.5**
STM2360		Amino acid transport (putative)	**100**
STM2359			**11**
STM2358			**18**
STM2357			**3.3**
STM2356	*ubiX*		**1.2**
STM2840		Nitric oxide reductase	**16**
STM2841	*ygbD*		**7.1**
STM2842	*hypF*	Hydrogenase maturation proteins	**13**
STM2843	*hydN*		**14**
STM2853	*hycA*	Hydrogenase 3	**26**
STM2852	*hycB*		**12**
STM2851	*hycC*		**170**
STM2850	*hycD*		**1.9**
STM2849	*hycE*		**96**
STM2848	*hycF*		**25**
STM2847	*hycG*		**30**
STM2846	*hycH*		**6.4**
STM2845	*hycI*		**12**
STM2844			**9.4**
STM2854	*hypA*	Formate-hydrogen lyase system	**5.6**
STM2855	*hypB*		**39**
STM2856	*hypC*		**6.3**
STM2857	*hypD*		**25**
STM2858	*hypE*		**31**
STM3521		RNA repair system (putative)	**71**

Locus Tag[a]	Gene Symbol	Function	WT/Δ*rpoN*[b]
STM3519	*rtcB*		**27**
STM3518	*rtcA*		**4.3**
STM3568	*rpoH*	Heat shock sigma factor (σ^{32})	**1.7**
STM3772		PTS system (putative)	**39**
STM3771			**37**
STM3770			**60**
STM3769.S			**31**
STM3768			**24**
STM3767			**41**
STM3766			**1.4**
STM4172	*zraP*	Zinc resistance-associated protein	**16**
STM4173	*hydH*	Zinc resistance two-component system	**3.7**
STM4174	*hydG*		**2.8**
STM4244	*pspG*	Phage shock protein	**1.4**
STM4285	*fdhF*	Formate dehydrogenase	**29**
STM4535		PTS system (putative)	**16**
STM4536			**1.7**
STM4537			**66**
STM4538			**13**
STM4539			**43**
STM4540.S			**25**
Novel σ^{54}-dependent operon:			
STM2944	*ygcB*	CRISPR-associated proteins	**1.6**
STM2943			**1.2**
STM2942			**1.3**
STM2941	*yghJ*		**1.2**
STM2940			**1.4**
STM2939	*ygcH*		**2.8**
STM2938			**4.1**
STM2937	*ygbF*		**0.8**
Possible indirectly regulated □^{54}-dependent operons:			
STM0144	*ppdD*	Pilin biogenesis (putative)	**0.78**
STM0143	*hofB*		**3.3**
STM0142	*hofC*		**2.2**

Locus Tag[a]	Gene Symbol	Function	WT/Δ*rpoN*[b]
STM0515	*allA*	Allantoin utilization	**1.1**
STM0516	*allR*		**1**
STM0517	*gcl*		**2**
STM0518	*gip*		**1.6**
STM0519	*glxR*		**2**
STM0520			**1.8**
STM0521	*ybbV*		**1.3**
STM0523	*allB*		**3.2**
STM0524	*ybbY*		**1.4**
STM0525	*glxK*		**0.99**
STM0791	*hutH*	Histidine-ammonia lyase	**3.8**
STM1252		Cytoplasmic protein (putative)	**3.1**
STM2038	*pduA*	Propanediol utilization	**0.75**
STM2039	*pudB*		**1.1**
STM2040	*pduC*		**2**
STM2041	*pduD*		**1.5**
STM2042	*pduE*		**0.46**
STM2043	*pduG*		**3.3**
STM2044	*pduH*		**0.46**
STM2045	*pduJ*		**0.58**
STM2046	*pduK*		**0.7**
STM2047	*pduL*		**4.3**[c]
STM2048	*pduM*		**2.5**
STM2049	*pduN*		**0.87**
STM2050	*pduO*		**1.4**
STM2051	*pduP*		**1.4**
STM2052	*pduQ*		**1.5**
STM2053	*pduS*		**1.7**
STM2054	*pduT*		**1.8**
STM2055	*pduU*		**0.94**
STM2056	*pduV*		**0.61**
STM2057	*pduW*		**2.1**
STM2058	*pduX*		**1.6**
STM2149	*stcD*	Chaperone-usher fimbriae (putative)	**1.5**
STM2150	*stcC*		**3.2**
STM2151	*stcB*		**1.9**
STM2152	*stcA*		**0.82**
STM2572	*yfhH*	Membrane transport protein (putative)	**3.1**
STM2773	*iroB*	Siderophore production	**2.1**
STM2774	*iroC*		**0.95**

Locus Tag[a]	Gene Symbol	Function	WT/Δ*rpoN*[b]
STM2775	*iroD*		**4**
STM2776	*iroE*		**1.2**
STM2839	*ygaA*	EBP activator for STM2840 operon (putative)	**3.2**
STM2859	*fhlA*	EBP activator for *fdhF/hyc/hyp* operons	**3.5**
STM2899	*invF*	Cell invasion proteins	**2**
STM2898	*invG*		**5.9**
STM2897	*invE*		**1.7**
STM2896	*invA*		**1.8**
STM2895	*invB*		**1**
STM2894	*invC*		**1.9**
STM2893	*invI*		**0.9**
STM2892	*invJ*		**2**
STM2891	*spaO*		**1.9**
STM2890	*spaP*		**1.6**
STM2889	*spaQ*		**0.69**
STM2888	*spaR*		**2.7**
STM2887	*spaS*		**1.4**
STM2886	*sicA*		**0.71**
STM2885	*sipB*		**1.1**
STM2884	*sipC*		**1.1**
STM2883	*sipD*		**1.2**
STM2882	*sipA*		**1.7**
STM3253		PTS system/sugar metabolism (putative)	**1.5**
STM3254			**1.8**
STM3255			**1**
STM3256			**3.7**
STM3257			**2.9**
STM3258			**1.4**
STM3259			**1.1**
STM3260			**2**
STM3261			**1.3**
STM3262			**1**

[a] Locus tags for all genes within operons in which at least one gene was up-regulated ≥3-fold (p-value <0.02) by RpoN in the microarray analysis; up-regulated genes are in **black** type and genes below the 3-fold cut off are in **grey** type). The up-regulated operons are grouped as known, predicted, novel, or possible indirectly regulated, as defined in Results.

[b] Both WT and Δ*rpoN* strains contain pPBHP92 expressing DctD249. Signal ratios below the 3-fold cutoff for up-regulation are displayed in **grey**.

[c] above 3-fold cutoff, but p-value >0.02

Additional File 2.2: Sequences used to generate the Position-Specific Score Matrix (PSSM)

Locus	Gene	Sequence
STM0368	*prpB*	TGGCATAGCCTTTGCTTT
STM0462	*glnK*	TGGCACATCCTTTGCAAT
STM0577		TGGCACGCCGTTTGCCAT
STM0649.S		TGGCACGCCTTTTGATTA
STM0665	*gltI*	TGGCACGTCTATTGCTTT
STM0830	*glnH*	TGGCATGATTTTTTCATT
STM1285	*yeaG*	TGGCATGAGAGTTGCTTT
STM1303	*argD*	TGGCACGAATGCTGCAAT
STM1690	*pspA*	TGGCACGCAAATTGTATT
STM2354*	*hisJ*	TGGCACGATAGTCGCATC
STM2355	*argT*	TGGCATAAGACCTGCATG
STM2360		TGGCATGCCTTTTGCTTT
STM_R0152	*glmY*	TGGCACAATTACTGCATA
STM2840		TGGCACACTAGCTGCAAT
STM2843	*hydN*	TGGCACGATTCGTGTATA
STM2853	*hycA*	TGGCATGGAAAATGCTTA
STM2854	*hypA*	TGGCATAAATATTGCTTT
STM3521		TGGCACGCTGGTTGCAAT
STM3568	*rpoH*	TGGCACGGTTGTTGCTCG
STM3772		TGGCACAACCTTTGCTCT
STM_R0167	*glmZ*	TGGCACGTTATGTGCAAT
STM4007	*glnA*	TGGCACAGATTTCGCTTT
STM4172	*zraP*	TGGCACGGAAGATGCAAG
STM4173	*hydH*	TGGCATGATCTCTGCTTA
STM4244	*pspG*	TGGCATGATTTTTGTAAG
STM4285	*fdhF*	TGGCATAAAACATGCATA
STM4535		TGGCACGCCGCTTGCTCT

*Although this sequence has been predicted to be a σ^{54}-dependent promoter, there is evidence indicating that this is not an active promoter [2].

Additional File 2.3: Oligonucleotides used in this study

Name	Forward	Reverse
pDS11MCS	CTAGAATTGAGCTCATTGGTACCATTG	GATCCAATGGTACCAATGAGCTCAATT
pDS12MCS	AGCTTAGATCTCTAGAGTCGACGGTACCATTG	GATCCAATGGTACCGTCGACTCTAGAGATCTA
LM-PCR linker	AGAAGCTTGAATTCGAGCAGTCAG	CTGCTCGAATTCAAGCTTCT
DctD	ATTGGATCCCACTCGACCGGAATTATCG	TTTTATCAGACCGCTTCTGC
STM0224 (*yaeT*)	GACATGAGTCCTTAGTCCG	AGGTAACATTACGCTATGGG
STM0334	CGATAGCGGAAACAAAACCG	TCAGGAAGCGAATATCTGGG
STM0462 (*glnK*)	TCGCCCATCATGCACCGTCG	TCCCTGAATGCCAATGGAAG
STM0504 (*ybbN*)	AGGGCGATGTCCTAGTCC	AGGACCTTCCAAAACGCG
STM0699	CTCTGCCCGTATGTTGTCCC	ATACGCACCACGCAAACCG
STM0961 (*lolA*)	GGTAGCCTGTTCTACAAACGG	GGAGTGTCGAAACCTGAGG
STM1285 (*yeaG)*	GTAGAGGCTCCCGGAAGAGG	ATCCACCAGCCTTTTCTACC
STM1586	GCGTTTTACCGTCTGCCG	GTAAGAAGTTGCTGGATGACGG
STM1594	CTCTGGATGTACTCGACGG	AGATACTCAAAACTACGCAGCG
STM1697	GGAAATAACGTGCCCTGGG	ATCTGTTGCGGACAATCGC
STM2016 (*cobT*)	TGGCTTTCTTTCCTACTCGG	GGCAGGACAATATTACTGGC
STM2430 (*cysK*)	AGCTGCAGGAAGATGAAAGC	GAATCAATGCCAGGTGAGG
STM_R0152 (*glmY)*	AAGGGGCTGACATAAGAAGG	TTAGGTGTTGCAGGTGTTGC
STM2939	TGAACACGATGGCTTAACGG	CATCAGCAAATCGTGACGC
STM2957 (*rumA*)	CGTCAATGTCGAACAGTGCC	CGACCATTTGCTGGTTTACCG
STM3127	GGACTGTTTTCATCGACCC	GGTAAATAACGTTGTGACGC
STM3521	AAGCGTAGAATCTAAAGGAAG	AGGGAATTGTCGCTTGCC
STM3568 (*rpoH*)	CAAATCCTCTCAATCAGTATTGC	GTTCGCAGGGAAAGAGTCC
STM_R0167 (*glmZ*)	TTCTGTCTCCACCGGGCGA	TCCAGGGTGTTTGATGAGG
STM4007	GCGCGTTATTGTACACGG	TGTACTCTCCCGGATTGG

CHAPTER 3:

CHARACTERIZATION OF THE σ^{54}-DEPENDENT RNA REPAIR OPERON IN *SALMONELLA*[1]

[1]Samuels, D.J. and A.C. Karls. To be submitted to *Nucleic Acids Research*

Abstract

The RtcB RNA ligase, which catalyzes the ligation of RNA substrates bearing uncommon 3'-PO_4 or 2',3'-cyclic phosphate and 5'-OH termini, is found as part of an operon in *Salmonella* along with a 2',3' phosphate cyclase and components of a ribonucleoprotein complex, Rsr and the YrlA and YrlB small RNAs, that has been shown to interact with damaged RNA molecules. RtcB homologs are found widely across all three domains of life. In Archaea and metazoans, RtcB is primarily responsible for ligating tRNA precursors after removal of an intron, but its physiological role in *Salmonella* and other bacteria (where intron-containing tRNAs are extremely rare) remains enigmatic. One reason for this is that transcription of the RtcB-encoding operon has never been reported in wild-type *Salmonella* or *E. coli.* Transcription of this operon is dependent on the alternative σ factor, σ^{54}, and the bacterial enhancer-binding protein RtcR, which requires an environmental stimulus to become activated. Here, we report conditions that result in expression of this operon in *Salmonella enterica* serovar Typhimurium. Using quantitative, reverse transcriptase PCR, we show that treatment with mitomycin C, a nucleic acid alkylating antibiotic, results in a 17-fold increase in transcript levels of this operon. This increase was dependent on both RtcR and RecA, indicating that expression of this operon may play a role in the SOS damage response. The fact that this operon is upregulated in response to nucleic acid damaged supports a model in which these gene products comprise an RNA repair system in which Rsr and YrlA (or YrlB) interacts with damaged RNA molecules as well as RNA repair/degradation enzymes (like RtcB, RtcA, or the ribonuclease PNPase) to direct appropriate decay or repair of these molecules.

Introduction

RtcB is an RNA ligase that is widely distributed in Bacteria, Archaea, and metazoans, but is not found in fungi or plants [1]. The activity of RtcB is distinguished from classical RNA ligases based on substrate specificity. While the latter enzymes ligate RNA molecules with 3'-OH and 5'-PO_4 termini [2, 3], RtcB catalyzes the formation of a 3'-5' phosphodiester linkage between molecules bearing unusual 2',3'-cyclic phosphate (2',3'>P) or 3'-PO_4 and 5'-OH termini [4, 5]. In humans and archaea, the primary role of RtcB homologs is to ligate tRNA precursors after a splicing endonuclease has removed an intron from the pre-tRNA molecule [6, 7]. However, introns are rare in bacterial tRNAs and those that have been described are all group I introns [8]. Group I introns are self-splicing ribozymes that remove themselves via two successive trans-esterifications and do not require a protein for ligation [9]. Therefore, routine processing of tRNA molecules seems like an unlikely function for bacterial RtcB.

The role of RtcB in eubacteria may be specialized for RNA ligation and repair in response to a particular stress. This idea is supported by the fact that expression of RtcB in eubacteria is commonly regulated by the alternative sigma factor σ^{54} [10]. While σ^{54} is like other alternative σ factors in that it interacts with RNA polymerase (RNAP) core and directs the holoenzyme (Eσ) to a specific promoter sequences, σ^{54} is distinct in structure, conserved promoter elements, and, most importantly, its requirement for an activator that hydrolyzes ATP to stimulate transcription initiation [11]. Each σ^{54}-dependent promoter has a cognate bacterial enhancer-binding protein (bEBP) that becomes activated in response to a specific environmental stimulus to multimerize, bind to its enhancer sequence, and interact with the $E\sigma^{54}$ closed complex through DNA looping, whereupon it hydrolyzes ATP to stimulate open complex formation and transcription initiation [12]. The bEBP generally has an N-terminal regulatory domain that responds to the environmental stimulus, a highly conserved AAA+ domain (ATPase superfamily) that interacts with $E\sigma^{54}$, and a C-terminal DNA binding domain that is specific for recognition of its associated enhancer sequence [12]. By containing multiple bEBPs within their genome, bacteria can adapt the response of σ^{54} to a variety of stresses.

The bEBP that regulates the *Salmonella enterica* subsp. *enterica serovar* Typhimurium (*S.* Typhimurium) and *Escherichia coli* operons containing *rtcB* is RtcR, which is encoded divergently, upstream of the operon it controls. The environmental conditions that activate RtcR are unknown, but work with constitutively-active bEBP variants that contain deletions of the N-terminal regulatory region—the promiscuous DctD250 [13] and the specific RtcR-ΔN [14, 15]—showed that the promoter upstream of these operons is dependent on both σ^{54} and RtcR.

In *E. coli* the *rtcB* operon encodes just one other gene, *rtcA*. The enzymatic activities of RtcA and RtcB from *E. coli* have been characterized *in vitro*, however the *in vivo* substrates are unknown. RtcB ligates a 2',3'-cyclic PO_4 or 3'- PO_4 with a 5'-OH end of RNA or single stranded DNA [5, 16] and RtcA is a 2', 3' PO_4 cyclase, converting 2'- or 3'-PO_4 ends on damaged or cleaved RNA to 2', 3'-cyclic PO_4; RtcA may function to repair 2'-PO_4 ends that cannot serve as substrates for RtcB activity [10].

In *Salmonella*, *rtcB* is the penultimate gene in a five-gene operon (Figure 3.1). RtcA is encoded downstream of *rtcB*, and upstream of *rtcB* are the genes for a homolog of the Ro sixty-related (*rsr*) protein and two Y RNAs (*yrlA* and *yrlB*) [15]. Ro orthologs, which are found in most metazoans and in ~5% of sequenced eubacteria, are bound by small Y RNAs; the Ro-Y RNA ribonucleoprotein (RNP) complex is thought to be involved in non-coding RNA (ncRNA) quality control in animal cells [15, 17]. The Ro ortholog, Rsr, from eubacterium *Deinococcus radiodurans* has been shown to function with RNase II and RNase PH in 23S rRNA maturation during heat stress and with polynucleotide phosphorylase (PNPase) in rRNA degradation during stationary phase [18, 19]; and Y RNA provides a scaffold for the interaction between Rsr and PNPase [15]. Association of the Rsr-Y RNA complex with PNPase directs PNPase specificity for degradation of structured RNAs [15]. To corroborate the association of Rsr-Y RNA with PNPase in eubacteria, *S.* Typhimurium was assayed for Rsr/Y RNA/PNPase complexes and the association of Rsr-YrlA with PNPase was confirmed; in addition Rsr was shown to associate with YrlB, but YrlB was not in the Rsr-YrlA-PNPase complex [15]. The functions of these ribonucleoprotein (RNP) complexes in *S.* Typhimurium are unknown. Although *D.*

radiodurans encodes an RtcB homologue, the *rtcB* gene is not in an operon with *rsr* and neither RpoN nor RtcR are encoded in its genome. Therefore, analogies between *S.* Typhimurium and *D. radiodurans* cannot be made for the regulation of RtcB-mediated RNA repair.

This chapter addresses the physiological function of the Rsr-YrlBA-RtcBA (RNA repair) system in *S.* Typhimurium, testing a model (Figure 3.1) in which Rsr-Y RNA and RtcB/RtcA control a cellular response to particular stresses through altered RNA repair/processing of specific structured RNAs. In this model, a ribonucleoprotein complex consisting of Rsr and either YrlA or YrlB binds certain damaged, structured RNAs. The Y RNA then targets this RNP complex to either PNPase for degradation or to RtcB (and/or RtcA) for repair. The initial step required to test this model is to determine the conditions that induce expression of the RNA repair operon. Utilizing quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) assays, conditions for up-regulation of the RNA repair operon were assessed. Treatment with the nucleic acid alkylating agent mitomycin C (MMC) results in increased transcript levels for the RNA repair operon. Further characterization of the RNA repair operon and the RNP complexes formed by its products are described in this chapter.

Results

RecA-dependent up-regulation of RNA repair operon expression upon treatment with mitomycin C

The RNA repair operon is regulated by σ^{54} and RtcR and its expression requires an unknown environmental signal to activate RtcR. The conditions that generate this signal are also likely to generate the predicted substrates for the products of the RNA repair operon, i.e. damaged RNA. Therefore, to determine conditions under which the RNA repair operon is expressed, *S.* Typhimurium cultures at mid-log phase growth were exposed to various treatments that result in damage to RNA, and quantitative, reverse-transcriptase PCR (qRT-PCR) was used to measure expression of *rtcA*, the last gene in this operon. Transcript levels for *rtcA* were normalized to transcript levels for *rpoD*, which encodes the primary sigma factor and is unchanged in its expression under these growth conditions (data not shown).

Cultures were subjected to the following stresses: mitomycin C (MMC; 1 μg/ml); nitrogen limitation (2.5 mM arginine as sole N source); carbon starvation (1% methyl α-D-glucopyranoside as sole C source); peroxide stress (3% H_2O_2); amino acid starvation (0.4 mg/ml serine hydroxamate); translation inhibition (30 μg/ml chloramphenicol); cell wall stress (2 μg/ml cefotaxime); iron limitation (250 μM 2,2'-dipyridyl). Cultures were exposed to these treatments for 90 min at 37°C, except for the peroxide stress which was only a 15 min exposure. All of these treatments have been shown to cause damage, directly or indirectly, to nucleic acids in bacterial cells (see Discussion).

Following a 90 min. treatment with mitomycin C, a 17-fold increase in *rtcA* expression was observed relative to untreated cells (Figure 3.2). Lack of this increased expression in a Δ*rtcR* mutant confirmed that the increased transcription of the RNA repair operon in the presence of mitomycin C is from the RtcR-dependent σ^{54}-dependent promoter. Cells containing pDS183, which harbors a constitutively-active RtcR variant, served as a positive control and showed high levels of *rtcA* expression in both the treated and untreated conditions. Preliminary results indicate that the other treatments examined do not lead to *rtcA* expression.

Mitomycin C, an antibiotic that alkylates nucleic acids, induces the SOS stress response pathway [20]. Therefore, to determine whether the SOS response is involved in the up-regulation of the RNA repair operon, a *recA*::Kan^R mutant was assessed in the qRT-PCR assay. Transcription of the RNA repair operon in a *recA*::Kan^R mutant was not up-regulated after treatment with mitomycin C, suggesting that expression of the operon involves the SOS response (Figure 3.2).

.

The RNA repair system does not appear to confer a survival advantage during exposure to mitomycin C

To address whether the significantly increased expression of the RNA repair operon affects cell survival following treatment with mitomycin C, liquid cultures of wild-type and Δ*rtcR S*. Typhimurium strains were treated with mitomycin C and their growth was assessed via optical density (OD_{600}). The

mitomycin C-treated cells grew to a lower density than the untreated cells, but the growth curves for the wild-type and Δ*rpoN* cultures were indistinguishable (Figure 3.3). Thus under the conditions tested (i.e. 1 μg/ml MMC treatment for 90 min. in minimal media at mid-log growth phase), expression of the RNA repair system does not appear to improve cell growth. Additionally, the growth curve for mitomycin C-treated wild-type *S.* Typhimurium expressing $RtcR^{con}$ from pDS183 essentially overlaps the growth curves for wild-type and Δ*rtcR* strains, indicating that the presence of the RNA repair system at the time of treatment did not alter the effect of mitomycin C treatment on cell growth.

To further assess the effect of the RNA repair operon on cell physiology following mitomycin C treatment, cells were checked for filamentation, as has been reported for *E. coli* and *Caulobacter crescentus* cells in response to treatment with MMC [21, 22], and for cell viability. Microscopic observations of treated versus untreated cells showed cell elongation after 60 min. of mitomycin C treatment for both wild-type and Δ*rtcR S.* Typhimurium strains (data not shown). Preliminary results from quantitative plate count assays with treated and untreated wild-type and Δ*rtcR S.* Typhimurium strains at 0, 30, 60, and 90 minutes post-treatment indicate similar effects of mitomycin C on cell viability for wild-type and Δ*rtcR* strains.

YrlA is not detected in a Δ*rsr* mutant

Since there are no predicted promoters for the RNA repair operon sRNAs other than the one upstream of *rsr*, YrlB/A are likely being transcribed as part of a polycistronic transcript. In order to carry out their functions as sRNAs, they must be processed from the full-length transcript by an as yet undescribed mechanism. It is possible that either RtcB or RtcA is required for proper maturation of these RNAs. Additionally, in eukaryotic cells deleted for Ro, Y RNAs were detected at lower levels than in wild-type cells [23, 24]. Therefore, interaction as part of an RNPc is believed to contribute to Y RNA stability. To determine whether any of the RNA operon components is involved in processing YrlA from

the full-length transcript, as well as whether the stability of YrlA is decreased in the absence of Rsr, Northern blot analysis was performed to detect YrlA in wild type cells and RNA repair system mutants.

To ensure transcription of this operon, the strains used in this experiment contained a plasmid expressing a constitutively-active variant of RtcR (pRtcRcon). Using a probe that hybridizes to YrlA, we were able to detect a band around the expected YrlA size (111 nt) in wild-type cells. As expected, this band was absent in a Δ*yrlA* mutant. This band is absent in Δ*rsr* but present in both the Δ*rtcB* and Δ*rtcA* mutants. The presence of YrlA in the RtcB and RtcA deletion mutants indicates that neither of these enzymes is needed for processing the YrlA product from the full-length transcript. Though it is possible that once removed from the full-length transcript, YrlA requires further RtcB- or RtcA-dependent modification (such as hydrolyzing or generating a 2',3'-cyclic phosphate terminal [4, 14]) which would not have been detected in this experiment. The most likely explanation for the absence of YrlA in the Rsr mutant is that it is unstable in the absence of Rsr. Alternatively, Rsr could facilitate processing of YrlA from the full-length transcript. It should be noted that the full-length transcript containing YrlA, which is ~2 kb in this mutant, is too large to be transferred from the polyacrylamide gel to the membrane. Preliminary RT-PCR experiments with this mutant showed that *rtcA* transcript was present, which argues against polar effects of the mutation as an explanation for the absence of YrlA. The presence of a band in the Δ*rtcR* cells indicates that the plasmid-borne $RtcR^{con}$ is capable of activating transcription from this promoter.

The reason for the YrlA product appearing as two bands is unknown. It is possible that the larger band represents a pre-YrlA processing intermediate, but the nature of these two species has not been investigated. It is also possible that the upper band is a predicted 114 nt pseudo-tRNA whose sequence partially overlaps *yrlA* in the region of the probe used for the Northern blots (see Discussion). In addition, the YrlA probe has moderate homology to cellular $tRNA^{Asn}$, thus the bands corresponding to ~75-80 nt in all lanes likely represent hybridization to $tRNA^{Asn}$. To control for RNA loading, the blot was stripped and re-probed against the 5S rRNA subunit (Figure 3.4; bottom panel).

Western blot Analysis of FLAG-Rtc constructs

When the *Salmonella* Rsr-Y RNA RNPc was initially identified [15], it was noted that some components of this complex sedimented in heavier fractions than the *Deinococcus* RNPc, indicating that *Salmonella*'s complex may contain additional components or differ in stoichiometry. However, the nature of these complexes was not investigated. Additionally, these complexes were observed after induction with a truncated RtcR activator. It is likely that under conditions that lead to activation of native RtcR that different proteins and RNAs will be present in the cell, which may result in Rsr-YrlA or –YrlB complexes with alternate compositions.

To determine the composition of the complexes formed under activation conditions for native RtcR, various components of these complexes (Rsr, RtcB, RtcA) have been tagged with a FLAG epitope so that RNPc's can be isolated by immunoprecipitation following exposure to mitomycin C. Partner proteins associating with these complexes will be identified via subtractive mass spectrometry, and substrate RNAs will be sequenced. FLAG-tagged variants of Rsr and RtcA under the control of a tightly-repressed P_{lac} promoter [25] have been generated. Western blot analysis shows that after induction with isopropyl β-D-1-thiogalctopyranoside (IPTG), 58- and 36-kDa bands representing FLAG-Rsr and FLAG–RtcA, respectively, can be detected with α-FLAG antibodies (Figure 3.5). FLAG-RtcB has been similarly constructed, but has not yet been examined by Western blotting.

RtcB gene neighborhood analysis

Components of the RNA repair operon have been identified and characterized throughout Archaea and metazoans. To determine the prevalence of the system and to draw insights about its regulation in bacteria, the neighborhoods surrounding ~400 homologs (>40% identity) of *rtcB* were analyzed using the Integrated Microbial Genomics (IMG) gene neighborhood browser. RtcB is widely-distributed across all three domains. Though results described here are only for bacteria, the initial search returned many eukaryotic and archaeal homologs. RtcB was present in multiple operon configurations

(Figure 3.6). The most common was an isolated copy of *rtcB* with none of the other components of the RNA repair system (i.e., RtcR, Rsr, or RtcA; the presence of Y RNAs was not examined). This configuration was seen in representatives of 19 different bacterial phyla. Co-localization with *rtcR*—either with or without *rtcA*—was limited largely to Proteobacteria. The one exception was an uncultured Acidobacteria isolate, "*Candidatus* Koribacter". The "complete" RNA repair system (i.e., *rtcR, rsr, rtcB,* and *rtcA*) is present in only five genera: *Salmonella, Acidovorax, Variovorax, Pseudomonas, Sphingobium,* and *Sphingomonas*.

Discussion

The Rsr-RtcBA system is expressed after treatment with Mitomycin C

In our previous work defining the global σ^{54} regulon of *Salmonella* Typhimurium LT2 [13], we were able to observe transcription from the RNA repair operon promoter at its chromosomal locus via microarray, as well as when the promoter sequence was cloned upstream of *lacZ* on a plasmid. Because we were interested in identifying all σ^{54}-dependent transcripts within the cell, this work was done in the presence of DctD250, a constitutively-active, promiscuous variant of a bacterial enhancer-binding protein (bEBP), which has the ability to indiscriminately activate transcription from any σ^{54}-dependent promoter in the genome under standard laboratory growth conditions [26, 27]. These findings are corroborated by studies showing that a constitutively-active variant of RtcR—the native activator of the RNA repair operon—was able to activate transcription in *Salmonella* and *E. coli* [14, 15]. Taken together, these results indicate that the promoter in front of this operon is fully functional. Additionally, *E. coli* RtcB and RtcA have been cloned, purified, and shown to have RNA ligase and 2',3'-phosphate cyclase activities, respectively *in vitro* [5, 14], suggesting that they are likely functional within the cell as well. *Salmonella* RtcB and RtcA, which are 88% and 68% identical, respectively, to the proteins in *E. coli* should function in a similar fashion.

Despite the presence of likely functional genes downstream of a functional promoter, to our knowledge, expression of this operon in wild-type *Salmonella* (or *E. coli*) has not been reported despite extensive analysis of the transcriptomes under various infection-related conditions [28]. Because this operon is σ^{54}-dependent, its transcription requires activation of the bEBP, RtcR. The active bEBP can contact the $E\sigma^{54}$ closed complex and hydrolyze ATP, providing the energy necessary to isomerize to open complex and allow transcription. As is the case for many bEBPs, the RtcR N-terminal domain is regulatory region [29] that is only activated under the proper environmental conditions. Comparison of the RtcR N-terminal domain to the other bEBPs whose activation is mediated by a two-component signal transduction system (e.g., NtrC, DctD) or binding of an effector molecule (e.g., NorR, FhlA) [12] does not suggest a particular type of regulation mechanism. The lack of similarity to other regulatory regions places the N-terminal RtcR domain into its own PFAM category (06956). The fact that expression of this operon from wild-type cells has yet to be reported indicates that the proper stimulus required to activate RtcR has not been determined.

The first step in characterizing the physiological role of the Rsr-RtcBA system is determining the conditions under which these genes are expressed. To do this we used quantitative, reverse transcriptase PCR to assess expression under a variety of stress conditions. Since this system appears to be involved in RNA repair, we used stresses that might generate damaged RNAs. Oxidative damage and ultraviolet (UV) light are known to damage nucleic acids [30-32]. Cefotaxime, an antibiotic which leads to cell wall damage, was used because recent reports indicate that treatment with bactericidal antibiotics—even at sub-lethal concentrations—can lead to substantial oxidative damage [33-35], and therefore may require RNA repair. Additional damage may come from the toxin component of toxin-antitoxin (TA) systems. These systems consist of a stable toxin molecule and labile antitoxin. Upon reduction in antitoxin levels, the toxin is no longer inhibited. *Salmonella* contains eleven type II TA systems [36], some of which are known to cleave RNA molecules, yielding 2',3'>P and 5'-OH ends [37, 38]. In a previous study, translation inhibition (chloramphenicol), amino acid starvation (serine hydroxamate), glucose starvation

(methyl-α-D-glycopyranoside), and nucleic acid damage (MMC) were shown to induce various TA systems in *E. coli* [39]. We therefore hypothesized that these conditions might be leading to RNA damage, which Rsr-RtcBA would be employed to repair.

Mitomycin C is a member of the mitomycin family of antibiotics, derived from multiple species of *Streptomyces,* including *S. caespitosus.* In humans, MMC displays broad-spectrum antitumor activity and is a commonly used chemotherapeutic agent in treatment of a variety of cancers. Once inside the cell, MMC is enzymatically reduced to its active form, a bifunctional alkylating agent, which reacts with dG to form MMC-mono-dG adducts, intrastrand bi-adducts at -GpG-, and interstrand cross-links within the sequence 5'-CpG-3' in DNA [40]. It is interesting that an antibiotic that causes DNA damage would induce expression of a putative RNA repair system. However, challenges to the dogma of "MMC damages DNA" have emerged. Early reports indicate the MMC leads to degradation of RNA and ribosome decomposition [41, 42], and a recent report presents evidence of RNA damage by MMC, including the formation of RNA-MMC adducts [43]. The target of MMC therefore may not be specifically DNA, but nucleic acids in general. Alternatively, since RtcB has been shown to partially repair specific DNA lesions *in vitro* , this system may play a role in DNA repair.

Surprisingly, the ability to express the Rsr-RtcBA system does not appear to provide a survival advantage after treatment with 1 μg/ml MMC. In terms of optical density, lower concentrations of MMC (100 pg and 1 pg/ml) affected growth of cells, although not as substantially as 1 μg/ml (data not shown). These lower concentrations are currently being examined in quantitative plate count assays to determine whether the Rsr-RtcBA system enhances survival after MMC treatment.

After establishing that the Rsr-RtcBA system is expressed in response to MMC treatment, the natural follow-up question is "what is this system doing in MMC-treated cells?" Our growth analysis indicates that the ability to express this system does not increase survival of treated cells at 1 μg/ml (lower concentrations are currently being examined.) To better understand the role of this RNA repair system, we wish to understand the global changes that occur in response to MMC at both the transcriptomic and

proteomic levels. To accomplish this, we will be analyzing total cell RNA and protein content via RNA sequencing and subtractive mass spectrometry, respectively. It should be noted that a similar analysis of the transcriptome has been done in *Salmonella.* After treatment with 2 μg/ml MMC, Benson, et al. used arbitrarily-primed PCR to show that 19 genes were up-regulated in *Salmonella* [44]. Most of these genes were located within prophages (Fels, Gifsy1, Gifsy2) and none of the genes in the *rsr-rtcBA* operon were detected. This could be due to a strain difference (they used DB7000), but since they failed to detect most of the SOS response genes known in the closely-related *E. coli* [45], it is more likely that this methodology was not highly sensitive. Using a high-throughput analysis like RNA-seq will overcome this limitation.

Rsr may affect Y RNA stability

Prior to the discovery that the region between *rsr* and *rtcB* encoded two small RNAs, genome annotation on the IMG server indicated that this region was a "pseudo-tRNA" (ψtRNA). Alignments of this sequence indicate homology to other tRNAs—particularly $tRNA^{Asn}$—with a ~40 nt intervening sequence near the anticodon loop. This annotated 114 nt ψtRNA is encoded upstream of and shares a 40 nt overlap with the 111 bp *yrlA*. Given that RtcB homologs in archaea and animals are responsible for ligating tRNA halves after intron removal, we hypothesized that *Salmonella* RtcB may be acting similarly on this ψtRNA. Northern blot analysis with a probe against this ψtRNA (which largely overlaps yrlA) showed that a ~114 nt product was generated in wild type cells but not in ΔψtRNA (Δ*yrlA*) mutants (Figure 3.4). This mutant has the entire intervening sequence between *rsr* and rtcB deleted. Therefore, it lacks the ψtRNA as well as *yrlA*. Interestingly this product was also absent in Δ*rsr*, but not Δ*rtcB* or Δ*rtcA* mutants, while preliminary RT PCR results with Δ*rsr* cells indicate that this mutation is non-polar (data not shown).

The reporting of YrlA as a small RNA helped explain these results. Previous studies with Ro (the eukaryotic homolog of Rsr) have shown that in $Ro^{-/-}$ mutants, Y RNAs are not observed [23, 24]. This is

likely due to a stabilizing effect of Ro/Rsr on the Y RNA. However, since the entire operon is transcribed as a monocistronic unit, a processing event must take place in order to generate the 114 nt product observed. As the mechanism for this processing has yet to be described, it is plausible that the lack of *rsr* in the full-length transcript has an effect on Y RNA processing rather than stability. While *yrlA* should be present in these experiments as part of the full-length transcript, at ~2 kB, these transcripts would be too large to detect with our methods.

While these results are likely to represent YrlA, we cannot rule out the possibility that the ~114 nt transcript observed is a ψtRNA processing intermediate (i.e., before removal of the intron). Subsequent blots with probes specific for the ψtRNA intron and *yrlA* will help distinguish between these possibilities.

RtcB homologs are found, in a variety of contexts, throughout eubacteria

In addition to being found in metazoans and archaea, homologs of RtcB are found in 19 different bacterial phyla. While metazoan and archaeal RtcB homologs play a role in splicing tRNA precursors after intron removal, the fact that the only known bacterial tRNAs are self-splicing indicates that these bacterial homologs have developed an alternative function. Interestingly, the prototypical RNA repair system, as seen in *Salmonella* (i.e., Rsr, Y RNA, RtcB, and RtcA under the control of RtcR) is not well maintained. Of all of the components of this system, the Rsr/Y RNA seem to be least conserved. This indicates that the RNPc component is either only important under specific circumstances that are likely to be encountered by the cells that encode it or that other cells have evolved an alternative system to account for this function. Examining differences in the repair functions between cells that encode the RNPc (e.g., *Salmonella*) and those that do not (e.g., *E. coli*) may provide an understanding of the role of this complex.

Another component of the RNA repair operon that is not widely-conserved outside of the proteobacteria is RtcR. This suggests that in other bacteria, RtcB may be constitutively expressed or that it is being regulated by some other mechanism. A deeper knowledge of the stimuli that RtcR responds to

in *Salmonella* may provide insight into why RtcR regulation has only been maintained in a small number of bacteria.

Materials and Methods

Growth media and conditions

Unless otherwise noted, bacteria were grown at 37° with aeration in either LB (Miller) [Fisher Scientific; Fair Lawn, NJ] or MOPS minimal media [46]. Unless otherwise indicated, all supplements and antibiotics were purchased from Sigma-Aldrich (St. Louis, MO). H_2O_2 was purchased from Kroger (Cincinnati, OH). Antibiotics were added at the following concentrations (LB/MOPS; in μg/ml): Ampicillin (Ap) 80/50; Spectinomycin (Spc) 50/50; Tetracycline (Tc) 12/--.

Strains and plasmids

The wild type strain used was *Salmonella enterica subspecies enterica serovar* Typhimurium 14028s. Deletion mutants were created via the λRED recombination method [47] using the primers described in Table 3.1. The *recA*::Kn^R mutant, JE10649 was generated in *S.* Typhimurium LT2 JE6583 [*metE*205 *ara*-9]; both strains were generously provided by Jorge Escalante. Cloning was done in *E. coli* DH5α and moved through a restriction$^-$/modification$^+$ mutant of *S.* Typhimurium LT2 [*leuA*414 *hsdL* Fels2] before introduction into *S.* Typhimurium 14028s..

All plasmids were generated using enzymes purchased from New England Biolabs (NEB; Ipswich, MA). *pDS171:* OneTaq was used to generate a PCR product from *S.* Typhimurium genomic DNA using *rpoD*-RT F/R, which was TOPO cloned into pCR2.1 (Invitrogen; Carlsbad, CA). *pDS185*: OneTaq was used to generate a PCR product from *S.* Typhimurium genomic DNA using *rtcA*-RT F/R, which was TOPO cloned into pCR4 (Invitrogen; Carlsbad, CA). *pDS183*: A fragment containing a truncated copy of RtcR beginning at Leu^{179} was PCR amplified using RtcRcon F2/RtcR R. This fragment was TOPO cloned into PCR4 and sub-cloned as a BamHI-HindIII fragment into pKH66 [48]. *pRtcR^{con}*: A

fragment containing a truncated copy of RtcR beginning at Leu^{179} was PCR amplified using RtcRconF1/RtcR R. This fragment was cloned into pGEM-T (Promega; Madison, WI) and sub-cloned as an EcoRI-HindIII fragment into pMALc (NEB). *pDS162:* A fragment containing an N-terminal FLAG-tagged copy of *rsr* was PCR amplified using FLAG-Rsr F/Rsr R. This fragment was TOPO cloned into pCR2.1 and sub-cloned as an NdeI-XbaI fragment into pSRKTc [25]. *pDS164:* A fragment containing an N-terminal FLAG-tagged copy of *rtcA* was PCR amplified using FLAG-RtcA F/RtcA R. This fragment was TOPO cloned into pCR2.1 and sub-cloned as an NdeI-XbaI fragment into pSRKTc. All plasmids were sequenced by Genewiz (South Plainfield, NJ).

RNA harvest

Overnight cultures were diluted 1:40 in fresh media, grown to mid-exponential phase OD_{600}= 0.4-0.6, and aliquoted to tubes with the appropriate treatment (see Results). If necessary (α-MG and Arginine treatments), cells were centrifuged 5' @ 10000xg. Cell pellets were suspended in an equivalent volume of treatment media. Cultures were grown for an additional 90'. RNA was harvested via the RNAsnap method [49]. Cell pellets were suspended in 1/10 volume of RNA extraction solution [18 mM ethylenediaminetetraacetic acid (EDTA), 0.025% sodium dodecyl sulfate (SDS), 1% β-mercaptoethanol (β-ME), 95% formamide] and incubated 5' @ 95°. Cellular debris was removed by centrifugation in a microcentrifuge at top speed for 10'. The supernatant was removed (taking care not to disturb the pellet) and ethanol precipitated. Pellets were rehydrated in RNase-free H_2O. RNA was quantified with either a Nanodrop or microplate reader and quality was assessed via visualization on a 1x Sodium Borate gel. Each condition described comprises at least three biological replicates.

cDNA preparation

To remove contaminating genomic DNA, RNA samples were treated for 1 hr with DNase I (NEB) according to the manufacturers' instructions. DNA removal was verified with PCR, as follows:

rpoD-RT F/R primers were used to amplify 1µl of RNA with OneTaq DNA polymerase (NEB) (1x—2' @ 94°; 30x—30" @ 94°, 30"@ 54.3°, 30" @ 68°; 1x—5'@ 68°). Samples that amplified the same 198bp band as pD171 were discarded. DNA-free RNA samples were purified with Zymo RNA clean and concentrate columns (Zymo; Irvine, CA) per the manufacturer's instructions. 1µg of RNA was treated with iScript cDNA synthesis kit (BioRad; Hercules, CA).

Quantitative, Reverse Transcriptase PCR (qRT-PCR)

qRT-PCRs were performed in a MyiQ thermocycler (BioRad). Reaction mixtures included 500 µg cDNA, 6 µM primers (*rpoD*-RT F/R; *rtcA*-RT F/R) (Table 3.1), and 2X SYBR Green Supermix (BioRad) in a volume of 20 µl. Cycling conditions were (1x—3' @ 95°; 35x—15" @ 95°, 30"@ 54.3°, 30" @ 72°). Presence of a single peak for each biological replicate was verified by agarose electrophoresis on a 2% SB gel. Standard curves were generated on each plate using 10-fold dilutions of pDS171 (*rpoD*) or pDS185 (*rtcA*). Data for each biological replicate indicates the average of 3 technical replicates. To determine relative *rtcA* expression, *rtcA* transcript levels were divided by *rpoD* transcript levels for each sample. The relative *rtcA* expression level was then compared between treated and untreated conditions. Data presented indicates the average ± standard deviation of three biological replicates. Results were analyzed using A Student's t-test 2-tailed, paired analysis.

Northern Blot Analysis

To detect the presence of a YrlA transcript, cells containing pRtcR^{con} were grown in LB^{Ap} until mid-logarithmic phase, induced with 1 mM IPTG (Gold Biotechnology; St. Louis, MO). RNA was harvested as described above 1 hr after induction. Northern blots were performed as previously described [50] 10 µg of RNA was separated via urea-polyacrylamide gel electrophoresis [urea-PAGE] and transferred onto a Nytran membrane (General Electric Healthcare; Little Chalfont, United Kingdom). RNA was linked to the membrane with UV light and baking 30' @ 80° C. A ^{32}P label was added to the "YrlA probe" oligonucleotide (Table 3.1) by incubating 20 pmol of DNA with 50 µC ATP [γ-^{32}P]

(Perkin-Elmer; Waltham, MA) and T4 polynucleotide kinase (NEB), per the manufacturer's instructions. Overnight hybridization was performed @ 44.4°C. The membrane was washed 4x in 2X SSC/0.2% SDS buffer. Imaging was performed on a GE Storm 840 phosphorimager. The blot was stripped and re-probed with "5S rRNA probe" (Table 3.1). All steps were as above except that hybridization was performed at 48.5°C. Results indicated represent one biological replicate.

Comparison of *rtcB* Gene Neighborhoods

Gene neighborhoods containing *rtcB* homologs were analyzed using the "IMG Top Homolog" tool on the Integrated Microbial Genomics (IMG) website (www.img.jgi.doe.gov). Using STM3519 (*rtcB*) as a reference, genes with >40% identity were selected. The >7500 returns were narrowed to ~400 via the following criteria: *(i)* exclusion of archaeal and eukaryotic sequences, *(ii)* inclusion of only genomes classified as "finished" (as opposed to "draft" or "permanent draft"), and *(iii)* elimination of multiple genomes representing the same species (multiple copies within the same strain were retained). These neighborhoods were visualized using the IMG neighborhood viewer and qualitatively assessed as described in the "Results" section.

Western Blot Analysis

Cells containing either pDS162 (FLAG-Rsr) or pDS164 (FLAG-RtcA) were grown in LB^{Tc} until mid-logarithmic phase and induced with 1 mM IPTG (Gold Biotechnology). Samples corresponding to 1 ml of cells at OD_{600}≈0.6 were harvested at T=0, 1, and 2 hrs post-induction. Cells were suspended in a 50:50 mixture of 2X SDS Sample Buffer (120 mM Tris pH 6.8, 10%β-mercaptoethanol, 4% SDS, 20% glycerol): Tris-buffered saline (TBS; 50 mM Tris-HCl, 150 mM NaCl, pH 7.5) . 10 µl of each sample was run on a 4%/12% SDS-PAGE [51] and transferred to a PVDF membrane (General Electric Healthcare). The membrane was blocked 90'@ room temperature with TBS + 0.1% Tween-20 (TBS-T) +5% dry milk (Kroger); incubated overnight at 4° with 1:2500 mouse monoclonal α-FLAG (Sigma-Aldrich; St. Louis,

MO) in TBS-T +3% dry milk; washed 4x with TBS-T @ room temperature; incubated 4 hr @ room temperature with 1/2500 rabbit α-mouse IgG-AP (Sigma) in TBS-T +1% milk; and washed with TBS-T 4x @ room temperature. Results were visualized with 25 ml of H_2O containing 1 NBT/BCIP substrate tablet (Roche; Indianapolis, IN) until color developed.

Acknowledgements

Josh Lewis constructed pRtcRcon. Ankit Agrawal prepared reagents for and assisted in performing the Northern blot analysis. Amber Enriquez and David Thoms generated some of the deletion mutant strains used, which Christine Hartman confirmed via PCR. Jorge Escalante provided the *Salmonella recA*::Kn^R strain and its parent. Funding for this research came from NSF grant MCB-1051175 (to Tim Hoover and Anna Karls) and from the UGA Microbiology Department.

References

1. Popow J, Schleiffer A, Martinez J: **Diversity and roles of (t)RNA ligases**. *Cellular and molecular life sciences : CMLS* 2012, **69**(16):2657-2670.

2. Pascal JM: **DNA and RNA ligases: structural variations and shared mechanisms**. *Current opinion in structural biology* 2008, **18**(1):96-105.

3. Huang SY, Pommier Y: **Cross-over of RNA 3'-phosphate ligase into the DNA world**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(51):20354-20355.

4. Tanaka N, Chakravarty AK, Maughan B, Shuman S: **Novel mechanism of RNA repair by RtcB via sequential 2',3'-cyclic phosphodiesterase and 3'-Phosphate/5'-hydroxyl ligation reactions**. *The Journal of biological chemistry* 2011, **286**(50):43134-43143.

5. Tanaka N, Shuman S: **RtcB is the RNA ligase component of an Escherichia coli RNA repair operon**. *The Journal of biological chemistry* 2011, **286**(10):7727-7731.

6. Englert M, Sheppard K, Aslanian A, Yates JR, 3rd, Soll D: **Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(4):1290-1295.

7. Popow J, Englert M, Weitzer S, Schleiffer A, Mierzwa B, Mechtler K, Trowitzsch S, Will CL, Luhrmann R, Soll D *et al*: **HSPC117 is the essential subunit of a human tRNA splicing ligase complex**. *Science* 2011, **331**(6018):760-764.

8. Reinhold-Hurek B, Shub DA: **Self-splicing introns in tRNA genes of widely divergent bacteria**. *Nature* 1992, **357**(6374):173-176.

9. Hausner G, Hafez M, Edgell DR: **Bacterial group I introns: mobile RNA catalysts**. *Mobile DNA* 2014, **5**(1):8.

10. Das U, Shuman S: **2'-Phosphate cyclase activity of RtcA: a potential rationale for the operon organization of RtcA with an RNA repair ligase RtcB in Escherichia coli and other bacterial taxa**. *Rna* 2013, **19**(10):1355-1362.

11. Buck M, Gallegos MT, Studholme DJ, Guo Y, Gralla JD: **The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor**. *J Bacteriol* 2000, **182**(15):4129-4136.

12. Bush M, Dixon R: **The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription**. *Microbiology and molecular biology reviews : MMBR* 2012, **76**(3):497-529.

13. Samuels DJ, Frye JG, Porwollik S, McClelland M, Mrazek J, Hoover TR, Karls AC: **Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the sigma(5)(4) (RpoN) regulon of Salmonella Typhimurium LT2**. *BMC genomics* 2013, **14**:602.

14. Genschik P, Drabikowski K, Filipowicz W: **Characterization of the Escherichia coli RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon**. *The Journal of biological chemistry* 1998, **273**(39):25516-25526.

15. Chen X, Taylor DW, Fowler CC, Galan JE, Wang HW, Wolin SL: **An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA**. *Cell* 2013, **153**(1):166-177.

16. Das U, Chakravarty AK, Remus BS, Shuman S: **Rewriting the rules for end joining via enzymatic splicing of DNA 3'-PO4 and 5'-OH ends**. *Proceedings of the National Academy of Sciences of the United States of America* 2013, **110**(51):20437-20442.

17. Kohn M, Pazaitis N, Huttelmaier S: **Why Y RNAs? About versatile RNAs and their functions**. *Biomolecules* 2013, **3**:143-156.

18. Chen X, Wurtmann EJ, Van Batavia J, Zybailov B, Washburn MP, Wolin SL: **An ortholog of the Ro autoantigen functions in 23S rRNA maturation in D. radiodurans**. *Genes & development* 2007, **21**(11):1328-1339.

19. Wurtmann EJ, Wolin SL: **A role for a bacterial ortholog of the Ro autoantigen in starvation-induced rRNA degradation**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(9):4022-4027.

20. Kenyon CJ, Walker GC: **DNA-damaging agents stimulate gene expression at specific loci in Escherichia coli**. *Proceedings of the National Academy of Sciences of the United States of America* 1980, **77**(5):2819-2823.

21. Suzuki H, Pangborn J, Kilgore WW: **Filamentous cells of Escherichia coli formed in the presence of mitomycin**. *J Bacteriol* 1967, **93**(2):683-688.

22. Modell JW, Hopkins AC, Laub MT: **A DNA damage checkpoint in Caulobacter crescentus inhibits cell division through a direct interaction with FtsW**. *Genes & development* 2011, **25**(12):1328-1343.

23. Labbe JC, Hekimi S, Rokeach LA: **The levels of the RoRNP-associated Y RNA are dependent upon the presence of ROP-1, the Caenorhabditis elegans Ro60 protein**. *Genetics* 1999, **151**(1):143-150.

24. Xue D, Shi H, Smith JD, Chen X, Noe DA, Cedervall T, Yang DD, Eynon E, Brash DE, Kashgarian M *et al*: **A lupus-like syndrome develops in mice lacking the Ro 60-kDa protein, a major lupus autoantigen**. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(13):7503-7508.

25. Khan SR, Gaines J, Roop RM, 2nd, Farrand SK: **Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing**. *Applied and environmental microbiology* 2008, **74**(16):5053-5062.

26. Lee JH, Scholl D, Nixon BT, Hoover TR: **Constitutive ATP hydrolysis and transcription activation by a stable, truncated form of Rhizobium meliloti DCTD, a sigma 54-dependent transcriptional activator**. *The Journal of biological chemistry* 1994, **269**(32):20401-20409.

27. Xu H, Gu B, Nixon BT, Hoover TR: **Purification and characterization of the AAA+ domain of Sinorhizobium meliloti DctD, a sigma54-dependent transcriptional activator**. *J Bacteriol* 2004, **186**(11):3499-3507.

28. Kroger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K *et al*: **An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium**. *Cell host & microbe* 2013, **14**(6):683-695.

29. Studholme DJ, Dixon R: **Domain architectures of sigma54-dependent transcriptional activators**. *J Bacteriol* 2003, **185**(6):1757-1767.

30. Bernard JJ, Cowing-Zitron C, Nakatsuji T, Muehleisen B, Muto J, Borkowski AW, Martinez L, Greidinger EL, Yu BD, Gallo RL: **Ultraviolet radiation damages self noncoding RNA and is detected by TLR3**. *Nature medicine* 2012, **18**(8):1286-1290.

31. Kladwang W, Hum J, Das R: **Ultraviolet shadowing of RNA can cause significant chemical damage in seconds**. *Scientific reports* 2012, **2**:517.

32. Doudney CO: **Ultraviolet light effects on the bacterial cell**. *Current topics in microbiology and immunology* 1968, **46**:116-175.

33. Kohanski MA, Dwyer DJ, Hayete B, Lawrence CA, Collins JJ: **A common mechanism of cellular death induced by bactericidal antibiotics**. *Cell* 2007, **130**(5):797-810.

34. Tkachenko AG, Akhova AV, Shumkov MS, Nesterova LY: **Polyamines reduce oxidative stress in Escherichia coli cells exposed to bactericidal antibiotics**. *Research in microbiology* 2012, **163**(2):83-91.

35. Wang X, Zhao X: **Contribution of oxidative damage to antimicrobial lethality**. *Antimicrobial agents and chemotherapy* 2009, **53**(4):1395-1402.

36. De la Cruz MA, Zhao W, Farenc C, Gimenez G, Raoult D, Cambillau C, Gorvel JP, Meresse S: **A toxin-antitoxin module of Salmonella promotes virulence in mice**. *PLoS pathogens* 2013, **9**(12):e1003827.

37. Zhang Y, Zhang J, Hara H, Kato I, Inouye M: **Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase**. *The Journal of biological chemistry* 2005, **280**(5):3143-3150.

38. Winther KS, Gerdes K: **Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(18):7403-7407.

39. Christensen-Dalsgaard M, Jorgensen MG, Gerdes K: **Three new RelE-homologous mRNA interferases of Escherichia coli differentially induced by environmental stresses**. *Molecular microbiology* 2010, **75**(2):333-348.

40. Bizanek R, McGuinness BF, Nakanishi K, Tomasz M: **Isolation and structure of an intrastrand cross-link adduct of mitomycin C and DNA**. *Biochemistry* 1992, **31**(12):3084-3091.

41. Suzuki H, Kilgore WW: **Decomposition of ribosomal particles in Escherichia coli treated with mitomycin C**. *J Bacteriol* 1967, **94**(3):666-676.

42. Suzuki H, Kilgore WW: **Effects of mitomycin C on macromolecular synthesis in Escherichia coli**. *J Bacteriol* 1967, **93**(2):675-682.

43. Snodgrass RG, Collier AC, Coon AE, Pritsos CA: **Mitomycin C inhibits ribosomal RNA: a novel cytotoxic mechanism for bioreductive drugs**. *The Journal of biological chemistry* 2010, **285**(25):19068-19075.

44. Benson NR, Wong RM, McClelland M: **Analysis of the SOS response in Salmonella enterica serovar typhimurium using RNA fingerprinting by arbitrarily primed PCR**. *J Bacteriol* 2000, **182**(12):3490-3497.

45. Courcelle J, Khodursky A, Peter B, Brown PO, Hanawalt PC: **Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient Escherichia coli**. *Genetics* 2001, **158**(1):41-64.

46. Maloy S, Stewart VJ, Taylor RK: **Genetic Analysis of Pathogenic Bacteria: A Laboratory Manual**. Plainview, NY: Cold Spring Harbor Laboratory Press; 1996.

47. Datsenko KA, Wanner BL: **One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**(12):6640-6645.

48. Hughes KT, Youderian P, Simon MI: **Phase variation in Salmonella: analysis of Hin recombinase and hix recombination site interaction in vivo**. *Genes & development* 1988, **2**(8):937-948.

49. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR: **RNAsnap: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria**. *Nucleic acids research* 2012, **40**(20):e156.

50. Mohanty BK, Giladi H, Maples VF, Kushner SR: **Analysis of RNA decay, processing, and polyadenylation in Escherichia coli and other prokaryotes**. *Methods in enzymology* 2008, **447**:3-29.

51. Sambrook J, Fritsch EF, Maniatis T: **Molecular Cloning: A Laboratory Manual**. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.

Table 3.1: Oligonucleotides used in this study

Name	Sequence
Deletion strain generation:	
rtcR-RED F	GCCTTTGGTTTTGTCGGTACGGTACTCGACTATGCATATGTAGGCTGGAGCTGCTTCG
rtcR-RED R	ATTCTGTAAAACGTCCCACGTCAGCCCAAAACGCGCCATATGAATATCCTCCTTAG
rsr-RED F	CATGGAGAATAACGGAAATGGGAAAAACAATGTGTAGGCTGGAGCTGCTTC
rsr-RED R	GGCCAGCATTCGCGCCGCGCTGAATACTCTCATATGAATATCCTCCTTA
yrlA-RED F	CTGGGTGGTCATTAATTCGTAATTCATCATTGTGTAGGCTGGAGCTGCTTC
yrlA-RED R	CTGGAAGAGATTGAAGCCGTGACGTTGTAACATATGAATATCCTCCTTA
rtcB-RED F	CAGAACGCACCGGTAAAAATGTGGACCAAAGGCGTATGTGTAGGCTGGAGCTGCTTCG
rtcB-RED R	TCCTTTAACGCACACCACCTGCCGCAGGGCGTACATCATATGAATATCCTCCTTAG
rtcA-RED F	TGTTTGCCTTGCAATGCTCGTAAAGATGACTGTGTAGGCTGGAGCTGCTTC
rtcA-RED R	CTGCGGCAGGTGTGTGCGTTAAAGGATAGCATATGAATATCCTCCTTA
qRT-PCR:	
rpoD-RT F	AACGAATAAGTGTGGATACCG
rpoD-RT R	TCTTCCATTACCTGAATACCC
rtcA-RT R	CTGGTTAGCTACCGCTTCCG
rtcA-RT F	CGAACGTGAAGTCGCAACGC
Cloning:	
rtcRcon-F1	GAATTCCTCAACTTCCTGAAGTCC
rtcRcon-F2	ATTATTGGATCCTAAAGAGGTATATATTAATGCTCAACTTCCTGAAGT
rtcR-R	AAGCTTAATTCTGTAAAACGTCCC
FLAG-Rsr F	ATTATTCATATGGACTACAAGGACGACGATGACAAAGCTAATCCACTTTTGTTCCG
Rsr R	GCATCGGGAAAGATGTAATCCC
FLAG-RtcB F	ATTATTCATATGGACTACAAGGACGACGATGACAAAATGAATTACGAATAAATG
RtcB R	ATTATTTCTAGAATCCAGCGCGATGATCCTTGCC
FLAG-RtcA F	ATTATTCATATGGACTACAAGGACGACGATGACAAAGCAAGGATCATCGCGCTGG
RtcA R	ATTATTTCTAGACTGAGACACACATACAAACCG
Northern blotting:	
yrlA probe	CTGACGGGTCTCGAACCC
5S rRNA probe	GGCGTTTCACTTCTGAGTTCGG

Figure 3.1: An Rsr-YrlA complex may direct damaged RNA to be degraded or repaired by different RNP complexes. Chen, et al. [15] proposed that YrlA tethers Rsr to PNPase; Rsr can then direct single-stranded ends of structured RNA to PNPase for processing. We propose that in *S.* Typhimurium the Rsr-YrlA complex may also interact with RtcB and/or RtcA (or other proteins also expressed under conditions that activate RtcR) and then direct other types of damaged RNA to be repaired.

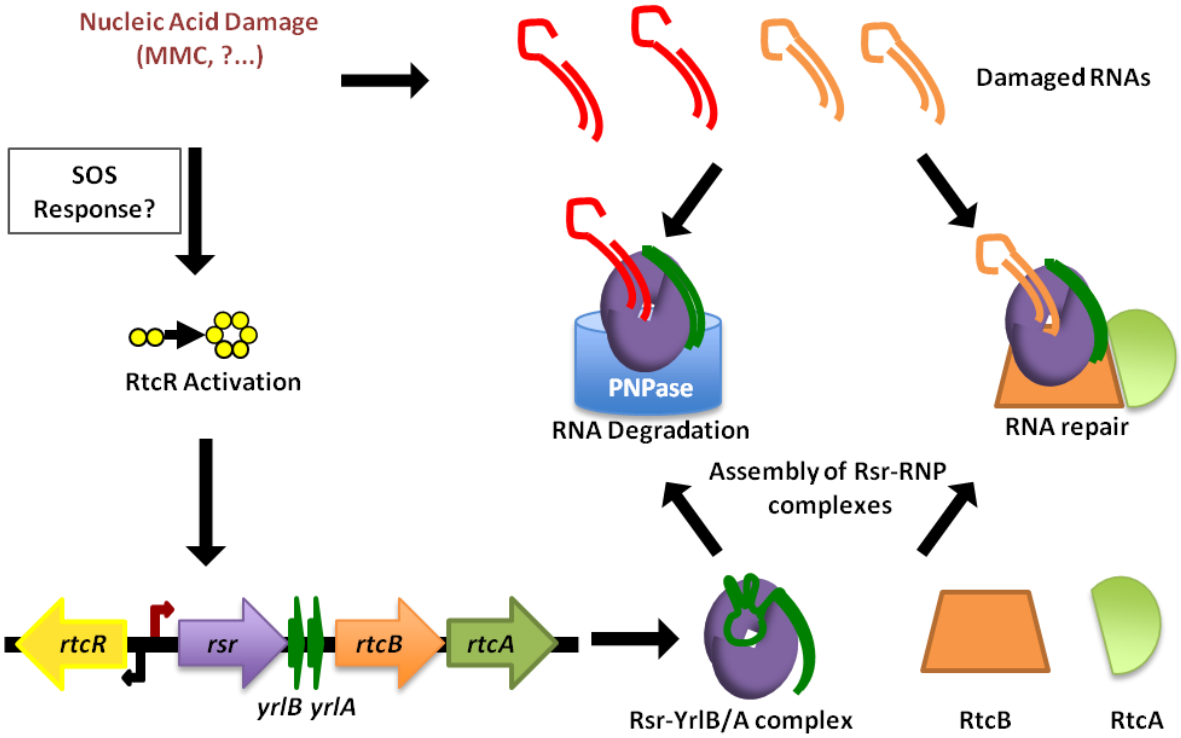

Nucleic Acid Damage
(MMC, ?...)
Damaged RNAs
SOS
Response?
RtcR Activation
PNPase
RNA Degradation
RNA repair
Assembly of Rsr-RNP
complexes
rtcR
rsr
rtcB
rtcA
yrlB yrlA
Rsr-YrlB/A complex
RtcB
RtcA

Figure 3.2: Mitomycin C treatment activates expression from P_{rtcBA}. Cultures grown to mid-logarithmic phase in MOPS minimal media were treated with 1μg/ml of Mitomycin C (MMC). After 90', total RNA was harvested and transcript levels of *rtcA* were assessed via quantitative reverse transcriptase PCR (A). Within each sample *rtcA* transcript levels were normalized to *rpoD* transcript levels. *rtcA/rpoD* for untreated WT cells was set equal to 1. Treatment with MMC caused 17-fold increase in WT cells. This transcription was dependent on RtcR and, since expression was not observed in a *recA* mutant, *rtcBA* expression may share common components with the SOS DNA damage response pathway. When a constitutively-active version of RtcR was provided, *rtcBA* expression was seen in both treated and untreated cells.

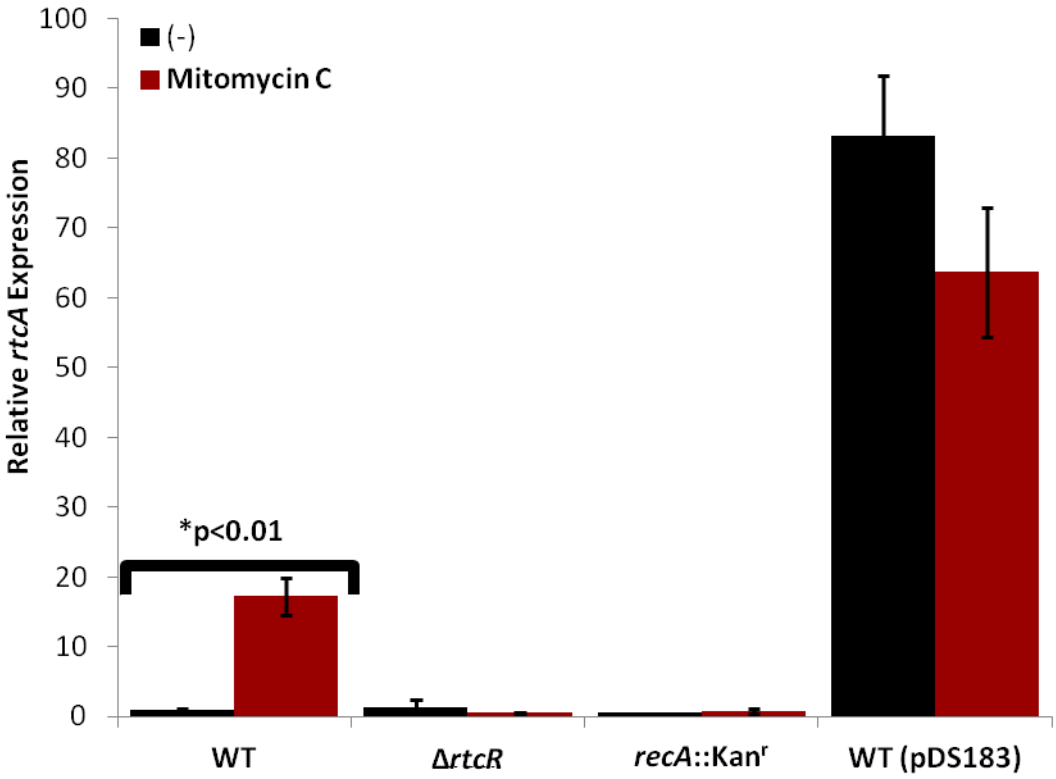

100
90
80
70
60
50
40
30
20
10
0
(-)
Mitomycin C
Relative rtcA Expression
*p<0.01
WT
ΔrtcR
recA::Kanr
WT (pDS183)

Figure 3.3: Wild type and Δ*rtcR* mutants are equally susceptible to MMC. Wild type, Δ*rtcR*, and wild type (pDS183) cells were grown in MOPS minimal medium +50 μM IPTG. At OD_{600}≈0.2 (indicated by the arrow), MMC was added to a concentration of 1 μg/ml. OD_{600} of cultures was measured at the times indicated.

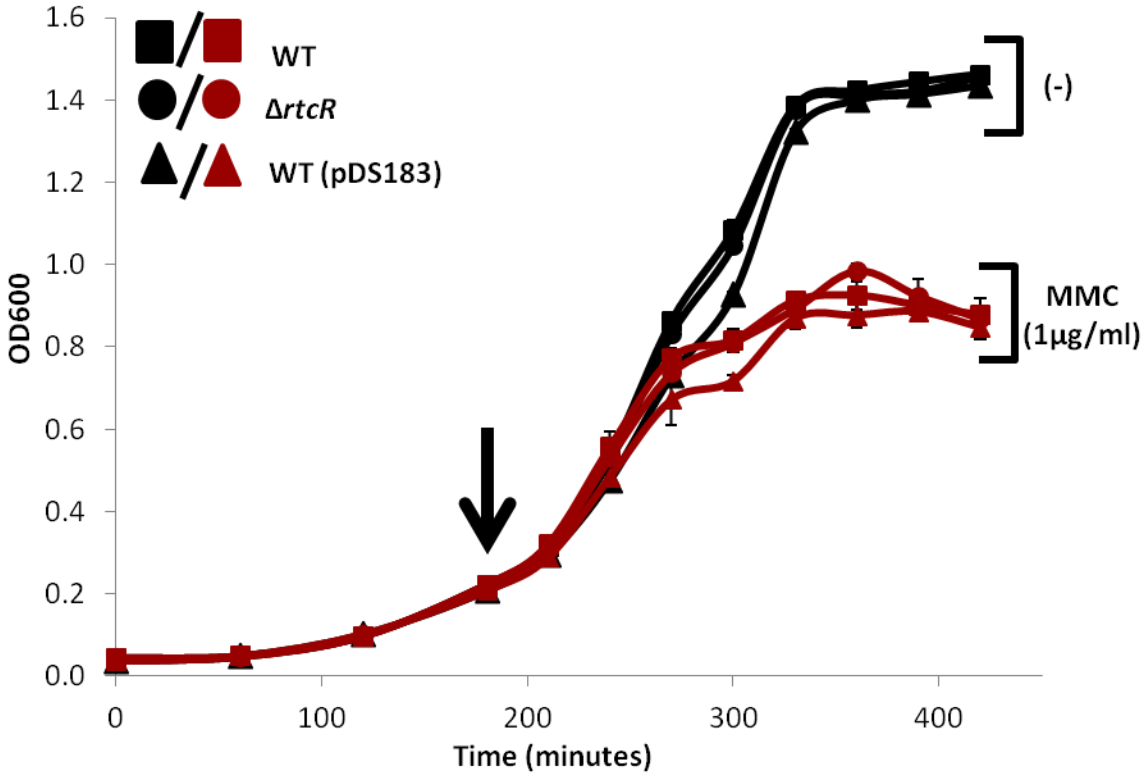

1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
OD600
0
100
200
300
400
Time (minutes)
WT
ΔrtcR
WT (pDS183)
(-)
MMC
(1μg/ml)

Figure 3.4: The YrlA Y RNA is not seen in cells lacking Rsr. Cells harboring pRtcRcon were grown to mid-logarithmic phase and induced with 1 mM IPTG. RNA samples harvested after 1hr were run on a Urea-SDS-PAGE and transferred to a Nytran membrane. Membranes were probed with a α-^{32}P-labeled YrlA probe and radiographic signals were detected on a storage phosphor screen (Top Panel). The expected size of YrlA is 110 nt. Due to the nature of the YrlA sequence, the probe had similarities to cellular tRNAs, which likely accounts for the ~70-80 nt band seen in each lane. The membrane was stripped and re-probed with a similarly-labeled probe against 5S rRNA (Bottom Panel).

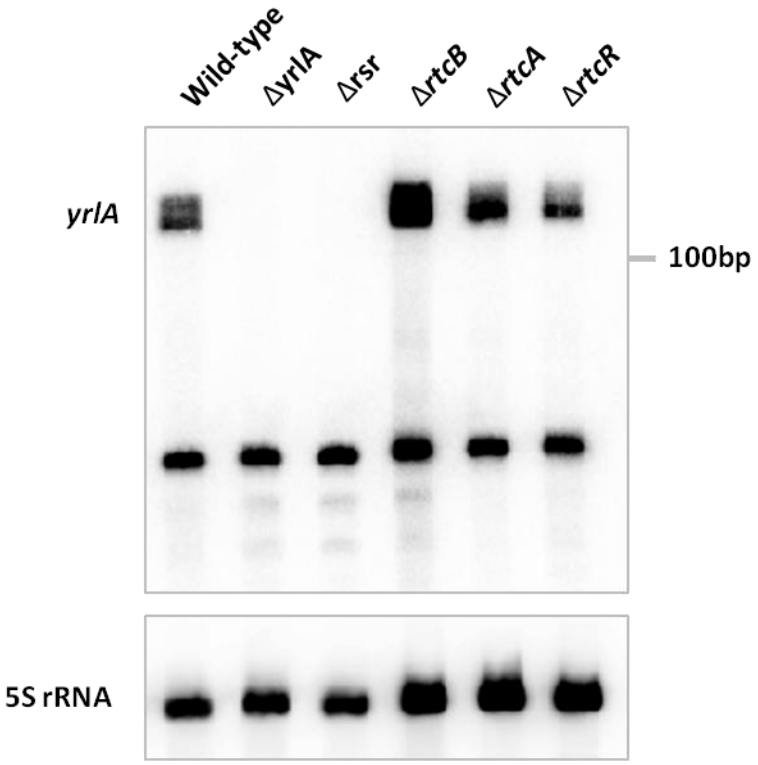

Wild-type
ΔyrlA
Δrsr
ΔrtcB
ΔrtcA
ΔrtcR
yrlA
100bp
5S rRNA

Figure 3.5: FLAG-tagged components of the RNA repair operon detected by Western blot. Cultures containing either pDS162 (FLAG-Rsr) or pDS164 (FLAG-RtcA) were grown to mid-logarithmic phase and induced with 1 mM IPTG. Samples were collected at T=0, 1, and 2 hours post-induction. Whole cell lysates were run on an SDS-PAGE and transferred to a PVDF membrane. The membrane was probed with α-FLAG antibodies and detected with an alkaline phosphatase substrate.

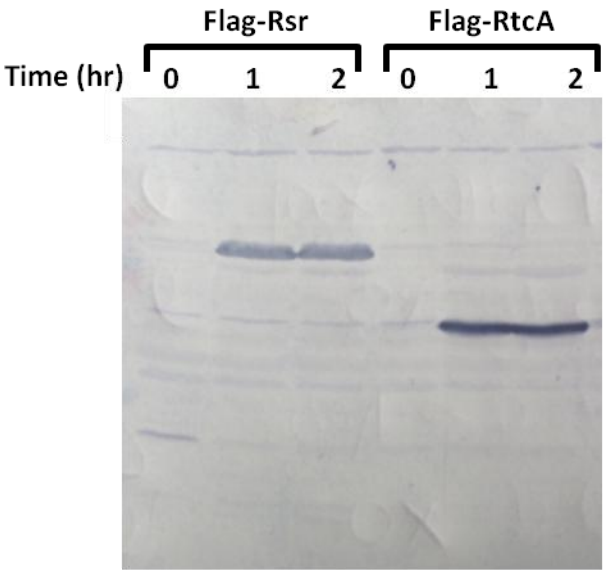
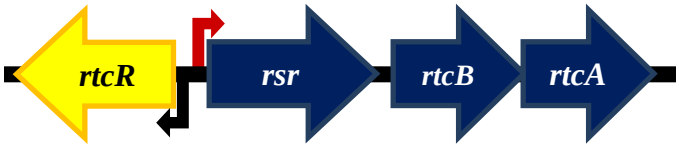

Flag-Rsr
Flag-RtcA
Time (hr)
0
1
2
0
1
2

Figure 3.6: *rtcB* gene neighborhoods. Genes with >40% amino acid identity to *Salmonella* RtcB were identified from the IMG database and their surrounding regions were analyzed using the Neighborhood Viewer tool. ~400 regions were chosen for analysis as described in Materials and Methods These neighborhoods were divided into four categories based on the genes surrounding *rtcB*. A) "Complete" neighborhoods contained orthologs of *rsr*, *rtcB,* and *rtcA* as well as the *rtcR* regulator. B) "RtcBA" neighborhoods contain orthologs of *rtcB*, *rtcA*, and *rtcR*, but lack a copy of *rsr*. C) "RtcB" neighborhoods contain orthologs of both *rtcB* and *rtcR,* but lack *rsr* and *rtcA.* D) "RtcB only" neighborhoods only contain an ortholog *rtcB*. The presence of *rtcR* at an alternative locus within these genomes was not determined. For each category, genera listed (grouped by phylum) contained at least one sequence with this organization. Genera marked with an (*) contained species with differing neighborhood organizations. The *rtcR* in *H. chejuensis*(†) is transcribed upstream of and on the same strand as the rest of the operon.

A

Complete *rsr-rtcBA* operon (+*rtcR*)

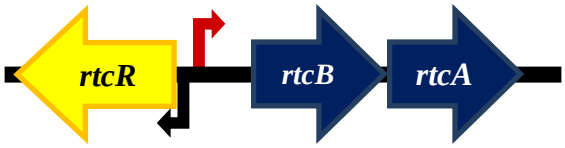


Proteobacteria:

Alpha: *Sphingobium, Sphingomonas* **Beta:** *Acidovorax*[*], *Variovorax*[*] **Gamma:** *Pseudomonas*[*], *Salmonella*[*]

B

***rtcBA* (+*rtcR*)**



Proteobacteria:

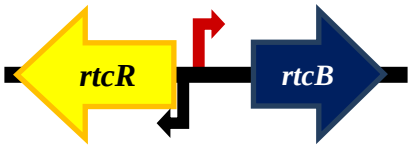
Beta: *Acidovorax*[*], Candidatus *Accumulibacter, Delftia, Polaromonas, Ralstonia*[*], *Variovorax*[*] **Gamma:** *Acinetobacter*[*], *Escherichia, Hahella*[†], *Methylomonas, Morganella, Pseudomonas*[*], *Salmonella*[*], *Shigella, Teredinibacter* **Delta:** *Haliangium, Myxococcus, Pelobacter, Sorangium, Stigmatella*

Acidobacteria:

Candidatus *Koribacter*

C

***rtcB* (+*rtcR*)**

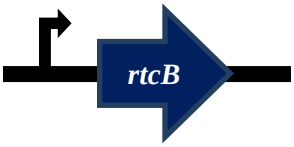


Proteobacteria:

Alpha: *Tistrella* **Beta:** *Azoarcus*[*], *Bordatella, Burkholderia*[*], *Chromobacterium* **Gamma:** *Citrobacter, Cronobacter, Dickeya, Enterobacter, Pectobacterium, Providencia, Serratia, Vibrio, Xenorhabdus*

D

rtcB only



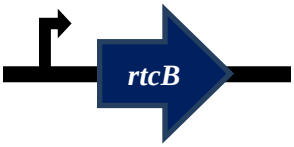
Proteobacteria:

Alpha: *Bradyrhizobium, Brevundimonas, Methylobacterium* **Beta:** *Advenella, Azoarcus*[*], *Burkholderia*[*], *Cupriavidus, Dechloromonas, Massilia, Nitrosomonas, Ralstonia*[*], *Thauera* **Gamma:** *Acinteobacter*[*], *Allochromatium, Halorhodospira, Nitrosococcus, Photorhabdus, Pseudomonas*[*], *Psychrobacter, Saccharophagus, Stenotrophomonas, Thiocystis, Xanthomonas* **Delta:** *Bdellovibrio, Desulfobulbus, Desulfomicrobium, Desulfotignum, Desulfovibrio, Geobacter, Hippea, Sorangium*

Bacteriodetes:

Bacteroides, Chitinophaga, Cytophaga, Dyadobacter, Flavobacterium, Haliscomenobacter, Niabella, Runella, Saprospira, Solitalea, Sphingobacterium, Spirosoma

rtcB only (cont.)



Fusobacteria:
Sebaldella

Firmicutes:
Bacillus, Butyrivibrio, Cellulosilyticum, Clostridium, Desulfitobacterium, Desulfotomaculum, Exiguobacterium, Halothermothrix, Listeria, Ruminococcus, Solibacillus, Thermacetogenium, Thermosediminibacter

Actinobacteria:
Actinoplanes, Actinosynnema, Amycolatopsis, Arthrobacter, Blastococcus, Brachybacterium, Catenulispora, Cellulomonas, Corynebacterium, Eggerthella, Frankia, Geodermatophilus, Gordonia, Intrasporangium, Jonesia, Kineococcus, Kitasatospora, Kribbella, Kutzneria, Microbacterium, Microlunatus, Micromonospora, Modestobacter, Mycobacterium, Nakamurella, Nocardia, Nocardiopsis, Propionibacterium, Rhodococcus, Saccharomonospora, Saccharopolyspora, Saccharothrix, Salinispora, Stackebrandtia, Streptomyces, Streptosporangium, Thermobifida, Thermomonospora, Tsukamurella, Verrucosispora

Cyanobacteria:
Acaryochloris, Anabaena, Calothrix, Chamaesiphon, Chroococcidiopsis, Crinalium, Cyanobacterium, Cyanothece, Cylindrospermum, Geitlerinema, Gloeobacter, Leptolyngbya, Microcoleus, Microcystis, Nodularia, Nostoc, Oscillatoria, Rivularia, Stanieria, Synechococcus

Acidobacteria:
Granulicella

Synergistetes:
Thermovirga

Chlamydiae:
Parachlamydia, Simkania

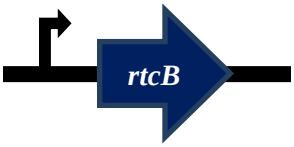
Planctomycetes:
Planctomyces

Deferribacteres:
Calditerrivibrio, Flexistipes

Chlorobi:
Chloroherpeton

Spirochaetes:
Brachyspira, Treponema

rtcB only (cont.)



Deinococcus-Thermus:
Deinococcus, Meiothermus

Chloroflexi:
Chloroflexus, Dehalococcoides, Herpetosiphon, Oscilloshloris, Rosieflexus

Thermotogae:
Kosmotoga

Aquificae:
Desulfurobacterium, Persephonella, Thermocrinis, Thermovibrio

Armatimonadetes:
Fimbriimonas, Chthonomonas

Chrysiogentes:
Desulfurispirillum

CHAPTER 4:

DISSERTATION DISCUSSION

σ^{54} is an alternative σ factor found widely throughout bacteria. Like all other σ factors, σ^{54} interacts with RNA polymerase (RNAP) core to form a holoenzyme that can catalyze transcription. Since the σ factor is the subunit of RNAP holoenzyme responsible for promoter interaction, different σ factors can direct the holoenzyme to different promoter sequences. Having multiple alternative σ factors—each recognizing their own subset of genes, or regulon—is one way that bacteria can regulate multiple genes at one time in response to a particular stimulus.

While σ^{54} has the same essential function as all of the other σ factors in a cell (known as the σ^{70}-family), substantial differences in amino acid sequence give rise to vastly different mechanisms. Compared to the σ^{70} family, σ^{54} recognizes promoter sequences that are in a different position relative to the transcription start site. σ^{54} has a much more stringent requirement for spacing between promoter elements. Whereas σ^{70} has some flexibility regarding the -10 and -35 regions it recognizes, addition or deletion of even one base pair between the -12 and -24 positions in a σ^{54}-dependent promoter will render it non-functional [1]. σ^{54} holoenzyme ($E\sigma^{54}$) forms a stable closed complex with promoter DNA. This means that unlike $E\sigma^{70}$, which can spontaneously isomerize into open complex, $E\sigma^{54}$ requires energy to generate the open complex and initiate transcription. This energy is provided through ATP hydrolysis by a bacterial enhancer-binding protein (bEBP). The N-terminal of most bEBPs contains a regulatory domain that responds to a particular environmental signal. Upon sensing this signal, the bEBP becomes active and, through a helix-turn-helix domain at its C-terminal, interacts with an upstream activator sequence (UAS) in the vicinity of the promoter (~75-100 bp away). DNA looping brings the central AAA+ ATPase domain of the bEBP into contact with the $E\sigma^{54}$-promoter closed complex, where it can carry out ATP

hydrolysis. The requirement for an upstream enhancer sequence, the DNA looping, and the energy provided by nucleotide hydrolysis make transcription activation by $E\sigma^{54}$ more similar to RNA polymerase II found in eukaryotes than other prokaryotic RNAP holoenzymes [2].

The fact that the N-terminal regulatory domain of most bEBPs is responsive to specific environmental conditions ensures that cells do not waste energy transcribing genes that are unnecessary or potentially deleterious. The presence of multiple bEBPs within many bacterial genomes indicates that these organisms have a robust capacity to respond to a variety of environmental stimuli [3, 4]. However, while this regulatory scheme is undoubtedly helpful to the bacteria, it can confound experimental attempts to study individual promoters in isolation or examine the σ^{54} regulon globally. Unless the proper environmental conditions are used, σ^{54}-dependent promoters will yield false negative results, owing to an inactive bEBP.

However, not all bEBPs require an upstream enhancer sequence. *Helicobacter pyori* and *Chlamydia trachomatis* are each only predicted to encode one bEBP: FlgR and CtrC, respectively [5]. Since these bEBPs need to activate transcription from every σ^{54}-dependent promoter in the cell, enhancer-mediated specificity is unnecessary. Consequently, these proteins lack a C-terminal DNA-binding domain. Experimental C-terminal truncation endows other bEBPs with similar promiscuity [6, 7].

Other bEBPs lack the N-terminal regulatory domain. In lieu of this domain, PspF relies on PspA to serve as an anti-activator [8]. Activating conditions result in PspB and PspC binding PspA, allowing PspF to interact with $E\sigma^{54}$. In cells deleted for PspA, PspF is constitutively active. Truncation of the regulatory regions from other bEBPs will also yield constitutively-active proteins.

Combining these ideas and truncating both the N- and C-terminal domains of a bEBP can lead to a constitutively-active, promiscuous activator: the central AAA+ ATPase domain of *Sinorhizobium meliloti* DctD (DctD250) was able to drive transcription from multiple σ^{54}-dependent promoters under non-activating conditions [9, 10]. In the work presented in Chapter 2, I used DctD250 to define the global σ^{54} regulon in *S.* Typhimurium. In *Salmonella*, DctD250 was not as effective as a native activator. Wild-

type cells (without DctD250) grown under nitrogen-limiting media (which activates NtrC) displayed 10-fold higher activity from the *glnKamtB* promoter than cells containing DctD250 grown in standard medium (Table 2.1). However, no transcription was observed in cells without DctD250 grown in standard medium. Similar DctD250-dependent transcription of the *rsr-rtcBA* promoter indicated that this bEBP variant was capable of functioning as a constitutively-active, promiscuous activator in *Salmonella.*

Using microarray analysis, we compared transcription in wild-type and Δ*rpoN* cells containing DctD250. Of the 20 *Salmonella* promoters that were predicted (by bioinformatics or homology to other organisms) to be σ^{54}-dependent, we were able to experimentally verify the σ^{54}-dependence of 16 (Table 2.2). Further analysis of some of the "missed" promoters showed transcription in both wild-type and mutant cells, indicating the presence of one or more additional promoters, which is supported by previous reports [11-13]. One interesting finding from this experiment was the σ^{54}-dependent transcription of STM2938, the penultimate gene in a nine-gene operon annotated as CRISPR-associated (*cas*) genes. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and *cas* genes constitute an adaptive immune system in Bacteria and Archaea that protects against invading mobile DNA, such as phage and plasmids [14]. Though σ^{54} does regulate other genes that respond to phage infection—the phage shock *psp* genes [7, 15, 16]—no connection between σ^{54} and CRISPR/*cas* has been described previously.

Another interesting feature of σ^{54} is its ability to bind to DNA on its own (as opposed to other σ factors, which can only bind DNA in the context of RNAP holoenzyme) [17]. Though σ^{54} has a lower affinity for DNA than $E\sigma^{54}$, the fact that σ^{54} can bind on its own indicates a potential regulatory function for this binding. Using chromatin-immunoprecipitation coupled with microarray analysis (ChIP-chip) analysis, we examined σ^{54} (or $E\sigma^{54}$) binding sites. We were able to identify 70 DNA regions that interact with σ^{54} or $E\sigma^{54}$. Because our microarray, a whole open reading frame array, could not precisely locate binding sites, we used a position-specific score matrix (PSSM) within the Motif Locator program [18], to predict the specific binding site. While some of these sites corresponded to the promoters defined with the

microarray experiments, more than half of the predicted sites were predicted to be located within open reading frames (ORFs). One of these intragenic sites was located within STM2938, upstream of STM2939, which we had identified as σ^{54}-dependent in our microarrays. These results indicated that intragenic binding may play a substantial role in σ^{54} regulation.

To determine whether binding sites could actually function as σ^{54}-dependent promoters, we cloned them into a vector containing a promoterless *lacZ*. In the presence of DctD250, we were able to detect promoter activity from ~40% (3/7) sequences that we examined (Figure 2.4). This indicates that these sites have the potential to act as σ^{54}-dependent promoters. If they *do* function as promoters, it will be interesting to examine the function of the transcripts being generated.

With these results in mind several new directions arise. The first involves the apparent σ^{54}-dependence of the *cas1* homolog, STM2939. We were able to demonstrate that an isolated sequence from within STM2938 could function as a σ^{54}-dependent promoter. What is the transcript being produced from this site? Our results indicate that the transcript hybridizes to STM2939 in a whole ORF array, but this does not preclude an alternative transcript that overlaps this ORF. Current research in our laboratory is attempting to identify the transcription start site from this promoter. In addition, determining which bEBP interacts with this site *in vivo* will improve our understanding of the role of this secondary promoter. Because of the role of CRISPR-*cas* systems in response to phage infections, the phage shock response regulator, PspF, is an attractive candidate. If this is the case, σ^{54} may play a broader role in response to phage stress than previously thought.

Another interesting question comes from the discovery of a large number of intragenic σ^{54} (or $E\sigma^{54}$) binding sites. Though we determined that several of these sites could function as promoters when cloned onto a plasmid, their function in the chromosomal context is still unknown. If these sites are able to serve as promoters, what transcript are they generating? Is the transcript an un-annotated ORF? Is it a small regulatory, perhaps anti-sense, RNA? Or does the act of transcription from an internal promoter, *per se*, regulate the gene that contains it [19, 20]? If these sites are not serving as promoters, what role does

binding play? Is it to regulate transcription of that gene? Does it sequester $E\sigma^{54}$ as a holoenzyme, which may be useful, given that σ^{54} has a lower affinity for core than σ^{70} does [21]? Also, if binding is playing an important role in regulation, is the binding by σ^{54} or $E\sigma^{54}$? Recent work in our laboratory utilizing high-density tiling arrays for ChIP-chip provided more precise localization of the $E\sigma^{54}$ chromosomal binding sites, and EMSA (Electrophoretic Mobility Shift Assays) with ~30 identified binding sites were used to determine whether the nature of the sites that interact with the sigma factor itself or the holoenzyme. Being able to distinguish between these two types of sites should allow the development of consensus sequences that can be used to make reliable predictions about σ^{54} regulation in other bacteria.

One final takeaway from this research on the global σ^{54} regulon is the potential portability of this σ^{54} regulon definition system. In this work, I used DctD250, a bEBP from *Sinorhizobium*, an α-Proteobacteria, to examine the regulon of *Salmonella*, a γ-Proteobacteria. This work indicates that DctD250 could be used to define the σ^{54} regulons in other bacteria, or at least other Proteobacteria. If DctD250 specifically does not work in a given organism, the approach can be modified by experimentally truncating one of that organism's bEBPs to generate a constitutively-active, promiscuous variant. Given advances in technology, current iterations of this approach should use high-throughput analyses, like RNA-seq or ChIP-seq to examine transcripts and σ^{54}/$E\sigma^{54}$ binding sites. Irrespective of the analytical methods, using a constitutively-active, promiscuous bEBP in defining global σ^{54} regulons will address the possibility of false negatives caused by the lack of proper environmental stimuli.

Had we attempted this global analysis in *Salmonella* without DctD250, one likely false negative result would have been the *rsr-yrlBA-rtcBA* operon. In our standard growth media, we were able to observe transcription from this promoter only in the presence of DctD250 (Table 2.1). In fact, expression of this operon in wild-type *Salmonella* has not been reported under any growth conditions. One simple explanation for this is that either the promoter or the gene products are non-functional. However, experiments using N-terminal truncations of RtcR, this operon's bEBP, have revealed transcripts, supporting our findings that the promoter upstream of *rsr-rtcBA* is functional [22, 23]. Additionally, the

products of these genes have been characterized. Rsr-Y RNA ribonucleoprotein complexes have been examined in *Deinococcus radiodurans* [24-26] and RtcB and RtcA from *E. coli* (which are 88% and 68% identical to their *Salmonella* homologs; Figure 1.3) have been cloned and characterized *in vitro* [22, 27]. Since these genes are ostensibly functional and under control of a functional promoter, they are likely to play some role in *Salmonella* physiology.

Based on the characterizations of these genes in other organisms, Rsr-RtcBA functions as an RNA repair system. RtcB is an RNA ligase that catalyzes reactions between unusual RNA termini: 5'-OH and either 3'-PO_4 or 2',3'-cyclic phosphate [27]. In bacteria, these termini are often generated by ribotoxins (e.g., toxin-antitoxin system toxins VapC [28] or MazF [29], colicins D or E5 [30], or the anticodon nuclease, PrrC [31]). RNA with a 2',3' cyclic phosphate may also be generated by RtcA, which can cyclize either 2'- or 3'-PO_4 moieties [22, 32]. Rsr (or its eukaryotic ortholog, Ro) form a ribonucleoprotein complex (RNPc) with small RNAs known as Y RNAs. These RNPc's are up-regulated in times of stress (e.g., UV [24, 33] or starvation [26]) and bind to damaged or mis-folded RNA molecules [33-36]. Rsr uses the Y RNA as a molecular tether to attach to other proteins, including the ribonuclease PNPase, thereby directing damaged RNA molecules to an appropriate processing enzyme [23, 37]. Our model (Figure 3.1) is that RNA damaging conditions will activate *Salmonella* RtcR and induce expression of Rsr-RtcBA (while also damaging RNA). Depending on the nature of the RNA damage, Rsr-Y RNA RNPc's will interact with RtcB, RtcA, PNPase, and other cellular proteins to properly repair (or degrade) these damaged RNA molecules.

In Chapter 3, I present the initial examination of this model in *Salmonella*. The first step was to determine the conditions that activate RtcR and lead to expression of Rsr, YrlB/A, and RtcB/A. We used quantitative reverse transcriptase PCR (qRT-PCR) to assess transcript levels after exposure to a variety of environmental stresses. Conditions were tested that are known to either *i)* directly damage RNA or *ii)* induce expressions of toxin-antitoxin systems, which can subsequently damage RNAs (and, importantly, generate the 2',3' cyclic phosphate and 5'-OH termini that are substrates for RtcB). Out of these

conditions, only one—mitomycin C—led to expression of this operon. Mitomycin C is an antibiotic that forms nucleic acid adducts, leading to inter- and intra-strand cross links. Since mitomycin C is known to activate the SOS DNA damage response [38], we were interested in determining whether this activation of Rsr-RtcBA was a part of this pathway. Experiments with a *recA* mutant (the master regulator of the SOS response) showed no increase in *rsr-rtcBA* expression after MMC treatment, indicating that this system was somehow related to the SOS response (Figure 3.2). In subsequent experiments, expression of this system did not give cells a survival advantage when treated with MMC.

Though this is the extent of the research we have accomplished on this system to date, there are several questions that need to be answered before we can understand the biological role of this system. The first step is to continue searching for conditions that stimulate transcription of the RNA repair operon. Knowing that nucleic acid damage by mitomycin C will turn this system on, other nucleic acid damaging agents should be examined, including methyl-methanesulfonate (MMS) and ultraviolet (UV) light. We have previously looked into the effects of a Δ*rtcR* mutation on survival after UV exposure but results were inconclusive.

Additionally, as discussed in Chapter 3, we need a better understanding of how the cell is reacting to MMC. Although the presence of Rsr-Y RNA and RtcB/RtcA does not confer a survival advantage after MMC treatment, it is plausible that our methods were flawed (i.e., measuring optical density after 1 μg/ml treatment). We are currently pursuing alternate approaches, including quantitative plate counts and lowering the MMC concentrations used in treatments. Additionally, to determine the changes that occur in the cell after MMC treatment and necessitate expression of Rsr-RtcBA, we are performing high-throughput transcriptomic and proteomic analysis (RNA-seq; subtractive mass spectrometry) of cells exposed to MMC. We are not only determining the changes that occur between treated and untreated cells, but also the differences that occur between wild-type cells and various RNA repair mutants (Δ*rtcR*, Δ*rsr*, Δ*rtcB*) after treatment. This will tell us what changes are occurring as well as whether or not any of those transcription/translation changes are dependent on components of the RNA repair system.

Now that at least one activating condition for RtcR is understood, we can attempt to pinpoint the specific stimulus to which RtcR responds. Characterized bEBPs are activated in one of three ways: *(i)* phosphorylation, *(ii)* interaction with a small molecule, *(iii)* direct protein-protein interaction. Most bEBPs activated by direct protein-protein interaction—like the phage shock regulator, PspF—lack an N-terminal domain, which is functionally replaced by an antiactivator protein that dissociates from the bEBP under activating conditions [8]. Since RtcR contains a full N-terminal domain, this mechanism seems unlikely. Other bEBPs, like ZraR [39] and NtrC [40] are response regulators of two-component systems (TCS). Under activating conditions, these proteins are phosphorylated by their cognate histidine kinase—ZraS or NtrB, respectively. Whereas many TCSs are co-transcribed (30/32 in *E. coli* [41]), *rtcR* is an isolated gene. This would mean that RtcR is either an orphan in an atypical TCS or that activation of RtcR is not via phosphorylation. This leaves small molecule ligand binding as the mostly activation route for RtcR. Similar bEBPs include XylR and NorR, which interact with certain aromatic hydrocarbons [42-44] and nitric oxide [45], respectively.

Another approach to determining the RtcR activation stimulus is to assess transcription in various *Salmonella* mutants. Since transcription of the RNA repair operon is abolished in a *recA*::Kn^R knockout strain, this system may be a part of the *Salmonella* SOS response pathway. Since the SOS response involves a transcriptional cascade, we could create various mutants in known SOS-dependent genes. By determining which mutations abolished MMC-induced expression of *rsr-rtcBA*, we could deduce the mechanism of RtcR activation. It should be noted that the apparent RecA-dependence of this operon is a preliminary result. The *recA*::kan^R allele is currently being transduced from its strain LT2 parent to a 14028s background so that results can be compared in an isogenic background. One of the best ways to determine the function of this system is to analyze the Rsr-Y RNA ribonucleoprotein complexes (RNPc's) themselves. In the initial analysis of the *Salmonella* RNPc, Chen, et al. identified complexes that sedimented at a larger mass than the *D. radiodurans* complexes containing only Rsr-YrlA-PNPase [23]. They interpreted this as indicative of additional components in the RNPc, which they did not

examine. Another important consideration with their results is that they expressed *rsr-rtcBA* with a truncated RtcR (as opposed to an environmental stress). Therefore, any RNPc-interacting proteins that are only expressed under stress conditions would not have been detected. To determine the nature of these RNPc's, I have generated components of this system (Rsr, RtcB, and RtcA) with FLAG-epitope tags. After treatment with MMC, these complexes can be immunoprecipitated and analyzed via mass spectrometry to determine any protein partners as well as RNA-seq to characterize substrate RNA molecules. Knowing which other proteins are interacting with this complex will help us understand the nature of the processing that the Rsr-Y RNA RNPc is directing damaged RNA toward; the associated enzymes may degrade or repair the damaged RNA. And knowing which RNA molecules are associated will yield insight into the types of RNA that are affected by this system, e.g. rRNA, tRNA, mRNA, sRNA, tmRNA, or a new RNA species.

The fact that well-conserved homologs (>40% identity) of RtcB are found in genomes representing 19 different bacterial genera argues for a widely-maintained function for this enzyme (Figure 3.6). Interestingly, RtcB is the only component of the *Salmonella* RNA repair operon that is this widely conserved. In contrast, RtcR and RtcA are present mostly in the Proteobacteria. Regulation-wise, this indicates that these genes are constitutively-expressed or that they have developed another form of regulation outside of the Proteobacteria. The lack of RtcA could indicate that this function is not important enough to have been maintained, or that there are other proteins in these organisms that possess complementary functions. The lack of *rsr* from the majority of bacterial genomes is even more perplexing; only five genera possess an *rtcBA* operon encoding a homolog of Rsr. The approach outlined below should yield at least basic answers as to why this ribonucleoprotein has been so sporadically maintained.

The functions of the RtcB RNA repair systems in three different bacteria, *S.* Typhimurium, *E. coli*, and *Pseudomonas aeruginosa*, will be compared. Genome analysis using the Integrated Microbial Genomes system from the Joint Genome Institute reveals that *E. coli* and *P. aeruginosa* have operons

similar to *S.* Typhimurium, except they do not include *rsr;* and *rsr* is not encoded elsewhere in their genomes. Additionally, *P. aeruginosa* encodes a second copy of *rtcB* at a separate locus. Though this second copy may still be σ^{54}- and RtcR-dependent, *rtcR* and *rtcA* are not duplicated. Instead, this second locus contains a homolog of PrfB, a protein involved in peptide release from ribosomes.

The Karls laboratory has already acquired mutants deficient each of the operon components from the Keio strain collection [46] and the University of Washington *Pseudomonas* mutant library [47]. The first step in this comparative analysis will be to examine transcription from this operon in response to MMC treatment. Does MMC treatment up-regulate the RtcBA operons in these organisms? Is it RtcR-dependent? Does MMC affect growth in these organisms in the same way it does *Salmonella*? If there are differences between these bacteria, Rsr would be a prime candidate. Do *E. coli/P. aeruginosa* have similar phenotypes to a *S.* Typhimurium Δ*rsr* mutant? Determining how the *rtcBA* operon is regulated in other bacteria will help us gain a broader insight into the nature of the conserved functions of this enigmatic system.

Additionally, understanding how this system functions could provide insight into human and animal health. If the stress conditions that induce expression of the RNA repair operon are found in a host environment, the components of the RNA repair operon could be adapting the cell for survival. If this is the case, then targeting these components would be a viable strategy toward the development of new antibiotic treatments.

References

1. Barrios H, Valderrama B, Morett E: **Compilation and analysis of sigma(54)-dependent promoter sequences**. *Nucleic acids research* 1999, **27**(22):4305-4313.

2. Kornberg RD: **Mechanism and regulation of yeast RNA polymerase II transcription**. *Cold Spring Harbor symposia on quantitative biology* 1998, **63**:229-232.

3. Bush M, Dixon R: **The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription**. *Microbiology and molecular biology reviews : MMBR* 2012, **76**(3):497-529.

4. Studholme DJ, Dixon R: **Domain architectures of sigma54-dependent transcriptional activators**. *J Bacteriol* 2003, **185**(6):1757-1767.

5. Beck LL, Smith TG, Hoover TR: **Look, no hands! Unconventional transcriptional activators in bacteria**. *Trends in microbiology* 2007, **15**(12):530-537.

6. North AK, Kustu S: **Mutant forms of the enhancer-binding protein NtrC can activate transcription from solution**. *Journal of molecular biology* 1997, **267**(1):17-36.

7. Jovanovic G, Rakonjac J, Model P: **In vivo and in vitro activities of the *Escherichia coli* sigma54 transcription activator, PspF, and its DNA-binding mutant, PspFDeltaHTH**. *Journal of molecular biology* 1999, **285**(2):469-483.

8. Joly N, Engl C, Jovanovic G, Huvet M, Toni T, Sheng X, Stumpf MP, Buck M: **Managing membrane stress: the phage shock protein (Psp) response, from molecular mechanisms to physiology**. *FEMS microbiology reviews* 2010, **34**(5):797-827.

9. Brahmachary P, Dashti MG, Olson JW, Hoover TR: ***Helicobacter pylori* FlgR is an enhancer-independent activator of sigma54-RNA polymerase holoenzyme**. *J Bacteriol* 2004, **186**(14):4535-4542.

10. Xu H, Gu B, Nixon BT, Hoover TR: **Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a sigma54-dependent transcriptional activator**. *J Bacteriol* 2004, **186**(11):3499-3507.

11. Gopel Y, Luttmann D, Heroven AK, Reichenbach B, Dersch P, Gorke B: **Common and divergent features in transcriptional control of the homologous small RNAs GlmY and GlmZ in Enterobacteriaceae**. *Nucleic acids research* 2011, **39**(4):1294-1309.

12. Janaszak A, Majczak W, Nadratowska B, Szalewska-Palasz A, Konopa G, Taylor A: **A sigma54-dependent promoter in the regulatory region of the *Escherichia coli rpoH* gene**. *Microbiology* 2007, **153**(Pt 1):111-123.

13. Ramirez-Santos J, Collado-Vides J, Garcia-Varela M, Gomez-Eichelmann MC: **Conserved regulatory elements of the promoter sequence of the gene *rpoH* of enteric bacteria**. *Nucleic acids research* 2001, **29**(2):380-386.

14. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P: **CRISPR provides acquired resistance against viruses in prokaryotes**. *Science* 2007, **315**(5819):1709-1712.

15. Lloyd LJ, Jones SE, Jovanovic G, Gyaneshwar P, Rolfe MD, Thompson A, Hinton JC, Buck M: **Identification of a new member of the phage shock protein response in *Escherichia coli*, the phage shock protein G (PspG)**. *The Journal of biological chemistry* 2004, **279**(53):55707-55714.

16. Weiner L, Brissette JL, Model P: **Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms**. *Genes & development* 1991, **5**(10):1912-1923.

17. Buck M, Cannon W: **Specific binding of the transcription factor sigma-54 to promoter DNA**. *Nature* 1992, **358**(6385):422-424.

18. Mrazek J: **Finding sequence motifs in prokaryotic genomes--a brief practical guide for a microbiologist**. *Briefings in bioinformatics* 2009, **10**(5):525-536.

19. Chatterjee A, Drews L, Mehra S, Takano E, Kaznessis YN, Hu WS: **Convergent transcription in the butyrolactone regulon in *Streptomyces coelicolor* confers a bistable genetic switch for antibiotic biosynthesis**. *PloS one* 2011, **6**(7):e21974.

20. Chatterjee A, Johnson CM, Shu CC, Kaznessis YN, Ramkrishna D, Dunny GM, Hu WS: **Convergent transcription confers a bistable switch in *Enterococcus faecalis* conjugation**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(23):9721-9726.

21. Maeda H, Fujita N, Ishihama A: **Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase**. *Nucleic acids research* 2000, **28**(18):3497-3503.

22. Genschik P, Drabikowski K, Filipowicz W: **Characterization of the *Escherichia coli* RNA 3'-terminal phosphate cyclase and its sigma54-regulated operon**. *The Journal of biological chemistry* 1998, **273**(39):25516-25526.

23. Chen X, Taylor DW, Fowler CC, Galan JE, Wang HW, Wolin SL: **An RNA degradation machine sculpted by Ro autoantigen and noncoding RNA**. *Cell* 2013, **153**(1):166-177.

24. Chen X, Quinn AM, Wolin SL: **Ro ribonucleoproteins contribute to the resistance of *Deinococcus radiodurans* to ultraviolet irradiation**. *Genes & development* 2000, **14**(7):777-782.

25. Chen X, Wurtmann EJ, Van Batavia J, Zybailov B, Washburn MP, Wolin SL: **An ortholog of the Ro autoantigen functions in 23S rRNA maturation in *D. radiodurans***. *Genes & development* 2007, **21**(11):1328-1339.

26. Wurtmann EJ, Wolin SL: **A role for a bacterial ortholog of the Ro autoantigen in starvation-induced rRNA degradation**. *Proceedings of the National Academy of Sciences of the United States of America* 2010, **107**(9):4022-4027.

27. Tanaka N, Shuman S: **RtcB is the RNA ligase component of an *Escherichia coli* RNA repair operon**. *The Journal of biological chemistry* 2011, **286**(10):7727-7731.

28. Winther KS, Gerdes K: **Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(18):7403-7407.

29. Zhang Y, Zhang J, Hara H, Kato I, Inouye M: **Insights into the mRNA cleavage mechanism by MazF, an mRNA interferase**. *The Journal of biological chemistry* 2005, **280**(5):3143-3150.

30. Papadakos G, Wojdyla JA, Kleanthous C: **Nuclease colicins and their immunity proteins**. *Quarterly reviews of biophysics* 2012, **45**(1):57-103.

31. Kaufmann G: **Anticodon nucleases**. *Trends in biochemical sciences* 2000, **25**(2):70-74.

32. Das U, Shuman S: **2'-Phosphate cyclase activity of RtcA: a potential rationale for the operon organization of RtcA with an RNA repair ligase RtcB in *Escherichia coli* and other bacterial taxa**. *Rna* 2013, **19**(10):1355-1362.

33. Chen X, Smith JD, Shi H, Yang DD, Flavell RA, Wolin SL: **The Ro autoantigen binds misfolded U2 small nuclear RNAs and assists mammalian cell survival after UV irradiation**. *Current biology : CB* 2003, **13**(24):2206-2211.

34. O'Brien CA, Wolin SL: **A possible role for the 60-kD Ro autoantigen in a discard pathway for defective 5S rRNA precursors**. *Genes & development* 1994, **8**(23):2891-2903.

35. Shi H, O'Brien CA, Van Horn DJ, Wolin SL: **A misfolded form of 5S rRNA is complexed with the Ro and La autoantigens**. *Rna* 1996, **2**(8):769-784.

36. Labbe JC, Hekimi S, Rokeach LA: **The levels of the RoRNP-associated Y RNA are dependent upon the presence of ROP-1, the *Caenorhabditis elegans* Ro60 protein**. *Genetics* 1999, **151**(1):143-150.

37. Wolin SL, Belair C, Boccitto M, Chen X, Sim S, Taylor DW, Wang HW: **Non-coding Y RNAs as tethers and gates: Insights from bacteria**. *RNA biology* 2013, **10**(10).

38. Benson NR, Wong RM, McClelland M: **Analysis of the SOS response in *Salmonella enterica* serovar typhimurium using RNA fingerprinting by arbitrarily primed PCR**. *J Bacteriol* 2000, **182**(12):3490-3497.

39. Leonhartsberger S, Huber A, Lottspeich F, Bock A: **The *hydH/G* Genes from *Escherichia coli* code for a zinc and lead responsive two-component regulatory system**. *Journal of molecular biology* 2001, **307**(1):93-105.

40. Dixon R, Kahn D: **Genetic regulation of biological nitrogen fixation**. *Nature reviews Microbiology* 2004, **2**(8):621-631.

41. Raghavan V, Groisman EA: **Orphan and hybrid two-component system proteins in health and disease**. *Current opinion in microbiology* 2010, **13**(2):226-231.

42. Delgado A, Ramos JL: **Genetic evidence for activation of the positive transcriptional regulator Xy1R, a member of the NtrC family of regulators, by effector binding**. *The Journal of biological chemistry* 1994, **269**(11):8059-8062.

43. Delgado A, Salto R, Marques S, Ramos JL: **Single amino acids changes in the signal receptor domain of XylR resulted in mutants that stimulate transcription in the absence of effectors**. *The Journal of biological chemistry* 1995, **270**(10):5144-5150.

44. Perez-Martin J, De Lorenzo V: **The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor**. *Proceedings of the National Academy of Sciences of the United States of America* 1995, **92**(20):9392-9396.

45. D'Autreaux B, Tucker N, Spiro S, Dixon R: **Characterization of the nitric oxide-reactive transcriptional activator NorR**. *Methods in enzymology* 2008, **437**:235-251.

46. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H: **Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection**. *Molecular systems biology* 2006, **2**:2006 0008.

47. Jacobs MA, Alwood A, Thaipisuttikul I, Spencer D, Haugen E, Ernst S, Will O, Kaul R, Raymond C, Levy R *et al*: **Comprehensive transposon mutant library of *Pseudomonas aeruginosa***. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(24):14339-14344.