CHARACTERIZATION OF NOVEL BINDING SITES AND REGULATORY ACTIVITIES FOR THE TRANSCRIPTION SIGMA FACTOR, RpoN, IN *SALMONELLA*

by

ASHLEY CHRISTINE BONO

(Under the Direction of ANNA KARLS)

ABSTRACT

During bacterial transcription, core RNA polymerase ($\alpha\alpha\beta\beta'\omega$) transiently interacts with a σ factor to identify a promoter and then isomerizes into transcriptionally active open complex. Sigma54 (σ^{54} or RpoN) is the lone member of an alternative sigma factor family that is highly conserved across diverse bacterial species. Sigma54-RNA polymerase holoenzyme ($E\sigma^{54}$) is unique among bacterial holoenzymes, and similar to Pol II of eukaryotes, in its requirement for physical interaction with a protein activator and hydrolysis of ATP for isomerization into open complex. The focus of the original research presented in this dissertation is to define the global σ^{54} regulon in the model organism *Salmonella enterica* subspecies *enterica* serovar Typhimurium 14028s, and characterize the regulatory roles of σ^{54}-dependent promoters and σ^{54} binding sites. Earlier work suggested that *S.* Typhimurium has a robust σ^{54}-dependent regulon of diverse genes regulated by at least 13 bacterial enhancer binding proteins (bEBPs) that are each responsive to different environmental signals.

To promote open complex formation by $E\sigma^{54}$ and stimulate expression of all σ^{54}-dependent genes, a previously vetted, constitutively-active, promiscuous bEBP, DctD250 was expressed in wild-type and Δ*rpoN* strains. Transcriptome profiling and identification of σ^{54} DNA binding sites from immunoprecipitated σ^{54}-chromosomal DNA (ChIP) were performed on tiling microarrays (chip). Three novel σ^{54}-dependent transcripts, in addition to the previously predicted/known σ^{54}-dependent transcripts, and 184 σ^{54} intergenic and intragenic DNA binding sites were defined. Thirteen $E\sigma^{54}/\sigma^{54}$ binding sites, identified by ChIP-chip, were confirmed using electrophoretic mobility shift assays, revealing two novel sites that bind σ^{54} alone, as well as $E\sigma^{54}$, and provide insight into DNA sequence features that impact σ^{54} binding. Selected novel σ^{54}-DNA binding sites and associated transcripts were further characterized by quantitative reverse-transcriptase polymerase chain reaction. From the perspective of recently published transcription factor regulon studies in *S.* Typhimurium and *Escherichia coli,* these results suggest regulatory mechanisms for σ^{54} binding site that integrate the σ^{54} regulon into the complex network of transcription factor regulons controlling the cellular response to the myriad of stressors that it encounters during the infection process.

INDEX WORDS: *Salmonella enterica* subspecies *enterica* serovar Typhimurium, transcription, sigma54, *rpoN*, regulon

CHARACTERIZATION OF NOVEL BINDING SITES AND REGULATORY ACTIVITIES FOR THE TRANSCRIPTION SIGMA FACTOR, RpoN, IN *SALMONELLA*

by

ASHLEY CHRISTINE BONO

BS Microbiology, University of Massachusetts Amherst, 2009

BS Biochemistry and Molecular Biology, University of Massachusetts Amherst, 2009

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

2015

CHARACTERIZATION OF NOVEL BINDING SITES AND REGULATORY ACTIVITIES FOR THE TRANSCRIPTION SIGMA FACTOR, RpoN, IN *SALMONELLA*

by

ASHLEY CHRISTINE BONO

Major Professor:	Anna Karls
Committee:	Timothy Hoover
	Lawrence Shimkets
	Claiborne Glover
	David J. Garfinkel

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2015

DEDICATION

I dedicate this to my father, Vincent P. Bono, for helping inspire an early love of science and nurturing it as it grew.

ACKNOWLEDGEMENTS

I would like to thank my wonderful mentor, Dr. Anna Karls, for the enormous amount of work she has done to improve my skills. Her dedication and support of her students is endless. I have always appreciated her open door and willingness to help.

I would like to thank my past and present lab mates, Dr. Dave Samuels and Christy Hartman. Your collaboration and help with problem solving has been immensely helpful. I would also like to thank the Dr. Tim Hoover and his lab for their helpful insight in many years of collaborative lab meetings.

I would like to thank the rest of my committee, Dr. Lawrence Shimkets, Dr. Claiborne Glover, and Dr. David J. Garfinkel for their recommendations and support. I really appreciate the support and assistance I have received from the Stabb lab, the Maier lab, and the Kushner lab. This project would not have been possible without the time, expertise, and assistance of Jan Mrázek.

I would like to thank my Yoshukai karate family, particularly Sensei Sherrie Hines, Sensei Erik Hofmeister, and my cohort, Conner Magill and Martha Vang. Karate has immensely helped me maintain balance in my life and my cohort continued to push me to be better than I was.

I would like to thank my family for their support and assistance. It hasn't been easy being away, but they helped me make it through the hard times.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF TABLES viii

LIST OF FIGURES ix

CHAPTER

1 Introduction 1

2 Review of Literature 4

Sigma54 Structure and DNA binding 6

Sigma54-dependent Transcription Initiation 10

Sigma54 Regulon Characterization 19

Interconnected Regulons: A Cooperative View of Regulation 22

3 Materials and Methods 37

Oligonucleotides, Enzymes, Media and Chemicals 37

Bacterial Strains and Plasmids 37

ChIP-chip and Transcriptome Profiling on Tiling Microarrays 39

Analyses of ChIP-chip and Microarray data 42

qRT-PCR with Selected RNA Samples from DNA Microarray Analysis 44

qRT-PCR Assays for Expression of *hypO* and the Intergenic Region between *hypO* and *yghW* 45

Purification of σ^{54} 46

Electrophoretic Mobility Shift Assay (EMSA).....47

4 Results.....59

Most Eσ^{54} Binding Sites in the *S.* Typhimurium Genome Are Positioned to Regulate Gene Expression through Roles other than the Promoter for Gene Transcription.....59

Sigma54 Both Positively and Negatively Regulates Transcription of Genes and Non-coding RNAs Involved in Diverse Cellular Processes.....62

Eσ^{54} DNA Binding Sites in *S.* Typhimurium 14028s Are Highly Conserved in Other *S.* enterica Serovars, but a Small Fraction of the Sites Are Fully Conserved in Other Enterobacteriaceae.....67

Purified σ^{54} Binds in the Absence of Core RNAP to Two Novel Eσ^{54} DNA Binding Sites.....69

5 Discussion.....105

Concluding Remarks.....112

REFERENCES.....114

LIST OF TABLES

Page

Table 2.1: Influence of Point Mutations on Transcription Activity from a Synthesized Consensus Promoter.....30

Table 2.2: Comparison of Effect of Site-directed Mutagenesis on Activation by NtrC and NifA on the *Klebsiella pneumonia nifH* and *nifL* Promoters.....32

Table 2.3: Summary of Characterized σ^{54} Regulons of Various Bacterial Species.....34

Table 3.1: Oligonucleotides.....50

Table 4.1: ChIP-chip results for Eσ^{54} DNA binding sites, grouped by position and chromosomal context.....73

Table 4.2: σ^{54}-regulated transcriptome in *S.* Typhimurium expressing DctD250. 86

Table 4.3: Summary of σ^{54}-dependent transcriptome in *S.* Typhimurium expressing DctD250.....92

Table 4.4: Summary of EMSA results for Eσ^{54} and σ^{54} binding to selected sites…….95

LIST OF FIGURES

Page

Figure 2.1: Activator-mediated σ^{54}-dependent transcription initiation.....28

Figure 4.1: Illustrations of contextual positions of $E\sigma^{54}$ DNA binding sites in the *S.* Typhimurium genome and examples of potential regulatory roles with corresponding plot of microarray data.....97

Figure 4.2: *In vitro* assays of $E\sigma^{54}$ and σ^{54} binding to DNA sequences identified in ChIP-chip analysis of $E\sigma^{54}$ genomic binding sites 99

Figure 4.3: Multiple sequence alignments for $E\sigma^{54}$ binding sites from ChIP-chip analysis and $E\sigma^{54}$ binding sites characterized by EMSA.....101

Figure 4.4: Alignment of σ^{54}-binding **s**ites and comparison to a similar $E\sigma^{54}$-binding sequence that does not bind σ^{54} in the absence of RNAP.....103

CHAPTER 1

INTRODUCTION

Two families of sigma factors determine promoter specificity of RNA polymerase in bacteria, the σ^{70} family and the σ^{54} family (1,2). The holoenzyme of RNA polymerase containing σ^{54} ($E\sigma^{54}$) is highly regulated; a closed complex of $E\sigma^{54}$ with a promoter requires ATP hydrolysis by an associated bacterial enhancer-binding protein (bEBP) for the transition to open complex and transcription initiation. The genome of *Salmonella enterica subspecies enterica serovar* Typhimurium encodes 13 known or predicted bEBPs (3,4), each responding to a unique signal that leads to multimerization, binding of ATP, and interaction with the DNA enhancer sequence and Eσ54 at the promoter (5). The focus of the original research presented in this dissertation is to define the regulon of σ^{54} in the model organism *S.*Typhimurium 14028s, and characterize the regulatory roles of identified σ^{54}-dependent promoters and $E\sigma^{54}$ binding sites. *S.* Typhimurium is a significant human pathogen causing approximately 1,000,000 illnesses, 19,000 hospitalizations, and nearly 400 deaths in the United States annually (6). S. Typhimurium is transmitted by the fecal-oral route, causing gastrointestinal disease and sepsis in humans. *S.* Typhimurium also infects a variety of host species that are important reservoirs for zoonotic transmission: chickens, pigs, and cattle. In a recent study of the *S.* Typhimurium genes that are important in infection and persistence in zoonotic transmission reservoirs, pools of random-insertion mutants of *S.* Typhimurium were used

to infect chickens, pigs, and bovine. Transposon-directed insertion-site sequencing of the recovered mutant strains from intestine and spleen showed that the *rpoN* gene product is important for survival in these animal systems; however, the data from this study did not specifically implicate any of the known σ^{54}-dependent coding sequences as being responsible for *S.* Typhimurium's broad host range (7).

Earlier studies in our lab characterizing the σ^{54} regulon of *S.* Typhimurium were conducted in the LT2 strain, which has been widely studied as the model organism for *Salmonella* for decades. LT2 has the same robust σ^{54}-dependent regulon under the control of 13 different bEBPs as seen in the more virulent strains (3,4), such as 14028s or SL1344; but the disadvantage to using LT2 for further characterization of the *S.* Typhimurium σ^{54} regulon is a point mutation in the start codon of the *rpoS* gene of LT2 that leads to very low production of RpoS and results in an avirulent strain (8,9). In characterizing the $E\sigma^{54}$ genomic binding sites and the σ^{54}-dependent global transcriptome it is important to ensure that the rest of the σ factor pool is intact because changes in σ^{54} competition for RNAP core may influence occupancy of low affinity promoters (10).

Studies of regulons have classically focused on a single regulatory molecule and its responsive transcripts. Recent studies have been able to integrate regulons of various activators and observe their overlap and competition; i.e., direct and indirect connections between regulons can have a significant impact on the transcriptional landscape (10,11). Sigma factor competition, occlusion of sites required for activation, and production of regulatory noncoding RNAs and proteins that alter availability or affinity of σ factors for interaction with core RNA polymerase alter dynamic expression in the cell (12-14). Further characterization of the σ^{54} regulon and its connections σ^{70} family regulons will

assist in continuing to paint this complicated picture of the large, interconnected regulatory systems that exist within the cells.

The dissertation research described in the following chapters (Materials and Methods, Results, Conclusions) refines the definition of the σ^{54} regulon and global binding sites in *S.* Typhimurium 14028s, reveals structural/sequence features of $E\sigma^{54}$ DNA binding sites that influence independent binding of σ^{54} in the absence of RNA polymerase core (RNAP), and addresses the roles of $E\sigma^{54}$ and σ^{54} binding sites in regulation of σ^{70}-type regulons.

CHAPTER 2

REVIEW OF LITERATURE

Fundamental processes of life: DNA Replication, Transcription, and Translation. The central dogma of biology summarizes the flow of biological information (15). One of the core tenants is production of RNA from DNA through the process of transcription. In bacterial systems a single type of RNA polymerase transcribes the information encoded by DNA into RNA. The RNA polymerase core complex (RNAP) comprises four different subunits: two identical α subunits and the β, β', and ω subunits. RNAP has the polymerization activity, but it cannot selectively identify promoter sequences or open double-stranded DNA to form the transcription bubble. An additional, variable subunit, called the σ (sigma) factor, transiently interacts with RNAP to form the holoenzyme (Eσ) which is capable of recognizing specific promoter sequences and initiating DNA strand opening for transcription (16). Sigma factors help to provide the flexibility required for the single RNA polymerase to transcribe from many different sets of promoters under changing growth conditions and different developmental/growth phases. After transcription initiation, the σ factor does not need to be maintained in complex with RNA polymerase core for elongation of the RNA molecule.

Nearly all bacteria except *Mycoplasma* encode multiple σ factors, ranging from 2 to over 100 (17). Bacterial cells have a primary housekeeping σ factor that recognizes promoters for most essential genes; the housekeeping σ factor for *Escherichia coli*, σ^{70},

was the first bacterial σ factor to be identified. Alternative σ factors direct RNAP to transcribe from a different set of promoters for genes that are needed in response to stress or developmental signals. Comparison of amino acid sequences and secondary/tertiary structures for housekeeping and alternative σ factors from diverse phyla revealed that there are two families of σ factors: the σ^{70} family, which comprises four groups of σ factors with varying levels of amino acid sequence similarity to σ^{70} (17); and the σ^{54} family, which has only one member, σ^{54}, and shares no sequence similarity with members of the σ^{70} family (1). Sigma54 (also called RpoN or GlnF) was first identified while characterizing the regulatory components associated with nitrogen metabolism in *Salmonella* (18). It has recently been shown that a σ^{54} ortholog is found in the majority of sequenced bacteria and in most phyla, but not all bacteria encode a σ^{54} subunit (19), and even a single species can show variability regarding the presence of σ^{54} among various isolates (20).

Although σ^{54} and the σ^{70}-type factors associate with the same RNAP to form holoenzymes, there is a significant difference in the activities of $E\sigma^{54}$ and $E\sigma^{70}$; $E\sigma^{70}$ is able to spontaneously melt DNA at the promoter to initiate transcription, but $E\sigma^{54}$ rests in an autorepressive state at the promoter until ATP hydrolysis by an associated activator remodels the structure of $E\sigma^{54}$ to promote open complex formation (Fig. 2.1). Like the eukaryotic transcription factors and enhancer sequences that control Pol II activity, activators of $E\sigma^{54}$ bind enhancer sequences that are distant from the promoter, and DNA looping allows the activator to contact $E\sigma^{54}$; thus, the $E\sigma^{54}$ activators are called bacterial enhancer binding proteins (bEBPs) (Fig. 2.1) (1,21,22). The ability of $E\sigma^{54}$ to interact

stably with promoter sequences without initiating transcription could play an important role in the rapid response of the σ^{54} regulon to environmental signals.

Sigma54 Structure and DNA Binding

There is a surprisingly strong conservation of amino acid sequence and structure for σ^{54} across bacterial families and classes, which is exemplified by reports of σ^{54} from a bacterium in one family/class binding or activating transcription from promoters in a bacterium from another family/class (18,23,24). The typical σ^{54} structure can be divided into three functional regions (Fig 2.1). Region I is a highly conserved sequence at the N-terminus and is responsible for holding the holoenzyme in closed complex, interfacing with the bEBP, and melting the DNA (25). Region II varies in length for σ^{54} homologs from diverse bacteria and functions as a flexible linker. Region III, the largest of the three, contains the amino acid sequence that interfaces with core RNAP and a helix-turn-helix motif that directs binding to the DNA promoter sequence (1,26,27).

The difference in structure between the σ^{70}-family members and σ^{54} is reflected in the conserved promoter elements that they recognize. The typical promoter at which $E\sigma^{70}$ initiates transcription has conserved recognition elements centered at -35 bp (TTGACA) and -10 bp (TATAAT) relative to the transcription start site (TSS), whereas the most conserved elements of a promoter that is recognized by $E\sigma^{54}$ are typically located at -24 bp (GG) and -12 bp (GC) relative to the transcription start site (Fig. 2.1) (2,23). Early studies of the σ^{54}-dependent promoters, such as the *Caulobacter crescentus* *flbG* and *flaN* promoters, indicated that that the sequence and spacing of the -24 (GG) and -12 (GC) elements are critical for promoter activity (24). Site-directed mutagenesis of a consensus promoter built from 16 σ^{54}-dependent promoters further demonstrated the

importance of the -12 region. The consensus, subsequent mutations, and their promoter activity are shown in Table 2.1. In this study, Wang and Gralla (28) observed that sequence modifications at positions -14 to -12 could lead to reduced promoter activity, while some modifications in the -17 to -15 region could somewhat increase detected transcript (28).

In addition to DNA sequence, some σ^{54}-dependent promoters have structural elements that are required for promoter function; for example, the *flbG* and *flaN* promoters require that the DNA template be supercoiled for *in vitro* transcription, while the *glnA* promoter gave similar activity from both supercoiled and linear substrates (24). The role for DNA supercoiling in the activity of the *flbG* and *flaN* promoters may be to facilitate DNA looping to bring the enhancer-bound bEBP into contact with $E\sigma^{54}$ at the promoter, or to provide torsional energy for stabilizing the DNA distortion associated with the -12 element in closed complex or opening the DNA duplex for formation of open complex. Identified transcription start sites indicate that there is some flexibility (of 1 to 2 bp) in the distance between the -24/-12 promoter elements and the transcription start site, but the correct spacing between the -24 and -12 components is essential for promoter function (29-31).

Although one of the primary functions of σ factors is to direct RNAP to bind specific promoter sequences, the σ^{70}-type subunits typically do not bind DNA independent of RNAP because the DNA-binding domain is inaccessible until the subunit undergoes a structural change upon interacting with RNAP (17). The only σ^{70}-type factor that has been demonstrated to bind DNA independently is σ^{D}, which can isomerize spontaneously to expose the DNA-binding domain (17,32,33). Sigma54, however, can

bind specifically to a subset of known σ^{54}-dependent promoters in its native state (23) [see Chapter 2]. For promoters that are recognized by both σ^{54} and $E\sigma^{54}$, the affinity of $E\sigma^{54}$ for the promoter on a double-stranded DNA probe is about 100-fold higher than the affinity of σ^{54} for the same probe (23,34). The poor specific DNA binding activity of σ^{54} in the absence of RNA polymerase suggests that the role of σ^{54} binding to promoter sequences is not to direct RNAP to the promoter for assembling holoenzyme. Nonetheless, binding of σ^{54} in the absence of RNAP is sequence-dependent and may play some regulatory role at some σ^{54}-dependent promoters or DNA binding sites that are not functional promoters. Buck and Cannon (23) observed that simply changing 3 bases in the -15 region of the *nifH* promoter native to *Klebsiella pneumoniae,* to more closely resemble the same promoter native to *Sinorhizobium meliloti,* allowed it to support σ^{54} binding while the wild-type sequence would not (23). This led to the proposal that the -17 to -15 thymine residues, which contribute to the major groove contacts made by holoenzyme, contribute significantly toward creating a structure capable of being bound by σ^{54}. Electrophoretic mobility shift assays (EMSA) with the -20 to -12 region of *nifH049* did not demonstrate binding by σ^{54}, indicating additional requirements for binding outside of this region (34). Footprints of holoenzyme and σ^{54} on *nifH049* were similar to those on *S. meliloti nifH* (35).

Extensive footprint analyses of binding at the *S. meliloti nifH* promoter by σ^{54} and $E\sigma^{54}$ have provided substantial information about the DNA elements that contribute to binding. DNase I and *ortho*-phenanthroline (*o*-CuP) evaluation of the protected DNA site showed that σ^{54} protects sequence from -34 to -5, while the $E\sigma^{54}$ footprint is only slightly larger, protecting to the -1 position (30,34). Contacts with DNA in the -24 region

of the promoter made by $E\sigma^{54}$ or σ^{54} are important in promoter recognition (36). *o*-CuP also indicated a distortion at the -5 position when the promoter was bound by σ^{54}, but not $E\sigma^{54}$. Probing for distortion in the double helix with potassium permanganate and diethylpyrocarbonate (DEPC) revealed that sequence 3' to the -12 promoter element is distorted in the closed complex with $E\sigma^{54}$ but not in the complex with σ^{54} alone; contacts made by $E\sigma^{54}$ with the bases on the bottom strand at -12 and -11 positions stabilize closed complex (30). Cryo-electron microscopy (EM) reconstructions revealed densities associated with Region I are in close proximity to the -12 position, as well as physically block admittance of DNA into the RNA polymerase active site (37). Using poorly hydrolysable analogs of ATP with Cryo-EM and DNA footprinting, the changes in structure of $E\sigma^{54}$ from closed complex to open complex were shown to shift the interactions of σ^{54} from the bottom strand at positions -11 and -12 to the top strand of the DNA distortion and extend the reach of σ^{54} domains to +1; in addition, the catalytic site of RNAP is moved into the transcription bubble (5,37,38). The absence of the distortion at positions -12 and -11 when bound by σ^{54} alone suggests that σ^{54} adopts a different conformation when incorporated into holoenzyme (25,30,36).

In vitro assessment of transcription from the *glnA* promoter derived from *Salmonella enterica* determined that the bEBP, NtrC, and the presence of ATP were required for isomerization from closed to open complex (39). The conformational change associated with open complex formation modified the binding footprint of $E\sigma^{54}$. At the *glnA* promoter, the observed closed complex footprint extended from -31 to -5, but in open complex it extended from -35 to +21 on the top strand (39). The closed complex

footprint was altered for the holoenzyme formed with RNAP containing truncated α subunits, but this variant was still capable of initiating transcription (40).

Sigma54-dependent Transcription Initiation

There is variation in the specifics of activation of σ^{54}-dependent promoters, depending on the specific bEBP and DNA structure of the promoter and enhancer sequences, but there are several general characteristics shared by many. The steps for activation are illustrated in Figure 2.1.

Step 1: Holoenzyme binds to the promoter sequence. Competition between σ factors for core RNAP and alteration of σ factor affinity for core by transcription factors influences which transcripts will be made in the cell at any moment in time (41,42). *In vitro* assessment of holoenzyme formation indicated substantial differences in the affinity of the seven *E. coli* σ factors for core RNAP; housekeeping σ^{70} has the highest affinity with a Kd of ~0.26 nM and σ^{S} has the lowest with a Kd of ~4.3 nM (43). Sigma54 affinity for core RNAP is close to that of σ^{70} and is the highest of the alternative σ factors, allowing it to compete strongly for core binding (27,43,44). Using gene expression dynamics as a proxy for holoenzyme formation, Hicks and Grossman (45) demonstrated that artificially altering the level of a single sigma factor in *Bacillus subtilis* not only changed expression of genes whose transcription is controlled by that sigma factor, but also changed expression of genes controlled by different sigma factors, reflecting a change in σ factor equilibrium. Mechanisms by which the σ factor equilibrium is altered under various growth conditions or developmental phases includes expression of anti-sigma factors that sequester or inactivate targeted σ factors (46).

The master regulator of stringent response, (p)ppGpp, has been shown to modulate promoter activity and competition for core RNA polymerase (14,47). Sigma54-dependent transcription changes in accordance with ppGpp levels. Observations in both *E. coli* and *Pseudomonas putida* note constant levels of σ^{54} from exponential to stationary phase, but an increase in σ^{54}-dependent transcription in accordance with increasing levels of ppGpp during the transition to stationary phase. ppGpp, in conjunction with DksA, enhances σ^{54}-dependent activity in stationary phase, but has little impact on transcription during exponential phase (48,49). DksA directly interacts with the secondary channel of RNA polymerase with or without ppGpp and destabilizes nascent RNA molecules, further promoting activity from secondary sigma factors (50). An additional protein product, YhbH, also appears to modulate σ^{54} activity. Identified as a factor associated with ribosomes in stationary phase, it appears to modulate σ^{54}-dependent gene expression in *Erwinia amylovora* in that deletion of *yhbH* presents the same avirulent phenotype as Δ*rpoN* (51,52).

Sigma54 may have an advantage in this competitive environment because it can sequester RNAP from the pool of σ factors since $E\sigma^{54}$ forms a stable closed complex at σ^{54}-dependent promoters; these closed complexes remain silent until the appropriate conditions stimulate a bEBP to activate transcription from its correlate promoters. Cross-linking studies demonstrated that σ^{54} and σ^{70} interact with the same regions of core RNAP, demonstrating that different structures and sequences can produce similar function, but the organization of the sigma factor components is not the same between the two families (53). Once the holoenzyme is formed, the binding of $E\sigma^{54}$ to a cognate promoter generally does not require assistance by transcription factors, but once it is

bound it cannot spontaneously isomerize to open complex without an associated bEBP and hydrolysis of ATP (35,54).

For decades, the mechanism by which RNA polymerase initially binds a specific promoter sequence within the vast surrounding non-promoter sequence on a DNA chromosome has been hypothesized to involve facilitated diffusion, whereby the holoenzyme binds nonspecifically to DNA and then slides in one dimension until it reaches a promoter (55,56). A recent study utilizing single-molecule technology to directly compare the rates of initial binding of $E\sigma^{54}$ to the *glnA* promoter or non-promoter DNA demonstrated that facilitated diffusion is not the mechanism by which $E\sigma^{54}$ locates and binds promoter sequences; the authors propose that a structural element of the promoter DNA may allow targeted binding by $E\sigma^{54}$ from solution (57). The proposed role for structural elements in the recruitment of $E\sigma^{54}$ to σ^{54}-dependent promoters is consistent with an earlier study that demonstrated a structural dependence for recruitment of $E\sigma^{54}$ to the Pu promoter in *Pseudomonas putida;* IHF-induced DNA bending allows the alpha subunits of $E\sigma^{54}$ to contact an UP-like element, thereby positioning $E\sigma^{54}$ for enhanced closed complex formation with the -24, -12 promoter elements (58). However, the role for IHF-induced or sequence-directed DNA bending in enhancing or reducing transcription from σ^{54}-dependent promoters has primarily been attributed to facilitating DNA looping to bring the enhancer-bound bEBP into contact with promoter-bound $E\sigma^{54}$ (10,59,60).

Step 2: Activation of the bEBP. bEBPs often, but not always, consist of three functional domains: Regulatory, AAA+ (ATPases associated with various cellular activities), and Helix-Turn-Helix DNA binding (Fig. 2.1). The structure of the AAA+

domain is essential and highly conserved among all bEBPs, including the Walker A and Walker B motifs associated with ATP binding and hydrolysis and the GAFTGA sequence motif for interaction with sigma54; and the DNA binding domain exhibits conservation of the helix-turn-helix binding motif in all bEBPs, with the exception of the few bEBPs that do not bind DNA enhancer sequences (reviewed in (5)).

The regulatory domain is variable in its structure and mechanism for controlling the on/off switch for bEBP activity. The mechanism by which bEBP-dependent activation is repressed varies, but there are generally three targets of repression: oligomerization of the bEBP, ATP hydrolysis, or interaction with σ^{54}. The regulatory domain may respond to phosphorylation state, ligand binding, or specific protein-protein interactions (reviewed in (5,61)). Phosphorylation within a two-component system, such as NtrBC, can serve to activate an inactive bEBP or relieve autorepressive activity. Histidine kinase NtrB phosphorylates the response regulator domain of NtrC under low nitrogen conditions. Phosphorylation alters the configuration of the inactive NtrC dimers allowing them to polymerize to hexamers or heptamers and form active AAA+ domains capable of hydrolyzing ATP. Other bEBPs can be activated directly by ligand binding typically mediated by a V4R (vinyl 4 reductase) or GAF (cyclic-GMP-specific phosphodiesterases, *Anabaena* adenylate cyclases, and FhlA) domain. bEBPs with a V4R domain, like XylR, are typically activated through binding hydrocarbons (5,61). XylR binds toluene and activates the Pu promoter to express genes involved in toluene catabolism in *P. putida.* GAF domains bind other small molecules. The FhlA GAF domain binds intracellular formate to activate transcription of the formate hydrogen lyase genes. NorR is another GAF-containing bEBP, but is unique in that it appears to be

repressed by a conformation that inhibits contact between the crucial σ^{54}-interacting motif, GAFTGA, and σ^{54}. The third mechanism of bEBP activation is to relieve repressive interactions with an anti-activator protein, whose binding to the bEBP is controlled by environmental factors. For example, PspF, which is a constitutively active bEBP controlling genes that respond to phage shock, is repressed *in trans* by interactions with PspA. In part, PspA inhibits PspF activation by interfering with the Walker A motif required for proper ATP hydrolysis (5,61,62). bEBPs responsive to diverse signals and their unique enhancer sequences are required for activation of σ^{54}-dependent promoters under most circumstances (24,54,63,64).

There is also evidence indicating that the promoter sequence offers some specificity with regard to which bEBP may be capable of inducing open complex formation. The *Klebisella pneumoniae nifH* promoter is activated by NifA in an enhancer-dependent fashion and by NtrC independent of the enhancer sequence (65,66). Site-directed mutagenesis and evaluation of promoter activity by *lacZ* fusion, summarized in Table 2.2, demonstrate that NifA and NtrC responded differently to changes in promoter sequence. Under activating conditions, a single T to A change at -18 increased NifA-dependent expression a modest 2%, but reduced NtrC-dependent expression to only 26 to 32% of wild-type (31,65). But if this point mutation was combined with a two base deletion of sequence at -19 and -20, NtrC-dependent activation is restored to 92% of wild-type function, while NifA-dependent expression is only 4% of wild-type (31). The decrease in NifA-dependent expression may be due to reducing promoter binding by $E\sigma^{54}$ because the spacing between the -12 and -24 promoter elements is altered and as a result of changing the relative position on the face of the

DNA helix for $E\sigma^{54}$ and NifA such that torsional stress prevents the enhancer-bound NifA from contacting $E\sigma^{54}$ at the promoter. The enhancer-independent activation by NtrC would not be affected by the change in enhancer position relative to the promoter. In combination with the T-18A mutation, an additional change from G to A at the -20 position further decreased NtrC-dependent activation from 32 to 23% of wild-type, but this single sequence change decreased NifA-dependent activation from 102 to 49% of wild-type expression (31). Even changes to the -12/-24 recognition sequences produced differing activation responses by NtrC and NifA. A nucleotide change from C to A in the -12 element reduces NtrC-dependent expression to 15% of wild-type under activating conditions, while NifA-dependent expression increased nearly 20% (65). Not all changes resulted in differing NifA/NtrC responses. Deletions not placing the upstream sequence in position to emulate the native -24 (1, 6, 7, and 12 bases between the -12/-24) and G to A at the -13 position abolished expression by both NtrC and NifA in similar ways (31,65). Interestingly the *nifL* promoter, which is also NtrC and NifA responsive, did not demonstrate the same altered responses between the two activators. Site-directed mutagenesis of the -12 and -13 position for *nifL* reduced induction by both of these activators equally (65). Promoter sequence specificity for bEBPs may add an additional mechanism of regulating and distinguishing $E\sigma^{54}$ dependent promoters.

Step 3: bEBP binds to an enhancer. Activation changes the affinity of the bEBP for itself, as well as for the enhancer DNA sequence, allowing it to polymerize into hexamers or heptamers at the enhancer (5). The enhancer is typically a dyad symmetric site located approximately 70-150 bp upstream of the promoter, but natural enhancers have also been found downstream of the target promoter and as far as 3kb from the target

promoter; in artificial systems, enhancers positioned several kilobases away from the promoter were still capable of activating transcription (54,66-69). However, this flexibility in distance from the promoter does have some restrictions. For example, enhancer activity may be affected by the distance between the enhancer and its target promoter if the added or deleted sequence alters the position of the enhancer-bound bEBP on the surface of the DNA helix relative to the promoter-bound $E\sigma^{54}$, i.e., enhancer activity can be altered by a change in distance that puts the bEBP on the incorrect face of the DNA relative to the RNA polymerase (70). The importance of the relative positions of the enhancer-bound bEBP and the promoter-bound $E\sigma^{54}$ on the DNA helix can be seen in the influence of DNA supercoiling at the *K. pneumoniae nifL* promoter on the facility of DNA looping to bring enhancer-bound bEBP in contact with the holoenzyme at the promoter (71).

Comparison of footprints for NifA (bEBP controlling nitrogen fixation genes), σ^{54}, $E\sigma^{54}$, and IHF, when added to the promoter DNA in various combinations, showed that binding of NifA and σ^{54} or $E\sigma^{54}$ were independent events at the *nifH* promoter (35). Binding of IHF, however, did appear to be influenced by the presence of holoenzyme possibly indicating some protein-protein interactions or alteration of the three dimensional structure of the DNA to better promote IHF binding (35). Later studies of other promoters found that in *E. coli* the bEBPs NtrC and PspF do not bind enhancer sequences as well in the absence of $E\sigma^{54}$ (54,59,72).

The usually low bEBP concentration in the cell affects the ability of bEBPs to activate transcription. Work in the early 1980s observed that cells maintaining multicopy plasmids carrying the *nifH* promoter region were not able to activate transcription

because of activator dilution (73). More recent evidence indicates that stoichiometry does not support simultaneous PspF-dependent activation at two promoters in *E. coli* (72). bEBPs in general are typically maintained in low concentration in the cell, thereby functioning as a limiting factor in activation, tempering the response to stimuli, and reducing the amount of spurious activation of stress response genes (48,72-74).

Step 4: Protein-protein interactions between bEBP and Eσ^{54} lead to ATP hydrolysis and open complex formation. bEBP binding to Eσ^{54} and subsequent hydrolysis of ATP serves three functions in the activation of transcription from σ^{54}-dependent promoters: 1) stimulate melting of -12 promoter element; 2) remodel Region I to relieve the blockage of the RNA polymerase active site; and 3) promote extension of the holoenzyme contacts with the DNA to properly position it for initiating transcription (5). The molecular mechanical actions that take place during activation have been extensively characterized (reviewed in (5)). A spontaneous or IHF-assisted DNA looping event brings the GAFTGA motif of the bEBP into close, unstable contact with Eσ^{54} at positions -15 to -1 of the promoter sequence (63,75-79). The GAFTGA motif is highly conserved, and intact sequence is required for most bEBPs to interact with σ^{54} within the holoenzyme-promoter complex and to remodel Region I of σ^{54} to relieve autorepression preventing DNA access to the RNAP active site (5,53,80,81). The AAA+ domain of bEBP hydrolyzes ATP to ADP and undergoes several conformational changes during the process. Hydrolysis does not occur simultaneously around the hexamer, but sequentially around the ring instead. The GAFTGA motif is responsible for the "power stroke" coupling the energy of ATP hydrolysis to conformational changes in Region I of σ^{54}, thereby remodeling the holoenzyme and melting the DNA to support transition to open

complex (25,37,53,82-84). Random mutagenesis of σ^{54} identified variants that are capable of binding to promoter sequence but not isomerizing to open complex; the altered amino acid residues in these variants pointed to Region III for playing a role in transcription initiation beyond binding to the promoter sequence (85). The initial transition to open complex melts the double-stranded DNA from the -12 to the +1 positions and allows single-stranded DNA to be loaded into RNA polymerase for initiating transcription (36,37,79). Open complexes do not require the continued presence of the bEBP to be maintained, so the bEBP is usually released from RNAP as it clears the promoter in the process of RNA polymerization (39,86).

Variations in activation of $E\sigma^{54}$-dependent transcription. There is substantial diversity in the naturally occurring bEBPs and how they function. Some of the naturally occurring bEBPs have altered domain content. CtcC of *Chlamydia trachomatis* and FlgR of *Helicobacter pylori* have both been shown to encode an N-terminal regulatory and AAA+ domain, but lack a DNA binding domain (87,88). FlgR was specifically demonstrated to activate transcription in the absence of an enhancer sequence (87). Two of the five bEBPs of *Rhodobacter sphaeroides* are believed to be constitutively active because they lack N-terminal regulatory domains and one of which is also lacking a C-terminal domain (reviewed in (89)). *E. coli* PspF lacks an N-terminal regulatory domain and is constitutively active in the absence of sequestration by PspA (90). These systems demonstrate some of the variability in σ^{54}-dependent systems.

Promiscuous activation can be accomplished with some bEBPs. High concentrations of NtrC can initiate transcription from promoter systems lacking enhancer sequences (54,71). Mutational analysis of FhlA found that some N-terminal changes in

amino acid sequence confer constitutive activity in the absence of activating conditions (91), and a PspF variant lacking the helix-turn-helix DNA binding motif has been used for activation of $E\sigma^{54}$ at isolated promoters *in vitro* (25). Our laboratory has demonstrated promiscuous and constitutive activity for a variant of DctD, a bEBP from *S. meliloti,* that lacks the N-terminal regulatory domain and the C-terminal DNA binding domain, but retains the full AAA+ domain (DctD250); ChIP-chip and transcriptome analyses using *S.* Typhimurium LT2 ORF arrays and promoter function assays demonstrated that DctD250 activates transcription from most previously characterized and predicted σ^{54}-dependent promoters that normally are specifically responsive to only one of 13 different bEBPs in *S.* Typhimurium LT2 (4).

Sigma54 Regulon Characterization

Sigma54 is found in very diverse bacteria, but it is not universally distributed (1). Unlike σ^{70}-type sigma factors, σ^{54} is almost always found in single copy in the genome. There are a few exceptions where two genes encoding σ^{54} have been identified (92). In some of these cases, divergent evolution of the duplicated genes has led to different, though sometimes overlapping, σ^{54} regulons within the same cell (93-95). Characterization of σ^{54}-dependent genes from bacteria across multiple phyla has identified a wide variety of functions under σ^{54} regulation. Each organism's regulon is unique, but some of the diverse cellular processes that require σ^{54}-dependent transcription include nitrogen cycling and fixation (18,96,97), carbon metabolism (98), zinc uptake and metabolism (99), propionate metabolism (100), arsenite metabolism (101), flagellar biogenesis (24,102,103), type III/VI secretion systems (52,104,105) and many others (1,64,106-109). An overall theme for the σ^{54} regulon has been proposed to be membrane

stress, but this may be an oversimplification of the diverse σ^{54} regulons that have been characterized for just a small subset of the bacteria that encode σ^{54} (19).

Early studies of σ^{54} regulation in a bacterium focused on single operons or single bEBP regulons, but as complete genome sequences became available, definition of global σ^{54} regulons in bacteria became a feasible goal. Many studies included *in silico* predictions for σ^{54}-dependent promoter sequences (3,110,111) (see Table 2.3), usually based on the consensus sequence determined by Barrios et al. (2). In some cases the computational analysis was restricted to intergenic sequences (see Table 2.3), but it soon became apparent that predicted σ^{54}-dependent promoter sequences are frequently found within coding sequences; *in silico* analyses of the full genome sequence of *Pirellula spp*. (112) identified more potential σ^{54}-dependent promoters within coding sequences than in intergenic regions. This report was an early glimpse of the potential σ^{54} regulatory roles for that go beyond the typical sigma factor role in directing RNAP to a promoter for expression of a downstream gene.

Fully sequenced bacterial genomes led to the development of microarrays for analysis of the level of expression of all genes under different growth conditions or in the presence/absence of various regulatory factors. Microarrays are slides or "chips" with ordered microscopic spots of DNA that correspond either to each open reading frame encoded in the genome (ORF arrays) or to short DNA sequences of ~60 bases that are designed to represent the full genome sequence with ~20 bp overlaps in the sequential DNA sequence (tiling arrays). Using microrrays to assess RNA changes under various growth conditions or with overexpression of *rpoN* has revealed modest global σ^{54} regulons ranging from 44 to 167 transcripts for *Listeria monocytogenes, Geobacter*

sulfurreducens, and *Campylobacter jejuni,* as listed in Table 2.3 (113-115). The *E. coli* global σ^{54} regulon was also analyzed by microarray, and an estimate of 70 σ^{54}-dependent promoters was determined from the microarray data; *in silico* analysis and targeted ChIP-PCR experiments was used to confirm a subset of the predicted $E\sigma^{54}$ binding sites (116). Recently, next-generation sequencing has been used to define the global σ^{54} regulon of *Vibrio chloerae*, including the transcriptome (RNA-seq) and the σ^{54}-binding sites (ChIP-seq); 68 σ^{54}-binding sites were identified, but 82 operons were positively regulated by σ^{54}, indicating the indirect control by σ^{54} on expression of some genes (104).

A limitation of all these global transcriptome studies, as noted by the investigators, is that the transcriptome was assayed under limited growth conditions, usually with over-expression of *rpoN* to facilitate detection of σ^{54}-dependent transcripts. This approach is unlikely to detect all σ^{54}-dependent transcripts since many or most of the bEBPs that control activation of transcription from the σ^{54}-dependent promoters are not active under the single growth condition used for the transcriptome analyses. This limitation is clearly demonstrated by the >3-fold increase in identified σ^{54}-dependent transcripts for wild-type *G. sulfurreducens* grown under multiple growth conditions reflecting environmental conditions that may be encountered by the bacterium (e.g., N_2 as the sole nitrogen source versus ammonium chloride and heat shock conditions versus growth at 30^{o}C) (11) as compared to the σ^{54}-dependent transcriptome determined under a single growth condition for a *G. sulfurreducens rpoN* diploid strain with IPTG-inducible overexpression of σ^{54} (113) (see Table 2.3). Without knowing all the conditions that are needed to activate bEBPs controlling σ^{54}-dependent transcription in *G. sulfurreducens* it is likely that the global σ^{54} regulon has still not been fully determined.

Previously published work from our laboratory (4) addressed this significant limitation in defining global σ^{54}-dependent regulons by utilizing the promiscuous and constitutively-active bEBP DctD250. Samuels et al. (4) utilized ORF arrays representing the *S.* Typhimurium LT2 genome for analysis of the σ^{54}-dependent transcriptome in the presence of DctD250 and for ChIP-chip to estimate the extent of genomic binding sites for σ^{54}. This study demonstrated that DctD250 was able to drive transcription from nearly all previously characterized σ^{54}-dependent promoters in *S.* Typhimurium LT2 under a single condition (exponential growth in nutrient broth). The expression of 33 operons, containing 76 genes, was shown to be positively regulated (directly or indirectly) by σ^{54} and at least 70 σ^{54} binding sites were revealed by ChIP-chip combined with *in silico* analysis. Transcriptional fusions with several of the novel σ^{54} binding sites demonstrated that many of these sites are capable of functioning as σ^{54}-dependent promoters, but the microarray analysis indicated that not all of these sites are associated with detectable transcripts (4). However, the use of ORF arrays in this proof-of-concept study prevented discovery of more than one transcript from a coding region or refined mapping of σ^{54} binding sites.

Interconnected Regulons: A Cooperative View of Regulation

In recent years, the study of regulatory systems has shifted from a focus on characterizing individual regulons to a more global perspective assessing interactions between various regulatory networks.

Transcription from Multiple Promoters. Interplay between sigma factor networks has been evident as early as 1984 when Dixon (117) and Reitzer and Magasanik (118) determined that the structural gene for glutamine synthetase (*glnA*) in *K.*

pneumoniae and *E. coli*, respectively, has tandem promoters; the gene-distal promoter is expressed by $E\sigma^{70}$ and repressed by binding of the bEBP (NtrC) to an overlapping enhancer sequence, while the gene-proximal promoter is σ^{54}-dependent and is activated by NtrC when bound to the enhancer sequences. The intermingling of σ factor regulons is sometimes fairly direct, as in the activation of σ^{S} synthesis by σ^{54} in *Borrelia burgdorferi* (119). Multiple σ regulons may intersect through the control of a single gene, such as *rpoH*, which encodes the heat shock sigma factor, σ^{32}. In *E. coli* strain MG1655, *rpoH* has five promoters, of which three are recognized by $E\sigma^{70}$, one is controlled by the stationary phase sigma factor σ^{s}, and the last promoter is σ^{54}-dependent (120); thus, the level of RpoH in the cell, and, in turn, expression of genes controlled by $E\sigma^{32}$ is modulated by how environmental conditions influence σ factor competition for RNAP core (10) and the binding of holoenzymes with associated regulatory proteins to the different promoter regions.

Several communication points between separate regulons have been observed, including through shared transcriptional regulators. cAMP receptor protein (CRP) typically is associated with activation of σ^{70}-dependent promoters, but in *S. meliloti* CRP has also been shown to prevent activation of the σ^{54}-dependent *dctA* promoter by occluding DctD binding to its enhancer DNA sequence and by interacting directly with $E\sigma^{54}$ (121). NtrC, which activates expression from an extensive array of σ^{54}-dependent promoters in *E. coli* and *S.* Typhimurium, has recently been shown to bind in close proximity to *relA*, the gene for ppGpp synthetase, in *E. coli* (122). Three transcription start sites for *relA* were mapped under N-limitation conditions by 5'RACE (Rapid Amplification of cDNA ends); one is a constitutively-expressed σ^{70}-dependent promoter

and the other two promoters are σ^{54}-dependent promoters that are activated by NtrC (122). These results suggest a mechanism by which σ^{54} assists in coupling nitrogen stress to activation of the stringent response (122).

Theoretical modeling of sigma factor competition in bacterial cells supports the notion that all genes are connected to each other and to the physiological state of the cell (10). More simply stated, transcriptional regulators do not work in isolation. There are complex and overlapping interactions of regulatory networks that reveal the balancing act of cellular responses to stimuli. An excellent example of the complexity and extent of interactions between the σ factor regulatory networks is found in *G. sulfurreducens* (11); under various growth conditions it was demonstrated that the three major alternative σ factors, σ^{N} (σ^{54}), σ^{S}, and σ^{H} regulate σ^{D}, and σ^{D} regulates expression of all three alternative σ factors. More than half of the essential genes of *G. sulfurreducens* are regulated by alternative σ factors in addition to σ^{D} and some essential genes are expressed only by $E\sigma^{N}$ (11). The extent of the σ^{N} regulon includes regulation of ~900 genes, of which 18 encode two component system genes, indicating the extended role of σ^{N} in cell signaling and regulation of transcription (11).

Non-coding RNA and Master Regulators. There are many ways to modulate expression and integrate regulons beyond classical protein transcriptional regulators. Perhaps the most rapidly expanding body of regulatory research involves trans-encoded small RNAs (sRNAs) and transcript-complementary antisense RNAs (asRNAs) that are found in all analyzed bacterial genomes (reviewed in (123,124)). RNA-seq and microarray data have shown expression of numerous sRNA and asRNA from intergenic and intragenic genomic sequences (125,126). Small RNAs, often in conjunction with the

RNA chaperone Hfq, can act on mRNAs to alter transcript stability and/or translation; in addition, sRNAs can bind and alter activity of proteins (reviewed in (123)). Antisense RNAs can also target mRNAs to regulate stability and/or translation; but in addition, asRNAs can regulate transcription initiation or cause early termination of transcription through interactions with the DNA in open complex/transcription bubbles or with elongating transcripts (reviewed in (124)). In *S.* Typhimurium the sRNAs *glmY* and *glmZ* function in a feedback cascade for GlmS (glucosamine-6-phosphate synthase); it has been demonstrated in *S.* Typhimurium that *glmY* and *glmZ* are controlled by σ^{54} and the bEBP GlrR, as well as σ^{70}, due to precisely overlapping σ^{70}- and σ^{54}-dependent promoters (127). The sRNA transcriptome of *S.* Typhimurium has been determined under various stress conditions that mimic those encountered in the infection process; no strictly σ^{54}-dependent sRNAs were reported (128). Although GlmY and GlmZ sRNAs were detected in this study, the transcription start sites are indistinguishable from the σ^{70}- and σ^{54}-dependent promoters (129,130). RNA-seq analysis of the *Campylobacter jejuni* identified potential σ^{54}-dependent noncoding RNAs, but these potential sRNAs/asRNAs have yet to be characterized (114). The small RNA OxyS is a characterized bridge point between regulatory systems. In *E. coli* expression of *oxyS* is induced under oxidative stress conditions and functions in conjunction with Hfq to reduce the intracellular levels of FhlA (a bEBP controlling genes involved in formate metabolism) and σ^{S}, and in turn, their activation targets in the cell (131-133).

Recent research has just started to reveal the diversity and elegance of the mechanisms by which $E\sigma^{54}$ regulates transcription of σ^{70}-type promoters. For example, σ^{54} regulates the σ^{S} regulon in *E. coli* by transcriptional occlusion of the promoter for *crl*,

which encodes a peptide that increases expression of the σ^S regulon by promoting the binding of σ^S to RNAP core. The σ^{70}-dependent promoter for *crl* overlaps a gene-proximal σ^{54}-dependent promoter; so under nitrogen-limiting conditions, NtrC activates dependent transcription of *crl*, thereby occluding $E\sigma^{70}$ from binding its promoter sequence. The transcript from the σ^{54}-dependent promoter is expressed at high levels, but it does not contain a ribosome binding site; thus no Crl is expressed and the σ^S regulon is efficiently down-regulated (134). Another example of a regulatory mechanism by which the process of transcription from a σ^{54}-promoter regulates expression of σ^{70}-dependent promoter is found in *P. putida* (135). In this case, the σ^{54}-dependent Po promoter expresses the *dmp* operon (allows growth on phenolic compounds) and is divergently oriented and non-overlapping with the σ^{70}-dependent Pr promoter that expresses the bEBP for the activation of transcription from Po. Johansson et al. (135) used mutational analysis of both promoter sequences and *in vitro* transcription assays to determine the mechanism by which Po positively regulates transcription from Pr; they demonstrated that it is the process of transcription initiation (open complex formation and the movement of $E\sigma^{54}$ from the promoter region) at Po that stimulates open complex formation by $E\sigma^{70}$ at the weak Pr promoter.

Transcription Interference. Another level of transcriptional regulation to consider is the configuration of the promoter region; the structure of DNA is not static and changes in conformation can influence recognition and utility of that sequence (136). Activators or repressors can serve to change the local DNA structure to either support or inhibit promoter activity. Promoters can be less accessible due to overlapping protein-binding sites or transcription elongation proceeding through that region of the DNA

(54,118,137). Once transcription has initiated, proteins bound to the downstream DNA may function as a roadblock to disrupt a fraction of the elongation complexes proceeding through depending on DNA-binding strength (138). Functional promoters oriented toward one another can lead to early termination due to steric hindrance from the physical collision (139). As the cellular regulatory activities are investigated at greater resolution, more intricate complex regulatory systems are observed.

Figure 2.1: Activator-mediated σ^{54}-dependent transcription initiation.

Fig. 2.1A illustrates the common domain composition of σ^{54} and bEBPs. Sigma54 is divided into Regions I, II, and III. Region I is primarily responsible for interfacing with bEBPs and specific interactions with the -12 element of the promoter. Region II is a flexible linker that is the most variable of the three domains across bacterial species. Region III exhibits DNA binding specificity at the -24 position, and the -12 region to a lesser extent. It is also responsible for interacting with RNA polymerase. bEBPs most often have three domains: the N-terminal domain, which is responsible for regulating activity of the bEBP; the AAA+ domain, which is required for hydrolyzing ATP and interacting with σ^{54}; and the C-terminal domain, which directs binding to the enhancer DNA.

Fig. 2.1B summarizes σ^{54}-dependent activation. ***Step 1:*** σ^{54}, which is complexed with RNAP core, directs binding to a promoter to form a stable closed complex. ***Step 2****:* An environmental stimulus activates bEBP, which is usually a dimer in solution, to polymerize into a heptamer or hexamer; polymerization creates the active binding sites for ATP. ***Step 3:*** The activated bEBP binds to the enhancer sequence. ***Step 4:*** A DNA looping event brings the activated bEBP in contact with $E\sigma^{54}$, and the AAA+ domain hydrolyzes ATP, which drives conformational changes through bEBP to σ^{54} to RNAP, resulting in open complex formation and transcription.

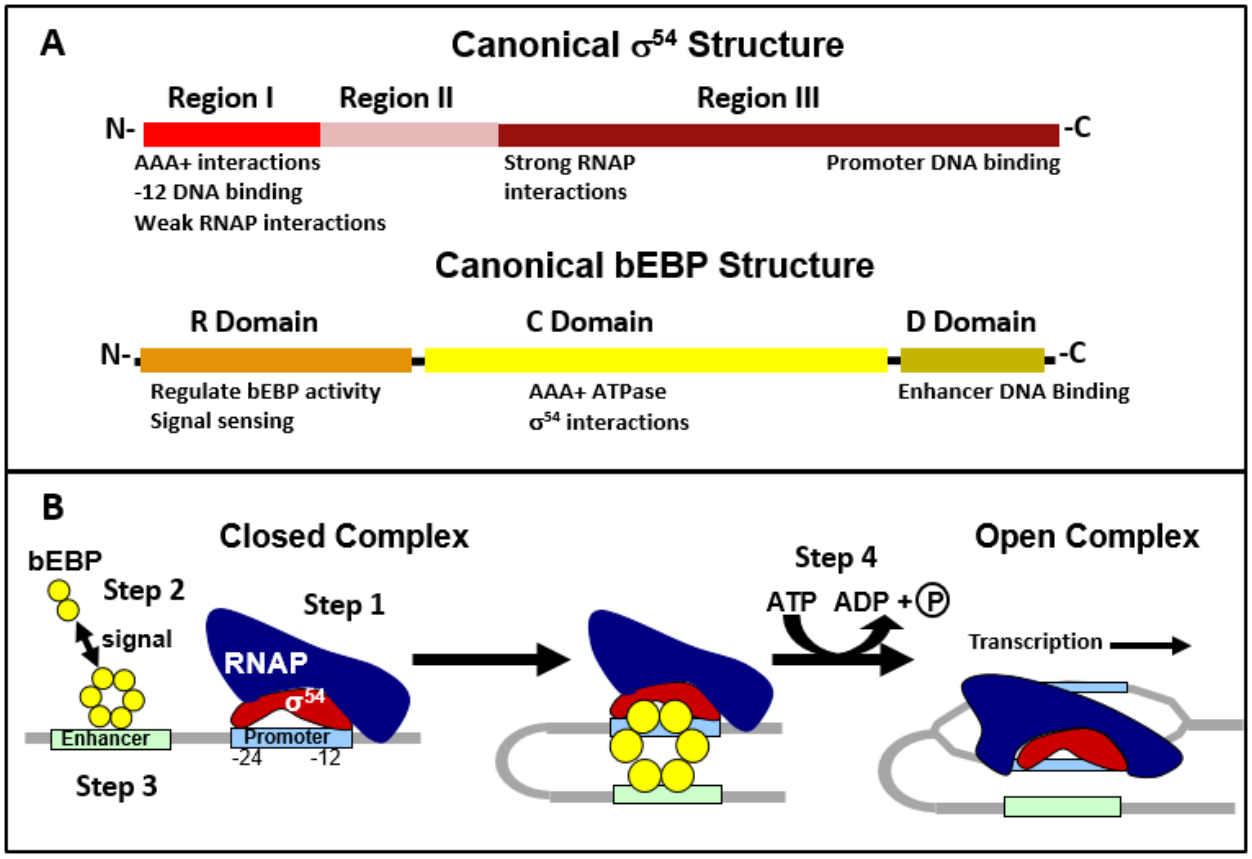
A
Canonical σ54 Structure
Region I
Region II
Region III
N-
-C
AAA+ interactions
-12 DNA binding
Weak RNAP interactions
Strong RNAP interactions
Promoter DNA binding
Canonical bEBP Structure
R Domain
C Domain
D Domain
N-
-C
Regulate bEBP activity
Signal sensing
AAA+ ATPase
σ54 interactions
Enhancer DNA Binding
B
Closed Complex
Open Complex
bEBP
Step 2
signal
Step 1
RNAP
σ54
Enhancer
Promoter
-24
-12
Step 3
Step 4
ATP
ADP +Ⓟ
Transcription

Table 2.1: Influence of Point Mutations on Transcription Activity from a Synthesized Consensus Promoter

Sequence Source	Sequence 5' → 3'[3]	Promoter Activity[1]	Ref
Consensus Assembled from 16 Promoters	CTGGCACAATATTTGCAT[2]	100%	(28)
	CTGGCACAATA**G**TTGCAT	180 ±10%	
	CTGGCACAATA**C**TTGCAT	90 ±20%	
	CTGGCACAATA**A**TTGCAT	100 ±10%	
	CTGGCACAATAT**G**TGCAT	160 ±20%	
	CTGGCACAATAT**C**TGCAT	110 ±30%	
	CTGGCACAATAT**A**TGCAT	100 ±30%	
	CTGGCACAATATT**G**GCAT	160 ±90%	
	CTGGCACAATATT**C**GCAT	180 ±30%	
	CTGGCACAATATT**A**GCAT	260 ±70%	
	CTGGCACAATATTT**C**CAT	50 ±40%	
	CTGGCACAATATTT**A**CAT	90 ±20%	
	CTGGCACAATATTT**T**CAT	60 ±40%	
	CTGGCACAATATTTG**G**AT	110 ±20%	
	CTGGCACAATATTTG**A**AT	110 ±10%	
	CTGGCACAATATTTG**T**AT	30 ±30%	
	CTGGCACAATATTTGC**G**T	40 ±30%	
	CTGGCACAATATTTGC**C**T	90 ±20%	

[1]Activity determined by ^{32}P-labeled primer extension from collected total RNA incorporating, expressed as the percent of activity from the unmodified consensus sequence.

[2]Unmodified consensus sequence.

[3]Bold and underlined nucleic acid residues indicate that there was a mutation at that position. The nucleic acid abbreviation at that position indicates the replacement residue.

Table 2.2: Comparison of Effect of Site-directed Mutagenesis on Activation by NtrC and NifA on the *Klebsiella pneumonia nifH* and *nifL* Promoters

Promoter	Sequence[1]	Relative NtrC-dependent Activation[2]	Relative NifA-dependent Activation[2]	Ref
nifH	ACGGCTGGTATGTTCCCTGCACTT	100%	100%	
	ACGGCTGGTATGT**A**CCCTGCACTT	26-33%	102%	(31,65)
	ACGGCTGGTATGT-CCCTGCACTT	1%	0.2%	(31)
	ACGGCTGGTAT--**A**CCCTGCACTT	92%	4%	
	ACGGCTGGT----**A**CCCTGCACTT	70%	0.2%	
	ACGGCTGGT-----CCCTGCACTT	1%	0.06%	
	ACGGCTGG------CCCTGCACTT	5%	0.06%	
	ACGGCTGGT----------CTT	1%	0.06%	
	ACGGCTGGTAT**A**T**A**CCCTGCACTT	23%	49%	
	ACGGCTGGTATGTTCCCTG**A**ACTT	15%	119%	(65)
	ACGGCTGGTATGTTCCCTG**T**ACTT	7%	152%	
	ACGGCTGGTATGTTCCCT**A**CACTT	4%	8%	
	ACGGCTGGTATGTTCCC**C**GCACTT	8%	143%	
	ACGGC**C**GGTATGTTCCCTGCACTT	18%	73%	
nifL	AGGGCGCACGGTTTGCA	100%	100%	
	AGGGCGCACGGTTTG**T**A	7%	7%	
	AGGGCGCACGGTTT**A**CA	7%	3%	

[1] Bold and underlined nucleic acid residues indicate that there was a mutation at that position. The nucleic acid abbreviation at that position indicates the replacement residue. Residues replaced with a dash (–) indicate a deletion without replacement at that position.

[2] Activity as determined by *lacZ* fusion to the various promoter sequences. Activation is expressed as the percent of activity from the wild-type promoter sequence.

Table 2.3: Summary of Characterized σ^{54} Regulons of Various Bacterial Species

Microbe	Transcribed regions differentially expressed in *rpoN* diploid or *ΔrpoN* strain	Sigma54 DNA binding Sites or Promoters	Detection Method	Ref
Agrobacter tumefaciens		31[1]	*In silico* Prediction	(110)
Bacillus thuringiensis	376	17[1,2]	Microarray	(140)
Borrelia burgdorferi	305	128-238[2]	Microarray	(141)
Bradyrhizobium japonicum		31[2]	Microarray	(96)
Brucella melitensis		12[1]	*In silico* Prediction	(110)
Campylobacter jejuni	44		RNA-seq	(114)
Escherichia coli	70	40 detected 70 estimated	Microarray and ChIP	(116)
Geobacter sulfurreducens	167[4]	110[1]	Microarray and *in silico* prediction	(113)
Geobacter	863	349	ChIP-Seq,	(142)

sulfurreducens			RNA-Seq	
Listeria monocytogenes	77		Microarray	(115)
Mesorhizobium loti		82[1]	*In silico* Prediction	(110)
Pirellula spp.		292[3]	*In silico* Prediction	(112)
Pseudomonas putida		46 detected, 55 estimated	*In silico* Prediction	(111)
Pseudomonas syringae	3635		Microarray	(143)
Salmonella enterica	95	70	Microarray	(4)
Salmonella enterica		36	*In silico* Prediction	(3)
Sinorhizobium meliloti		43[1]	*In silico* Prediction	(110)
Vibrio cholerae	144	68	ChIP-seq, RNA-seq	(104)
Xylella fastidiosa	448[5]	44[1]	Microarray	(144)

[1]Analysis focused on identification of intergenic promoters associated with protein-coding sequences only.

[2]As determined by bioinformatic analysis of proximal sequence.

[3]Reported for a moderately stringent cutoff.

[4]Comparison between wild-type and a strain overexpressing *rpoN*.

[5]Assessment of multiple growth conditions.

CHAPTER 3

MATERIALS AND METHODS

Oligonucleotides, Enzymes, Media and Chemicals

Oligonucleotides were synthesized by Integrated DNA Technologies (Table 3.1). Enzymes were purchased from New England Biolabs, unless otherwise indicated, and used according to manufacturer's recommendations. Cells were grown in Lennox broth (LB, Thermo-Fisher), nutrient broth (NB, Difco), InSPI2 medium (Inducing Salmonella Pathogenicity Island 2 medium, (125)), or MOPS medium with ammonium chloride as the nitrogen source and glucose as the carbon source (145,146). Media supplements where noted are as follows: 5 mM glutamine (Sigma-Aldrich), 60-120 μg/mL ampicillin (Amp, Thermo-Fisher), 50 μg/mL kanamycin (Kan, Roche Life Science), 100 μg/mL rifampicin (Rif, Thermo-Fisher), 30 mM potassium formate (formate, Sigma-Aldrich), and 250 μM Spermine NONOate (NO, Sigma-Aldrich).

Bacterial Strains and Plasmids

Salmonella enterica serovar Typhimurium (*S*. Typhimurium) strain ATCC 14028s is the wildtype strain (WT) in these studies. S. Typhimurium strain 14028s Δ*rpoN* (AB01), 14028s Δ*fhlA* (AB03), and 14028s Δ*ntrC* (CH01) were created via Lambda Red recombineering system (147) as adapted from KA Miller, et al (146). Oligonucleotides 1 and 2 (for Δ*rpoN*), 3 and 4 (for Δ*fhlA*), and 77 and 78 (for Δ*ntrC*) were used to amplify the kanamycin resistance cassette (*kanR*), which encodes *aph*

flanked by FRT (FLP recombination target) sites and a downstream ribosome binding site (rbs) with an appropriately placed methionine codon, from pKD4 (147). The amplicons were transformed by electroporation into 14028s WT competent cells containing the λ *red*-expression plasmid pKD46 (147). The Δ*rpoN*::*kanR*, Δ*fhlA*::*kanR*, and Δ*ntrC*::*kanR* mutant strains were confirmed by PCR with primer pairs, 55 and 56, 3 and 4, and 79 and 80, respectively. The substitution mutations were transduced by P22 HT *int* into a clean genetic background; the *aph* gene was removed by introducing the FLP-expression plasmid pCP20 (148). Excision of *aph* in each of the mutants, AB01 (Δ*rpoN*), AB03 (Δ*fhlA*), and CH01 (Δ*ntrC*) was confirmed by kanamycin sensitivity and sequencing (Genewiz, Inc.) PCR products from amplification of chromosomal DNA with the same primers used to confirm the insertion of the cassette. Plasmid pCP20 has a temperature-sensitive replication origin so the deletion strains were cured of the plasmid by multiple rounds of growth at 42^{o}C in the absence of selective pressure to retain the plasmid. After FLP-mediated recombination these mutant strains have the cassette scar, encoding the rbs and translation start codon, which ensures that the mutation is not polar on downstream genes.

All plasmids were maintained in *E. coli* DH5α (149). pJES937, which is a pET28a(+)-derived plasmid expressing His6-tagged RpoN (29), was introduced by heat shock into One Shot chemically competent *E. coli* BL21 DE3 (Life Technologies). Plasmids that were introduced into *S.* Typhimurium 14028s by electroporation were first passaged through the restriction^{-} modification^{+} strain *S. enterica* MS1868 (150).

ChIP-chip and Transcriptome Profiling on Tiling Microarrays

Wild type and Δ*rpoN S*. Typhimurium 14028s cells containing pPBHP192 (WT+DctD250; Δ*rpoN*+DctD250), which expresses DctD250 at a low level without induction (151), were grown overnight with aeration at 37°C from single colonies in NB-Amp^{100}. Three biological replicates were prepared for each strain. One ml of overnight culture was used to inoculate 100 ml NB-Amp^{100} and the cultures were incubated with aeration at 37°C until they reached early-to-mid exponential growth phase (OD_{600}≈0.5). Fourty ml of the culture was centrifuged to pellet cells and decanted; cell pellets were stored on dry ice for subsequent RNA isolation for microarray analysis (see below). Fifty ml of culture was prepared for chromatin immunoprecipitation (ChIP) using the ChIP Assay Kit (USB Corporation) as described in detail by Samuels et. al. (4), with the modifications noted in the following description.

Each 50 ml culture for ChIP assay was first treated with rifampicin (100μg/mL) for 10 min at 37°C with shaking to hold holoenzymes in open complex at promoters before the treatment with formaldehyde (1.1%) for 10 min. Lysate was prepared from the treated cells by French press, cleared of debris by centrifugation, incubated with Protein A bead slurry (USB Corporation) to reduce background due to nonspecific interactions, and then incubated with rabbit polyclonal antiserum against *S*. Typhimurium σ^{54} (α-σ^{54})(152) as described in (4). Antibody-bound protein-DNA complexes were incubated with Protein A beads, washed, and disrupted per the manufacturer's instruction. The cross-link between the protein and DNA was reversed, the DNA was isolated and subjected to Ligation-Mediated PCR (LM-PCR) using primers 61 and 62 to non-specifically amplify all enriched DNA fragments as described in Samuels, et al. (4). The

size range of the LM-PCR products was 0.15-1.0 kb, with most products in the 0.3-0.5 kb range. For biological replicates 1 (R01), the full range of LM-PCR products were used for labeling WT+DctD250 and Δ*rpoN* +DctD250 with the fluorescent dyes Cy3 and Cy5, respectively. For biological replicates 2 (R02) and 3 (R03), size-selection by agarose gel electrophoresis was used to isolate LM-PCR products of 0.15-0.45 kb for labeling with Cy3 and Cy5. For the labeling reactions, 500-750 ng LM-PCR products were incubated at 98^{o}C for 10 min, quick-chilled in ice-water for 2 min, and the following reagents were added: 50 mM Tris pH 6.8, 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, 0.36 ng/μL random hexamers, 0.12 mM each dATP, dGTP, dTTP, 0.06 mM dCTP, 0.4 mM Cy3- or Cy5-labeled dCTP (GE Healthcare), and 10 U Klenow fragment. The reactions were incubated overnight at 37°C. Excess hexamers and nucleotides were removed using the PCR Purification Kit (Qiagen) according to the manufacturer's instructions. The efficiency of incorporation of the Cy3- or Cy5-labeled dCTP was determined as described in (153) by measuring absorbance on a Nanodrop 2000 (Thermo Scientific) for the purified labeled DNA at 260 nm and 320 nm to determine yield, and at 550 nm and 650 nm (or 750 nm) to measure dye incorporation for Cy3 (or Cy5). The expected base-to-dye ratios for both Cy3 and Cy5 are 40-80; all labeled samples had ratios greater than 40.

RNA for microarray analysis was extracted from the frozen cell pellets, described above, using the RNA SNAP method described by Stead, *et al.* (154). In brief, cell pellets were resuspended in 4 mL of RNA extraction solution [18 mM EDTA, 0.025% SDS, 1% 2-mercaptoethanol, 95% RNA-grade formamide], divided into 4 aliquots and incubated in a 95°C water bath for 10 min. Cell debris was removed by centrifugation,

supernatant was transferred to a fresh tube, and RNA was ethanol precipitated. RNA was treated with RNase-free DNase I (Ambion) and subjected to a second ethanol precipitation for DNase I removal. RNA concentration and purity was determined from the absorbance at 230 nm, 260 nm and 280 nm on a Nanodrop 2000 Spectrophotometer. Absence of residual DNA was confirmed by PCR using primers to amplify *rpoD* sequence (primers 29 and 30). Reverse transcription of the RNA with Superscript II (Invitrogen) and labeling of cDNA with Cy3 and Cy5 were performed essentially as described in (155). Briefly, 25 µg of RNA was combined with 2.4 µg random hexamers, denatured at 70°C, and quick-chilled in ice water. The remaining components of the reverse transcription reaction were added to the denatured RNA and random hexamers to the following concentrations in a 60 µl reaction: 0.01 M DTT, 1X Superscript II first strand reaction buffer, 0.5 mM dATP, dTTP, dGTP, 0.2 mM dCTP, 0.066 mM Cy3- or Cy5-labeled dCTP (GE Healthcare), RNasin (Roche), and 800 U Superscript II. After 1 hour at 42°C, an additional 400 U of Superscript II was added and incubated another hour. Excess random hexamers and nucleotides were removed with a PCR Purification Kit (Qiagen) per manufacturer's instructions. The efficiency of incorporation of the Cy3- or Cy5-labeled dCTP was assessed as described for the ChIP samples above. In the reverse transcription reactions for the 3 biological replicates, the dyes used to label WT+DctD250 and Δ*rpoN* +DctD250 were flipped for the second biological replicates, to assess whether labeling efficiency for Cy3 or Cy5 impacts the microarray results.

The labeled ChIP-enriched DNA samples and the labeled cDNA were hybridized, as described in (156), to NimbleGen tiling microarrays for *S.* Typhimurium strain 14028s, which are described in (157). Arrays were scanned with a GenePix 4000B lasar

scanner (Molecular Devices) at 5 µm resolution. Signals were quantified by NimbleScan software v2.4 (Roche NimbleGen).

Analyses of ChIP-chip and Microarray Data

Background calculations for each array and normalization of intensity values within and between arrays was performed as described (157) using WebarrayDB (158). Differential analyses for all the paired biological replicates of WT+DctD250 and Δ*rpoN* +DctD250, utilizing normalized and averaged intensity $\log_2$ values corresponding to each probe on the arrays, were performed in WebarrayDB to obtain the M-values ($\log_2$-ratio of WT+DctD250 and Δ*rpoN*+DctD250) and *p*-values (from LIMMA analysis) (158). The M-values for ChIP-chip and microarray results are converted to signal ratios (WT+DctD250/Δ*rpoN*+DctD250) as follows: 2^M-value = signal ratio for Δ*rpoN*+DctD250/WT+DctD250; this value is inverted to give the signal ratio for WT+DctD250/Δ*rpoN*+DctD250.

Analysis of the ChIP-chip data was accomplished with original software, ChIPeak, designed for this study in collaboration with Jan Mrazek, Hao Tong and Sina Solaimanpour. ChIPeak calculates statistical significance of the peaks (*p*-value), maps peak positions in the genome, and predicts associated binding sites using the standard PSSM (Position-Specific Scoring Matrix) method as previously described for the Motif Locator software (159). The application and source codes can be downloaded at http://www.cmbl.uga.edu/downloads/programs/. The source codes are distributed under the terms of the GNU General Public License (http://www.gnu.org/licenses/gpl.html).

The input data for ChIPeak analysis was the start and end positions for each probe (from NCBI GenBank: CP001363.1) and the associated M-values from the WebarrayDB

differential analysis for the combined biological and technical replicates. A .txt file containing the sequence of 75 $E\sigma^{54}$ previously characterized promoters and highly-enriched DNA binding sites was also provided to generate the PSSM. ChIPeak determines significant peaks for the enriched ChIP DNA using a sliding window average; window size is set to the average size of the ChIP'd DNA fragments, ~500 bp, and the window is moved along genomic sequence in 10 bp steps. The background is determined using the signal values (M-values) for probes located within sliding windows that have the lowest values from *r*% of the genome, where *r*% is a conservative estimate for the percent of the genome that is devoid of binding sites. For this study *r*% was set to 80%, estimating that there would be no more that 20% of the genome with peaks from $E\sigma^{54}$ binding. The statistical significance of each potential peak is determined using the student t-test for two-sample unequal sample size and unequal variance to compare the M-values of probes within the window with the highest average signal value to the M-values of the background samples. The program then provides plots of potential peaks with various *p*-values below 10^{-10} and a *p*-value cutoff is selected based on visual analysis of the quality of the peaks with different p-values; a *p*-value of $<10^{-17}$ was selected as the cut-off for this study. The output from ChIPeak analysis includes: peak position (genome location of the maximum in the sliding window average plot); *p*-value; peak intensity (the average M-value from all probes within the window that yields the local maximum); number of probes in that window; maximum single-probe value in that window; left and right boundary of the peak region (determined by sliding window analysis of all probes flanking the peak using the *p*-value cut-off); distance to the nearest gene on the left- and right-side of the peak; orientation of the gene (coding strand + or -); gene description and

other information obtained from the annotation; and a list of up to three predicted binding sites with the highest PSSM scores within a user-specified distance from the peak, including the PSSM score, distance from the peak, orientation (strand), whether the binding site is located inside a gene or intergenic, and the sequence of the predicted binding site.

Possible binding sites identified by the program were curated individually to confirm assignment of the binding sequence associated with the peak maximum in ChIPeak. If ChIPeak identified more than one high-scoring site close to the peak maximum, the highest scoring site is reported, unless the PSSM scores differ by less than 10%, and then the site closest to the peak maximum is reported. Binding sites were evaluated for their chromosomal context by BLAST (NCBI) to determine their position relative to annotated open reading frames (ORFs).

qRT-PCR with Slected RNA Samples from DNA Microarray Analysis

RNA samples from cultures prepared for the transcriptome microarray experiment were reverse transcribed with random hexamers using Superscript II (Invitrogen) as recommended by the manufacturer. This cDNA was used for qPCR performed with SYBR Supermix (BioRad) using oligonucleotides 29-54, 69-70 as primer pairs indicated in Table 3.1. Resulting cycle thresholds were compared to a standard curve generated by amplification of serial 10-fold dilutions of genomic DNA (20 ng to 0.002 ng per PCR reaction) to determine transcript levels, which were then normalized to *rpoD* transcript levels (*rpoD* transcript levels were constant in WT and mutant strains). Primers and annealing conditions were optimized to give 90-100% PCR efficiency (slope of standard curve is between -3.6 and -3.3). Reactions were performed in triplicate. Technical

triplicates were averaged and then normalized to *rpoD* for each biological replicate. The averages of the three biological replicates were used for the reported ratios WT+DctD250 to Δ*rpoN*+DctD250, and calculation of standard deviation and *p*-values.

qRT-PCR Assays for Expression of *hypO* and the Intergenic Region between *hypO* and *yghW*

Expression from σ^{54}-dependent $P_{inter\text{-}hypO\text{-}yghW}$ and from the primary/secondary σ^{70}-dependent promoters for the *hypO_hybA-G* operon was assessed in cells grown aerobically in LB or anaerobically in LB in the presence or absence of 30 mM formate. For the anaerobic growth bottles with a fused extension for Klett measurements were aseptically prepared with LB containing 5 mM glutamine with or without formate, then sealed and sparged with nitrogen for 10 min followed by 10 min of anaerobic mixture (10% CO_2, 10% H_2, 80% N_2). The anaerobic media were inoculated with *S.* Typhimurium WT, AB01 (Δ*rpoN*), and AB03 (Δ*fhlA*) overnight cultures (3 biological replicates for each), which were grown aerobically overnight in LB containing 5 mM glutamine, to obtain a starting OD_{600} of ~0.01. Cultures were grown anaerobically with shaking to mid-log phase (a prior growth curve under these growth conditions was used to determine the Klett units at which mid-log growth was achieved). Cultures of S. Typhimurium WT and AB01 were grown aerobically to mid-log in LB. Cells were harvested in 2 mL aliquots and immediately pelleted and stored at -80°C. Frozen cell pellets from the different growth conditions were resuspended in 100 µL RNA SNAP boil solution (154) and incubated at 95°C for seven minutes. Debris was pelleted for five minutes and the supernatant transferred to a fresh tube. RNA was ethanol precipitated, treated with RQ1 DNase (Promega), and reverse-transcribed with Superscript II, as

described above. cDNA was quantified by qRT-PCR as described above using primer pairs 29/30 (*rpoD* reference gene), 45/46 (*hypO,* first gene in *hyb* operon), 69/70 (*fdhF,* known FhlA- and σ^{54}-dependent gene) and 43/44 (inter-*hypO-yghW*); the primers for the inter-*hypO-yghW* amplify a region downstream and upstream of the primary σ^{70} promoter, so that the product is specific for cDNA from the σ^{54}-dependent transcript. Three technical replicates per biological replicate were utilized in the qRT-PCR; averages of the three biological replicates were used for the reported ratios WT+DctD250 to Δ*rpoN*+DctD250, and calculation of standard deviation and *p*-values.

Purification of σ^{54}

Sigma54 was purified from BL21 DE3 *E. coli* containing pJES937, encoding N-terminally His-tagged σ^{54} (*his*6-*rpoN*). Purification was performed as described in Cannon, *et al*. with the modifications in Kelly and Hoover (85,160). Protein expression was induced by 1 mM IPTG for 3 hours. Cells were pelleted and subjected to one freeze/thaw cycle, resuspended in 0.1x volume breakage buffer (50 mM Tris-Acetate pH 8.2, 200 mM KCl, 1 mM EDTA, 1 mM DTT, 1% Triton X-100), and 1x protease inhibitor cocktail (Sigma-Aldrich) and freeze/thawed again. Protein was pelleted and washed in 1M NaCl, 1% Triton-X 100. Protein was solubilized in 50 mM Tris-HCl pH 8, 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 5% glycerol, 1% sarkosyl and 1x protease inhibitor cocktail. Insoluble debris was removed by centrifugation. Protein was dialyzed against 50 mM Tris-HCl pH8.0, 250 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 5% glycerol, and 1x protease inhibitor cocktail overnight at 4°C, and then for an additional four hours against 50 mM Tris-HCl pH8.0, 100 mM NaCl, 0.1 mM EDTA, 1 mM DTT, 5% glycerol. The soluble protein was then exchanged into 25 mM $Na_2HPO_4 \bullet H_2O$, 0.5 M

NaCl pH 7.0. Protein was applied to Ni-NTA resin (BioRad) and washed three times with $Na_2HPO_4•H_2O$, 0.5 M NaCl, 20 mM imidazole pH 7.0. His-tagged σ^{54} was eluted from the resin with $Na_2HPO_4•H_2O$, 0.5 M NaCl, 250 mM imidazole pH 7. Collected protein was dialyzed into storage buffer (50 mM Tris-HCl pH 8.0, 0.5 M NaCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol. Protein was quantified via Bradford assay (Biorad) per manufacturer's instructions. Sigma54 purity (>95%) and quality (~85% full length σ^{54}) were assessed by Coomassie stain and Western blot, as described (85).

Electrophoretic Gel Mobility Shift Assay (EMSA)

The affinity of $E\sigma^{54}$ or σ^{54} for selected predicted binding sites from the ChIP-chip analysis was assessed using EMSA as described in Gallegos and Buck (161) with the exception of the length of the heteroduplex DNA used for the binding substrate. The oligonucleotides utilized for EMSA (5 through 28 and 71 through 74, Table 3.1) are 50 nucleotides long and the annealed oligonucleotides (pairings indicated in Table 3.1) form a heteroduplex in which the top strand sequence has mismatches at the -11 and -10 positions of the promoter sequence to produce the DNA distortion where $E\sigma^{54}$ or σ^{54} interact with the bases on the bottom strand in a stable closed complex (161). Fifty pmol of top- or bottom-strand oligonucleotides (depending on the 5' nucleotide (162)) were labeled with γ-^{32}P ATP (Perkin-Elmer Life Sciences) using T4 polynucleotide kinase for 30 minutes before heat inactivation per manufacturer's instruction. Unincorporated isotope was removed using Centri-Sep spin columns (Princeton Separations).

Each binding reaction was performed in STA buffer (25 mM Tris-acetate (pH 8), 8 mM magnesium acetate, 10 mM KCl, 1 mM DTT, 3.5% (w/v) PEG 8000) with 16 nM DNA probe and 5 μM sonicated calf thymus DNA. Shifts were performed with 10, 50,

100, and 200 nM holoenzyme (2:1 RpoN:core RNA Polymerase (Epicentre)), 100 nM, 500 nM, 1 μM, or 2 μM RpoN protein, or 100 nM core RNA polymerase. This range of $E\sigma^{54}$ and σ^{54} concentrations overlap calculated estimates for the intracellular concentrations of core RNAP (~260 nM) and σ^{54} (~140 nM) in *E. coli* during exponential growth in LB; for calculating these estimated intracellular concentrations we used the ratio of σ^{54} to σ^{70} in the cell as 0.16 (163), the level of σ^{70} in the cell as 60 -170 fmole/μg total protein (13,163), the level of RNAP core as 46 fmol/μg total protein (164), 450 μg total protein/10^9 cells (13), and cell volume 0.8 x 10^{-15} L (165). Reactions were incubated at 30°C for 10 min, then loaded into a pre-chilled 5 or 6.5% native polyacrylamide gels (29:1 acrylamide to bis-acrylamide; Bio-Rad) as it was running at 60 mV in 1X TBE. After samples were loaded, voltage was increased to 120 mV for 5 min to allow protein-DNA complexes to enter the gel before returning the voltage to 60 mV for 90-100 min at 4°C. Gels were wrapped in plastic wrap, exposed to a phosphor screen (GE Healthcare Life Sciences), and imaged (Typhoon, GE Healthcare). GelAnalyzer software (166) was used to analyze gel images; background was subtracted by rolling ball (set to 25 pixels) and the intensity of the protein- probe complexes (shifted bands) and the unbound probe was quantified for each reaction (each gel lane). The ratios of bound probe to total probe (unbound + bound) were plotted versus $E\sigma^{54}$ or σ^{54} concentrations in the binding reactions using Excel and the line of best fit was calculated, from which the estimated concentration of protein at which 50% of probe is bound at equilibrium was determined. These values are used for designating the level of binding affinity of $E\sigma^{54}$ or σ^{54} for the assayed binding sites, but they are not binding constants for the native DNA binding sites because the probes were heteroduplex DNA.

Competition assays were performed to determine the specificity of σ^{54} and holoenzyme binding to target DNA. Assays were performed as described above with the following modifications: each binding reaction contained 200 nM core RNA polymerase, 200 nM holoenzyme, or 2 μM σ^{54}, which were empirically-determined concentrations that give ~50% bound DNA target; and nonspecific competitor DNA (sonicated calf thymus DNA, GE Healthcare) or specific, unlabeled target DNA was added at 1, 10, 50, 100 and 500 molar excess relative to the labeled target DNA before addition of protein to the binding reaction. Protein-DNA complex formation was measured as described above.

Table 3.1: Oligonucleotides

No.	Oligonucleotide Name	Sequence 5' – 3'	Purpose
1	RpoNRed F	GAAAACGACTCTGAATATGAAGCAAGGTTTGCAACTCAGGTGT GTAGGCTGGAGCTGCTTC	Generate Δ*rpoN* with 2
2	RpoNRed R	GCGGAATGGATAAAGACTCTCGGTACTTCGCAACAGTGCGCATA TGAATATCCTCCTTA	Generate Δ*rpoN* with 1
3	FhlARedF	GTGTGTGCCGCGGTTTGCGGTCAGCCATCCGGTGTGTAGGCTGG AGCTGCTTC	Generate Δ*fhlA* with 4
4	FhlARedR	TTACGCTAACGCGTCTTTATCAATACCCAATCGCTCATATGAAT ATCCTCCTTA	Generate Δ*fhlA* with 3
5	RpoHShiftTop	TTGCTCATCGGCTTTGGCACGGTTGTTGCGAGCTGACGGTGCCA GGCAAT	EMSA with 6
6	RpoHShiftBot	ATTGCCTGGCACCGTCAGCGAGCAACAACCGTGCCAAAGCCGA TGAGCAA	EMSA with 5

7	clpPShiftTop	TGACGGGTTTTTTTTGTCACGTATTTTGCCGGGTAAGGGTGCGA AAACCG	EMSA with 8
8	clpPShiftBot	CGGTTTTCGCACCCTTACCATGCAAAATACGTGACAAAAAAAAC CCGTCA	EMSA with 7
9	wecDShiftTop	CATCGCGACCAGCATGGCGCGGAAATTGCCAACCTTACGGACAT TCCTGC	EMSA with 10
10	wecDShiftBot	GCAGGAATGTCCGTAAGGTGTGCAATTTCCGCGCCATGCTGGTC GCGATG	EMSA with 9
11	0816ShiftTop	TTTCGCCACCGGACTGGCATCGATATTGCCCACGCGCGAGGAGA TGCGCT	EMSA with 12
12	0816ShiftBot	AGCGCATCTCCTCGCGCGTTTGCAATATCGATGCCAGTCCGGTG GCGAAA	EMSA with 11
13	1057ShiftTop	TATTTAAGTCAGTATGGCCTGAATCTTGCGCAATTTGAAACTAA AACCGG	EMSA with 14
14	1057ShiftBot	CCGGTTTTAGTTTCAAATTTAGCAAGATTCAGGCCATACTGACTT	EMSA with 13

		AAATA	
15	2985ShiftTop	CGTTTTTAACATGCTGGCATCACTGTTGCCTTTCTTTCTCAGTAA AGAGA	EMSA with 16
16	2985ShiftBot	TCTCTTTACTGAGAAAGAACTGCAACAGTGATGCCAGCATGTTA AAAACG	EMSA with 15
17	3814ShiftTop	CCGTTACGAAGACCTGGCTTTTATTTTGCCATGTTCGCGAAGAA GTTATT	EMSA with 18
18	3814ShiftBot	AATAACTTCTTCGCGAACAGTGCAAAATAAAAGCCAGGTCTTCG TAACGG	EMSA with 17
19	5080ShiftTop	TTGCTGGCTTGCTCAGGCGCGAATAATGCCGCCGCCAGATGAAC CGTCTC	EMSA with 20
20	5080ShiftBot	GAGACGGTTCATCTGGCGGATGCATTATTCGCGCCTGAGCAAGC CAGCAA	EMSA with 19
21	glnAShiftTop	GGCAATGTGAAAGTTGGCACAGATTTCGCGGTATATTTTTACGG CGACAC	EMSA with 22

22	glnAShiftBot	GTGTCGCCGTAAAAATATAAAGCGAAATCTGTGCCAACTTTCAC ATTGCC	EMSA with 21
23	otsABShiftTop	AAATCTCCGGGGTCGGGAATGGAATATGCCTGAAAAAACCAAT TCGGTTG	EMSA with 24
24	otsABShiftBot	CAACCGAATTGGTTTTTTCCTGCATATTCCATTCCCGACCCCGGA GATTT	EMSA with 23
25	proP50mer-Top	AACAGTAACGTTATTGGCCTGATTTTTGCCAGTTTGTTGATGCTG GCGGT	EMSA with 15
26	proP50mer-Bot	ACCGCCAGCATCAACAAACCTGCAAAAATCAGGCCAATAACGT TACTGTT	EMSA with 18
27	proPControlTop	AACAGTAACGTTATTTTCCTGATTTTTTTCAGTTTGTTGATGCTG GCGGT	EMSA with 17
28	proPControlBot	ACCGCCAGCATCAACAAACCTAAAAAAATCAGGAAAATAACGT TACTGTT	EMSA with 20
29	RpoD-RT F	AACGAATAAGTGTGGATACCG	qRT-PCR with 30

30	RpoD-RT R	TCTTCCATTACCTGAATACCC	qRT-PCR with 29
31	2985qPCRRpoN-F2	TCTGATGGAAGAGGAAGGTAT	qRT-PCR with 32
32	2985qPCRRpoN-R2	AAGAGATCGGCGAACAAC	qRT-PCR with 31
33	1057qPCRAnti-F2	GCCAATTACTGACCAATCTC	qRT-PCR with 34
34	1057qPCRAnti-R2	ACGATGTAGCCGAATAGG	qRT-PCR with 33
35	0816qPCRRpoN-F2	AGTTATCGGGAAGATATGATGAAA	qRT-PCR with 36
36	0816qPCRRpoN-R2	ATCTGACTTATTCGCTCTGAG	qRT-PCR with 35
37	0816qPCRpotE-F2	TATCTCCTGGCTGGTCAC	qRT-PCR with 38
38	0816qPCRpotE-R2	AAGGCATATTCGGCGTAAC	qRT-PCR with 37
39	1057qPCRvirK-F2	TAACCAACAACGCACACTATT	qRT-PCR with 40
40	1057qPCRvirK-R2	GAGAGCCTCCATCACGAT	qRT-PCR with 39
41	2985qPCRptsH-F2	GGATTGCGTAGAACACCTT	qRT-PCR with 42
42	2985qPCRptsH-R2	CGGAGCGGTAATGGTAAC	qRT-PCR with 41
43	3814qPCRRpoN-F2	AATTACACTATCGCCTCTTCC	qRT-PCR with 44
44	3814qPCRRpoN-R2	TACGCCATTTACCACACTTT	qRT-PCR with 43

45	3814qPCRhypO-F2	GGCATTAACCGTCGTGAT	qRT-PCR with 46
46	3814qPCRhypO-R2	CGCTGTGGATTGGATACC	qRT-PCR with 45
47	5080qPCRmaIE-F2	TCGCCGTTGATCCAGATA	qRT-PCR with 48
48	5080qPCRmaIE-R2	TCGCATTATCCGCACTTAC	qRT-PCR with 47
49	5080qPCRmalK-F2	GGAGTTCGTGGTGTTTGT	qRT-PCR with 50
50	5080qPCRmalK-R2	GGCAGGCGGAATATCATT	qRT-PCR with 49
51	qPCR5161proP-F	GCGGTGATAGGCTTGATT	qRT-PCR with 52
52	qPCR5161proP-R	CTTCGCTTCCTGGATGTC	qRT-PCR with 51
53	qPCR4723wecE-F	TGTGGCAGACTTACTACGA	qRT-PCR with 54
54	qPCR4723wecE-R	GCAGCGGAATGTAATGGA	qRT-PCR with 53
55	DrpoN-F	TTCCGTTATCGACATCAAACGC	Confirm deletion with 57
56	DrpoN-R	AATGGCGGCATACATATCTTGG	Confirm deletion with 56
57	glnAP F	GCGCGTTATTGTACACGG	Confirm ChIP with 59

58	glnAP R	TGTACTCTCCCGGATTGG	Confirm ChIP with 58
59	glnA 6F	GAACGTCTGGAAGCGGG	Confirm ChIP with 61
60	glnA 6R	TTCAGACATGCTCACTCC	Confirm ChIP with 60
61	LMPCR-1	AGAAGCTTGAATTCAGCAGTCAG	Amplify ChIP with 63
62	LMPCR-2	CTGCTCGAATTCAAGCTTCT	Amplify ChIP with 62
63	glnA RACE 2 PCR	CGGATAATCAGGGTGGAGTCCGC	Amplify RACE product
64	glnA RACE 3 Nested	GCGAAGAACGGGTCGATAACCGC	Nested RACE PCR
65	glnA TSS 2	TACCAAAGGCAAAGAACAGC	cDNA for RACE
66	proP cDNA 3	AGAGATTAAGCGAACCN	cDNA for RACE
67	proP RACE 2 PCR	GCGGATGTTGCTGTACCAGACGC	Amplify RACE product
68	proP RACE 3 Nested	TGCAGCTCCGCAATTTCCTGATCG	Nested RACE PCR
69	fdhF qPCR For	AATCTGGTGGTCGATAACG	qPCR with 71
70	fdhF qPCR Rev	GGTCAGGATCTGGGTATC	qPCR with 70

71	LowEnrich978Top	ACCGGAGAGGACTTTGGCAGGCTTTTTGAGCTCATCCTGGCTAA ATAGCA	EMSA with 72
72	LowEnrich978Bot	TGCTATTTAGCCAGGATGATATCAAAAAGCCTGCCAAAGTCCTC TCCGGT	EMSA with 71
73	LowPSSM0238Top	AAAACGTTATAACCTGGCCGCTGTTTTGCGTCTTTTGCTGGGCG GCTACG	EMSA with 74
74	LowPSSM0238Bot	CGTAGCCGCCCAGCAAAAGCAGCAAAACAGCGGCCAGGTTATA ACGTTTT	EMSA with 73
75	KpnifHmod-Top	TAAACAGGCACGGCTGGTATGTTTTTTGCCATTCTCTGCTGGCA AACACT	EMSA with 76
76	KpnifHmod-Bot	AGTGTTTGCCAGCAGAGAAGTGCAAAAAACATACCAGCCGTGC CTGTTTA	EMSA with 77
77	NtrC::KanF	CGAGTTTTCGGTTTACCTGCCGATTCGGAAATAGAGGTGTTGTG TAGGCTGGAGCTGCTTC	Generate Δ*ntrC* with 76
78	NtrC::KanR	CGCGGGTAATGTTTACTCCATTCCCAGCTCTTTCAACTTCCATAT	Generate Δ*ntrC* with 75

		GAATATCCTCCTTA	
79	ntrC-Fwd1	CGAGTTTTCGGTTTACCTGCCG	Confirm Δ*ntrC* with 78
80	ntrC-Rev	CGGGTAATGTTTACTCCATTCCC	Confirm Δ*ntrC* with 77

CHAPTER 4

RESULTS

Most $E\sigma^{54}$ Binding Sites in the *S.* Typhimurium Genome Are Positioned to Regulate Gene Expression through Roles other than the Promoter for Gene Transcription

ChIP-chip assays were used to characterize $E\sigma^{54}$ genomic DNA binding sites in *S.* Typhimurium 14028s in the presence of DctD250. DctD250 activates open complex formation supporting stronger and more stable interactions with the DNA. Protein-DNA complexes were cross-linked *in vivo* with glutaraldehyde in the presence of rifampicin to capture $E\sigma^{54}$ associated with promoter sequences from which it initiates transcription (116). $E\sigma^{54}$- and σ^{54}-DNA complexes were enriched from sheared genomic DNA by immunopreciptation with α-σ^{54}. The DNA released from the complex when the cross-link was reversed was then amplified by ligation-mediated PCR, labeled using nucleotides linked to fluorescent dyes Cy3 or Cy5, and applied to the tiling microarrays. The WebarrayDB (158) database system and cross-platform analysis suite was utilized for initial processing of the ChIP-chip data for the three biological replicates of WT+DctD250 and Δ*rpoN* +DctD250; $\log_2$ values for Cy3- and Cy5-intensities corresponding to each probe on the microarrays were normalized and averaged for calculation of M-values ($\log_2$-ratio of WT+DctD250 and Δ*rpoN*+DctD250). Comparison of WT and Δ*rpoN* enrichment eliminated possible false positives stemming from any α-σ^{54} cross-reactivity with other substrates. These M-values and the genomic positions of

the corresponding probes were the input data for ChIPeak (ChIP-chip analysis software developed for this study, Supplemental Materials) to identify peaks of enriched DNA sequence, calculate statistical significance of the peaks of (p-value), and map peak positions in the genome. The most likely $E\sigma^{54}$ DNA binding site associated with each peak maximum was predicted using a standard Position-Specific Scoring Matrix (PSSM) method (167).

One hundred eighty-four DNA regions were enriched at least three-fold in WT+DctD250 relative to Δ*rpoN*+DctD250, with p-values $<10^{-17}$ S2). Sigma54 has been shown to specifically bind to a naturally occurring promoter in the absence of RNA polymerase (23), so all enriched DNAs in the α-σ^{54} pull-down potentially encode a DNA site that binds σ^{54} and/or $E\sigma^{54}$. Since no DNA sequence that binds σ^{54}, but not $E\sigma^{54}$, has been reported, the binding sites are referred to in general as $E\sigma^{54}$ DNA binding sites; σ^{54} binding to selected enriched DNA sequences was assessed by EMSA, as described below.

The resolution afforded by the tiling arrays did not allow clear peak resolution for $E\sigma^{54}$ DNA binding sites that are less than 500 bp apart; the closest peak maxima identified by ChIPeak that met the criteria for three-fold enrichment and p-value $<10^{-17}$ were 550 bp apart (peak maxima at 14028s genomic positions 4484310 and 4484860 in Table 4.1). The closely-spaced divergently-transcribing σ^{54}-dependent promoters between the *hyc* and *hyp* operons (168) and between *zraSR* and *zraP* (99), which were previously described in *E. coli,* are associated with a single peak from the ChIP-chip analysis for *S.* Typhimurium 14028s. Twenty-eight of the 184 enriched sequences encode two potential $E\sigma^{54}$ DNA binding sites with similar PSSM scores and within 250

bp of the peak maximum; where the PSSM scores differ by less than 10%, the site closest to the peak maximum is reported in Table 4.1, otherwise the site with the highest PSSM score is provided.

The identified $E\sigma^{54}$ DNA binding sites were divided into intergenic or intragenic groupings based on their position and orientation relative to open reading frames (ORFs) annotated in *S.* Typhimurium 14028s, as illustrated in Figure 4.1A. A caveat to this classification scheme is the assignment of $E\sigma^{54}$ DNA binding sites located within short hypothetical ORFs that are annotated in the 14028s genome, but not in the LT2 genome deposited in NCBI. *S.* Typhimurium 14028s is highly related to the LT2 strain and the numerous hypothetical ORFs annotated only in 14028s do not reflect the actual differences in the LT2 and 14028s genomes (169). Therefore, $E\sigma^{54}$ DNA binding sites within these short, hypothetical ORFs in the 14028s genome annotation were designated as intergenic or intragenic according to the LT2 genome annotation.

The $E\sigma^{54}$ DNA binding site position is suggestive of its potential role in gene regulation; examples of these roles are depicted in Figure 4.1B and described in the Discussion. Forty of the 184 $E\sigma^{54}$ DNA binding sites (21.7%) are located in intergenic regions. The 33 intergenic binding sites that are oriented toward the 5' end of the nearest coding sequence, are positioned to act as promoters of the downstream gene(s) and include all 26 previously identified intergenic σ^{54}-dependent promoters in *S.* Typhimurium (4). The other 7 intergenic sites are oriented toward the 3' end of the nearest annotated ORF; potential regulatory roles for these seven sites include transcription interference by various mechanisms (12), such as promoter competition and polymerase collision. The remaining 144 (78.2% of total) enriched DNA regions in the

ChIP-chip assay have potential $E\sigma^{54}$ DNA binding sites located within annotated coding sequences (intragenic sites). Eight of the intragenic sites are within 250 bp of the 3′ end of the ORF and oriented outward toward the 5′ end of the adjacent ORF such that the sites could act as promoters for the adjacent genes or sRNAs that include the rho-independent terminator (125,170). Eleven intragenic sites are in the 5′ end of an ORF and oriented outward, ten sites face the 3′ end of an adjacent ORF and one site is oriented toward a large intergenic space (Fig. 4.1A), consistent with potential roles in transcriptional interference (12). One hundred and twelve of the intragenic sites are oriented across the ORF in which they reside; 70 of these intragenic sites are positioned in the sense direction and the remaining 42 are antisense to the coding sequence (Fig. 4.1A). These intragenic sites may have regulatory roles involving multiple mechanisms for transcriptional interference, including sRNA or antisense RNA (12,124,171). Twelve inter- and intragenic $E\sigma^{54}$ DNA binding sites overlap σ^{70}-type promoters associated with TSSs mapped by Kröger et al. (125), suggesting regulation by promoter competition; these sites are indicated by footnote in Table 4.1.

Sigma54 Both Positively and Negatively Regulates Transcription of Genes and Non-coding RNAs Involved in Diverse Cellular Processes

The σ^{54}-dependent transcriptome of *S.* Typhimurium 14028s was assayed for the same cultures (WT+DctD250 and Δ*rpoN*+DctD250) used in the ChIP-chip analysis. RNA was isolated from early log-phase cultures, converted to cDNA and labeled with Cy3 or Cy5, and then applied to the Affymetrix tiling arrays for *S.* Typhimurium 14028s. Differential analyses for the paired biological replicates of WT+DctD250 and Δ*rpoN* +DctD250 was performed using WebArrayDB and LIMMA, which normalized and

averaged intensity $\log_2$ values corresponding to each probe on the arrays and calculated M-values ($\log_2$-ratio of WT+DctD250 and Δ*rpoN*+DctD250) and *p*-values (158). For comparison of expression of full ORFs, the M-values for all probes across the ORF were averaged and the *p*-value calculated (158). The M-values were used to calculate the fold-change in expression levels (WT+DctD250/Δ*rpoN*+DctD250), reported as signal ratios in Tables 4.2.

Up-regulated operon transcripts. Sixty-five genes from twenty-two operons were expressed at >two-fold higher levels in WT+Dct250 than in Δ*rpoN*+DctD250, with *p*-values $<10^{-17}$ (Table 4.2). The first gene in each operon that was up-regulated in the presence of σ^{54} and DctD250 has an Eσ^{54} binding site immediately upstream (Table 4.2). Most of these σ^{54}-dependent transcripts were previously reported for *S.* Typhimurium LT2, as indicated in Table 4.2, but σ^{54}-dependent transcription of *gltI_sroC_gltJ-L, yeaGH, hisJ*, and *pspG,* which was previously reported in *E. coli* (172,173)*,* was confirmed for *S.* Typhimurium in this microarray analysis. The transcription start sites (TSS) for mRNA from 10 of the 22 up-regulated operons were mapped to predicted σ^{54}-dependent promoters in RNA-seq and dRNA-seq analyses performed with *S.* Typhimurium strain 4/74 under infection-relevant conditions by Kröger et al. (125), as indicated in Table 4.2. Detection in the Kröger et al. study of only 10 out the 22 σ^{54}-dependent transcripts that were identified in this microarray assay reflects the diverse conditions required to activate the 13 bEBPs that control σ^{54}-dependent promoters in *S.* Typhimurium and the utility of the constitutively active, promiscuous bEBP (DctD250) to assess the full σ^{54} regulon (4). The microarray and ChIP-chip data for the STM14_0673-0668 and *gltI_sroC_gltJ-L* operons are illustrated in Figure 4.1B. The

STM14_0673-0668 (STM0577-0572) operon, which is predicted to encode a mannose-family phosphotransferase system (20) that may be required in colonization by *S.* Typhimurium of reservoir hosts (7,174), is not expressed under infection-relevant conditions assayed by Kröger et al. (125). It is likely that this operon is expressed only from the σ^{54}-dependent promoter, which is responsive to DctD250 in our microarray assays [Table 4.2; (4)]. The regulation by the σ^{54}- and σ^{s}-dependent promoters of the *gltI_sroC_gltJ-L* operon impacts cellular metabolism far beyond the transport of glutamate and aspartate because SroC, a sRNA processed from the transcript that terminates between *gltI* and *gltJ,* negatively regulates Gcv, which has one of the largest sRNA regulons in *S.* Typhimurium (128) (see Discussion).

Up-regulated intragenic transcripts. Sigma54-dependent transcripts that originate from intragenic promoters were considered differentially expressed in WT+DctD250 versus Δ*rpoN*+DctD250 strains if four or more adjacent probes from the same DNA strand exhibited a minimum two-fold change in expression with *p*-values less than 0.05 (Table 4.2). Two novel σ^{54}-dependent transcripts were identified that initiate from intragenic promoters in *wecD* and *proP* (Table 4.2, Fig. 4.1B). Sigma54-dependent expression from the intragenic promoters was confirmed by qRT-PCR; the relative transcript levels (WT+DctD250/ Δ*rpoN*+DctD250) from the intragenic *wecD* and *proP* promoters were 13 ±10 (0.0068 *p*-value) and 5.3 ±3.2 (*p*-value 0.0015), respectively (Fig. 4.1B). The TSS for the intragenic *proP* transcript was determined by 5'RACE (data not shown) and mapped to the $E\sigma^{54}$ binding site. The transcript from this intragenic promoter in *proP* may be processed to a sRNA, as suggested by enrichment of a *proP* transcript in Hfq immunoprecipitation (175) (see Discussion).

Indirectly down-regulated transcripts. Transcripts from four ORFs were >two-fold down regulated in the presence of σ^{54} and DctD250: STM14_1795; *lysA* (STM14_3638); *malE* (STM14_5085); and *lamB* (STM24_5087) (Table 4.2). The down regulation of *malE* was confirmed by qRT-PCR; the relative transcript level (WT+DctD250/ Δ*rpoN*+DctD250) for *malE* was 0.49 ±0.18 (*p*-value 0.038). Based on the ChIP-chip analysis there is an $E\sigma^{54}$ DNA binding site in the gene just downstream of the *malE* operon (STM14_5080, *yjbA*), but it is oriented away from the *malE* operon (Table 4.1), and there are no $E\sigma^{54}$ binding sites within several kilobases of STM14_1795 or STM14_3638 (*lysA*). Therefore, the reduced expression of these genes in the presence of σ^{54} and DctD250 is most likely to be by an indirect mechanism (see Discussion).

Down-regulated operon transcripts. Several down-regulated operon transcripts (*speF_potE*_STM14_0699, *ybjX*, *cbiA-Q_ cobU-T, glmY*, and *hypO_hybA-E*) that are associated with $E\sigma^{54}$ binding sites did not meet the stringent two-fold cut-off for differential expression, but transcript levels for genes in each of these five operons were reduced by 10-43% (with *p*-values <0.01 for all three biological replicates) in the WT+DctD250 strain as compared to the Δ*rpoN*+DctD250 strain (Table 4.2). Small changes in transcript were viewed as indicators of possible true physiological differences, but requiring additional confirmation. For the sRNA, *glmY*, in *S.* Typhimurium and *E. coli* it has previously been shown that transcription from the σ^{54}-dependent promoter, which is stimulated by the bEBP GlrR under glucosamine-6-phosphate-limiting conditions, represses expression from a σ^{70}-dependent promoter that precisely overlaps the σ^{54}-dependent promoter (127,129). The potential transcriptional interference mechanisms for the $E\sigma^{54}$ DNA binding sites associated with the

*speF_potE*_STM14_0816 and *hypO_hybA-E* operons are illustrated in Figure 4.1B. Relative transcript levels in WT+DctD250 versus Δ*rpoN*+DctD250 strains, as indicated by the microarray analysis, were confirmed by qRT-PCR for the σ^{54}-dependent promoter upstream of *hypO* (3.2 ±1.1; *p*-value 0.0025) and for the σ^{70}-dependent *hypO* (0.74 ±0.071; p-value 0.16) transcripts (Fig. 4.1B).

For the novel Eσ^{54} DNA binding sites identified as having potential regulatory roles in gene transcription in the *S*. Typhimurium strains expressing DctD250, the growth conditions and bEBP that activate the predicted σ^{54}-dependent promoters are generally not evident. However, the previously demonstrated differential expression of Hydrogenase 2 (*hypO_hybA-E*) and Hydrogenase 1 (*hyaA-F*) operons under conditions that favor fermentation (176,177) indirectly suggested that the bEBP FhlA, which is activated by the fermentation product formate, might stimulate transcription from $P_{inter\text{-}hypO\text{-}yghW}$ to modulate expression of the Hydrogenase 2 operon (Fig. 4.1B). In response to formate, FhlA stimulates expression of four σ^{54}-dependent operons of *S.* Typhimurium encoding the structural and assembly proteins for formate hydrogen lyase (*fdhF, hycA-I, hydNA*, and *hypA-E*), which breaks down formate to hydrogen and carbon dioxide (4,178). Transcript levels associated with $P_{inter\text{-}hypO\text{-}yghW}$, the first gene of the Hydrogenase 2 operon (*hypO*), and *fdhF* were assayed by qRT-PCR for *S*. Typhimurium 14028s WT, Δ*rpoN*, and Δ*fhlA* strains grown under aerobic conditions in LB and under anaerobic conditions in LB and LB+30 mM formate. Transcription of *hypO* and from $P_{inter\text{-}hypO\text{-}yghW}$ was unaltered in WT as compared to the Δ*rpoN* and Δ*fhlA* strains under all three growth conditions (data not shown); transcript levels for the positive control, *fdhF,* were significantly reduced in both the Δ*rpoN* and Δ*fhlA* strains compared to in WT, as

expected (data not shown). Thus, the condition under which $P_{inter\text{-}hypO\text{-}yghW}$ is expressed, and whether the Hydrogenase 2 operon is regulated by $P_{inter\text{-}hypO\text{-}yghW}$, is still unknown.

$E\sigma^{54}$-DNA Binding Sites in *S.* Typhimurium 14028s Are Highly Conserved in Other *S. enterica* serovars, but a Small Fraction of the Sites Are Fully Conserved in Other Enterobacteriaceae

A recent analysis of the interactions of $E\sigma^{54}$ with the genome of *E. coli* strain BW25113 by Schaefer et al. (179) also found $E\sigma^{54}$ binding sites associated with genes that are down-regulated in a σ^{54}-dependent manner. A BLAST search of the *S.* Typhimurium 14028s and LT2 genomes (in NCBI) for each of the twenty-six *E. coli* $E\sigma^{54}$ binding sites with potential roles in transcription interference, revealed that only two of these sites are conserved between *E. coli* and *S.* Typhimurium: the intergenic site associated with *argT* and the intragenic site in *glgA*. Conservation of the $E\sigma^{54}$ DNA binding site is defined here as the presence of the identical 18 bp $E\sigma^{54}$ core DNA binding sequence in the same genome context; thus, a non-conserved binding site may differ in single base pair or may be completely absent in the genome of *S.* Typhimurium. The *argT* and *glgA* sites in *S.* Typhimurium were detected as binding sites in our ChIP-chip analysis; however, unlike in *E. coli*, relative expression in WT *S.* Typhimurium versus the Δ*rpoN* strain for *argT* was 3.6-fold higher (Table 4.2) and for *glgA* was 1.2-fold higher (below 2-fold cutoff for differential expression). The negative versus positive regulation of *argT* and *glgA* by the σ^{54}-dependent promoters in *E. coli* versus *S.* Typhmiurium probably reflects a key difference in the methodology for detecting σ^{54}-dependent transcription from promoters requiring different bEBPs and different environmental signals, i.e. the *E. coli* study utilized overexpression of σ^{54} from a plasmid

to take advantage of basal level activation of bEBPs, while our study assayed σ^{54}-dependent transcription in *S.* Typhimurium strains expressing a constitutive, promiscuous bEBP.

Although there are *E. coli* $E\sigma^{54}$ binding sites reported to down regulate gene transcript (179) differing by as little as one or two base pairs from their corresponding site in *S.* Typhimiurium, none of these sites were enriched in our ChIP-chip analysis or were associated with σ^{54}-dependent changes in gene transcripts in the microarray analysis. This observation led us to assess the conservation of the 184 *S.* Typhimurium $E\sigma^{54}$ DNA binding sites in other Enterobacteriaceae. The dataset for this analysis comprised complete genomes from the NCBI database for 43 *S. enterica*, 2 *S. bongori*, 57 *E. coli*, and 10 *Shigella sp*. strains; the number of strains from each species in which the $E\sigma^{54}$ binding site is conserved, as defined above, is reported in Table 4.1. All of the 184 $E\sigma^{54}$ binding sites are conserved in multiple *S. enterica* serovars; 75% of the 184 sites are conserved only in *S. enterica* strains, 18% are conserved in *S. enterica* and *S. bongori* strains, 1.6% are conserved in *S. enterica, E. coli* and *Shigella* strains, and 4.9% are conserved in *S. enterica, S. bongori, E. coli* and *Shigella* strains. If we look at conservation within the intergenic and intragenic groupings of the $E\sigma^{54}$ sites, the intergenic sites are somewhat more broadly conserved than the intragenic sites; 7.5% intergenic sites versus 4.2% intragenic sites are in all assessed species, and 35% intergenic sites versus 14% intragenic sites are in both *S. enterica* and *S. bongori*. Comparison of conservation for intragenic sites in the sense versus antisense orientations reveals virtually no differences; e.g., for both sense and antisense intragenic sites, 4% are conserved in all assessed species and 2% are conserved in *S. enterica, E. coli,* and

Shigella species, and 15% of sense intragenic sites versus 11% of antisense intragenic sites are conserved in *S. enterica* and *S. bongori*.

Purified σ^{54} Binds in the Absence of Core RNAP to Two Novel $E\sigma^{54}$ DNA Binding Sites

Thirteen $E\sigma^{54}$ genomic binding sites that were identified in the ChIP-chip analysis were selected for *in vitro* confirmation of $E\sigma^{54}$-DNA interactions and analysis of σ^{54} binding to the same DNA sequences (Table 4.4). The selected binding sites are representative of the range of PSSM scores (7.36 to 21.1) for the significantly enriched sites (>3-fold enriched, p-value $<10^{-17}$) in the ChIP-chip assay (see Table 4.1) and include two sites that were chosen to assess the arbitrary 3-fold cutoff for delineation of probable binding sites (the *stfD* and *ybiT* associated sites were >2-fold enriched, but <3-fold enriched, in ChIP-chip). A derivative of the *Klebsiella pneumoniae nifH* σ^{54}-dependent promoter, designated *nifH*049 (23), was used as a positive control for both $E\sigma^{54}$ and σ^{54} *in vitro* binding; *nifH*049 and the *Sinorhizobium meliloti nifH* promoter are the only two promoters that have previously been shown to bind σ^{54} protein in the absence of core RNAP (23,180). EMSAs for binding of $E\sigma^{54}$ and σ^{54} to heteroduplex oligonucleotide probes were performed essentially as described by Gallegos and Buck (161), with the only significant exception being the replacement of the 88 bp heteroduplex probe with a 50 bp heteroduplex probe; the shorter probe did not affect the affinity of either $E\sigma^{54}$ or σ^{54} for the *nifH*049 probe (data not shown). The use of heteroduplex oligonucleotide probes was based on the finding by Gallegos and Buck (25) that σ^{54} has a 6-fold higher affinity for the *S. meliloti nifH* promoter sequence with a DNA distortion of two unpaired bases immediately downstream of the conserved -12 GC motif than for the *nifH* homoduplex

promoter sequence; this distortion allows σ^{54} to interact with the two exposed bases on the bottom strand, which mimics the "early melted DNA" in the stable closed complex of $E\sigma^{54}$ with the same promoter DNA (161,180).

Figure 4.2 shows examples of the EMSA results with the control *nifH*049 promoter sequence and three of the novel $E\sigma^{54}$ DNA binding sites identified by ChIP-chip analysis. Table 4.4 summarizes the EMSA results (averaged from three replicate assays) for all 13 sites, the positive control *nifH*049, and the negative control -24TT,-12TT*proP*, which is mutated in the -24GG and -12GC promoter elements of the *proP* intragenic site. The specificity of $E\sigma^{54}$ and σ^{54} binding to the *proP* site was assessed by competition with 10- to 500-fold molar excess of non-specific DNA (sonicated calf thymus DNA) or specific unlabeled probe DNA; the $E\sigma^{54}$-*proP* complex and $E\sigma^{54}$-*proP* were resistant to 500-fold molar excess of non-specific DNA, but were reduced by 10-50% at 10-fold molar excess of specific DNA competitor and by >95% at 500-fold molar excess of specific competitor. All the assessed sites, with the exception of *glnA* (discussed below), bound $E\sigma^{54}$ with varying levels of affinity that did not correlate with PSSM scores for the sites. $E\sigma^{54}$ binding to the *stfD*- and *ybiT*-associated sites, which were less than 3-fold enriched in the ChIP-chip analysis of genomic $E\sigma^{54}$ binding sites, suggests that the 3-fold enrichment cutoff excluded some authentic $E\sigma^{54}$ genomic binding sites. The *rpoH* and *glnA* σ^{54}-dependent promoters in *E. coli* were previously assessed for *in vitro* $E\sigma^{54}$ DNA binding by EMSA (120,181). The affinity of $E\sigma^{54}$ for the *S.* Typhimurium *rpoH* promoter (Table 4.4) closely matches that previously determined for the E. coli *rpoH* promoter (120), which differs at three positions of the 18 bp core $E\sigma^{54}$ DNA binding sequence. However, the *glnA* σ^{54}-dependent promoter from *S.*

Typhimurium, whose 18 bp core sequence is conserved in *E. coli,* does not bind $E\sigma^{54}$ in EMSA even at a >1000-fold higher $E\sigma^{54}$ concentration than used for 50% binding of a 43 bp homoduplex probe containing the *E. coli glnA* σ^{54}-dependent promoter sequence (181). The poor binding of $E\sigma^{54}$ to the *glnA* σ^{54}-dependent promoter is most likely due to the sequence modification to create the DNA distortion bubble, which required changing the sequence at the two bases adjacent to the -12 GC motif (position relative to the TTS) in the oligonucleotide for the top strand while keeping the wild-type sequence in the oligonucleotide for the bottom strand where σ^{54} contacts the bases in closed complex. Comparison of the sites that exhibited weak or no binding of $E\sigma^{54}$ revealed all four sites normally have a T at the -11 position of the promoter (Fig. 4.3). In the design of the oligonucleotide for the top strand of the heteroduplex, a T at position -10 or -11 was substituted with a G; the *glnA* promoter sequence has a T at both positions, so it has GG on the top strand and AA on the bottom strand at -10 -11. Sequence comparison of the ChIP-chip-identified 184 genomic $E\sigma^{54}$ DNA binding sites revealed only 12 sites (6.5%) with a G at the -11 position and no sites with G at both the -10 and -11 positions, which is the case for the heteroduplex *glnA* probe. The EMSA results suggest that even though σ^{54} (within the $E\sigma^{54}$-promoter closed complex) contacts the bottom strand at the -10 -11 DNA distortion, a G in the -11 position on the top strand appears to impair $E\sigma^{54}$ binding to the promoter and GG at -10 -11 severely reduces binding to the promoter.

Two $E\sigma^{54}$ DNA binding sites, the intragenic sites in *proP* and STM14_0816, also bound σ^{54} in the absence of core RNAP. Buck and Cannon (23) previously characterized $E\sigma^{54}$ and σ^{54} binding to wild-type *Sinorhizobium meliloti nifH* promoter and a modified *K. pneumoniae nifH* promoter (*nifH*049), demonstrating that a T-tract from -14 to -17 is

critical for σ^{54} binding to these sites. A sequence alignment of wild-type *S.meliloti nifH, K. pneumoniae nifH049, proP*, and STM14_0816 produced a consensus sequence, shown in Figure 4.4. Subsequent comparison to just one site that does not demonstrate detectable σ^{54} binding (upstream of *hypO*) eliminates most of the conserved sequence features as the determinant for σ^{54} binding (Fig. 4.4), suggesting that a composite of sequence determinants is required for σ^{54} binding must extend beyond T residues in the -17 to -14 position.

Table 4.1. ChIP-chip results for Eσ^{54} DNA binding sites, grouped by position and chromosomal context

Peak Position[1]	Associated 14028s ORF	Equivalent LT2 ORF	Gene Name	Signal Ratio[2]	PSSM Score	Conserved[3] *Se/Sb/Ec/Ss*	Distance to Peak[4]	Predicted Eσ54 DNA Binding Site[5]
Intergenic binding sites oriented toward 5' end of associated ORF								
418770	STM14_0431	STM0368[6]	*prpB*	17.9	17.7	43/2/0/0	146	ttgccTGGCATAGCCTTTGCTTTggtga
503670	STM14_0530	STM0448[6]	*clpP*	12.5	16.9	42/2/0/0	-51	tttttTGTCACGTATTTTGCATGgtaag
521100	STM14_0546	STM0462[6]	*glnK*	21.7	17.5	41/0/0/0	-39	gcttcTGGCACATCCTTTGCAATatcca
637570	STM14_0673	STM0577[6]		14.6	16.4	36/2/0/0	0	aaaacTGGCACGCCGTTTGCCATatctg
712450	STM14_0757	STM0649.S[6]		14	16	37/0/0/0	0	aaaaaTGGCACGCCTTTTGATTAgggaa
730680	STM14_0773	STM0665[6]	*gltI*	19.2	19.9	38/0/0/0	76	cattcTGGCACGTCTATTGCTTTgttaa
898110	STM14_0964	STM0830[6]	*glnH*	17.4	19.5	41/2/0/0	-4	ataatTGGCATGATTTTTTCATTacgct
1373930	STM14_1558	STM1285[6]	*yeaG*	15	16.9	38/0/0/0	69	taaaaTGGCATGAGAGTTGCTTTttttt
1392120	STM14_1582[8]	STM1303[6]	*astC*	12.7	19	41/0/0/0	22	ctggcTGGCACGAATGCTGCAATctaca
1793230	STM14_2040	STM1690[6]	*pspA*	18.6	16	37/2/54/10	54	aaagtTGGCACGCAAATTGTATTaatca
1823320	STM14_2075	STM1714	*topA*	4.47	12.4	43/2/1/0	14	caatcTGGCAACAGAATTGCTTGacatt
2091070	STM14_2420	STM1996	*cspB*	5.39	14.3	17/0/0/0	-82	aaatgTGGCGCGTTTATTGCCCGgcagg
2516940	STM14_2900[8]	STM2354	*hisJ*	6.19	15	38/0/0/0	52	attaaTGGCACGATAGTCGCATCggatc

2517910	STM14_2901	STM2355[6]	*argT*	14.2	14.6	43/0/56/10	-1	gaaaaTGGCATAAGACCTGCATGaaaaa
2524240	STM14_2907	STM2360[6]		17.1	21.2	41/0/0/0	-26	aatagTGGCATGCCTTTTGCTTTatccc
2637350	STM14_3035[8]	STM2475		12.2	12.5	41/0/0/0	-49	atttaTGGTATAGAGATTGCTACgttac
2759250	STM14_3143.P[8]	STM_R0152[6]	*glmY*	15	19.3	41/0/0/0	-160	aatgtTGGCACAATTACTGCATAatggt
3005310	STM14_3431	STM2840[6]	*norV*	14.3	15.4	17/0/0/0	49	aaagtTGGCACACTAGCTGCAATaagca
3010840	STM14_3436	STM2843[6]	*hydN*	16.1	14.7	41/0/0/0	-99	aaaatTGGCACGATTCGTGTATAtatcg
			hypA/					
3019780	STM14_3450	STM2854[6]	*hycA*[7]	20.2	21.1	37/2/0/0	-208	aatatTGGCATAAATATTGCTTTacagg
3410490	STM13_3905	STM3224	*ygjT*	8.35	15.5	43/2/0/0	21	ccgaaTGGCAAGGTCTATGCATAaaagc
3698110	STM14_4239	STM3521[6]	*rsr*	16.6	18.1	43/1/56/10	-305	ttttcTGGCACGCTGGTTGCAATaacca
3750440	STM14_4295	STM3568[6]	*rpoH*	14.1	18.8	43/2/0/0	-77	ggcttTGGCACGGTTGTTGCTCGctgac
3986090	STM14_4548	STM3772[6]		18	16.5	42/0/0/0	-69	aattcTGGCACAACCTTTGCTCTtctga
4155380	STM14_4733.P[8]	STM_R0167[6]	*glmZ*	14.7	17.4	41/2/0/0	52	aaagaTGGCACGTTATGTGCAATaatgt
4230730	STM14_4820	STM4007[6]	*glnA*	14.4	17.1	43/1/45/8	36	aaagtTGGCACAGATTTCGCTTTatatt
			zraSR/					
4402030	STM14_5014[8]	STM4173[6]	*zraP*[7]	18.5	17.8	43/0/0/0	47	aagaaTGGCATGATCTCTGCTTAagtaa
4478500	STM14_5102	STM4244[6]	*pspG*	19.5	19.7	14/0/0/0	-84	aatatTGGCATGATTTTTGTAAGggctt

4484860	STM14_5108	STM4250	*yjbQ*	3.31	13.9	41/0/0/0	-108	cttttTTGCATGATTTTTGCACAaactt
4541120	STM14_5155	STM4285[6]	*fdhF*	15.3	17	43/2/0/0	14	aaaagTGGCATAAAACATGCATActgag
4623230	STM14_5249	STM4367[6]	*nsrR*	6.13	16.6	35/0/0/0	-7	acgccTGGCAGATATTTTGCTTGccggg
4655920	STM14_5290[8]	STM4405		4.99	11.6	40/2/0/0	-50	caaacTGGCATCCCCTTTGCGGGaaaaa
4807650	STM14_5449	STM4535[6]		20.3	15	33/0/0/0	-64	ttttcTGGCACGCCGCTTGCTCTctttt
Intergenic binding sites oriented toward 3' end of associated ORF								
1011870	STM14_1097	STM0971		3.01	8.58	43/2/0/10	-39	cacccTGGTAAGGGGGTTGCATCctgta
1254030	STM14_1385[8]	STM1211	*ndh*	3.85	13.5	42/2/0/0	-29	gctatTGGTAACAATTTTGCAACcagta
1363540	STM14_1539	STM1274	*yeaQ*	3.51	13	43/2/0/0	0	attttTGGCATGGTTTTTTATTGgcggt
1640160	STM14_1877	STM1554		7.82	12.6	16/0/0/0	-43	ccagtTAGCACAGCGTTTGCATGaactg
1920590	STM14_2188	STM1810		4.48	12.8	37/0/0/0	0	ccgccTGGCACAGACGGTGCATCaggcg
2975970	STM14_3390	STM2808[6]	*nrdF*	8.93	18.9	43/0/0/0	-87	gcttcTGGCATGAATATTGCGAGgtgct
3332000	STM14_3816[8]	STM3151	*yghW*	11	13.3	39/0/0/0	-40	agaccTGGCTTTTATTTTGCACTgttcg
Intragenic sites in sense orientation relative to associated ORF								
23680	STM14_0025	STM0020		3.01	12.5	42/0/0/0	3	gcgacCGGAACACTTTTTGCTCGcagaa
154320	STM14_0157	STM01315	*ftsQ*	10.6	14.2	42/2/0/0	0	aataaTGGAACGCGTCTTGCAGGaattt
183520	STM14_0187	STM0155[5]		5.92	14.2	41/0/0/0	23	atttcCGGCATGGCATTTGCCAGcgatc

313410	STM14_0319	STM0272		6.57	12.1	35/0/0/0	0	ggaacTGGTACATACGTTGCAGGaactg
322100	STM14_0330	STM0282		3.85	10.1	42/0/0/0	7	tctgcTGGTACGCGAGATGCATCtgtac
506500	STM14_0533	STM0450	*lon*	4.71	14.5	43/2/0/0	28	gaccgTGGCCTCTATTTTGCAGAtgctg
593040	STM14_0619	STM0529	*fdrA*	3.8	14	39/0/0/0	38	ccgacTGGCATGTTTATTGTCTCgcgtg
668310	STM14_0704	STM0605	*ybdN*	3.49	9.45	43/1/0/0	-96	aacgtTGGCGCGCAAAATGAATAaaaag
802590	STM14_0855	STM073[6]	*kgd*	4.62	7.36	43/0/0/0	26	ttccgTGGCCATCAACATGCAAAcctcg
952380	STM14_1030	STM0879	*potH*	7.74	14	23/0/0/0	17	tatatCGGCATTGTTTATGCTTAtctgc
956570	STM14_1035	STM0884[6]	*ulaA1*	3.42	13.1	43/0/0/0	-27	actttCGGCACGATTTTTTCCATgggca
976360	STM14_1057[8]	STM0940[6]	*ybjX*	13.3	17.1	43/2/0/0	-49	cagtaTGGCCTGAATCTTGCTAAatttg
993240	STM14_1079	STM0957	*cydD*	6.28	13.4	42/0/0/0	-53	gatcaTGGCGCGTATCCTGCAACatatg
1041350	STM14_1124	STM0994	*mukB*	3.7	11.3	43/0/0/0	43	ggaagCGGCGCGATTTGTGCAGCaatat
1103010	STM14_1199	STM1057	*pepN*	3.86	10.8	43/0/0/0	34	tccaaTGGCAACCGTGTTGCACAgggcg
1210780	STM14_1336	STM1167	*rimJ*	11.7	15.4	42/0/0/0	39	tctggTGGTACGTTTAGTGCATGagcgt
1426040	STM14_1624	STM1337	*pheS*	3.54	10.1	43/0/0/0	0	gcgacTGGCCCGGAAATTGAAGAtgact
1453820	STM14_1654	STM1361[6]	*ydiM*	4.94	15.3	35/0/0/0	-42	ttcatTGGCATTCTTTATGCTCAgacga
1471950	STM14_1673	STM1379	*orf48*	4.66	14.9	43/0/0/0	-16	aacttTGGCGCGCTTTTCGCTTTttttg
1485540	STM14_1684	STM1390[6]	*orf242*	13.7	17.9	32/0/0/0	1	ctggaTGGCATCATTATTGCCTAcacct

1492550	STM14_1690	STM1395	*ssaD*	5.12	10.8	43/0/0/0	56	cttatTGGTACGAGAAGTGCAGGatgtt
1500360	STM14_1705	STM1409	*ssaJ*	4.32	15.5	42/0/0/0	0	aatgcTGGCATTACTTATGCAGCatcat
1502730	STM14_1708	STM1412	*ssaL*	3	8.2	43/0/0/0	-91	tgaacGGGTCCGTATTTTGCTAAgagca
1507970	STM14_1714	STM1418	*ssaQ*	6.24	16.2	42/0/0/0	10	ggggcTGGCGCCGTTATTGCAAGccagt
1531370	STM14_1744	STM1446	*anmK*	5.54	9.83	43/0/0/0	64	cagttAGGCCGGCTTTTTGCACAggcgg
1554370	STM14_1772	STM1469	*fumC*	3.22	11.1	36/0/0/0	6	cggtaTGGAACGCAAGGTGCATCccaat
1656230	STM14_1889	STM1566	*sfcA*	3.65	8.87	43/0/0/0	22	caataTGGACGACATTTTGCAGAatgtg
1694780	STM14_1930	STM1595	*srfC*	4.68	14.3	39/0/0/0	4	ggcgcTGGCGCATATGTTGCAACaaacc
1783150	STM14_2027	STM1679	*mppA*	3.49	15.1	13/0/0/0	-8	ctggcTGGCATCACTAATGCACAagcca
1882700	STM14_2143	STM1773	*ychA*	7.17	13	43/0/0/0	-33	tgtgaTGGCATGATCCTGGCATCagaga
2017480	STM14_2324	STM1910		3.54	13.3	27/0/0/0	0	atgtaTGGCATGGCGAATGCACCcaacg
2084440	STM14_2412	STM1990[6]	*yedA*	13.8	17.5	42/0/0/0	20	cttttTGGCGCGCTTTTTGCCTTgtaca
2099150	STM14_2433	n/a[9]		6.45	12	43/0/0/0	-81	ttatcTGGCCCGGAATATGCTGCctgat
2274430	STM14_2627	STM2129	*yegB*	7.58	15	42/0/0/0	63	gccgcTGGTACTGCTTTTGCAGGgaatg
2329530	STM14_2688	STM2181[6]	*yohJ*	14.9	13.9	42/0/0/0	-16	tacgcAGGCATTTTTCTTGCATCactgc
2394090	STM14_2769	STM2241	*sspH2*	4.48	10.2	43/0/0/0	-126	agcagTGGTACAGAAAATGCGTGcctgc
2582100	STM14_2969	STM2415	*gltX*	3.1	10.1	40/0/0/0	-58	atatgTGGCGACGCATTTGCAGTggcat

2595680	STM14_2985	STM2430[6]	*cysK*	18.8	17	43/2/1/0	-62	catgcTGGCATCACTGTTGCAGTtcttt
2616160	STM14_3012	STM2454	*eutR*	6.44	10.4	43/0/0/0	-33	ccgtcAGGCCACGTTTTTGCATAacccg
2939140	STM14_3345	STM2774	*iroC*	3.12	12.3	28/0/0/0	53	cgcgtTGGCGTAGTTTTTGAAGAcgcgt
3009890	STM14_3434	STM2842	*hypF*	6.34	9.68	38/0/0/0	-220	gcagcTGGCGCAGCAGTTGCGATtacac
3100250	STM14_3543	STM2939[6]	*ygcH*	15.4	15	30/0/0/0	13	ttcgcCGGCACAGCTCTTGCATCtggtg
3125610	STM14_3565	STM2957[6]	*rumA*	19.4	15.8	43/2/39/10	0	aaaacTGGAACGCTTTTCGCATTctgag
3376800	STM14_3869	STM3195	*ribB*	4.97	12.2	42/0/0/0	69	tctttTGGTACGCCTTTTGAACGtgttg
3408660	STM14_3902	STM3222	*ygjQ*	4.29	16.1	38/0/0/0	104	aacagTGGCGTGTTTTCTGCTTGtcgca
3476020	STM14_3977	STM3294	*glmM*	3.63	12.3	40/0/0/0	-1	cagctTGGCATTCCGTTTGCCCGggcga
3486000	STM14_3989	STM3302[6]	*yhbE*	10.5	12.4	43/0/0/0	62	ccggtTGGCATGATGGTCGCCAGcgtct
3522940	STM14_4030	STM3339	*nanA*	4.52	13.6	43/2/0/0	33	aggtaTGGCAAAAGCATTGCAAGgcgta
3608880	STM14_4141	STM3438	*rplW*	4.53	12.7	43/2/56/10	74	gatgcTGGCATGATTCGTGAAGAacgtc
3716020	STM14_4255	STM3535[6]	*glgA*	14.4	16.4	42/2/0/0	6	tatcaAGGCATGTTTTATGCAAAgcata
3746640	STM14_4290	STM3564	*livK*	4.16	10.9	41/0/0/0	-14	aggacTGGCGCGCTCCGTGCAGGatggc
3848670	STM14_4397	STM3647	*yiaF*	3.43	12.8	43/2/0/0	-64	cgtgcTGGCGCAACAGTTGCAGAatgca
3929980	STM14_4482	STM3721[6]	*rfaP*	10.7	12.6	42/0/0/0	0	gactaTGGTACGTAAAATGCACGccggg
3996450	STM14_4561	STM3782		3.17	13.6	35/0/0/0	-8	gcgtaTGGTACGAATTTTGATGGaaggt

4049150	STM14_4625	STM3832		4.64	16.1	32/0/0/0	19	tttatTGGAATATTTATTGCTTAtctcg
4112560	STM14_4684	STM3887	*yieO*	3.47	13.4	43/0/0/0	0	ggcgaTGGCCTTCTTTATGCAGGcgcta
4146960	STM14_4722	STM3924[6]	*wecD*	15.6	14.8	38/2/0/0	20	cagcaTGGCGCGGAAATTGCACAcctta
4158750	STM14_4736	STM3937	*hemD*	4.98	7.73	43/0/0/0	80	gcaggTGGCCTGGAGTTTTCCACtgatt
4206460	STM14_4790	STM3984	*pepQ*	5.73	12.3	42/0/0/0	21	cgatgTGGCCGGGTTTATGCAGGatgat
4236140	STM14_4826	STM4013.S		4.64	11.9	43/0/0/0	-66	cgtaaAGGCCCGAAAAATGCATTtgcct
4257340	STM14_4850	STM4035	*fdoI*	3.99	14.9	41/0/38/10	64	gcaacTGGCGCGAATTCTGCACCcgttt
4450450	STM14_5070	STM4217		3.18	15.2	24/0/2/0	4	gggtaTGGCGCAATTTATGCCTGcgacg
4484310	STM14_5107	STM4249	*aphA*	4.22	12.7	39/0/0/0	-13	aatccGGGCACTAATGTTGCAAAactcg
4545620	STM14_5161	STM4290[6]	*proP*	16.1	19.2	13/0/0/0	39	gttatTGGCCTGATTTTTGCAGGtttgt
4576560	STM14_5193[8]	STM4319	*phoN*	3.07	11.6	43/0/0/0	-112	gcttaTGGTACACTTCTGGCATTagtat
4717060	STM14_5353	STM4460	*pyrB*	4.54	12.7	41/2/0/0	-39	agggaTGGCCTGGAGCCTGCATGgttct
4767620	STM14_5400	STM4500	*yjhP*	3.79	11.8	43/0/0/0	9	ccggtCGGCGAGACTTTTGCAGGcaagc
4801250	STM14_5445	STM4532	*yjiY*	3.03	11	43/0/0/0	32	ttcccCGGAACAAATCCTGCAAActgcg
4841050	STM14_5492	STM4572	*stjB*	7.41	10.6	43/0/0/0	0	atttcAGGCACAAAAAATGCGTGcgcaa
4851050	STM14_5502	STM4581	*yjjK*	8.08	13.5	42/2/0/0	0	gtggcTGGAACGTTTCCTGCACGacttc

Intragenic sites in antisense orientation relative to the associated ORF								
10190	STM14_0010	STM0010	*htgA*	5.78	13.6	19/0/0/0	50	gtcgcAGGCAGGCTATTTGCAAGaccgc
187000	STM14_0190	STM0158	*acnB*	3.81	13.9	36/0/0/0	9	ctgttTGGCATATTCATTGCCCGctttg
251480	STM14_0254	STM0214	*glnD*	3.58	10.1	43/0/0/0	23	acggtTGGCGCAGATCGTGCTGCaataa
265270	STM14_0266	STM0244	*yeaT*	9.02	15.3	33/0/0/0	17	cgagtTGGCATACTTTTCGCTGAtgaac
368750	STM14_0379	STM0322[6]	*proA*	9.26	16.5	34/0/0/0	7	gctcgCGGCACAGTTTATGCAAGcctgc
377020	STM14_0388	STM0332[6]		3.42	12.1	34/0/0/0	49	ttagcTGGCCAGAAATATGCTTAcgcag
378210	STM14_0389	STM0333		6.52	16.3	34/0/0/0	-59	ccgccTGGCAAGCATTGTGCAAGatgct
395770	STM14_0409	STM0351		3.59	12.1	36/0/0/0	0	gaatgTGGCTAAATTTGTGCATTagctt
450790	STM14_0468	STM0396	*sbcD*	11.1	11.8	43/0/0/0	8	gtagtTGGCACGCTTCCTGATACtgctg
510130	STM14_0535	STM0452	*cypD*	4.15	11.7	43/2/0/0	-42	cgtttTGGTCTGAATAATGCTGTaacga
516450	STM14_0542	STM0458		3.65	13.8	42/0/0/0	0	ggaggCGGCATAAATTTCGCAGGcgctc
519050	STM14_0544	STM0460	*mdlA*	4.27	16.5	43/0/0/0	42	cccacTGGCGCAAATTATGCAAAatctg
722570	STM14_0766	STM0658	*ybeV*	3.28	10.7	43/0/0/0	-239	cttttTGTCACGAATATCGCCTGgcttc
839020	STM14_0899	STM0774	*galK*	6.28	12.3	42/0/0/0	16	tcagtTGGCCCATACGTTGCAAAtcacc
891000	STM14_0954	STM0822	*ybiB*	3.15	12.9	43/0/0/0	-69	cgggtTGGCATAGACTTCGCCTTcggta
1045230	STM14_1125[8]	STM0995	*ycbB*	3.36	14.2	43/2/0/0	34	accttTGGCACAATTTGTGCGCTggatc

1355490	STM14_1512	STM1261		8.4	11.8	43/0/0/0	-19	ctcttCGGCACAGGCTTTGCAGAtatcc
1622960	STM14_1856	STM1538		4.09	8.94	43/0/0/0	-10	agaacTGGCGAGCTTTTTGGATGatgga
1674660	STM14_1906	STM1578	*narY*	4.09	14.7	39/0/0/0	0	cgggtTGGCGAAAATTTTGCTAAttacc
1682830	STM14_1918	STM1586		8.08	17.3	39/0/0/0	0	gatttTGGCAAGAATATTGCCATtacgg
1691470	STM14_1929	STM1594	*srfB*	6.27	15.4	39/0/0/0	-19	gccgcAGGCATATTTTTTGCCAGctgac
1769060	STM14_2012	STM1665[6]		6.1	16.5	33/0/0/0	-85	aaaccTGGCATCATTTTTTCAAGgccgt
1799990	STM14_2047	STM1697		5.2	13.9	28/0/0/0	28	actaaTGGCGCAGCAGTTGCATTcccta
1951920	STM14_2231	STM1844	*htpX*	3.32	9.25	20/0/0/0	-37	cgtcaGGGCTCATATTTTGCAACagacc
2035660	STM14_2345	STM1928	*otsA*	5.84	8.62	43/0/0/0	0	gggtcGGGAATGGAATATGCAGGaaaaa
2087570	STM14_2415	STM1993	*yedJ*	3.75	10.5	43/0/0/0	-19	gtggtAGGCGCAATTTTTGCGCTgaagc
2111930	STM14_2450	n/a[9]		3.02	9.78	19/0/0/0	12	tttgtAGGCATAAAAGTTGCCCCgcagg
2187850	STM14_2552	STM2062	*dacD*	4.45	11.4	43/0/0/0	88	ctatcCAGCACATATTTTGCTTTgatat
2232410	STM14_2593	STM2099	*wcaM*	3.49	13.1	14/0/0/0	73	gccatTGGTACAATTAATGCGTTcgatg
2452340	STM14_2830	STM2293		3.52	12.3	41/0/0/0	72	cagacCGGCATAATGATTGCTGCgccgt
2555600	STM14_2943	STM2392		3.1	9.77	43/0/0/0	-13	tgccgTAGAACGGAAGTTGCATAtaagg
2702030	STM14_3085	STM2517[6]	*sinH*	12.9	15.7	40/0/0/0	0	caggcTGGTACGGATCTTGCCATaaggt
2757400	STM14_3142	STM2563[6]	*yfhG*	9.66	15	42/0/0/0	44	tccggCGGCGTAATTTTTGCATCcgcca

3253330	STM14_3713	STM3072[6]		4.54	10.9	43/0/0/0	10	ggtttTGGCCCATTGAATGCATCcaggc
3275070	STM14_3737	STM3093	*endA*	3.92	10.8	43/0/56/10	-130	gcgccTGGCAAGCGCGTTGCACAtaggg
3305860	STM14_3781	STM3128		8.57	12.8	40/0/0/0	-51	tccttCGGCCAGGATTTTGCATAcaata
3393680	STM14_3887	STM3210	*dnaG*	3.09	10.6	43/2/0/0	26	tcgccCGGAATAACATATGCATAtggtc
3654190	STM14_4197	STM3485	*damX*	3.83	13.3	40/0/0/0	-6	cgaacTGGCGCGACGGTTGCAGGctctg
4444130	STM14_5062	STM4211		4.9	12.1	24/0/0/0	49	taataCGGCATACATTATGCGTActccc
4598110	STM14_5223	STM4343	*frdA*	3.1	8.56	43/2/0/0	-69	aagtcTGGAACAGGGTATGCAGCatatg
4714930	STM14_5349	STM4456	*mgtA*	3.44	12.8	30/0/0/0	0	cggacTGGAACAAAGTTTGCGTTctgg
4814370	STM14_5456	STM4541	*mdoB*	10.5	14.2	38/0/0/0	3	ggattAGGAATGGTTTTTGCTGGttctt
Intragenic site in 3' end of associated ORF and oriented toward intergenic space								
2919560	STM14_3325	STM2759		10.9	13.5	16/0/1/0	7	aacttTGGCTCGAATAATGCTACgcccg
Intragenic site in 3' end of associated ORF and oriented toward 5' end of adjacent ORF								
306420	STM14_0314	STM0267		3.71	12.9	32/0/0/0	-33	gtgacCGGCATGATTTTTGAACCtgaac
589380	STM14_0617	STM0527	*allC*	12.9	15.5	34/0/0/0	-57	aactcTGGCATTAATGCTGCATCaactt
812830	STM14_0863	STM0742	*ybgT*	5.86	10.1	43/2/57/10	26	attctGGGAACGCTTCTTGCCTGtgcat
1381780	STM14_1571	STM1294	*ansA*	3.83	12.3	43/2/0/0	-17	cgcgaTGGCGCAAAACTTGCGTGgtgaa
2009870	STM14_2315	STM1903	*yecE*	5.81	15.3	41/0/0/0	-49	cgcttTGGCATGATTTACGCAGCgtctt

2148990	STM14_2504	STM2016	*cobT*	6.46	15.1	38/0/0/0	-4	gtctaTGGAACCCTATTTGCATAtggcg
3694410	STM14_4237	STM3518	*rtcA*	4.55	16.2	30/0/0/0	99	ggtggTGGAACGGTTTTTGCCGGtcaga
4141830	STM14_4716	STM3918	*rfe*	10.5	12.7	43/2/56/10	79	tatgtTGGTGCTCTTTTTGCTAGcattc
Intragenic site in 5' end of associated ORF and oriented toward 5' end of adjacent ORF								
4421830	STM14_5035	STM4188.S	*metH*	3.92	12.8	43/0/0/0	0	catacTGGCACAAAGCGTGCTGGcggaa
Intragenic site in 3' end of associated ORF and oriented toward 3' end of adjacent ORF								
564850	STM14_0592	STM0504	*ybbM*	4.76	12.6	18/0/0/0	-32	gcaacTGGCATTGCAGTTGCACCaggtc
762120	STM14_0816	STM0699[6]		15.4	19	42/0/0/0	-51	cggacTGGCATCGATATTGCAAAcgcgc
901070	STM14_0971	STM0833	*ompX*	4.59	11.5	43/0/0/0	0	gacgtTGGCACCTGGATTGCTGGcgtag
1535860	STM14_1749	STM1451	*gst*	4.83	11.5	43/0/0/0	-59	cggtcTGGAAAAGAAATTGCAGTatgtg
2256600	STM14_2616	STM2120	*asmA*	3.37	9.03	43/0/0/0	-48	gcaagTGGATCAGTTATTGCGTAagcat
2578320	STM14_2964	STM2410	*yfeA*	3.33	11.9	24/2/0/0	-37	gatttCGGCACCGGTTATGCAAActatg
2765360	STM14_3147	STM2567	*yfhD*	5.36	9.67	43/0/0/0	42	aagttTGGTGGGGTATTTGCAGGaaaaa
3628930	STM14_4171	STM3465	*yhfA*	3.15	13.3	39/0/0/0	-95	tgatgTGGTATCAATCCTGCAAAagggc
3747980	STM14_4292	STM3565	*yhhK*	3.49	9.73	42/0/0/0	19	cgtgaTGGCCGCGTTTATGCAGGcgtta
3760170	STM14_4306	STM3578	*yhhP*	5.15	11.8	43/2/0/0	-90	ctttaTGGAACACGATTTGCTGGcgcaa
3802330	STM14_4349	STM3613	*yhjJ*	16.3	17.7	41/0/0/0	-71	ggagaTGGCATTAATTTTGCTGCaaccg

4086610	STM14_4659	STM3863[6]		9.46	17.1	42/0/0/0	-25	gctttTGGCGCGATTATTGCCAGcgtca
Intragenic site in 5' end of associated ORF and oriented toward 3' end of adjacent ORF								
566310	STM14_0593	STM0504	*ybbO*	4.22	9.5	43/2/0/0	89	atcggCGGCACGATCCACGCTTTcaggg
1000030	STM14_1085	STM0961[6]	*lolA*	8.32	14.4	42/2/0/0	0	aaaggTGGCATGAAAGCTGCTCActtta
1429140	STM14_1626	STM1339	*ihfA*	5.94	13.4	43/2/57/10	26	ttcttTGGCATCCCGCTTGCTAAgccca
2053450	STM14_2370	STM1952	*yecS*	4.97	7.62	43/2/0/0	15	ccagtTGGATACTTTCTTGCATTcgcgt
2162750	STM14_2521	STM2033[6]	*cbiC*	4.57	11.2	43/0/0/0	-22	ggtatCGGTATAAATAATGCACGgatgg
2224820	STM14_2585	STM2091	*rfbG*	6.37	10.6	43/0/0/0	37	acagtTGGCGCATCAAGTGCATAgcctt
2250330	STM14_2610	STM2115[6]	*wcaA*	3.96	9.68	43/0/0/0	0	caggtTGGCATATAAATTGAGATcagcg
3599740	STM14_4122	STM3419	*rpmJ*	3.7	13.1	43/2/56/10	-4	agcctTGGCGCTGTTTATGCTTCggctc
4459840	STM14_5080	STM4226[6]	*yjbA*	6	12.6	22/0/0/0	-2	tgctcAGGCGCGAATAATGCATCcgcca
4805060	STM14_5448	STM4534		6.17	9.74	43/0/0/0	-3	gggacGGGTATTGATTTTGCAGAgcacg

[1] Position (in 14028s genome) of peak maximum (from sliding window average plot by Chipeak)

[2] Signal Ratio for WT+DctD250/ΔrpoN+DctD250 (Materials and Methods); all p-values are <10-17

Conservation of the Eσ54 DNA binding site is defined here as the presence of the identical 18 bp Eσ54 core DNA binding sequence in the same genome context. The dataset for this analysis comprised the complete genomes from NCBI database for 43 S. enterica (Se), 2 S. bongori (Sb), 57 E. coli (Ec), and 10 Shigella species (Ss) strains. The number of strains with the

[3] conserved binding site for each indicated species are in order Se/Sb/Ec/Ss.

[4] Distance in bp upstream (positive values) or downstream (negative values) of binding site to peak maximum

[5] Predicted Eσ54 DNA binding site (core sequence in upper-case letters) based on PSSM score and proximity to the peak max

[6] Same binding site was predicted by Samuels, et al. (4)

[7] Site encompasses two previously identified promoters

[8] σ54 binding site overlaps (sense or antisense) σ70-type promoter for which TSS was determined by Kröger et al. (124)

[9] Locus is a phage gene not present in LT2

Table 4.2. σ^{54}-regulated transcriptome in *S.* Typhimurium expressing DctD250.

14028s Locus Tag	LT2 Locus Tag	Gene Name	Signal Ratio
Full ORFs with >2-fold increase in signal ratio[1]			
STM14_0431[2]	STM0368	*prpB*	3.09
STM14_0432	STM0369	*prpC*	2.68
STM14_0433	STM0370	*prpD*	2.35
STM14_0546[2]	STM0462	*glnK*	5.17
STM14_0668	STM0572		2.76
STM14_0669	STM0573		2.67
STM14_0670	STM0574		5.28
STM14_0671	STM0575		4.26
STM14_0672	STM0576		3.67
STM14_0673[2]	STM0577		6.86
STM14_0757[2]	STM0649.S		7.70
STM14_0758	STM0650	*uxaA*	6.02
STM14_0770	STM0662	*gltL*	2.36
STM14_0771	STM0663	*gltK*	2.79
STM14_0772	STM0664	*gltJ*	2.41
STM14_0772.P	STMsR025	*sroC*	2.75
STM14_0773[2]	STM0665	*gltI*	6.62
STM14_0962	STM0828	*glnQ*	2.11
STM14_0963	STM0829	*glnP*	2.38
STM14_0964[2]	STM0830	*glnH*	4.39

STM14_1556	STM1284	*yeaH*	2.15
STM14_1558[2]	STM1285	*yeaG*	4.10
STM14_1582[2]	STM1303	*astC*	2.77
STM14_1583	STM1304	*astA*	2.30
STM14_1584	STM1305	*astD*	2.25
STM14_2037	STM1687	*pspD*	2.84
STM14_2038	STM1688	*pspC*	6.76
STM14_2039	STM1689	*pspB*	5.71
STM14_2040[2]	STM1680	*pspA*	9.95
STM14_2900[2]	STM2354	*hisJ*	2.71
STM14_2901[2]	STM2355	*argT*	3.64
STM14_2905	STM2358		3.94
STM14_2906	STM2359		9.50
STM14_2907[2]	STM2360		11.0
STM14_3431[2]	STM2840	*norV*	9.65
STM14_3432	STM2841	*ygbD*	3.77
STM14_3436[2]	STM2843	*hydN*	3.45
STM14_3438	STM2844		2.07
STM14_3439	STM2845	*hycI*	3.73
STM14_3441	STM2847	*hycG*	2.51
STM14_3442	STM2848	*hycF*	3.47
STM14_3443	STM2849	*hycE*	2.58
STM14_3444	STM2850	*hycD*	2.28

STM14_3446	STM2852	*hycB*	2.26
STM14_3447			2.43
STM14_3448[2]	STM2853	*hycA*	2.68
STM14_3451[2]	STM2855	*hypB*	2.09
STM14_4238	STM3522	*rtcB*	9.05
STM14_4238.LP		*yrlA/yrlB*	14.6
STM14_4239[2]	STM3521	*rsr*	4.44
STM14_4543	STM3767	*dgaF*	2.88
STM14_4544	STM3768	*dgaE*	5.37
STM14_4545	STM3769	*dgaD*	9.18
STM14_4546	STM3770	*dgaC*	5.68
STM14_4547	STM3771	*dgaB*	5.57
STM14_4548[2]	STM3772	*dgaA*	9.23
STM14_4820[2]	STM4007	*glnA*	12.8
STM14_5013[2]	STM4172	*zraP*	12.2
STM14_5102[2]	STM4244	*pspG*	2.23
STM14_5449[2]	STM4535	*gfrA*	16.3
STM14_5450	STM4536	*gfrB*	15.9
STM14_5451	STM4537	*gfrC*	11.7
STM14_5452	STM4538	*gfrD*	9.19
STM14_5453	STM4539	*gfrE*	7.33
STM14_5454	STM4540.s	*gfrF*	9.32

Novel Intragenic transcripts with >2-fold increase in signal ratio

STM14_4722/4723		*wecD/ wecE*[3]	2.11
STM14_5161		*proP*[4]	3.05
Full ORFs with >2-fold decrease in signal [5]			
STM14_1795	STM1485		0.491
STM14_3638	STM3103	*lysA*	0.425
STM14_5085	STM4229	*malE*	0.475
STM14_5087	STM4231	*lamB*	0.464
Full ORFs with significant decrease in signal, but <2-fold [6]			
STM14_0816[7]	STM0699		0.63
STM14_0817	STM0700	*potE*	0.59
STM14_0818	STM0701	*speF*	0.74
STM14_1057[7]	STM0940	*ybjX*	0.78
STM14_2504[7]		*cobT*	0.89
STM14_2505		*cobS*	0.86
STM14_2506		*cobU*	0.86
STM14_2507		*cbiP*	0.68
STM14_2508		*cbiO*	0.66
STM14_2509.J		*cbiQ*	0.74
STM14_2510		*cbiN*	0.69
STM14_2511.RJ		*cbiM*	0.71
STM14_2512		*cbiL*	0.68
STM14_2513		*cbiK*	0.61
STM14_2514		*cbiJ*	0.66

STM14_2515	*cbiH*	0.57
STM14_2516	*cbiG*	0.68
STM14_2517	*cbiF*	0.67
STM14_2518	*cbiT*	0.65
STM14_2519	*cbiE*	0.66
STM14_2520	*cbiD*	0.78
STM14_2521[7]	*cbiC*	0.79
STM14_2523	*cbiA*	0.88
STM14_3143.P[7]	*glmY*	0.85
STM14_3809	*hybE*	0.81
STM14_3810	*hybD*	0.72
STM14_3811	*hybC*	0.75
STM14_3812	*hybB*	0.90
STM14_3813	*hybA*	0.85
STM14_3814[7]	*hypO*	0.83

[1]Signal Ratios (WT+DctD250/Δ*rpoN*+DctD250) based on averaged M-values for all probes across the ORF; all *p*-values are $<10^{-7}$ (LIMMA analysis). Alternating shading of rows delimits operons.

[2]Immediately downstream of identified $E\sigma^{54}$ binding site; see Table 4.1.

[3]Signal ratio based on average of M-values for probes immediately downstream of intragenic $E\sigma^{54}$ binding site (from position 4146957 to 4147350 with *p*-values <0.03) and all probes for *wecE* (LIMMA analysis *p*-value $<10^{-19}$).

[4]Signal ratio based on average of M-values for probes immediately downstream of intragenic $E\sigma^{54}$ binding site (from position 4545579 to 4545780 with *p*-values <0.01).

[5]Signal Ratios (WT+DctD250/Δ*rpoN*+DctD250) based on averaged M-values for all probes across the ORF; *p*-values <0.002 (LIMMA analysis). There are no $E\sigma^{54}$ binding sites associated with the ORFs.

[6]Signal Ratios (WT+DctD250/Δ*rpoN*+DctD250) based on averaged M-values for all probes across the ORF; *p*-values ≤0.01 (LIMMA analysis).

[7]$E\sigma^{54}$ binding site associated with ORF; see Table 4.1.

Table 4.3. Summary of σ^{54}-dependent transcriptome in *S.* Typhimurium expressing DctD250

Locus Tag 14028s	Locus Tag LT2	Gene Name	Signal Ratio[a]	Confirmed in *Salmonella*[b]	Conditions for TSS Detection in (182)[c]
First gene in operon exhibiting σ^{54}-dependent expression					
STM14_0431[d]	STM0368	*prpB*	3.09	(183)	ND
STM14_0546[d]	STM0462	*glnK*	5.17	(184)	H_2O_2 Shock
STM14_0673[d]	STM0577		6.86	(184)	ND
STM14_0757[d]	STM0649.S		7.70	(184)	ND
STM14_0773[d]	STM0665	*gltI*	6.62	This work	H_2O_2 Shock
STM14_0964[d]	STM0830	*glnH*	4.39	(185)	Mid-Exponential
STM14_1558[d]	STM1285	*yeaG*	4.10	This work	ND
STM14_1582[d]	STM1303	*astC*	2.77	(184)	H_2O_2 Shock
STM14_2040[d]	STM1680	*pspA*	9.95	(184)	0_2 shock
STM14_2900[d]	STM2354	*hisJ*	2.71	This work	ND
STM14_2901[d]	STM2355	*argT*	3.64	(184)	H_2O_2 Shock

STM14_2907[d]	STM2360		11.0	(184)	ND
STM14_3431[d]	STM2840	*norV*	9.65	(184)	ND
STM14_3436[d]	STM2843	*hydN*	3.45	(184)	Anaerobic
STM14_3448[d]	STM2853	*hycA*	2.68	(184)	ND
STM14_3450[d]	STM2854	*hypA*	1.75[e]	(184)	ND
STM14_4239[d]	STM3521	*rsr*	4.44	(184)	H_2O_2 Shock
STM14_4548[d]	STM3772	*dgaA*	9.23	(184,186)	ND
STM14_4820[d]	STM4007	*glnA*	12.8	(184)	H_2O_2 Shock
STM14_5013[d]	STM4172	*zraP*	12.2	(184)	ND
STM14_5102[d]	STM4244	*pspG*	2.23	This work	NaCl shock
STM14_5449[d]	STM4535	*gfrA*	16.3	(184,187)	ND
σ^{54}-dependent intragenic transcripts					
STM14_4722/	STM3924/	*wecD/*	2.11	This work	ND
STM14_4723[d]	STM3925	*wecE*			
STM14_5161	STM4290	*proP*	3.05	This work	ND

Genes exhibiting σ^{54}-dependent down-regulation					
STM14_1795	STM1485		0.491	This work	NA
STM14_3638	STM3103	*lysA*	0.425	This work	NA
STM14_5085	STM4229	*malE*	0.475	This work	NA
STM14_5087	STM4231	*lam*	0.464	This work	NA

[a]Microarray Signal Ratios (WT+DctD250/Δ*rpoN*+DctD250) for these genes, and for other genes in the same operon, are described in Table 4.1.

[b]References for demonstration of σ^{54}-dependent expression of the operon in *Salmonella.*

[c]Growth conditions under which Kröger et al. (182) detected a transcription start site (TSS) correlating to the σ^{54}-dependent promoter (ND, no correlating TSS was detected; NA, transcript is not from a σ^{54}-dependent promoter); if more than one growth condition activated the promoter, the condition resulting in the highest level of expression is given.

[d]Directly downstream of identified Eσ^{54} binding site; see Table 4.1 for binding sites.

[e]Second gene in this operon, *hypB*, has a signal ratio >two (Table 4.2).

Table 4.4. Summary of EMSA results for $E\sigma^{54}$ and σ^{54} binding to selected sites

Binding Site Associated ORFs (genes)	PSSM Score	$E\sigma^{54}$ Binding[a]	σ^{54} Binding[a]	$E\sigma^{54}$ DNA Binding Site Core Sequence[b]
STM14_0236 *(stfD*)	6.32	+	-	TGGCCGCTGTTTTGCTGC
STM14_0530 *(clpP)*	16.9	+++	-	TGTCACGTATTTTGCATG
STM14_0816	12.6	+++	+	TGGCATCGATATTGCAAA
STM14_0978 *(ybiT)*	13.2	++	-	TGGCAGGCTTTTTGATAT
STM14_1057 *(ybjX)*	17.1	+	-	TGGCCTGAATCTTGCTAA
STM14_2345 *(otsA)*	8.62	+++	-	GGGAATGGAATATGCAGG
STM14_2985 *(cysK)*	17.0	+++	-	TGGCATCACTGTTGCAGT
STM14_3815/3816 *(hypO/yghW)*	13.3	+++	-	TGGCTTTTATTTTGCACT
STM14_4295 *(rpoH)*	18.8	+	-	TGGCACGGTTGTTGCTCG
STM14_4722 *(wecD)*	14.8	+++	-	TGGCGCGGAAATTGCACA
STM14_4820 *(glnA)*	17.1	-	-	TGGCACAGATTTCGCTTT

STM14_5080 ***(yjbA)***	12.6	++	-	AGGCGCGAATAATGCATC
STM14_5161 ***(proP)***	19.2	+++	+	TGGCCTGATTTTTGCAGG
Control binding sites				
K. pneumoniae nifH049		+++	+	TGGTATGTTTTTTGCACT
-24TT,-12TT proP		-	-	TTTCCTGATTTTTTTAGG

[a] $E\sigma^{54}$ binding affinity for DNA site based on estimated protein concentration required for 50% binding of probe (see Materials and Methods): +++ (≤0.3 μM), ++ (>0.3 μM and ≤0.8 μM), + (>0.8 μM and ≤2 μM), - (>2 μM).

[b] Core 18 bp sequence for σ^{54}-dependent promoter (positions -9 to -26 from the TSS)

Figure 4.1: Illustrations of contextual positions of $E\sigma^{54}$ DNA binding sites in the *S.* Typhimurium genome and examples of potential regulatory roles with corresponding plot of microarray data.

Fig. 4.1A. The 184 $E\sigma^{54}$ binding sites identified by ChIP-chip analysis are grouped by position and orientation (blue arrows) relative to annotated ORFs (grey arrows).

Fig 4.1B. Six examples of potential regulatory roles for the $E\sigma^{54}$ binding sites are illustrated with the relative transcript levels (WT+DctD250/Δ*rpoN*+DctD250), as determined by microarray (red numbers and arrows for σ^{54}-dependent gene transcripts, green numbers and arrows for $\sigma^{70/38}$-dependent transcripts; see Table 4.2) and qRT-PCR (red or green numbers in brackets). The fold-enrichment in WT+DctD250 vs Δ*rpoN*+DctD250 in ChIP-chip analysis for each $E\sigma^{54}$ binding site (Table 4.1) is shown in blue, and an asterisk indicates that the binding site was confirmed by EMSA (see Table 4.4). Primary and secondary promoter designations are from Kröger et al. (125). Adjacent to each example is the WebArrayDB plot of microarray data for the genes whose transcription is positively or negatively regulated by σ^{54}. Each dot, which is the log base 2 transformed ratios (Δ*rpoN*+DctD250/WT+DctD250) for each probe (for all 3 biological replicates), is plotted on the x-axis by genome position (nucleotide positions are not shown). The dot colors indicate probe orientation and significance of *p*-values; red is positive strand with significant *p*-value, pink is positive strand without significant *p*-value, dark blue is negative strand with significant *p*-value, and light blue is negative strand without significant *p* value. The upward- and downward-carats designate the start and end, respectively, of a gene (name or 14028s locus number is given).

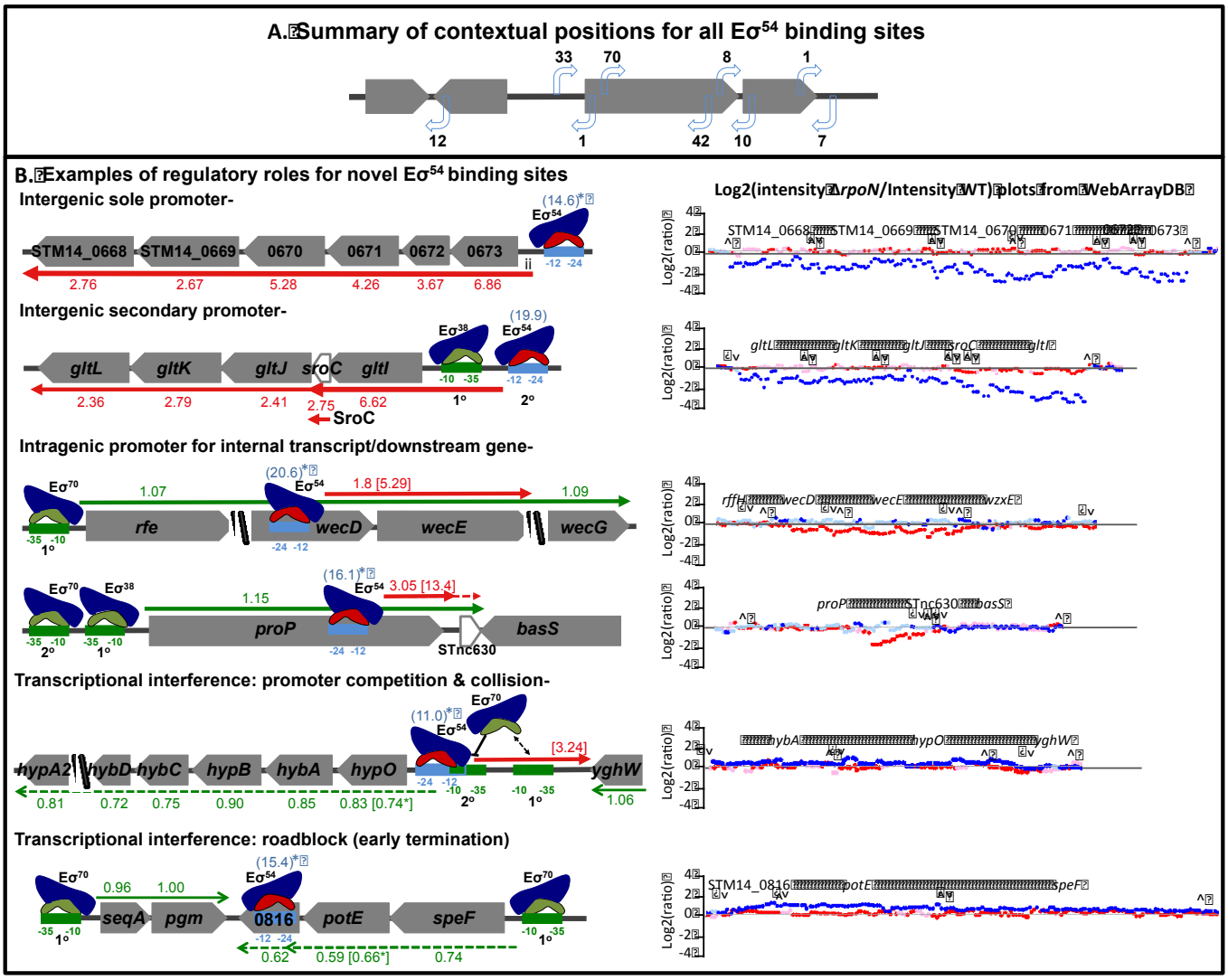
A. Summary of contextual positions for all Eσ54 binding sites
33
70
8
1
12
1
42
10
7
B. Examples of regulatory roles for novel Eσ54 binding sites
Log2(intensity ΔrpoN/Intensity WT) plots from WebArrayDB
Intergenic sole promoter-
(14.6)*
Eσ54
STM14_0668
STM14_0669
0670
0671
0672
0673
2.76
2.67
5.28
4.26
3.67
6.86
Intergenic secondary promoter-
(19.9)
Eσ38
Eσ54
gltL
gltK
gltJ
sroC
gltI
2.36
2.79
2.41
2.75
6.62
SroC
1°
2°
Intragenic promoter for internal transcript/downstream gene-
Eσ70
(20.6)*
Eσ54
1.07
1.8 [5.29]
1.09
rfe
wecD
wecE
wecG
1°
(16.1)*
3.05 [13.4]
1.15
proP
basS
STnc630
2°
Transcriptional interference: promoter competition & collision-
(11.0)*
[3.24]
hypA2
hybD
hybC
hypB
hybA
hypO
yghW
0.81
0.72
0.75
0.90
0.85
0.83 [0.74*]
1.06
Transcriptional interference: roadblock (early termination)
(15.4)*
0.96
1.00
seqA
pgm
0816
potE
speF
0.62
0.59 [0.66*]
0.74
Log2(ratio)

Figure 4.2: *In vitro* assays of $E\sigma^{54}$ and σ^{54} binding to DNA sequences identified in ChIP-chip analysis of $E\sigma^{54}$ genomic binding sites.

Fig. 4.2A. Representative EMSAs for binding reactions containing 16 nM ^{32}P-labeled 50 bp heteroduplex oligonucleotide probes (*P) for the positive control (*nifH*049 promoter), intragenic $E\sigma^{54}$ binding sites in *proP* and STM14_0816, or the intergenic $E\sigma^{54}$ binding site between *hypO* and *yghW,* with 0, 10, 50, 100, 200 nM $E\sigma^{54}$, 100 nM core RNAP, or 100 nM, 500 nM, 1 μM, 2 μM σ^{54} protein. Unbound probe and protein-DNA complexes, $E\sigma^{54}$-*P, σ^{54}-*P, or Core-*P (marked by arrows), were separated by native PAGE (6.5% acrylamide). Images are from Typhoon scans of gel-exposed phosphor-imager screens.

Fig. 4.2B. Examples of ChIPeak output from analysis of $E\sigma^{54}$ ChIP-chip data show the peaks for enriched probes within *proP* and STM14_0816 and in the intergenic regulatory region between *hypO* and *yghW*.

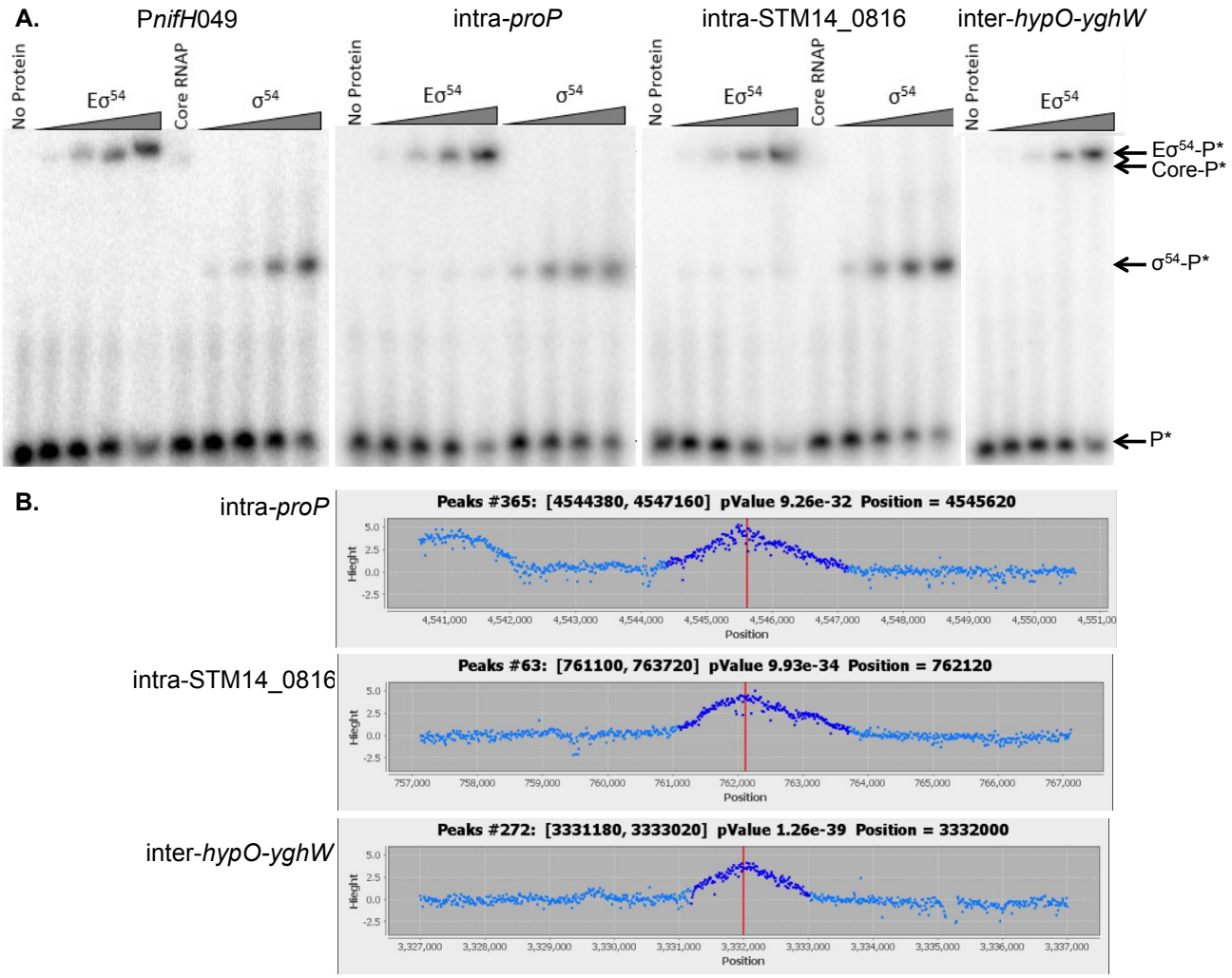
A.
P*nifH*049
intra-*proP*
intra-STM14_0816
inter-*hypO-yghW*
No Protein
Eσ54
Core RNAP
σ54
Eσ54-P*
Core-P*
σ54-P*
P*
B.
intra-*proP*
Peaks #365: [4544380, 4547160] pValue 9.26e-32 Position = 4545620
intra-STM14_0816
Peaks #63: [761100, 763720] pValue 9.93e-34 Position = 762120
inter-*hypO-yghW*
Peaks #272: [3331180, 3333020] pValue 1.26e-39 Position = 3332000
Hieght
Position

Figure 4.3: Multiple sequence alignments for $E\sigma^{54}$ binding sites from ChIP-chip analysis and $E\sigma^{54}$ binding sites characterized by EMSA. The relative frequency of bases at a given position based on multiple sequence alignments is illustrated by Weblogo (188,189).

Fig. 4.3A. Multiple sequence alignment for the 184 binding sites identified by ChIP-chip analysis.

Fig. 4.3B. Alignment of sequences for which $E\sigma^{54}$ has weak affinity (>800nM $E\sigma^{54}$ required for 50% binding of heteroduplex probes in EMSA).

Fig. 4.3C. Alignment of sequences for which $E\sigma^{54}$ has strong affinity (<350 nM $E\sigma^{54}$ required for 50% binding of heteroduplex probes in EMSA).

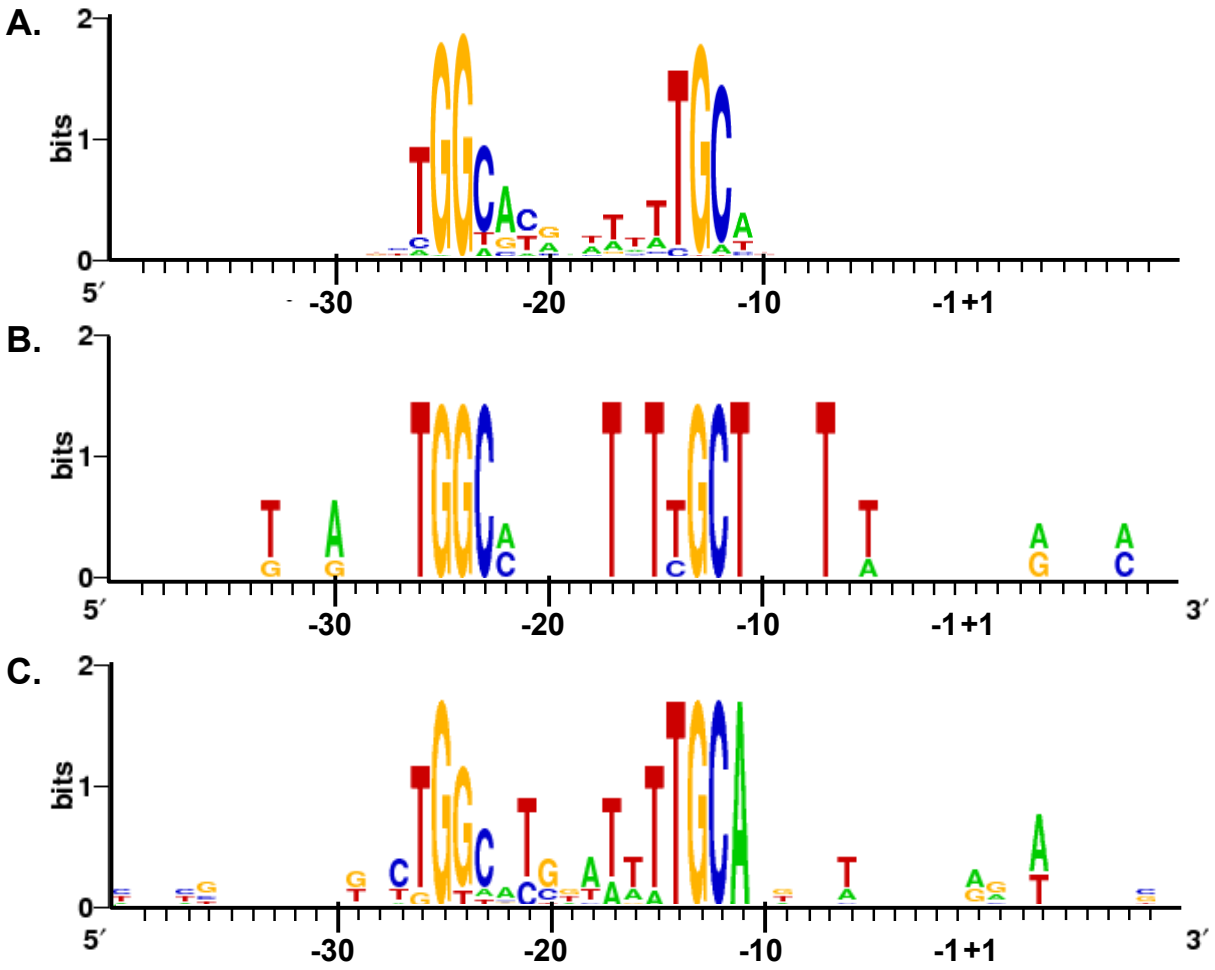
A.
B.
C.
bits
2
1
0
5′
3′
-30
-20
-10
-1+1

Figure 4.4: Alignment of σ^{54}-binding-**s**ites and comparison to a similar Eσ^{54}-binding sequence that does not bind σ^{54} in the absence of RNAP. Alignment of previously characterized sites, *K. p.* (*K. pneumoniae*) *nifH*049, *S. m.* (*S. meliloti*) *nifH* (23,30), with the newly identified *S.* T. (*S.* Typhimurium) *proP* and STM14_0816 sites that are bound by σ^{54} with high affinity. Bold nucleic acid residues indicate the -24 and -12 positions. The DNAseI footprint for σ^{54} in closed complex with *nifH*049 and *S. meliloti nifH* (23,30)is indicated by the black bar underneath the sequence. The consensus sequence was generated for the four high affinity sites using the single letter codes for nucleotides as defined by NCBI: M (A/C), R (A/G), W (A/T), S (C/G), Y (C/T), K (G/T), V (not T), H (not G), D (not C), B (not A), and N (any nucleotide). The consensus σ^{54} binding sequence is aligned with the inter-*hypO-yghW* sequence, which does not bind σ^{54}, but has the -14 to -17 T-tract that is shared by the known σ^{54} binding sites. Nucleic acid residues shared between the consensus sequence and non-binding sequence are struck through and colored red.


K. p. nifH049 TAAACAGGCACGGCTGGTATGTTTTTTGCACTTCTCTGCTGGCAAACACT
S. m. nifH TTATTTCAGACGGCTGGCACGACTTTTGCACGATCAGCCCTGGGCGCGCA
S. T. proP AACAGTAACGTTATTGGCCTGATTTTTGCacGTTTGTTGATGCTGGCGGT
STM14_0816 TTTCGCCACCGGACTGGCATCGATATTGCAAACGCGCGAGGAGATGCGCT
Consensus WWHHBHVRSVBKRYTGGYMYSDHTWTTGCAMDHBYVBBVNKRSDNRCRSW
WWHHBHVRSVBKRYTGGYMYSDHTWTTGCAMDHBYVBBVNKRSDNRCRSW
S. T. hypO-yghW CCGTTACGAAGACCTGGCTTTTATTTTGCACTGTTCGCGAAGAAGTTATT

CHAPTER 5

DISCUSSION

As an excellent model system for bacterial pathogen-host interactions, regulation of the *S.* Typhimurium transcriptome and proteome by both protein- and sRNA-mediated mechansims has been extensively studied (125,128,170,175,190,191). However, an essential regulon for *S.* Typhimurium pathogenesis, the RpoN (σ^{54}) regulon (7), has mostly gone uncharacterized due to the diverse (and sometimes unknown) conditions needed to activate the bacterial enhancer-binding proteins (bEBPs) required to interact with promoter-bound σ^{54}-holoenyzme ($E\sigma^{54}$) and hydrolyze ATP to initiate transcription (5). Utilizing the constitutively-active and promiscuous bEBP variant, DctD250, to promote open complex formation at σ^{54}-dependent promoters in the *S.* Typhimurium 14028s genome, ChIP-chip and microarray analyses revealed 184 $E\sigma^{54}$ DNA binding sites and 24 σ^{54}-dependent operon transcripts (including 65 genes) and intragenic transcripts, as well as, nine transcripts that are down-regulated through σ^{54}-dependent processes (see Tables 4.1 and 4.2).

The 184 $E\sigma^{54}$ genomic binding sites are predominantly within annotated ORFs and in the sense orientation relative to the genes in which they are encoded (49% of total sites); 29% of the sites are also intragenic, but orientated antisense to the gene, and 22% of sites are in intergenic sequence (see Fig. 4.1A). The position and context of the $E\sigma^{54}$ binding sites suggest potential roles in transcription regulation, ranging from a promoter

expressing sRNA to directing transcription interference. Transcription interference mechanisms include competition for binding at overlapping promoter sequences (competition), convergent transcription from active promoters (collision), elongating transcription complexes inhibiting binding to a promoter (occlusion), or transcription disruption by collision with a holoenzyme that is bound to the DNA, but has not initiated transcription (sitting duck interference or roadblock mechanisms) (12); these regulatory mechanisms allow the σ^{54} regulon to intersect and impact the regulons of σ^{70} and other alternative sigma factors. Antagonistic or synergistic interactions between sigma regulons can occur on many levels, from sigma factor association with core RNAP to transcript stability, and influence gene expression under changing growth conditions (10,11,13,17,192,193). Previously described examples for mechanisms by which σ^{54}-dependent promoters directly control expression from σ^{70}-type promoters include: 1) σ^{54}-dependent promoter competition with an overlapping σ^{70} promoter for the sRNA gene *glmY* in *S.* Typhimurium and *E. coli* (127) and for *crl* whose product controls σ^{38} (RpoS) interactions with core RNA polymerase in *E. coli* (134) and 2) σ^{54}-dependent promoter activation of an adjacent, non-overlapping σ^{70}-dependent promoter for *dmpR* in *Pseudomonas putida* (135). Supporting evidence for the regulatory activities of novel or confirmed $E\sigma^{54}$ binding sites from this study are summarized in Figure 4.1B and a few examples are discussed here.

Intergenic secondary promoter controlling mRNA and sRNA expression: A σ^{54}-dependent promoter was previously predicted for the *gltIJKL* operon in *S.* Typhimurium (3,4), and early microarray studies in *E. coli* indicated that *gltIJKL* is regulated by NtrC (172). Our ChIP-chip and microarray results for *S.* Typhimurium expressing DctD250

identified the σ^{54}-dependent promoter and confirmed σ^{54}-dependent expression of the *gltI_sroC_gltJKL* operon and SroC sRNA. Two recently published studies demonstrate that in *S*. Typhimurium SroC is processed from a transcript that terminates between *gltI* and *gltJ* (194) and that the level of SroC is positively regulated by Eσ^{38} following late-exponential phase growth (195). The RNA-seq and differential RNA-seq (dRNA-seq) analyses by Kröger et al. (125) of the *S*. Typhimurium infection-related transcriptomes showed that the *gltI-sroC-gltJ*KL operon has a primary transcript expressed during stationary phase, which maps to the σ^{38}-dependent promoter, and has a secondary transcript that maps to the σ^{54}-dependent promoter and is expressed under conditions that activate genes in the *Salmonella* Pathogenicity Island 2 (SPI2), including low phosphate/low pH medium (PCN medium), peroxide shock, and nitric oxide shock. The co-regulation of the operons by σ^{54} and σ^{38} is common in *E. coli*; in microarray analyses of σ^{54} and σ^{38} regulated genes in *E. coli* by Dong et al. (192), ~60% of genes in the σ^{54}-regulon are also controlled by σ^{38}, and σ^{54} negatively regulates the level of σ^{s} in the cell. The regulation of *gltI-sroC-gltJKL* is not only linked with the σ^{38} regulon, but also with the GcvB posttranscriptional regulon; Miyakoshi et al. (194) demonstrated that SroC acts as a "RNA sponge" by sequestering the sRNA, GcvB, which posttranscriptionally represses numerous mRNAs encoding amino acid transporters and metabolism genes in *S.* Typhimurium, including the *gltI-sroC-gltJKL* mRNA (128). Thus, σ^{54}-regulated SroC positively regulates expression from *gltI-sroC-gltJKL* mRNA *in cis* and relieves GcvB-mediated repression *in trans* for other mRNAs encoding amino acid transporters and biosynthetic enzymes.

Intragenic promoter for a downstream gene or an intragenic transcript: The σ^{54}-dependent transcript generated from the newly identified intragenic $E\sigma^{54}$ binding site in *proP* (Table 4.1, Table 4.2, Fig. 4.1B, Fig. 4.2) may be processed to generate a sRNA, which is supported by the enrichment of *proP* transcript in Hfq immunoprecipitation (175). However, in the analysis of the *S.* Typhimurium transcriptome by Kröger et al. (125), an intergenic TSS is identified for a sRNA STnc630, which maps to the 3´ UTR of the *proP* transcript. There is substantial evidence that many sRNAs are processed from the 3´ UTRs of mRNAs in *Salmonella, E. coli, Vibrio cholera,* and *Streptomyces coelicolor* (171). It is possible that two sRNAs are generated from this region or, more likely, the same sRNA is generated by two mechanisms: transcription from the intragenic promoter and processing of the 3' UTR of the full-length *proP* transcript.

Transcriptional Interference by promoter competition or transcription collision: The Hydrogenase 2 operon (*hypOhybA-E*) is expressed from two σ^{70} promoters, which are regulated by multiple transcription factors under anaerobic and aerobic growth (125,176). The $E\sigma^{54}$ DNA binding site identified upstream of the *hypOhybA-E* operon (Table 4.1, Table 4.4, Fig. 4.1B, Fig. 4.2) operon overlaps the -10 region of the secondary σ^{70} promoter for the operon and is oriented toward the intergenic space such that its transcription is on a collision course with the primary σ^{70} promoter for the *hypOhybA-E*. Sigma54-dependent transcription from the $E\sigma^{54}$ binding site was demonstrated by qRT-PCR in the presence of DctD250 and expression of the *hypOhybA-E* operon was significantly down-regulated in the WT strain expressing DctD250 as compared to the Δ*rpoN* strain (Fig. 4.1B). These results suggest that transcription interference of the secondary and primary promoters for *hypOhybA-E* may occur by promoter competition

or collision, respectively, with the intergenic σ^{54}-dependent promoter. Growth conditions that activate transcription of the intergenic σ^{54}-dependent promoter in the absence of DctD250 have not yet been identified.

The four genes that showed greater than two-fold down regulation in WT S. Typhimurium (+DctD250) as compared to the Δ*rpoN* strain (+DctD250) were STM14_1795, *lysA*, *malE*, and *lamB* (Table 4.1). Since none of these genes is associated with an $E\sigma^{54}$ DNA binding site, the σ^{54}-dependent regulation must be indirect. STM14_1795 is annotated in NCBI as encoding an acid shock protein precursor that is required for growth in moderately acid conditions, but the mechanism for regulation of this gene has not be reported so the impact of the σ^{54} regulon on expression of STM14_1795 cannot be predicted. A potential mechanism for the σ^{54}-dependent decrease in *lysA* transcription is more evident. The *lysA* gene encodes a diaminopimelate decarboxylase that catalyzes decarboxylation of diaminopimelate to lysine. Expression of *lysA* is activated by LysR in the presence of diaminopimelate and is repressed in the presence of lysine (172). The σ^{54}-dependent gene *argT* (STM14_2901), which is a lysine/arginine/ornithine transport protein, is highly expressed in the WT+DctD250 strain, due to both direct transcription of *argT* from a σ^{54}-dependent promoter (Table 4.2), and relief of GcvB negative regulation by SroC (194), which is also expressed from a σ^{54}-dependent promoter (Table 4.2). The increased levels of lysine due to ArgT lysine transporter activity in the lysine-rich Nutrient Broth medium is likely to result in repression of *lysA* expression. It should be noted that there is another σ^{54}-dependent gene that is annotated as a diaminopimelate decarboxylase, STM14_2907 (Table 4.2), but the product of this gene was recently shown to have deaminase activity rather than

decarboxylase activity (K. Miller and T. Hoover, personal communication). The remaining two genes that exhibited σ^{54}-dependent down regulation are *malE*, which encodes a periplasmic protein involved in maltose transport, and *lamB*, which encodes a porin involved in transport of maltose and maltodextrins (173). These genes are encoded in divergently transcribed operons whose σ^{70}-dependent promoters are activated by MalT when bound by ATP and maltotriose; transcription of *malT* is positively regulated by cAMP and CRP (173). Three σ^{54}-dependent operons encode different mannose family phosphosphotransferase systems (PTS) and associated enzymes, and substrates have been identified for two of these mannose family PTS, D-glucosamine [*dgaABCDEF*; (145))] and fructoselysine/glucoselysine [*gfrABCDEF*; (166)]. Expression of one or more of these PTS operons may result in decreased expression of *malE* and *lamB* through a catabolite repression mechanism (174).

Upon noting significant differences in the potential regulatory $E\sigma^{54}$ binding sites recently characterized in *E. coli* (179) and those identified in this study, the extent of shared sites between *S.* Typhimurium and some of the closest relatives among the Enterobacteriaceae was determined. Our assessment of conservation of $E\sigma^{54}$ DNA binding sites in the Enterobacteriaceae, which was limited to 112 members of the family and required identity between orthologous sites, suggests that it is rare for the regulatory sites to be conserved in multiple species (~5% of the 184 $E\sigma^{54}$ binding sites are conserved in *S. enterica, S. bongori, E. coli,* and *Shigella* strains). A single base pair change in the sequence of a promoter, or associated enhancer sequences for the cognate bEBP, can drastically change regulation. J. Mrázek and Y. Huang, collaborators on this project, are developing software to assess the evolution of regulatory motifs in prokaryotic genomes

and determine the selective constraints on regulatory sites evolving to respond to changes in its environment. This analysis of the $E\sigma^{54}$ binding sites may provide some insight into the acquisition or loss of promoter elements or contextual features that allow σ^{54} binding in the absence of RNAP to a σ^{54}-dependent promoter, as well as the functional significance of this unusual sigma factor property.

Concluding Remarks

Regulation of transcription is one of the primary mechanisms of controlling response to stimuli. Our work has identified a large number of σ^{54}-binding sites and postulated some potential regulatory roles for these sites. We also provided further understanding of sequence elements required for $E\sigma^{54}$ binding and expanded the number of identified sites capable of binding σ^{54} in the absence of core RNA polymerase.

Further studies analyzing σ^{54} in the presence or absence of core RNA polymerase would improve our understanding of protein-DNA interactions. Despite 85% sequence identity and 94% amino acid identity, there is little overlap in studies assessing $E\sigma^{54}$ binding in *S.* Typhimurium and *E. coli*. This small number of sequence differences appears to strongly effect binding sequence selection either through differences in holoenzyme conformation or direct interaction with the DNA. Site-directed mutagenesis and assessment of binding to different sequences could provide futher insight into what makes these very similar proteins behave differently.

Improvements in the techniques used to assess bacterial transcriptomes and proteomes has dramatically expanded our understanding of regulatory elements and their complex interaction, cooperation, and redundancies. These data assist in building a more complete picture of *Salmonella* regulatory systems. Future studies will explore the physiologically relevant conditions under which the novel σ^{54}-promoters regulate transcription in overlapping sigma regulons. In the absence of enhancer sequences or a bEBP able to activate transcription directly from the cytoplasm, $E\sigma^{54}$ interacting with the DNA will remain transcriptionally silent. The primary task is to locate enhancer

sequecnes across the genome and their position relative to $E\sigma^{54}$ binding sites, most easily by ChIP-seq. Binding sites in close proximity to enhancer are likely activated by the corresponding bEBP under physiologically relevant conditions. Those binding sites lacking a near-by enhancer will be further assessed for their use in non-transcriptionally active tasks: inhibition of transcription from other holoenzymes and sequestration of core RNA polymerase. These future studies will improve our understanding of transcription regulation and bacterial mechanisms of response to stimuli.

REFERENCES

1. Merrick, M. (1993) In a class of its own - the RNA polymerase sigma factor s^{54} (s^{N}). *Mol Microbiol*, **10**, 903-909.
2. Barrios, H., Valderrama, B. and Morett, E. (1999) Compilation and analysis of s^{54}-dependent promoter sequences. *Nuc Acids Res*, **27**, 4305-4313.
3. Studholme, D. (2002) Enhancer-dependent Transcription in *Salmonella enterica* Typhimurium: New Members of the s^{N} Regulon Inferred from Protein Sequence Homology and Predicted Promoter Sites. *J Mol Microbiol Biotechnol*, **4**, 367-374.
4. Samuels, D., Frye, J., Porwollik, S., McClelland, M., Mrazek, J., Hoover, T. and Karls, A. (2013) Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the s^{54} (RpoN) regulon of *Salmonella* Typhimurium LT2. *BMC Genomics*, **14**.
5. Bush, M. and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of s^{54}-dependent transcription. *Microbiol Mol Biol Rev*, **76**, 497-529.
6. CDC. (2014) Pathogens causing US foodborne illnesses, hospitalizations, and deaths, 2000–2008.
7. Chaudhuri, R., Morgan, E., Peters, S., Pleasance, S., Hudson, D., Davies, H., Wang, J., vanDiemen, P., Buckley, A., Bowen, A. *et al.* (2013) Comprehensive Assignment of Roles for *Salmonella* Typhimurium Genes in Intestinal Colonization of Food-Producing Animals. *PLoS Genet*, **9**, e1003456.

8. Swords, W., Cannon, B. and WH Benjamin, J. (1997) Avirulence of LT2 Strains of *Salmonella typhimurium* Results from a Defective *rpoS* Gene. *Infect Immun*, **65**, 2451-2453.

9. Wilmes-Riesenberg, M., Foster, J. and III, R.C. (1997) An altered *rpoS* allele contributes to the avirulence of *Salmonella typhimurium* LT2. *Infect Immun*, **65**, 203-210.

10. Mauri, M. and Klmpp, S. (2014) A Model for Sigma Factor Competition in Bacterial Cells. *PLoS Comput Biol*, **10**, e1003845.

11. Qiu, Y., Nagarajan, H., Embree, M., Shieu, W., Abate, E., Juarez, K., Cho, B.K., Elkins, J.G., Nevin, K.P., Barrett, C.L. *et al.* (2013) Characterizing the interplay between multiple levels of organization within bacterial sigma factor regulatory networks. *Nat Commun*, **4**, 1755.

12. Shearwin, K., Callen, B. and Egan, J. (2005) Transcriptional Interference - A Crash Course. *Trends Genet*, **21**, 339-345.

13. Grigorova, I., Phleger, N., Mutalik, V. and Gross, C. (2006) Insights into transcriptional regulation and s competition from an equilibrium model of RNA polymerase binding to DNA. *Proc Natl Acad Sci USA*, **103**, 5332-5337.

14. Laurie, A., Bernardo, L., Sze, C., Skarfstad, E., Szalewska-Palasz, A., Nystrom, T. and Shingler, V. (2003) The Role of the Alarmone (p)ppGpp in s^N Competition for Core RNA Polymerase*. *J Biol Chem*, **278**, 1494-1503.

15. Crick, F. (1970) Central Dogma of Molecular Biology. *Nature*, **227**, 561-563.

16. deHaseth, P., Lohman, T., Burgess, R. and Jr., M.R. (1978) Nonspecific interactions of *Escherichia coli* RNA polymerase with native and denatured

DNA: differences in the binding behavior of core and holoenzyme. *Biochemistry*, **17**, 1612-1622.

17. Feklistov, A., Sharon, B., Darst, S. and Gross, C. (2014) Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective. *Annu Rev Microbiol*, **68**, 357-376.

18. Garcia, E., Bancroft, S., Rhee, S. and Kustu, S. (1977) The product of a newly identified gene, *glnF*, is required for synthesis of glutamine synthetase in *Salmonella*. *Proc Natl Acad Sci USA*, **74**, 1662-1666.

19. Francke, C., Kormelink, T., Hagemeijer, Y., Overmars, L., Sluijter, V., Moezelaar, R. and Siezen, R. (2011) Comparative analyses imply that the enigmatic sigma factor 54 is a central controller of the bacterial exterior. *BMC Genomics*, **12**.

20. Studholme, D. and Buck, M. (2000) The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences. *FEMS Microbiol Lett*, **189**, 1-9.

21. Struhl, K. (1999) Fundamentally Different Logic of Gene Regulation in Eukaryotes and Prokaryotes. *Cell*, **98**.

22. Wang, W., Carey, M. and Gralla, J. (1992) Polymerase II Promoter Activation: Closed Complex Formation and ATP-Driven Start Site Opening. *Science*, **255**, 450-453.

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

24. Ninfa, A., Mullin, D., Ramakrishnan, G. and Newton, A. (1989) *Escherichia coli* sigma 54 RNA polymerase recognizes *Caulobacter crescentus flbG* and *flaN* flagellar gene promoters *in vitro*. *J Bacteriol*, **171**, 383-391.

25. Cannon, W., Gallegos, M., Casaz, P. and Buck, M. (1999) Amino-terminal sequences of sigmaN (sigma54) inhibit RNA polymerase isomerization. *Genes Dev*, **13**, 357-370.

26. Cannon, W., Missailidris, S., Smith, C., Cottier, A., Austin, S., Moore, M. and Buck, M. (1995) Core RNA polymerase and promoter DNA interactions of purified domains of sigma N: bipartite functions. *J Mol Biol*, **248**, 781-803.

27. Gallegos, M.-T. and Buck, M. (1999) Sequences in s^N Determining Holoenzyme Formation and Properties. *J Mol Biol*, **288**, 539-553.

28. Wang, L. and Gralla, J. (1998) Multiple in vivo Roles for the -12-region Elements of Sigma 54 Promoters. *J Bacteriol*, **180**, 5626-5631.

29. Popham, D., Keener, J. and Kustu, S. (1991) Purification of the Alternative s Factor, s^{54}, from *Salmonella typhimurium* and Characterization of s^{54}-Holoenzyme*. *J Biol Chem*, **266**, 19510-19518.

30. Morris, L., Cannon, W., Claverie-Martin, F., Austin, S. and Buck, M. (1994) DNA Distortion and Nucleation of Local DNA Unwinding within Sigma-45 (s^N) Holoenzyme Closed Promoter Complexes. *J Biol Chem*, **269**, 11563-11571.

31. Buck, M. (1986) Deletion Analysis of the *Klebsiella pneumoniae* Nitrogenase Promoter: Importance of Spacing between Conserved Sequences around Positions -12 and -24 for Activation by the *nifA* and *ntrC (glnG)* Products. *J Bacteriol*, **166**, 545-551.

32. Chen, Y. and Helmann, J. (1995) The *Bacillus subtilis* flagerllar regulatory protein s^D: overproduction, domain analysis and DNA-binding properties. *J Mol Biol*, **249**, 743-753.

33. Sevim, E., Gaballa, A., Belduz, A. and Helmann, J. (2011) DNA-binding Properties of the *Bacillus subtilis* and *Aeribacillus pallidus* AC6 s^D Proteins. *J Bact*, **193**, 575-579.

34. Cannon, W., Claverie-Martin, F., Austin, S. and Buck, M. (1993) Core RNA Polymerase assists binding of the transcription factor s^{54} to promoter DNA. *Mol Microbiol*, **8**, 287-298.

35. Wang, X., Kolb, A., Cannon, W. and Buck, M. (1997) Nucleoprotein complex formation by the enhancer binding protein NifA. *Nuc Acids Res*, **25**, 3478-3485.

36. Cannon, W., Austin, S., Moore, M. and Buck, M. (1995) Identification of close contacts between the s^N (s^{54}) protein and promoter DNA in closed promoter complexes. *Nuc Acids Res*, **23**, 351-356.

37. Bose, D., Pape, T., Burrows, P., Rappas, M., Wigneshweraraj, S., Buck, M. and Zhang, X. (2008) Organization of an Activator-Bound RNA Polymerase Holoenzyme. *Mol Cell*, **32**, 337-346.

38. Cannon, W., Wigneshweraraj, S.R. and Buck, M. (2002) Interactions of regulated and deregulated forms of the sigma54 holoenzyme with heteroduplex promoter DNA. *Nucleic acids research*, **30**, 886-893.

39. Popham, D., Szeto, D., Keener, J. and Kustu, S. (1989) Function of a Bacterial Activator Protein that Binds to Transcriptional Enhancers. *Science*, **243**, 629-635.

40. Lee, H., Ishihama, A. and Kustu, S. (1993) The C terminus of the alpha subunit of RNA polymerase is not essential for transcriptional activation of sigma 54 holoenzyme. *J Bacteriol*, **175**, 2479-2482.

41. Ishihama, A. (2000) Functional Modulation of *Escherichia coli* RNA Polymerase. *Annu Rev Microbiol*, **54**, 499-518.

42. Gruber, T. and Gross, C. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol*, **57**, 441-466.

43. Maeda, H., Fujita, N. and Ishihama, A. (2000) Competition among seven *Escherichia coli* s subunits: relative binding affinities to the core RNA polymerase. *Nuc Acids Res*, **28**, 3497-3503.

44. Scott, D., Ferguson, A., Gallegos, M., Pitt, M., Buck, M. and Hoggett, J. (2000) Interaction of sigma factor sN with *Escherichia coli* RNA polymerase core enzyme. *Biochem J*, **352**, 539-547.

45. Hicks, K. and Grossman, A. (1996) Altering the level and regulation of the major sigma subunit of RNA polymerase affects gene expression and development in *Bacillus subtilis*. *Mol Microbiol*, **20**, 201-212.

46. Brown, K. and Hughes, K. (1995) The role of anti-sigma factors in gene regulation. *Mol Microbiol*, **16**, 397-404.

47. Chatterji, D. and Ojha, A. (2001) Revisiting the stringent response, ppGpp, and starvation signaling. *Curr Opin Microbiol*, **4**, 160-165.

48. Bernardo, L., Johansson, L., Skärfstad, E. and Shingler, V. (2009) s^{54}-Promoter Discrimination and Regulation by ppGpp and DksA. *J Biol Chem*, **284**, 828-838.

49. Jishage, M., Kvint, K., Shingler, V. and Nyström, T. (2002) Regulation of s factor competition by the alarmone ppGpp. *Genes Dev*, **16**, 1260-1270.

50. Perederina, A., Svetlov, C., Vassylyeva, M., Tahirov, T., Yokoyama, S., Artsimovitch, I. and Vassylyev, D. (2004) Regulation through the Secondary Channel-Structural Framework for ppGpp-DksA Synergism during Transcription. *Cell*, **118**, 297-309.

51. Maki, Y., Yoshida, H. and Wada, A. (2000) Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase *Escherichia coli*. *Genes Cells*, **5**, 965-974.

52. Ancona, V., Li, W. and Zhao, Y. (2014) Alternative sigma factor RpoN and its modulation protein YhbH are indispensable for *Erwinia amylovora* virulence. *Mol Plant Pathol*, **15**, 58-66.

53. Yang, Y., Darbari, V., Zhang, N., Lu, D., Glyde, R., Wang, Y., Winkelman, J., Gourse, R., Murakami, K., Buck, M. *et al.* (2015) Structures of the RNA polymerase-s^{54} reveal new and conserved regulatory strategies. *Science*, **349**, 882-885.

54. Ninfa, A., Reitzer, L. and Magasanik, B. (1987) Initiation of transcription at the bacterial *glnAp2* promoter by purified E. coli components is facilitated by enhancers. *Cell*, **50**, 1039-1046.

55. Wang, F. and Greene, E.C. (2011) Single-molecule studies of transcription: from one RNA polymerase at a time to the gene expression profile of a cell. *Journal of molecular biology*, **412**, 814-831.

56. Berg, O.G., Winter, R.B. and von Hippel, P.H. (1981) Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*, **20**, 6929-6948.

57. Friedman, L., Mumm, J. and Gelles, J. (2013) RNA polymerasse approaches its promoter without long-range sliding along DNA. *Proc Natl Acad Sci USA*, **110**, 9740-9745.

58. Macchi, R., Montesissa, L., Murakami, K., Ishihama, A., De Lorenzo, V. and Bertoni, G. (2003) Recruitment of sigma54-RNA polymerase to the Pu promoter of Pseudomonas putida through integration host factor-mediated positioning switch of alpha subunit carboxyl-terminal domain on an UP-like element. *The Journal of biological chemistry*, **278**, 27695-27702.

59. Bertoni, G., Fujita, N., Ishihama, A. and Lorenzo, V.d. (1998) Active recruitment of s^{54}-RNA polymerase to the *Pu* promoter of *Pseudomonas putida*: role of IHF and the aCTD. *EMBO*, **17**, 5120-5128.

60. Wassem, R., Souza, E.d., Yates, M., Pedrosa, F. and Buck, M. (2000) Two roles for integration host factor at an enhancer-dependent *nifA* promoter. *Mol Microbiol*, **35**, 756-764.

61. Shingler, V. (2011) Signal sensory systems that impact s^{54}-dependent transcription. *FEMS Microbiol Rev*, **35**, 425-440.

62. Hooper, S. and Böck, A. (1995) Effector-mediated stimulation of ATPase activity by the sigma 54-dependent transcriptional activator FHLA from *Escherichia coli*. *J Bacteriol*, **177**, 2798-2803.

63. Su, W., Porter, S., Kustu, S. and Echols, H. (1990) DNA-looping and enhancer activity: association between DNA-bound NtrC activator and RNA polymerase at the bacterial glnA promoter. *Proc Natl Acad Sci USA*, **87**, 5504-5508.

64. Kustu, S., Santero, E., Keener, J., Popham, D. and Weiss, D. (1989) Expression of s^{54} (*ntrA*)-Dependent Genes is Probably United by a Common Mechanism. *Microbiol Revs*, **53**, 367-376.

65. Buck, M., Khan, H. and Dixon, R. (1985) Site-directed mutagenesis of the *Klebsiella pneumoniae nifL* and *nifH* promoters and *in vivo* analysis of promoter activity. *Nuc Acids Res*, **13**, 7621-7638.

66. Buck, M., Miller, S., Drummond, M. and Dixon, R. (1986) Upstream activator sequences are present in the promoters of nitrogen fixation genes. *Nature*, **320**, 374-378.

67. Belitsky, B. and Sonenshein, A. (1999) An enhancer element located downstream of the major glutamate dehydrogenase gene of *Bacillus subtilis*. *Proc Natl Acad Sci USA*, **96**, 10290-10295.

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

69. Ray, V., Eddy, J., Hussa, E., Misale, M. and Visick, K. (2013) The *syp* Enhancer Sequence Plays a Key Role in Transcriptional Activation by the s54-Dependent Response Regulator SypG and in Biofilm Formation and Host Colonization by *Vibrio fischeri*. *J Bacteriol*, **195**, 5402-5412.

70. Buck, M., Cannon, W. and Woodcock, J. (1987) Transcriptional activation of the *Klebsiella pneumoniae* nitrogenase promoter may involve DNA loop formation. *Mol Microbiol*, **1**, 243-249.

71. Whitehall, S., Austin, S. and Dixon, R. (1993) The function of the upstream region of the s^{54}-dependent *Klebsiella pnuemoniae nifL* promoter is sensitive to DNA supercoiling. *Mol Microbiol*, **9**, 1107-1117.

72. Mehta, P., Jovanovic, G., Lenn, T., Bruckbauer, A., Engl, C., Ying, L. and Buck, M. (2013) Dynamics and stoichiometry of a regulated enhancer-binding protein in live *Escherichia coli* cells. *Nat Commun*, **4**.

73. Reidel, G., Brown, S. and Ausubel, F. (1983) Nitrogen fixation by *Klebsiella pneumoniae* is inhibited by certain multicopy hybrid *nif* plasmids. *J Bacteriol*, **153**, 45-56.

74. Beynon, J., Williams, M. and Cannon, F. (1988) Expression and functional analysis of the *Rhizobium meliloti nifA* gene. *EMBO*, **7**, 7-14.

75. Santero, E., Hoover, T., North, A., Berger, D., Porter, S. and Kustu, S. (1992) Role of integration host factor in stimulating transcription from the sigma 54-dependent *nifH* promoter. *J Mol Biol*, **227**, 602-620.

76. Matthews, K. (1992) DNA Looping. *Microbiol Revs*, **56**, 123-136.

77. Weiner, L., Brissette, J. and Model, P. (1991) Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on s54 and modulated by positive and negative feedback mechanisms. *Genes Dev*, **5**, 1912-1923.

78. Bordes, P., Wignseshweraraj, S., Schumacher, J., Zhang, X., Chaney, M. and Buck, M. (2003) The ATP hydrolyzing transcription activator phage shock

protein F of *Escherichia coli*: identifying a surface that binds sigma 54. *Proc Natl Acad Sci USA*, **100**, 2278-2283.

79. Burrows, P., Severinov, K., Buck, M. and Wigneshweraraj, S. (2004) Reorganization of an RNA polymerase-promoter DNA complex for DNA melting. *EMBO*, **23**, 4253-4563.

80. Zhang, N., Joly, N., Burrows, P., Jovanovic, M., Wigneshweraraj, S. and Buck, M. (2009) The role of the conserved phenylalanine in the sigma54-interacting GAFTGA motif of bacterial enhancer binding proteins. *Nuc Acids Res*, **37**, 5981-5992.

81. Li, J., Passaglia, L., Rombel, I., Yan, D. and Kustu, S. (1999) Mutations affecting motifs of unknown function in the central domain of nitrogen regulatory protein C. *J Bacteriol*, **181**, 5443-5454.

82. Huala, E. and Ausubel, F. (1989) The central domain of *Rhizobium meliloti* NifA is sufficient to activate transcription from the *R. meliloti nifH* promoter. *J Bacteriol*, **171**, 3354-3365.

83. Casaz, P. and Buck, M. (1997) Probing the assembly of transcription initiation complexes through changes in s^N protease sensitivity. *Proc Natl Acad Sci USA*, **94**, 12145-12150.

84. Chen, B., Sysoeva, T., Chowdhury, S., Guo, L., Carlo, S.D., Hanson, J., Yang, H. and Nixon, B. (2010) Engagement of argenine finger to ATP triggers large conformational changes in NtrC1 AAA+ ATPase for remodeling bacterial RNA polymerase. *Structure*, **18**, 1420-1430.

85. Kelly, M. and Hoover, T. (1999) Mutant Forms of *Salmonella typhimurium* s^{54} Defective in Transcription Initiation but Not Promoter Binding Activity. *J Bacteriol*, **181**, 3351-3357.

86. Sasse-Dwight, S. and Gralla, J. (1988) Probing the *Escherichia coli glnALG* upstream activation mechanism *in vivo*. *Proc Natl Acad Sci USA*, **85**, 8934-8938.

87. Brahmachary, P., Dashti, M., Olson, J. and Hoover, T. (2004) *Helicobacter pylori* FlgR is an enhancer-independent activator of sigma54-RNA polymerase holoenzyme. *J Bacteriol*, **186**, 4535-4542.

88. Koo, I. and Stephens, R. (2003) A developmentally regulated two-component signal transduction system in *Chlamydia*. *J Biol Chem*, **278**, 17314-17319.

89. Beck, L., Smith, T. and Hoover, T. (2007) Look, no hands! Unvonventional transcriptional activators in bacteria. *Trends Microbiol*, **15**, 530-537.

90. Jovanovic, G., Weiner, L. and Model, P. (1996) Identification, Nucleotide Sequence, and Characterization of PspF, the Transcriptional Activator of the *Escherichia coli* Stress-Induced *psp* Operon. *J Bacteriol*, **178**, 1936-1945.

91. Korsa, I. and Böck, A. (1997) Characterization of *fhlA* mutations resulting in ligand-independent transcriptional activation and ATP hydrolysis. *J Bacteriol*, **179**, 41-45.

92. Kill, K., Binnewies, T., Siicheritz-Ponten, T., Willenbrock, H., Hallin, P., Wassenaar, T. and Ussery, D. (2005) Genome update: sigma factors in 240 bacterial genomes. *Microbiology*, **151**, 3147-3150.

93. Michiels, J., Moris, M., Dombrecht, B., Verreth, C. and Vanderleyden, J. (1998) Differential Regulation of *Rhizobium etli rpoN2* Gene Expression during Symbiosis and Free-living Growth. *J Bacteriol*, **180**, 3620-3628.

94. Poggio, S., Osorio, A., Dreyfus, G. and Camarena, L. (2006) Transcriptional specificity of RpoN1 and RpoN2 involves differential recognition of the promoter sequences and specific interaction with the cognate activator proteins. *J Biol Chem*, **281**, 27205-27215.

95. Ray, S., Kumar, R., Peeters, N., Boucher, C. and Genin, S. (2015) rpoN1, but not rpoN2, is required for twitching motility, natural competence, growth on nitrate, and virulence of *Ralstonia solanacearum*. *Front Microbiol*, **6**, 229.

96. Hauser, F., Pessi, G., Friberg, M., Weber, C., Rusca, N., Lindemann, A., Fischer, H. and Hennecke, H. (2007) Dissection of the *Bradyrhizobium japonicum* NifA+sigma54 regulon, and identification of a ferredoxin gene (*fdxN*) for symbiotic nitrogen fixation. *Mol Gen Genet*, **278**, 255-271.

97. Lardi, M., Aguilar, C., Pedrioli, A., Omasits, U., Suppiger, A., Cárcamo-Oyarce, G., Schmid, N., Ahrens, C., Eberl, L. and Pessi, G. (2015) s^{54}-Dependent Response to Nitrogen Limitation and Virulence in *Burkholderia cenocepacia* Strain H111. *Appl Enivorn Microbiol*, **81**, 4077-4089.

98. Lundgren, B., Villegas-Penaranda, L., Harris, J., Mottern, A., Dunn, D., Boddy, C. and Nomura, C. (2014) Genetic Analysis of the Assimilation of C_5-Dicarboxylic Acids in *Pseudomonas aeruginosa* PAO1. *J Bacteriol*, **196**, 2543-2551.

99. Leonhartsberger, S., Huber, A., Lottspeich, F. and Bock, A. (2001) The *hydH/G* Genes from *Escherichia coli* Code for a Zinc and Lead Responsive Two-component Regulatory System. *J Mol Biol*, **307**, 93-105.

100. Horswill, A. and Escalante-Semerena, J. (1997) Propionate Catabolism in *Salmonella typhimurium* LT2: Two Divergently Transcribed Units Comprise the *prp* Locus at 8.5 Centisomes, *prpR* Encodes a Member of the Sigma-54 Family of Activators, and the *prpBCDE* Genes Constitute an Operon. *J Bacteriol*, **179**, 928-940.

101. Kang, Y.-S., Bothner, B., Rensing, C. and McDermott, T. (2012) Involvement of RpoN in Regulating Bacterial Aresenite Oxidation. *Appl Enivorn Microbiol*, **78**, 5638-5645.

102. Millikan, D. and Ruby, E. (2003) FlrA, a sigma54-dependent transcriptional activator in *Vibrio fischeri*, is required for motility and symbiotic light-organ colonization. *J Bacteriol*, **185**, 3547-3557.

103. Wösten, M., Wagenaar, J. and Putten, J.v. (2004) The FlgS/FlgR two-component signal transduction system regulates the *fla* regulon in *Campylobacter jejuni*. *J Biol Chem*, **279**, 16214-16222.

104. Dong, T. and Mekalanos, J. (2012) Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholarae* O37 strain V52. *Nucleic Acids Res*, **40**, 7766-7775.

105. Bernard, C.S., Brunet, Y.R., Gavioli, M., Lloubes, R. and Cascales, E. (2011) Regulation of type VI secretion gene clusters by sigma54 and cognate enhancer binding proteins. *J Bacteriol*, **193**, 2158-2167.

106. Kazmierczak, M., Wiedmann, M. and Boor, K. (2005) Alternative Sigma Factors and Their Roles in Bacterial Virulence. *Microbiol Mol Biol Rev*, **69**, 527-543.

107. Potvin, E., Sanschagrin, F. and Levesque, R. (2008) Sigma factors in *Pseudomonas aeruginosa*. *FEMS Microbiol Revs*, **32**, 38-55.

108. Wolfe, A., Millikan, D., Campbell, J. and Visick, K. (2004) *Vibrio fischeri* sigma54 controls motility, biofilm formation, luminescence, and colonization. *Appl Enivorn Microbiol*, **70**, 2520-2524.

109. Damron, F., Owings, J., Okkotsu, Y., Varga, J., Schurr, J., Godberg, J., Schurr, M. and Yu, H. (2012) Analysis of the *Pseudomonas aeruginosa* regulon controlled by the sensor kinase KinB and sigma factor RpoN. *J Bacteriol*, **194**, 1317-1330.

110. Dombrecht, B., Marchal, K., Vanderleyden, J. and Michiels, J. (2002) Prediction and overview of the RpoN-regulon in closely related species of the *Rhizobiales*. *Genome Biol*, **3**.

111. Cases, I., Ussery, D. and Lorenzo, V.d. (2003) The s^{54} regulon (sigmulon) of *Pseudomonas putida*. *Environ Microbiol*, **5**, 1281-1293.

112. Studholme, D. and Dixon, R. (2004) In silico analysis of the s54-dependent enhanacer-binding proteins in *Pirellula* species strain 1. *FEMS Microbiol Lett*, **230**, 215-225.

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

114. Chaudhuri, R., Yu, L., Kanji, A., Perkins, T., Gardner, P., Choudhary, J., Maskell, D. and Grant, A. (2011) Quantitative RNA-seq analysis of the *Campylobacter jejuni* transcriptome. *Microbiology*, **157**, 2922-2932.

115. Arous, S., Buchrieser, C., Folio, P., Glaser, P., Namane, A., Hébraud, M. and Héchard, Y. (2004) Global Analysis of Gene Expression in an *rpoN* mutant of *Listeria monocytogenes*. *Microbiology*, **150**, 1581-1590.

116. Zhao, K., Mingzhu, L. and Burgess, R. (2010) Promoter and regulon analysis of nitrogen assimilation factor, s^{54}, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. *Nucleic Acids Res*, **38**, 1273-1283.

117. Dixon, R. (1984) Tandem promoters determine regulation of the Klebsiella pneumoniae glutamine synthetase (glnA) gene. *Nucleic acids research*, **12**, 7811-7830.

118. Reitzer, L. and Magasanik, B. (1985) Expression of *glnA* in *Escherichia coli* is regulated at tandem promoters. *Proc Natl Acad Sci USA*, **82**, 1979-1983.

119. Ouyang, Z., Blevins, J. and Norgard, M. (2008) Transcriptional interplay among the regulators Rrp2, RpoN and RpoS in *Borrelia burgdorferi*. *Microbiology*, **154**, 2641-2658.

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

121. Wang, Y., Kolb, A., Buck, M., Wen, J., O'Gara, F. and Buc, H. (1998) CRP interacts with promoter-bound s^{54} RNA polymerase and blocks transcriptional activation of the *dctA* promoter. *EMBO*, **17**, 786-796.

122. Brown, D., Barton, G., Pan, Z., Buck, M. and Wigneshweraraj, S. (2014) Nitrogen stress response and stringent response are coupled in *Escherichia coli*. *Nat Commun*, **5**.

123. Storz, G., Vogel, J. and Wassarman, K.M. (2011) Regulation by small RNAs in bacteria: expanding frontiers. *Molecular cell*, **43**, 880-891.

124. Georg, J. and Hess, W.R. (2011) cis-antisense RNA, another level of gene regulation in bacteria. *Microbiology and molecular biology reviews : MMBR*, **75**, 286-300.

125. Kroger, C., Colgan, A., Srikumar, S., Handler, K., Sivasankaran, S., Hammarlof, D., Canals, R., Grissom, J., Conway, T., Hokamp, K. *et al.* (2013) An Infection-Relevant Transcriptomic Compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host Microbe*, **14**, 683-695.

126. Schlüter, J., Reinkensmeier, J., Barnett, M., Lang, C., Krol, E., Giegerich, R., Long, S. and Becker, A. (2013) Global mapping of transcription start sites and promoter motifs in the symbiotic a-proteobacterium *Sinorhizobium meliloti* 1021. *BMC Genomics*, **14**, 153.

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

128. Hebrard, M., Kroger, C., Srikumar, S., Colgan, A., Handler, K. and Hinton, J.C. (2012) sRNAs and the virulence of Salmonella enterica serovar Typhimurium. *RNA biology*, **9**, 437-445.

129. Reichenbach, B., Gopel, Y. and Gorke, B. (2009) Dual control by perfectly overlapping sigma 54- and sigma 70- promoters adjusts small RNA GlmY expression to different environmental signals. *Molecular microbiology*, **74**, 1054-1070.

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

131. Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R. and Storz, G. (1998) The *oxyS* regulatory RNA represses *rpoS* translation by binding Hfq (HF-1) protein. *EMBO*, **17**, 6061-6068.

132. Altuvia, S., Zhang, A., Argaman, L., Tiwari, A. and Storz, G. (1998) The *Escherichia coli oxyS* regulatory RNA represses *fhlA* translation by blocking ribosome binding. *EMBO*, **17**, 6069-6075.

133. Altuvia, S., Weinstein-Fischer, D., Zhang, A., Postow, L. and Storz, G. (1997) A small stable RNA induced by oxidative stress: rola as a pleiotropic regulator and antimutator. *Cell*, **90**, 43-53.

134. Zafar, M.A., Carabetta, V.J., Mandel, M.J. and Silhavy, T.J. (2014) Transcriptional occlusion caused by overlapping promoters. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 1557-1561.

135. Johansson, L.U., Solera, D., Bernardo, L.M., Moscoso, J.A. and Shingler, V. (2008) sigma54-RNA polymerase controls sigma70-dependent transcription from a non-overlapping divergent promoter. *Molecular microbiology*, **70**, 709-723.

136. Drew, H. and Travers, A. (1984) DNA Structural Variations in the *E. coli tyrT* Promoter. *Cell*, **37**, 491-502.

137. Adhya, S. and Gottesman, M. (1982) Promoter occulusion: transcription through a promoter may inhibit its activity. *Cell*, **29**, 939-944.

138. Callen, B. (2004) Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol Cell*, **14**, 647-656.

139. Prescott, E. and Proudfoot, N. (2002) Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci USA*, **99**, 8796-8801.

140. Peng, Q., Wang, G., Liu, G., Zhang, J. and Song, F. (2015) Identification of metabolism pathways directly regulated by sigma^{54} factor in *Bacillus thuringiensis*. *Front Microbiol*, **6**, 407.

141. Fisher, M., Grimm, D., Henion, A., Elias, A., Stewart, P., Rosa, P. and Gherardini, F. (2005) *Borrelia burgdorferi* sigma54 is required for mammalian infection and vector transmission but not for tick colonization. *Proc Natl Acad Sci USA*, **102**, 5162-5167.

142. Qui, Y., Nagarajan, H., Embree, M., Shieu, W., Abate, E., Juarez, K., Cho, B.-K., Elkins, J., Nevin, K., Barrett, C. *et al.* (2013) Characterizing the interplay between multiple levels of organization within bacterial sigma factor regulatory networks. *Nature Communications*, **4**.

143. Yu, X., Lund, S., Greenwald, J., Records, A., Scott, R., Nettleton, D., Lindow, S., Gross, D. and Beattie, G. (2014) Transcriptional Analysis of the Global Regulatory Networks Active in *Pseudomona syringae* during Leaf Colonization. *mBio*, **5**, e01683-01614.

144. Neto, J.d.S., Koide, T., Gomes, S. and Marques, M. (2010) Global gene expression under nitrogen starvation in *Xylella fastidiosa*: contribution of the s^{54} regulon. *BMC Microbiol*, **10**, 231.

145. Maloy, S., Stewart, V. and Taylor, R. (1996) *Genetic Analysis of Pathogenic Bacteria: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Plainview, NY.

146. Miller, K., Phillips, R., Mrázek, J. and Hoover, T. (2013) *Salmonella* utilizes D-glucosaminate via a mannose family phosphotransferase system permease and associated enzymes. *J Bacteriol*, **195**, 4057-4066.

147. Datsenko, K. and Wanner, B. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA*, **97**, 6640-6645.

148. Cherepanov, P. and Wachernagel, W. (1995) Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene*, **158**, 9-14.

149. Perkins-Balding, D., Duval-Valentin, G. and Glasgow, A. (1999) Excision of IS*492* Requires Flanking Target Sequences and Results in Circle Formation in *Pseudoalteromonas atlantica*. *J Bacteriol*, **181**, 4937-4948.

150. Grana, D., Youderian, P. and Susskind, M. (1985) Mutations that improve the *ant* promoter of *Salmonella* phage P22. *Genetics*, **110**, 1-16.

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

152. Kelly, M. and Hoover, T. (2000) The amino terminus of *Salmonella enterica* serovar Typhimurium sigma54 is required for interactions with an enhancer-binding protein and binding to fork junction DNA. *J Bacteriol*, **182**, 513-517.

153. Lieu, P.T., Jozsi, P., Gilles, P. and Peterson, T. (2005) Development of a DNA-labeling system for array-based comparative genomic hybridization. *J Biomol Tech*, **16**, 104-111.

154. Stead, M., Agrawal, A., Bowden, K., Nasir, R., Mohanty, B., Meagher, R. and Kushner, S. (2012) RNA*snap*^{TM}: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. *Nucleic Acids Res*, **40**, e156.

155. Wang, Q., Frye, J.G., McClelland, M. and Harshey, R.M. (2004) Gene expression patterns during swarming in Salmonella typhimurium: genes specific to surface growth and putative new motility and pathogenicity genes. *Molecular microbiology*, **52**, 169-187.

156. Santiviago, C.A., Reynolds, M.M., Porwollik, S., Choi, S.H., Long, F., Andrews-Polymenis, H.L. and McClelland, M. (2009) Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLoS pathogens*, **5**, e1000477.

157. Morales, E.H., Collao, B., Desai, P.T., Calderon, I.L., Gil, F., Luraschi, R., Porwollik, S., McClelland, M. and Saavedra, C.P. (2013) Probing the ArcA regulon under aerobic/ROS conditions in Salmonella enterica serovar Typhimurium. *BMC genomics*, **14**, 626.

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

159. Mrazek, J., Xie, S., Guo, X. and Srivastava, A. (2008) AIMIE: a web-based environment for detection and interpretation of significant sequence motifs in prokaryotic genomes. *Bioinformatics*, **24**, 1041-1048.

160. Cannon, W., Chaney, M., Wang, X.-Y. and Buck, M. (1997) Two domains within s^N (s^{54}) cooperate for DNA binding. *Proc Natl Acad Sci USA*, **94**, 5006-5011.

161. Gallegos, M.-T. and Buck, M. (2000) Sequences in s^{54} Region I Required for Binding to Early Melted DNA and their Involvement in Sigma-DNA Isomerisation. *J Mol Biol*, **297**, 849-859.

162. Houten, V.v., Denkers, F., Dijk, M.v., Brekel, M.v.d. and Brakenhoff, T. (1998) Labeling efficiency of oligonucleotides by T4 polynucleotide kinase depends on 5'-nucleotide. *Anal Biochem*, **265**, 386-389.

163. Jishage, M., Iwata, A., Ueda, S. and Ishihama, A. (1996) Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of four species of sigma subunit under various growth conditions. *J Bacteriol*, **178**, 5447-5451.

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

165. Kubitschek, H.E. and Friske, J.A. (1986) Determination of bacterial cell volume with the Coulter Counter. *J Bacteriol*, **168**, 1466-1467.

166. Skosyrev, V.S., Vasil'eva, G.V., Lomaeva, M.G., Malakhova, L.V., Antipova, V.N. and Bezlepkin, V.G. (2013) [Specialized software product for comparative analysis of multicomponent DNA fingerprints]. *Genetika*, **49**, 531-537.

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

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

169. Jarvik, T., Smillie, C., Groisman, E. and Ochman, H. (2010) Short-Term Signatures of Evolutionary Change in the *Salmonella enterica* Serovar Typhimurium 14028 Genome. *J Bacteriol*, **192**, 560-567.

170. Kroger, C., Dillon, S., Cameron, A., Papenfort, K., Sivasanharan, S., Hokamp, K., Chao, Y., Sittka, A., Hebrard, M., Handler, K. *et al.* (2012) The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Tyhphimurium. *Proc Natl Acad Sci USA*, **109**, e1277-e1286.

171. Miyakoshi, M., Chao, Y. and Vogel, J. (2015) Regulatory small RNAs from the 3' regions of bacterial mRNAs. *Curr Opin Microbiol*, **24**, 132-139.

172. Zimmer, D.P., Soupene, E., Lee, H.L., Wendisch, V.F., Khodursky, A.B., Peter, B.J., Bender, R.A. and Kustu, S. (2000) Nitrogen regulatory protein C-controlled genes of Escherichia coli: scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci U S A*, **97**, 14674-14679.

173. Lloyd, L.J., Jones, S.E., Jovanovic, G., Gyaneshwar, P., Rolfe, M.D., Thompson, A., Hinton, J.C. and Buck, M. (2004) 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*, **279**, 55707-55714.

174. Miller, K., Phillips, R., Kilgore, P., Smith, G. and Hoover, T. (2015) A mannose family phosphotransferase system permease and associated enzymes are required for utilization of fructoselysine and glucoselysine in *Salmonella enterica* serovar Typhimurium. *J Bacteriol*.

175. Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M. and Vogel, J. (2012) An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. *EMBO J*, **31**, 4005-4019.

176. Zbell, A., Benoit, S. and Maier, R. (2007) Differential expression of NiFe uptake-type hydrogenase genes in *Salmonella enterica* serovar Typhimurium. *Microbiology*, **153**, 3508-3516.

177. Sawers, R.G., Jamieson, D.J., Higgins, C.F. and Boxer, D.H. (1986) Characterization and physiological roles of membrane-bound hydrogenase isoenzymes from Salmonella typhimurium. *J Bacteriol*, **168**, 398-404.

178. Maupin, J.A. and Shanmugam, K.T. (1990) Genetic regulation of formate hydrogenlyase of Escherichia coli: role of the fhlA gene product as a transcriptional activator for a new regulatory gene, fhlB. *J Bacteriol*, **172**, 4798-4806.

179. Schaefer, J., Engl, C., Zhang, N., Lawton, E. and Buck, M. (2015) Genome wide interactions of wild-type and activator bypass forms of sigma54. *Nucleic Acids Res*, **43**, 7280-7291.

180. Morris, L., Cannon, W., Claverie-Martin, F., Austin, S. and Buck, M. (1994) DNA distortion and nucleation of local DNA unwinding within sigma-54 (sigma N) holoenzyme closed promoter complexes. *J Biol Chem*, **269**, 11563-11571.

181. Vogel, S., Schulz, A. and Rippe, K. (2002) Binding affinity of *Escherichia coli* RNA polymerase s^{54} holoenzyme for the *glnA*p2, *nifH*, and *nifL* promoters. *Nuc Acids Res*, **30**, 4094-4101.

182. Kroger, C., Colgan, A., Srikumar, S., Handler, K., Sivasankaran, S.K., Hammarlof, D.L., Canals, R., Grissom, J.E., Conway, T., Hokamp, K. *et al.* (2013) An infection-relevant transcriptomic compendium for Salmonella enterica Serovar Typhimurium. *Cell Host Microbe*, **14**, 683-695.

183. Palacios, S. and Escalante-Semerena, J.C. (2000) 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*, **182**, 905-910.

184. Samuels, D.J., Frye, J.G., Porwollik, S., McClelland, M., Mrazek, J., Hoover, T.R. and Karls, A.C. (2013) Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the sigma(5)(4) (RpoN) regulon of Salmonella Typhimurium LT2. *BMC genomics*, **14**, 602.

185. Klose, K.E. and Mekalanos, J.J. (1997) Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates Salmonella typhimurium virulence. *Infect Immun*, **65**, 587-596.

186. Miller, K.A., Phillips, R.S., Mrazek, J. and Hoover, T.R. (2013) Salmonella utilizes D-glucosaminate via a mannose family phosphotransferase system permease and associated enzymes. *J Bacteriol*, **195**, 4057-4066.

187. Miller, K.A., Phillips, R.S., Kilgore, P.B., Smith, G.L. and Hoover, T.R. (2015) A Mannose Family Phosphotransferase System Permease and Associated Enzymes Are Required for Utilization of Fructoselysine and Glucoselysine in Salmonella enterica Serovar Typhimurium. *J Bacteriol*, **197**, 2831-2839.

188. Schneider, T.D. and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic acids research*, **18**, 6097-6100.

189. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res*, **14**, 1188-1190.

190. Adkins, J.N., Mottaz, H.M., Norbeck, A.D., Gustin, J.K., Rue, J., Clauss, T.R., Purvine, S.O., Rodland, K.D., Heffron, F. and Smith, R.D. (2006) Analysis of the Salmonella typhimurium proteome through environmental response toward infectious conditions. *Mol Cell Proteomics*, **5**, 1450-1461.

191. Shi, L., Ansong, C., Smallwood, H., Rommereim, L., McDermott, J.E., Brewer, H.M., Norbeck, A.D., Taylor, R.C., Gustin, J.K., Heffron, F. *et al.* (2009) Proteome of Salmonella Enterica Serotype Typhimurium Grown in a Low Mg/pH Medium. *J Proteomics Bioinform*, **2**, 388-397.

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

193. Bang, I.S., Frye, J.G., McClelland, M., Velayudhan, J. and Fang, F.C. (2005) Alternative sigma factor interactions in Salmonella: sigma and sigma promote antioxidant defences by enhancing sigma levels. *Molecular microbiology*, **56**, 811-823.

194. Miyakoshi, M., Chao, Y. and Vogel, J. (2015) Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA. *EMBO J*, **34**, 1478-1492.

195. Fuentes, D.N., Calderon, P.F., Acuna, L.G., Rodas, P.I., Paredes-Sabja, D., Fuentes, J.A., Gil, F. and Calderon, I.L. (2015) Motility modulation by the small non-coding RNA SroC in Salmonella Typhimurium. *FEMS Microbiol Lett*, **362**.