Title
Novel DNA Binding and Regulatory Activities for σ54 (RpoN) in Salmonella enterica Serovar Typhimurium 14028s

Journal
Journal of Bacteriology, 199(12)

ISSN
0021-9193

Authors
Bono, Ashley C
Hartman, Christine E
Solaimanpour, Sina
et al.

Publication Date
2017-06-15

DOI
10.1128/jb.00816-16

Peer reviewed

AMERICAN SOCIETY FOR MICROBIOLOGY | Journal of Bacteriology®

RESEARCH ARTICLE

Check for updates

Novel DNA Binding and Regulatory Activities for σ^{54} (RpoN) in *Salmonella enterica* Serovar Typhimurium 14028s

Ashley C. Bono,[a]* Christine E. Hartman,[b]* Sina Solaimanpour,[c] Hao Tong,[d] Steffen Porwollik,[f] Michael McClelland,[f] Jonathan G. Frye,[g] Jan Mrázek,[a,e] Anna C. Karls[a]

Department of Microbiology, University of Georgia, Athens, Georgia, USA[a]; Department of Genetics, University of Georgia, Athens, Georgia, USA[b]; Department of Computer Science, University of Georgia, Athens, Georgia, USA[c]; Department of Statistics, University of Georgia, Athens, Georgia, USA[d]; Institute of Bioinformatics, University of Georgia, Athens, Georgia, USA[e]; Department of Microbiology and Molecular Genetics, University of California—Irvine School of Medicine, Irvine, California, USA[f]; Bacterial Epidemiology and Antimicrobial Resistance Research Unit, U.S. Department of Agriculture, Athens, Georgia, USA[g]

ABSTRACT The variable sigma (σ) subunit of the bacterial RNA polymerase (RNAP) holoenzyme, which is responsible for promoter specificity and open complex formation, plays a strategic role in the response to environmental changes. *Salmonella enterica* serovar Typhimurium utilizes the housekeeping σ^{70} and five alternative sigma factors, including σ^{54}. The σ^{54}-RNAP differs from other σ-RNAP holoenzymes in that it forms a stable closed complex with the promoter and requires ATP hydrolysis by an activated cognate bacterial enhancer binding protein (bEBP) to transition to an open complex and initiate transcription. In *S.* Typhimurium, σ^{54}-dependent promoters normally respond to one of 13 different bEBPs, each of which is activated under a specific growth condition. Here, we utilized a constitutively active, promiscuous bEBP to perform a genome-wide identification of σ^{54}-RNAP DNA binding sites and the transcriptome of the σ^{54} regulon of *S.* Typhimurium. The position and context of many of the identified σ^{54} RNAP DNA binding sites suggest regulatory roles for σ^{54}-RNAP that connect the σ^{54} regulon to regulons of other σ factors to provide a dynamic response to rapidly changing environmental conditions.

IMPORTANCE The alternative sigma factor σ^{54} (RpoN) is required for expression of genes involved in processes with significance in agriculture, bioenergy production, bioremediation, and host-microbe interactions. The characterization of the σ^{54} regulon of the versatile pathogen *S.* Typhimurium has expanded our understanding of the scope of the σ^{54} regulon and how it links to other σ regulons within the complex regulatory network for gene expression in bacteria.

KEYWORDS RpoN, *Salmonella enterica* serovar Typhimurium, σ^{54}, bEBP, promoter, regulon

Received 26 November 2016 **Accepted** 27 March 2017

Accepted manuscript posted online 3 April 2017

Citation Bono AC, Hartman CE, Solaimanpour S, Tong H, Porwollik S, McClelland M, Frye JG, Mrázek J, Karls AC. 2017. Novel DNA binding and regulatory activities for σ54 (RpoN) in *Salmonella enterica* serovar Typhimurium 14028s. J Bacteriol 199:e00816-16. https://doi.org/10.1128/JB.00816-16.

Editor George O'Toole, Geisel School of Medicine at Dartmouth

Address correspondence to Anna C. Karls, akarls@uga.edu.

* Present address: Ashley C. Bono, Shire Pharmaceuticals, Social Circle, Georgia, USA; Christine E. Hartman, Office for Teaching and Learning, Wayne State University, Detroit, Michigan, USA.

Salmonella enterica subsp. *enterica* serovar Typhimurium is the most common serotype of *Salmonella* associated with gastrointestinal disease in humans and has been extensively studied to reveal the virulence factors that lead to morbidity and mortality. As an excellent model system for bacterial pathogen-host interactions, *S.* Typhimurium has been extensively characterized for regulation of its transcriptome and proteome by both protein- and small RNA (sRNA)-mediated mechanisms (1–6). Coordination of the complex, overlapping regulatory networks that control bacterial gene expression in response to the diverse stresses encountered in a host or in the environment commonly occurs at the level of sigma factors (7–9). In bacterial systems, a single core RNA polymerase (RNAP), comprising two α subunits and one β, β', and ω subunit each, transcribes genes into RNA. A variable σ (sigma) factor subunit, which transiently interacts with RNAP to form the holoenzyme (Eσ), is required for recognition of specific promoter sequences and open complex formation during transcription initiation (10). There are two families of σ factors: the σ^{70} family, which comprises four groups of σ factors with various levels of amino acid sequence similarity to σ^{70} (RpoD) (11), and the σ^{54} family, which has only one member, σ^{54} (RpoN), and shares no sequence similarity with members of the σ^{70} family (12). The σ^{54} affinity for the core RNAP is close to that of the housekeeping sigma factor σ^{70} and is higher than most of the σ^{70}-related alternative σ factors, allowing it to compete strongly for RNAP binding (13–15). Although σ^{54} and the σ^{70}-type factors associate with the same RNAP to form holoenzymes, there are significant differences in the activities of Eσ^{54} and Eσ^{70}: Eσ^{70} recognizes and binds loosely conserved promoter elements at the -10 and -35 regions relative to the transcription start site (TSS [$+1$ position]) and spontaneously melts DNA at the promoter to initiate transcription, while Eσ^{54} recognizes and binds highly conserved -12 (GC) and -24 (GG) promoter elements and remains in an autorepressive state at the promoter until ATP hydrolysis by an associated activator remodels the structure of Eσ^{54} to promote open complex formation (reviewed in references 16 and 17). Like the transcription factors and enhancer sequences that control eukaryotic polymerase II (Pol II) activity (18), activators of Eσ^{54} bind DNA sequences that can be distant from the promoter, and DNA looping allows the activator to contact Eσ^{54}; thus, the Eσ^{54} activators are termed bacterial enhancer binding proteins (bEBPs) (17). The ability of Eσ^{54} to interact stably with promoter sequences without initiating transcription could give σ^{54} an advantage over σ^{70}-type factors in the competition for RNAP binding (9, 19) and could play an important role in the rapid response of the σ^{54} regulon to environmental signals.

There are six sigma factors controlling gene expression in *S.* Typhimurium: the housekeeping factor σ^{70}, and five alternative sigma factors, σ^{24} (*rpoE*), σ^{32} (*rpoH*), σ^{38} (*rpoS*), σ^{28} (*fliA*), and σ^{54} (*rpoN*) (8, 20). The regulons of these σ factors in *S.* Typhimurium have been characterized to various extents (8, 21–25), and a comprehensive analysis of the transcriptome of *S.* Typhimurium has been conducted under growth conditions that mimic infection-related environments (2). The σ^{54} regulon of *S.* Typhimurium is known to contribute to diverse cellular processes, including response to envelope stress (26, 27), detoxification of nitric oxide under anaerobic conditions (28), and uptake/utilization of alternative carbon/nitrogen sources (29–33). However, it has been difficult to characterize this regulon due to the diverse (and sometimes unknown) conditions needed to activate the individual bEBPs required for σ^{54}-dependent transcription (for reviews, see references 17 and 34).

In our previous study of the σ^{54} regulon in *S.* Typhimurium LT2 (25), we showed that the bEBP variant DctD250, which is active in the absence of an environmental signal and activates without enhancer binding, stimulated transcription from 20 σ^{54}-dependent promoters that are known or predicted to be responsive to 12 of the 13 bEBPs in *S.* Typhimurium. In addition, 70 Eσ^{54} genomic binding sites were identified by chromatin immunoprecipitation with microarray technology (ChIP-chip) combined with *in silico* genome sequence analyses (25). While providing an effective approach to defining the Eσ^{54} transcriptome of *S.* Typhimurium, this earlier work was limited in determining the full complement of Eσ^{54} binding sites and σ^{54}-dependent transcripts due to the utilization of open reading frame (ORF) arrays of the *S.* Typhimurium LT2 genome for the ChIP-chip and gene expression microarray analyses and the use of the LT2 strain, which has a point mutation in the start codon of the *rpoS* gene leading to very low production of RpoS (σ^{38}) and reduced virulence (35, 36). In the present work, we aimed to ensure that the full σ factor pool is intact, because changes in σ^{54} competition for the RNAP core may influence occupancy of low-affinity promoters (9).

The ChIP-chip and microarray assays presented here were performed with the virulent 14028s strain of *S.* Typhimurium (37) and utilized high-density tiling arrays of the *S.* Typhimurium 14028s genome (38). Sixty-four of the 70 intergenic and intragenic Eσ^{54} DNA binding sites predicted in our earlier work (25) were confirmed, and 122 additional binding sites were identified, most of which are intragenic. The transcriptome of the σ^{54} regulon was expanded from 21 to 24 σ^{54}-dependent transcripts, including two novel transcripts originating from intragenic promoters. In addition, nine transcripts exhibit downregulation potentially through σ^{54}-dependent mechanisms. We propose regulatory roles for the Eσ^{54} DNA binding sites based on genomic context, bioinformatics, and transcription analyses. Characterization of *in vitro* binding of Eσ^{54} and σ^{54} in the absence of RNAP to 11 identified Eσ^{54} genomic binding sites suggests sequence determinants for stable closed complex formation.

RESULTS AND DISCUSSION

Identification of Eσ^{54} genomic DNA binding sites in *S.* Typhimurium. ChIP-chip assays with high-density tiling arrays, described in detail in Materials and Methods, were used to characterize Eσ^{54} genomic DNA binding sites in *S.* Typhimurium 14028s. In the ChIP-chip assays, the wild-type (WT) and Δ*rpoN* 14028s strains expressed the constitutively active, promiscuous bEBP DctD250, which activates most σ^{54}-dependent promoters (25); therefore, the *in vivo* cross-linking of protein-DNA complexes was performed in the presence of rifampin, which blocks progression of RNAP from the promoter region (39), allowing capture of Eσ^{54} associated with active promoters. Since σ^{54} has been shown to specifically bind to a naturally occurring promoter in the absence of RNAP (40), all enriched DNAs in the α-σ^{54} pulldown encode a DNA site that potentially binds σ^{54} and/or Eσ^{54}.

ChIPeak analysis identified 184 peaks of enriched DNA sequence that met a conservative *P* value cutoff of 10^{-19} and a mean binding signal ratio of >3. The most likely Eσ^{54} DNA binding site within 350 bp of each peak maximum was predicted using a standard position-specific scoring matrix (PSSM) method (see Materials and Methods); the predicted 18-bp core DNA binding sequence includes the highly conserved -12 and -24 sequence elements of σ^{54}-dependent promoters and spans positions -9 to -26 (relative to a $+1$ transcription start site). One hundred eighty-six probable Eσ^{54} DNA binding sites were identified within the 184 peaks (listed in Table S2 in the supplemental material; a graphical representation of the multiple-sequence alignment is shown in Fig. 1C). The closely spaced ($<$155 bp apart) divergently transcribing σ^{54}-dependent promoters between the *hyc* and *hyp* operons (41) and between *zraSR* and *zraP* (42) were associated with single peaks at 3019780 and 4402030, respectively (Table S2). The resolution afforded by the ChIP-chip methodology employed did not allow clear peak separation for Eσ^{54} DNA binding sites that are less than $\sim$500 bp apart; the closest peak maxima identified by ChIPeak that met the criteria for 3-fold enrichment and a *P* value of $<10^{-19}$ were 550 bp apart (peak maxima at 14028s genomic positions 4484310 and 4484860 [Table S2]). The 186 identified Eσ^{54} DNA binding sites include 64 of the 70 Eσ^{54} DNA binding sites predicted in our earlier work using the LT2 ORF array hybridization data (25) (footnoted in Table S2), including all 26 previously predicted or identified intergenic σ^{54}-dependent promoters in *Salmonella*, and the 4 intragenic promoters described in reference 25 (italicized LT2 locus tags in Table S2).

Inferences from PSSM scores of identified Eσ^{54} DNA binding sites in the *Salmonella* genome. The identified Eσ^{54} binding sites from our ChIP-chip analysis have PSSM scores ranging from 4.40 to 22.3 and binding signal ratios ranging from 3.01 to 21.7 (Table S2). To obtain benchmarks for interpretation of the PSSM scores associated with each significant ChIP-enriched sequence, we compared the counts of 18-mers with PSSM scores above a given score cutoff in the *S.* Typhimurium 14028s genome and in 1,000 randomized *S.* Typhimurium genome sequences (Fig. 2). Each randomized genome sequence, generated as described in Materials and Methods, reproduces the characteristic nucleotide, dinucleotide, and codon usage biases of the whole genome, as well as every single gene and intergenic region, and serves as a more accurate null model to estimate expected counts of longer oligonucleotides or sequence patterns in DNA sequences than the commonly used Bernoulli model of independent trials (43). As seen in Fig. 2, the distribution of PSSM scores in the *S.* Typhimurium genome is quite similar to that in the randomized genomes, particularly at low PSSM scores, and begins

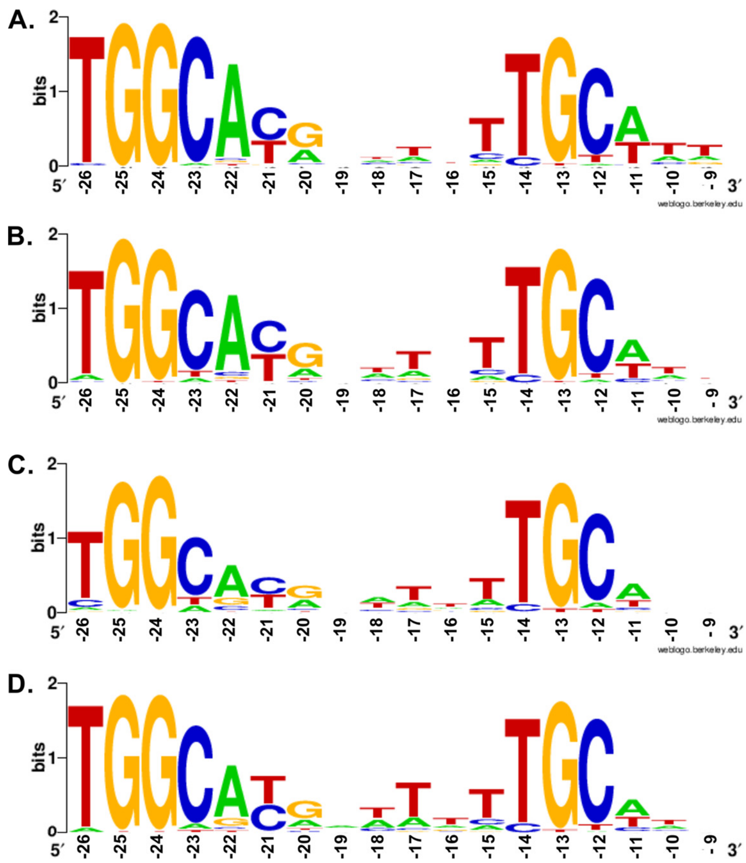

FIG 1 Graphical representations of multiple sequence alignments for *S.* Typhimurium 14028s Eσ^{54} binding sites. The relative frequency of bases at a given position is illustrated by WebLogo (100, 101) for multiple sequence alignments of (A) the 18-bp core sequences (-9 to -26 from the transcriptional start) of 33 σ^{54}-dependent promoters in *Salmonella* (25, 50; this study), (B) the 52 Eσ^{54} binding sites used for generating the PSSM for Eσ^{54} binding site prediction (see Materials and Methods); (C) the 186 predicted Eσ^{54} binding sites from the ChIP-chip analysis; and (D) the 68 predicted Eσ^{54} binding sites from the ChIP-chip analysis that have PSSM scores of $>$14.

to deviate more for PSSM scores of $>$14. This result suggests that most of the Eσ^{54} DNA binding sites with PSSM scores below ~14 (and some of the sites with PSSM scores of $>$14) may have arisen in the *S.* Typhimurium genome stochastically in the absence of selective constraints since sites with these PSSM scores occur with approximately equal frequency in the randomized genomes. This result does not indicate whether a DNA site with a specific PSSM score binds Eσ^{54} or not: rather it suggests the likelihood of physiological relevance for a binding site.

The occupancy of a predicted genomic binding site by Eσ^{54} is reflected in the binding signal ratio from the ChIP-chip assay (44); the correlation between position weight matrix scores (such as PSSM) and binding signal ratios from ChIP-chip assays varies for different transcription factors in bacteria and eukaryotes and may depend on various experimental and physiological factors (45–47). A plot of peak intensity ($\log_2$ binding signal ratio) for the 184 identified ChIP-chip peaks versus the PSSM score of the identified Eσ^{54} DNA binding sites within 350 bp of the peak maximum (see Table S2) indicates that the PSSM score positively correlates with the peak intensity (Pearson product-moment correlation coefficient r of 0.71 with a P value of $<$0.0001 [see Fig. S1 in the supplemental material]), similar to binding studies for the *Escherichia coli* transcription factor LexA (46).

Table 1 lists all ChIP-chip-identified Eσ^{54} DNA binding sites with PSSM scores of $>$14 (a total of 68 sites [a graphical representation of the multiple-sequence alignment in Fig. 1C]), which presumably have a greater likelihood of physiologically relevant roles

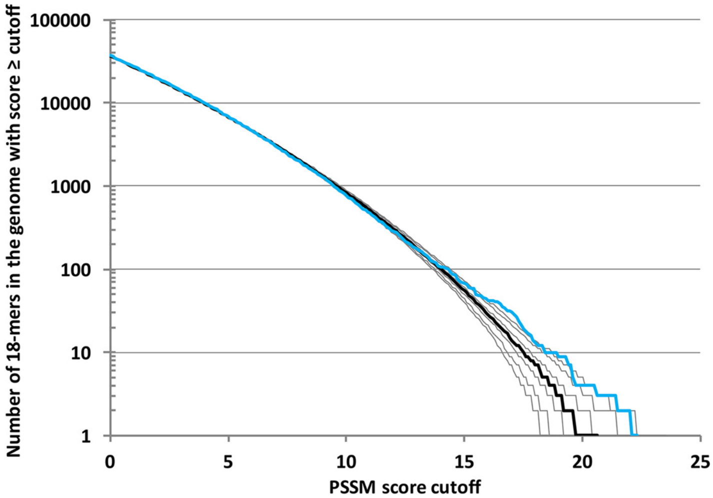

FIG 2 PSSM score distribution for 18-mer potential Eσ^{54} binding sites in *S.* Typhimurium and randomized genome sequences. The ordinate shows the number of 18-mers found in the sequence with PSSM scores greater than or equal to the cutoff shown on the abscissa. Note that the plot displays only the right tail of the whole PSSM score distribution (scores of ≥0) and the 55,840 18-mers that yield scores of ≥0 represent only 0.57% of all 9,740,530 18-mers in the genome (in both DNA strands). The blue line refers to the *S.* Typhimurium genome. The thick black line signifies the median value among 1,000 randomized genome sequences, and the thin black lines correspond to the 1st, 5th, 25th, 75th, 95th, and 99th percentiles.

in *Salmonella*. The remaining 118 Eσ^{54} binding sites (with PSSM scores of <14) that were enriched more than 3-fold in the ChIP-chip analysis include 16 sites that were identified in *S.* Typhimurium LT2 (25), four sites with orthologues in *E. coli* (48), one site that was demonstrated to have σ^{54}-dependent promoter activity (25), and five sites that overlap σ^{70}-type promoters (Table S2), which suggests the potential for functional relevance for these 16 sites. For any of the identified Eσ^{54} DNA binding sites, the physiological function must be individually characterized, but certain features of a site may suggest possible function, such as genomic context and promoter activity.

Genomic context of ChIP-chip-identified Eσ^{54} binding sites. The 186 identified Eσ^{54} DNA binding sites (sequences listed in Table S2) were divided into intergenic or intragenic groupings (A through F) based on their position and orientation relative to open reading frames (ORFs) annotated in *S.* Typhimurium 14028s, as illustrated in Fig. 3A and detailed in Table S2. Eσ^{54} DNA binding sites were designated intragenic only if the corresponding ORF was annotated in both the 14028s and LT2 genomes, with the exception of the prophage genes that are not shared by these strains (37). Eσ^{54} DNA binding site positions (Fig. 3A) are indicative of potential roles for the sites in gene regulation; examples of these roles are depicted in Fig. 3B.

(i) Class A and B intergenic sites. Forty-two of the 186 Eσ^{54} DNA binding sites (22.6%) are located in intergenic regions. The 36 class A intergenic binding sites, which are oriented toward the 5′ end of the nearest coding sequence, are positioned to act as promoters for the downstream gene(s) and include all 26 previously identified intergenic σ^{54}-dependent promoters in *S.* Typhimurium (25). Nine of the class A sites overlap σ^{70}-type promoters (Table S2) and potentially regulate transcription by promoter competition (49); notably, two of these class A sites have previously been shown to regulate expression of *glmY* and *glmZ* by promoter competition (50). Class B intergenic sites (6 sites) are oriented toward the 3′ end of the nearest annotated ORF and thus are positioned for transcription interference by polymerase collision (49); one of these class B sites overlaps a σ^{70}-type promoter (Fig. 3B; Table S2) and therefore may regulate transcription by promoter competition (49).

(ii) Class C, D, E, and F intragenic sites. The remaining 144 enriched sites in the ChIP-chip assay (77.4% of total sites) are within annotated coding sequences. Classifications of these intragenic sites are based on both orientation relative to the gene

TABLE 1 Summary of the highest-scoring Eσ^{54} binding sites identified in ChIP-chip analysis, grouped by position and orientation relative to the associated ORF

Class and peak position[a]	Associated 14028s ORF[b]	Equivalent LT2 ORF[b]	Gene name(s)	Signal ratio[c]	PSSM score	Distance to peak (bp)[d]	Identified Eσ^{54} DNA binding site[e]
Class A: intergenic sites oriented toward 5′ end of associated ORF							
418770	***STM14_0431***	*STM0368*[f]	*prpB*	17.9	17.8	146	**TGGCATAGCCTTTGCTTT**
503670	STM14_0530	*STM0448*[f]	*clpP*	12.5	17.6	−51	**TGTCACGTATTTTGCATG**
521100	***STM14_0546***	*STM0462*[f]	*glnK*	21.7	17.8	−39	**TGGCACATCCTTTGCAAT**
637570	***STM14_0673***	*STM0577*[f]		14.6	17.3	0	**TGGCACGCCGTTTGCCAT**
712450	***STM14_0757***	*STM0649.S*[f]		14.0	17.1	0	**TGGCACGCCTTTTGATTA**
730680	***STM14_0773***	*STM0665*[f]	*gltI*	19.2	21.5	76	**TGGCACGTCTATTGCTTT**
898110	***STM14_0964***	*STM0830*[f]	*glnH*	17.4	19.6	−4	**TGGCATGATTTTTTCATT**
1373930	***STM14_1558***	*STM1285*[f]	*yeaG*	15.0	18.5	69	**TGGCATGAGAGTTGCTTT**
1392120	***STM14_1582***[h]	*STM1303*[f]	*astC*	12.7	20.6	22	**TGGCACGAATGCTGCAAT**
1793230	***STM14_2040***	*STM1690*[f]	*pspA*	18.6	18.0	54	**TGGCACGCAAATTGTATT**
2516940	***STM14_2900***[h]	STM2354	*hisJ*	6.19	17.2	52	**TGGCACGATAGTCGCATC**
2517910	***STM14_2901***	*STM2355*[f]	*argT*	14.2	14.7	−1	**TGGCATAAGACCTGCATG**
2524240	***STM14_2907***	*STM2360*[f]		17.1	22.3	−26	**TGGCATGCCTTTTGCTTT**
2759250	STM14_3143.P[h]	*STM_R0152*[f]	*glmY*	15.0	19.4	−160	**TGGCACAATTACTGCATA**
3005310	***STM14_3431***	*STM2840*[f]	*norV*	14.3	17.1	49	**TGGCACACTAGCTGCAAT**
3010840	***STM14_3436***	*STM2843*[f]	*hydN*	16.1	15.4	−99	**TGGCACGATTCGTGTATA**
3019780[g]	***STM14_3448***	*STM2853*[f]	*hycA*	20.2	22.1	−77	**TGGCATGGAAAATGCTTA**
3019780[g]	***STM14_3450***	*STM2854*[f]	*hypA*	20.2	22.1	−208	**TGGCATAAATATTGCTTT**
3698110	***STM14_4239***	*STM3521*[f]	*rsr*	16.6	19.6	−305	**TGGCACGCTGGTTGCAAT**
3750440	STM14_4295	*STM3568*[f]	*rpoH*	14.1	19.4	−77	**TGGCACGGTTGTTGCTCG**
3986090	***STM14_4548***	*STM3772*[f]	*dgaA*	18.0	16.9	−69	**TGGCACAACCTTTGCTCT**
4155380	STM14_4733.P[h]	*STM_R0167*[f]	*glmZ*	14.7	18.3	52	**TGGCACGTTATGTGCAAT**
4230730	***STM14_4820***	*STM4007*[f]	*glnA*	14.4	18.3	36	**TGGCACAGATTTCGCTTT**
4402030[g]	***STM14_5013***[h]	*STM4172*[f]	*zraP*	18.5	17.8	215	**TGGCACGGAAGATGCAAG**
4402030[g]	STM14_5014[h]	*STM4173*[f]	*zraSR*	18.5	17.8	47	**TGGCATGATCTCTGCTTA**
4478500	***STM14_5102***	*STM4244*[f]	*pspG*	19.5	19.5	−84	**TGGCATGATTTTTGTAAG**
4484860	STM14_5108	STM4250	*yjbQ*	3.31	14.3	−108	TTGCATGATTTTTGCACA
4541120	STM14_5155	*STM4285*[f]	*fdhF*	15.3	17.3	14	**TGGCATAAAACATGCATA**
4623230	STM14_5249	STM4367[f]	*nsrR*	6.13	16.5	−7	**TGGCAGATATTTTGCTTG**
4807650	***STM14_5449***	*STM4535*[f]	*gfrA*	20.3	16.6	−64	**TGGCACGCCGCTTGCTCT**
Class B: intergenic sites oriented toward 3′ end of associated ORF							
2975970	STM14_3390	STM2808[f]	*nrdF*	8.93	15.5	−87	TGGCATGAATATTGCGAG
3332000	***STM14_3816***[h]	STM3151	*yghW*	11.0	14.6	−40	**TGGCTTTTATTTTGCACT**
Class C: intragenic sites directed to transcribe ORF in sense orientation							
593040	STM14_0619	STM0529	*fdrA*	3.80	15.5	38	TGGCATGTTTATTGTCTC
976360	STM14_1057[h]	STM0940[f]	*ybjX*	13.3	17.5	−49	**TGGCCTGAATCTTGCTAA**
1210780	STM14_1336	STM1167	*rimJ*	11.7	15.3	39	**TGGTACGTTTAGTGCATG**
1453820	STM14_1654	STM1361[f]	*ydiM*	4.94	15.1	−42	TGGCATTCTTTATGCTCA
1471950	STM14_1673	STM1379	*orf48*	4.66	15.5	−16	TGGCGCGCTTTTCGCTTT
1485540	STM14_1684	STM1390[f]	*orf242*	13.7	17.4	1	**TGGCATCATTATTGCCTA**
1500360	STM14_1705	STM1409	*ssaJ*	4.32	14.3	0	TGGCATTACTTATGCAGC
1691470	STM14_1929	STM1594[f]	*srfB*	6.27	14.7	−19	AGGCATATTTTTTGCCAG
1694780	STM14_1930	STM1595	*srfC*	4.68	14.7	4	TGGCGCATATGTTGCAAC
2084440	STM14_2412	STM1990[f]	*yedA*	13.8	17.4	20	**TGGCGCGCTTTTTGCCTT**
2329530	STM14_2688	STM2181[f]	*yohJ*	14.9	14.4	−16	**AGGCATTTTTCTTGCATC**
2595680	STM14_2985	*STM2430*[f]	*cysK*	18.8	16.8	−62	**TGGCATCACTGTTGCAGT**
3125610	STM14_3565	*STM2957*[f]	*rumA*	19.4	16.8	0	**TGGAACGCTTTTCGCATT**
3408660	STM14_3902	STM3222	*ygjQ*	4.29	15.4	104	TGGCGTGTTTTCTGCTTG
3716020	STM14_4255	*STM3535*[f]	*glgA*	14.4	15.8	6	**AGGCATGTTTTATGCAAA**
4049150	STM14_4625	STM3832		4.64	16.5	19	TGGAATATTTATTGCTTA
4141830	STM14_4717	STM3919[f]	*wzzE*	10.5	14.5	−156	**TGGCCTGCTATTTGCCCT**

(Continued on next page)

TABLE 1 (Continued)

Class and peak position[a]	Associated 14028s ORF[b]	Equivalent LT2 ORF[b]	Gene name(s)	Signal ratio[c]	PSSM score	Distance to peak (bp)[d]	Identified Eσ^{54} DNA binding site[e]
4146960	***STM14_4722***	STM3924[f]	*wecD*	15.6	14.3	20	**TGGCGCGGAAATTGCACA**
4257340	STM14_4850	STM4035	*fdoI*	3.99	14.4	64	TGGCGCGAATTCTGCACC
4545620	***STM14_5161***	STM4290[f]	*proP*	16.1	17.6	39	**TGGCCTGATTTTTGCAGG**
Class D: intragenic sites directed to transcribe within ORF in antisense orientation							
265270	STM14_0266	STM0244	*yeaT*	9.02	15.1	17	TGGCATACTTTTCGCTGA
519050	STM14_0544	STM0460	*mdlA*	4.27	14.8	42	TGGCGCAAATTATGCAAA
1682830	STM14_1918	STM1586[f]		8.08	14.4	0	TGGCAAGAATATTGCCAT
1769060	STM14_2012	STM1665[f]		6.10	14.6	−85	TGGCATCATTTTTTCAAG
1799990	STM14_2047	STM1697		5.20	14.4	28	TGGCGCAGCAGTTGCATT
2702030	STM14_3085	STM2517[f]	*sinH*	12.9	16.4	0	**TGGTACGGATCTTGCCAT**
Class E: intragenic site (sense) at 2919560 in 3′ end of associated ORF and oriented toward long intergenic space	STM14_3325	STM2759		10.9	14.4	7	**TGGCTCGAATAATGCTAC**
Intragenic sites (sense) in 3′ end of associated ORF and oriented toward 5′ end of adjacent ORF							
589380	STM14_0617	STM0527[f]	*allC*	12.9	16.7	−57	**TGGCATTAATGCTGCATC**
2009870	STM14_2315	STM1903[f]	*yecE*	5.81	14.7	−49	TGGCATGATTTACGCAGC
2148990	STM14_2504	STM2016	*cobT*	6.46	14.7	−4	TGGAACCCTATTTGCATA
3694410	STM14_4237	STM3518	*rtcA*	4.55	14.6	99	TGGAACGGTTTTTGCCGG
Intragenic sites (sense) in 3′ end of associated ORF and oriented toward 3′ end of adjacent ORF							
762120	STM14_0816	*STM0699*[f]		15.4	18.0	−51	**TGGCATCGATATTGCAAA**
3802330	STM14_4349	STM3613	*yhjJ*	16.3	17.4	−71	**TGGCATTAATTTTGCTGC**
4086610	STM14_4659	STM3863[f]		9.46	15.8	−25	TGGCGCGATTATTGCCAG
Class F: intragenic sites (antisense) in 5′ end of associated ORF and oriented toward 3′ end of adjacent ORF							
1000030	STM14_1085	STM0961[f]	*lolA*	8.32	16.2	0	TGGCATGAAAGCTGCTCA
2224820	STM14_2585	STM2091	*rfbG*	6.37	15.9	148	TGGCTTAAATTCTGCAAT

[a]Position (in 14028s genome) of the peak maximum (from sliding window average plot by ChIPeak).
[b]All listed Eσ^{54} binding site sequences are identical in *S.* Typhimurium strains 14028s and LT2. The Eσ^{54} binding sites associated with the boldface and italic 14028s locus tags produced transcripts identified in the expression microarray assays or by qRT-PCR (Table 2 and Table S3), the italic LT2 locus tags were previously predicted/confirmed σ^{54}-dependent promoters in *Salmonella* (25), and the underlined LT2 locus tags have *E. coli* homologues that were also shown to bind Eσ^{54} (52).
[c]The signal ratio for WT DctD250 to the Δ*rpoN* DctD250 mutant is $2^{\text{peak intensity}}$ (see Materials and Methods); all *P* values are $<10^{-19}$.
[d]Distance in base pairs upstream (positive values) or downstream (negative values) of the identified Eσ^{54} binding site from peak maximum.
[e]Identified Eσ^{54} DNA binding site based on PSSM score and proximity to the peak maximum (see Materials and Methods). Sequences in boldface were included in the 52 sites used to create the PSSM (see Materials and Methods).
[f]The same binding site was predicted by Samuels et al. (25).
[g]The peak encompasses two previously identified promoters.
[h]The Eσ^{54} binding site overlaps the σ^{70}-type promoter for which the TSS was determined by Kröger et al. (2).

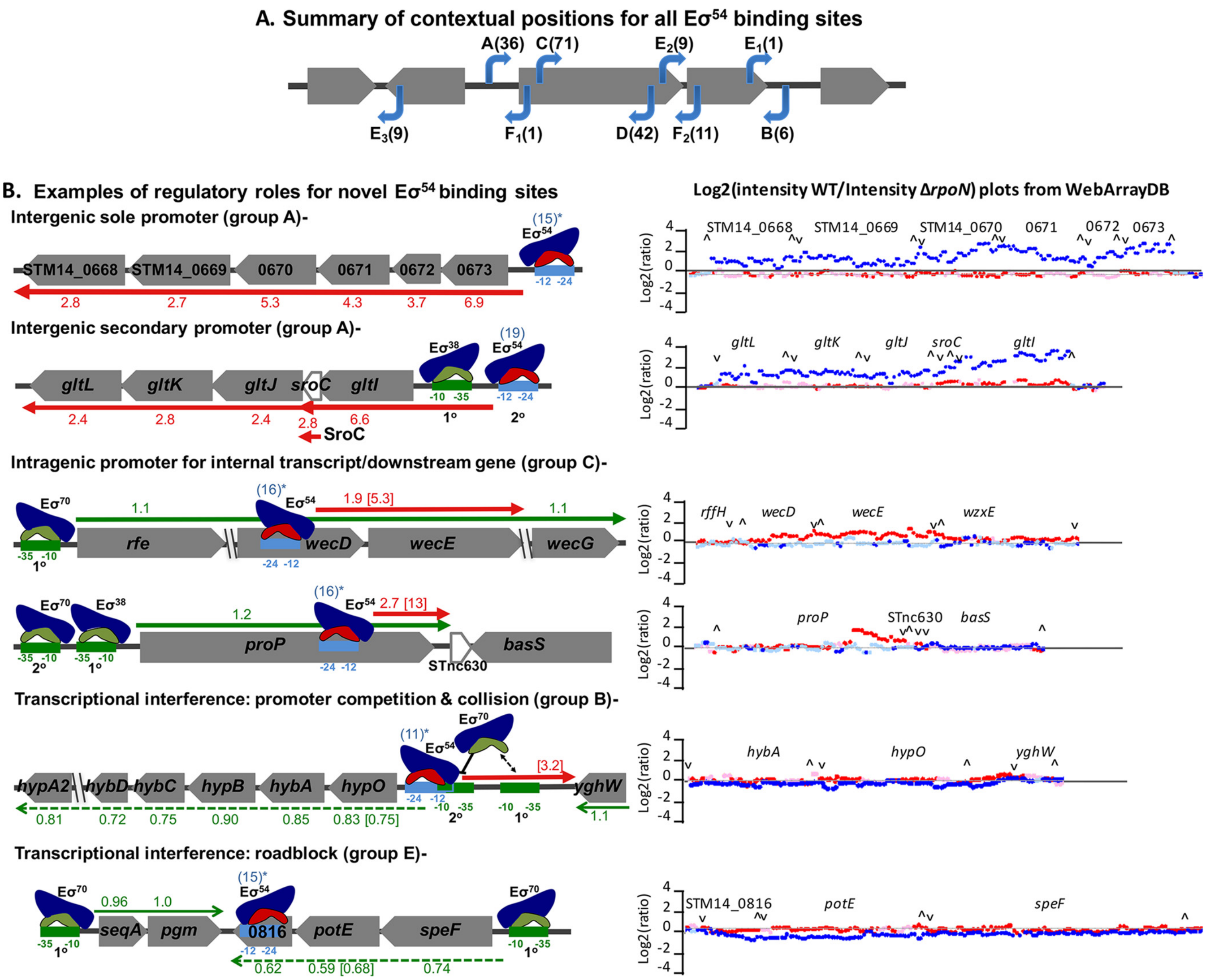

FIG 3 Illustrations of contextual positions of Eσ^{54} DNA binding sites in the *S.* Typhimurium genome and examples of potential regulatory roles with corresponding plots of microarray data. (A) The 186 Eσ^{54} binding sites that were identified in the ChIP-chip analysis are grouped into classes A to F by position and orientation (blue arrows) relative to annotated ORFs (gray arrows), as described in the Results section. Sequences of these binding sites are given in Table S2. (B) Six examples of potential regulatory roles for the Eσ^{54} binding sites are illustrated with the relative transcript levels (ratio of WT DctD250 to Δ*rpoN* DctD250 mutant), as determined by microarray (red numbers and arrows for σ^{54}-dependent gene transcripts, green numbers and arrows for $\sigma^{70/38}$-dependent transcripts) and qRT-PCR (σ^{54}-dependent gene transcripts in red brackets, σ^{70}-dependent gene transcripts in green brackets). The dashed green arrows indicate the σ^{70}-dependent gene transcripts that are significantly downregulated, but less than 2-fold, in the presence of RpoN. The fold enrichment in WT DctD250 versus the Δ*rpoN* DctD250 mutant in ChIP-chip analysis for each Eσ^{54} binding site is shown in blue, and an asterisk indicates that the binding site was confirmed by EMSA (Table 3). Primary and secondary promoter designations are from Kröger et al. (2). Adjacent to each example is the WebArrayDB plot of expression microarray data for the genes whose transcription is positively or negatively regulated by σ^{54}. Each dot, which is the $\log_2$-transformed ratio (WT DctD250 to Δ*rpoN* DctD250 mutant) 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 a significant *P* value, light pink is a positive strand without a significant *P* value, dark blue is a negative strand with a significant *P* value, and light blue is a negative strand without a significant *P* value. The upward and downward carets designate the start and end, respectively, of a gene. (The name or 14028s locus number is given.)

(sense or antisense) and position within the gene (oriented to transcribe within the gene in either the sense or antisense orientation or located within the first 250 bp or terminal 250 bp of the ORF and oriented outward). One hundred thirteen of the intragenic sites are oriented to transcribe within the ORF in which they reside; 71 of these intragenic sites (class C) are positioned in the sense direction, and 42 are antisense to the coding sequence (class D). Class C intragenic sites might have regulatory roles involving expression of sRNAs (4) or roadblock transcriptional interference (49), while class D sites have the potential to express antisense RNA (asRNA) or interfere with transcription through collision (51). Two class C sites and one class D site

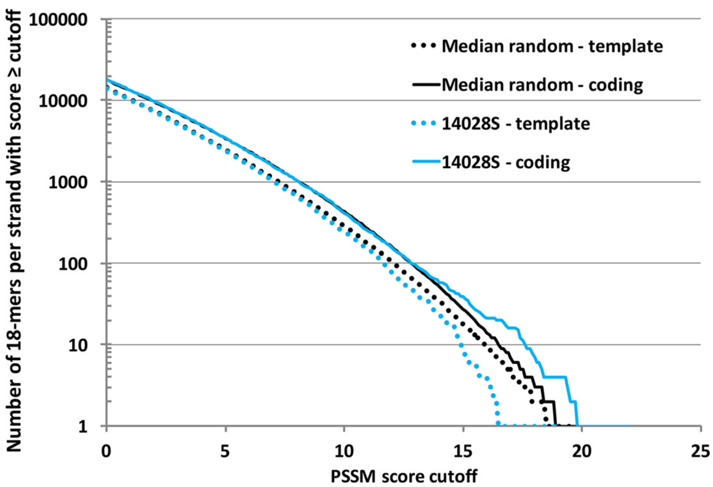

FIG 4 PSSM score distribution for intragenic 18-mer potential Eσ^{54} binding sites on the coding and template strands of *S.* Typhimurium and randomized gene sequences. The ordinate shows the number of 18-mers with PSSM scores greater than or equal to the cutoff shown on the abscissa for the coding strands (solid lines) and template strands (dashed lines) of *S.* Typhimurium protein-coding sequences (blue lines) or randomized protein-coding sequences (median values, black lines).

overlap σ^{70}-type promoters (Table S2) and thus may be involved in promoter competition (49). Class E intragenic sites (19 sites) are within 250 bp of the 3′ terminus of the ORF and oriented outward (sense orientation). Class E sites that are oriented toward a long intergenic space (1 site, subclass E_1) or the 5′ end of the adjacent ORF (9 sites, subclass E_2) conceivably act as promoters for adjacent genes or sRNAs, particularly those that include the rho-independent terminator (1, 4). Class E sites oriented toward the 3′ termini of adjacent genes (9 sites, subclass E_3) potentially interfere with transcription through polymerase collision (49). Class F intragenic sites, which are in the 5′ end of an ORF and oriented outward (antisense), face intergenic space or the 5′ end (1 site, subclass F_1) or 3′ end (11 sites, subclass F_2) of the adjacent ORF; these sites have the potential to interfere with transcription through multiple mechanisms, including collision, occlusion, and roadblock (49).

Of the 144 Eσ^{54} intragenic binding sites (Table S2), 62% are in the sense orientation (encoded on the coding strand; classes C and E), with the remaining intragenic sites in classes D and F in the antisense orientation (encoded on the template strand); for the sites with PSSM scores of >14 (Table 1), 78% of intragenic binding sites are in the sense orientation. To assess the apparent strand bias for intragenic binding sites, we compared the counts of 18-mers with PSSM scores above a given score cutoff that are encoded on the coding strand or on the template strand within protein-coding genes of *S.* Typhimurium 14028s and 1,000 sets of randomized genes (Fig. 4); the sets of randomized gene sequences contain the same number of genes as *S.* Typhimurium, each mimicking the length, codon usage, and dinucleotide frequency of the original gene. The plots reflecting the distribution of sites on the coding and template strands for the randomized gene sequences show slightly more 18-mers with high PSSM scores in the coding strand than in the template strand; this difference is due to biased codon usage and the resulting asymmetry between the coding and template strands, which is reflected in the randomized genes. For the 14028s genome, there are more potential binding sites (higher PSSM scores) in the coding strand than expected based on the random sequences (Fig. 4); the excess of high-scoring sites relative to the randomized sequences is close to that observed in the analysis of the whole genome (Fig. 2). However, sites with high PSSM scores occur much less frequently in the template strand than expected based on the randomized sequences (Fig. 4), suggesting suppression of high-scoring intragenic Eσ^{54} binding sites in the antisense orientation. This result is consistent with the potential of antisense Eσ^{54} intragenic binding sites that have promoter activity to disrupt gene transcription by collision.

A recent study of the Eσ^{54} genomic binding sites in *E. coli* by Bonocora et al. (52) revealed a similar distribution of the binding sites to that seen in this analysis for *S.* Typhimurium. The majority of the 135 identified *E. coli* Eσ^{54} binding sites are intragenic (62%), and of these intragenic sites, most are in the sense orientation relative to the overlapping gene (68%). Twenty-nine of the *E. coli* Eσ^{54} DNA binding sites are orthologues of the *S.* Typhimurium identified binding sites (underlined in Table S2); these orthologous sites are not all identical in sequence but are located in the same position relative to annotated genes. Twenty-one of the orthologous sites are intergenic, including 19 known σ^{54}-dependent promoters, and 8 are intragenic, including the *rumA* and *proP* intragenic σ^{54}-dependent promoters whose potential regulatory roles are described below.

σ^{54} both positively and negatively regulates transcription of genes and noncoding RNAs involved in diverse cellular processes. Strand-specific analysis of the σ^{54}-dependent transcriptome of *S.* Typhimurium 14028s was performed on the high-density tiled microarrays with the same cultures (WT expressing DctD250 and Δ*rpoN* mutant expressing DctD250) used in the ChIP-chip analysis, thereby facilitating the comparison of Eσ^{54} binding sites with σ^{54}-dependent transcripts generated from those sites. Differential expression of transcripts was assessed as described in Materials and Methods. Table 2 lists significantly upregulated operons, upregulated intragenic transcripts, and downregulated genes/operons in the WT DctD250 versus Δ*rpoN* DctD250 strains from the microarray analysis. The associated Eσ^{54} binding sites (from ChIP-chip assay), those transcripts that were previously known to be σ^{54} dependent, and the conditions under which some σ^{54}-dependent transcripts were detected in the global transcriptome of *S.* Typhimurium under infection-related conditions (2) are also provided in Table 2.

(i) σ^{54}-dependent expression of operons associated with intergenic Eσ^{54} DNA binding sites. Sixty-six genes from 22 operons were expressed at >2-fold higher levels in the WT Dct250 strain than in the Δ*rpoN* DctD250 strain, with *P* values of $<10^{-17}$ (see Table S3 in the supplemental material). The first gene in each operon that was upregulated in the presence of σ^{54} and DctD250 has an Eσ^{54} binding site immediately upstream (Table 2). Most of these σ^{54}-dependent transcripts were previously reported for *S.* Typhimurium LT2, as indicated in Table 2, but σ^{54}-dependent transcription of *gltI-sroC-gltJKL*, *yeaGH*, *hisJ*, and *pspG* was first confirmed for *S.* Typhimurium in this microarray analysis. The transcription start sites (TSSs) for mRNA from 10 of the 22 upregulated operons were mapped to predicted σ^{54}-dependent promoters by Kröger et al. (2) in transcriptome sequencing (RNA-seq) and differential RNA sequencing (dRNA-seq) analyses performed with *S.* Typhimurium strain 4/74 under 22 different infection-relevant conditions (indicated in Table 2). Detection in the Kröger et al. study of only 10 of the 22 σ^{54}-dependent operon 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 (25).

As illustrated in Fig. 3B, some σ^{54}-regulated operons appear to be transcribed solely from a σ^{54}-dependent promoter (STM14_0673-0668); while other σ^{54}-regulated operons are also expressed from one or more promoters recognized by RNAP associated with σ^{70}-related factors (*gltI-sroC-gltJKL*). The analysis by Kröger et al. (2) of the *S.* Typhimurium infection-related transcriptomes showed that the *gltI-sroC-gltJKL* operon has a primary transcript expressed during stationary phase, which maps to the σ^{38}-dependent promoter, and has a secondary transcript that is expressed under conditions that activate genes in *Salmonella* pathogenicity island 2 (SPI2), including low-phosphate, low-pH medium (PCN medium), peroxide shock, and nitric oxide shock. The TSS for this secondary *gltI-sroC-gltJKL* transcript maps to the Eσ^{54} DNA binding site identified by ChIP-chip (Table 2; Fig. 3B) and confirms the previously predicted σ^{54}-dependent promoter for the *gltIJKL* operon in *S.* Typhimurium (25, 53). Two recently published studies demonstrate that in *S.* Typhimurium the sRNA SroC is processed from a transcript that terminates between *gltI* and *gltJ* and that the level of SroC is positively

TABLE 2 Summary of the σ^{54}-dependent transcriptome in *S.* Typhimurium expressing DctD250

Type and locus tag(s) in 14028s	Gene name(s)	Expression signal ratio[a]	PSSM score for associated Eσ^{54} binding site	Reference(s) confirming expression in *Salmonella*[b]	Conditions for TSS detection in reference 2[c]
First genes in operon exhibiting σ^{54}-dependent expression					
STM14_0431[d]	*prpB*	3.09	17.8	29	ND
STM14_0546[d]	*glnK*	5.17	17.8	25	H_2O_2 shock
STM14_0673[d]		6.86	17.3	25	ND
STM14_0757[d]		7.70	17.1	25	ND
STM14_0773[d]	*gltI*	6.62	21.5	This work	H_2O_2 shock
STM14_0964[d]	*glnH*	4.39	19.6	102	Mid-exponential phase
STM14_1558[d]	*yeaG*	4.10	18.5	This work	ND
STM14_1582[d]	*astC*	2.77	20.6	25	H_2O_2 shock
STM14_2040[d]	*pspA*	9.95	18.0	25	0_2 shock
STM14_2900[d]	*hisJ*	2.71	17.2	This work	ND
STM14_2901[d]	*argT*	3.64	14.7	25	H_2O_2 shock
STM14_2907[d]		11.0	22.3	25	ND
STM14_3431[d]	*norV*	9.65	17.1	25	ND
STM14_3436[d]	*hydN*	3.45	15.4	25	Anaerobic
STM14_3448[d]	*hycA*	2.68	22.1	25	ND
STM14_3450[d]	*hypA*	1.76[e]	22.1	25	ND
STM14_4239[d]	*rsr*	4.44	19.6	25	H_2O_2 shock
STM14_4548[d]	*dgaA*	9.23	16.9	25, 32	ND
STM14_4820[d]	*glnA*	12.8	18.3	25	H_2O_2 shock
STM14_5013[d]	*zraP*	12.2	17.8	25	ND
STM14_5102[d]	*pspG*	2.23	19.5	This work	NaCl shock
STM14_5449[d]	*gfrA*	16.3	16.6	25, 33	ND
σ^{54}-dependent intragenic transcripts					
STM14_4722[f]/STM14_4723[d]	*wecD/wecE*	1.90	14.3	This work	ND
STM14_5161[f]	*proP*	2.67	17.6	This work	ND
Genes exhibiting indirect σ^{54}-dependent downregulation					
STM14_1795		0.491	NA	This work	NA
STM14_3638	*lysA*	0.425	NA	This work	NA
STM14_5085	*malE*	0.474	NA	This work	NA
STM14_5087	*lamB*	0.464	NA	This work	NA
Genes associated with Eσ^{54} binding site within operon exhibiting σ^{54}-dependent downregulation[g]					
STM14_0816[f,h]		0.628	18.0	This work	NA
STM14_1057[f]	*ybjX*	0.778	17.5	This work	NA
STM14_2504[f]	*cobT*	0.891	14.7	This work	NA
STM14_2521[f]	*cbiC*	0.796	8.38	This work	NA
STM14_3143.P[f,h]	*glmY*	0.848	19.4	This work	NA
STM14_3814[f,h]	*hypO*	0.833	14.6	This work	NA

[a]The microarray expression signal ratio (WT DctD250 to Δ*rpoN* DctD250 mutant) is $1/2^{(\text{LIMMA-generated average } M \text{ value})}$. *M* values and *P* values for transcripts from these genes and other genes in the same operon, as well as for the intragenic transcripts, are presented in Table S3. The *M* value ($\log_2$ ratio of the Δ*rpoN* DctD250 mutant to WT DctD250) for each gene is the average of *M* values for all tiling array probes corresponding to the gene (Data Set S1). The *M* value for the intragenic transcript starting in *wecD* is the average of *M* values for tiling array probes immediately downstream of the intragenic Eσ^{54} binding site in *wecD* from positions 4146957 to 4147350 with *P* values of <0.05 (-0.9403 [Data Set S2]) averaged with the median *M* value for all tiling array probes corresponding to *wecE* (-0.9160; $P < 10^{-19}$ [Data Set S1]). For intragenic transcript in *proP*, the *M* values (with $P < 0.05$) for tiling array probes immediately downstream of the intragenic Eσ^{54} binding site (positions 4545579 to 4545979 [Data Set S2]) were averaged.

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

[c]Growth conditions under which Kröger et al. (2) detected a transcription start site (TSS) correlating to the σ^{54}-dependent promoter. ND, no correlating TSS was detected; NA, not applicable (reported 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]Gene downstream of identified Eσ^{54} binding site (see Table S2 for the predicted binding site sequence).

[e]The second gene in this operon, *hypB*, has a signal ratio of >2 (Table S3).

[f]Gene that contains or is immediately downstream of the Eσ^{54} binding site (see Table S2 for the binding site).

[g]Genes within the operon associated with the indicated gene exhibit σ^{54}-dependent downregulation (10 to 50% downregulated with *P* values of <0.01 [Table S3]).

[h]The Eσ^{54} binding site associated with the downregulated gene(s) has been shown to have σ^{54}-dependent promoter activity (25, 50) (Fig. 3B).

regulated by Eσ^{38} following late-exponential-phase growth (31, 54). Dual regulation by σ^{54} (RpoN) and σ^{38} (RpoS), as seen for the *gltI-sroC-gltJKL* operon, is common in *E. coli*; in microarray analyses of σ^{54}- and σ^{38}-regulated genes in *E. coli* by Dong et al. (55), ~60% of genes in the σ^{54} regulon are also controlled by σ^{38}, and σ^{54} negatively regulates the level of σ^{38} in the cell. The regulation of *gltI-sroC-gltJKL* is linked not only with the RpoS and RpoN regulons but also with the GcvB posttranscriptional regulon; Miyakoshi et al. (31) demonstrated that SroC acts as a "RNA sponge" by base pairing with GcvB, which is an sRNA that posttranscriptionally represses numerous mRNAs of genes coding for amino acid transporters, regulatory proteins, and catabolic enzymes in *S.* Typhimurium, including the *gltI-sroC-gltJKL* mRNA (3). Thus, σ^{54}-regulated SroC positively regulates expression from its parental mRNA (*gltI-sroC-gltJKL*) and relieves GcvB-mediated repression for other mRNAs of genes coding for amino acid transporters (e.g., *tppB*, *livJ*, *livK*, and *metQ*), regulatory proteins (e.g., *lrp* and *iciA*), and biosynthetic enzymes (e.g., *ilvC*, *gdhA*, and *ndk*) (3).

(ii) σ^{54}-dependent expression of transcripts associated with intragenic Eσ^{54} DNA binding sites. Two novel σ^{54}-dependent transcripts were detected that initiate from ChIP-chip-identified intragenic Eσ^{54} binding sites; these intragenic transcripts are associated with the Eσ^{54} binding sites in *wecD* and *proP* (Table 2; Fig. 3B). σ^{54}-dependent expression from the intragenic promoters was confirmed by quantitative reverse transcription-PCR (qRT-PCR); the relative transcript levels (ratio of WT DctD250 to Δ*rpoN* DctD250 mutant) from the intragenic *wecD* and *proP* promoters were 5.3 ± 3.2 ($P = 0.0015$) and 13 ± 10 ($P = 0.0068$), respectively (Fig. 3B). The start of the σ^{54}-dependent transcript within *proP* was mapped by 5′ rapid amplification of cDNA ends (RACE) to the +1 position relative to the intragenic Eσ^{54} binding site; the 3′ end of the transcript was not mapped but extends at least 30 nucleotides after the translational stop of *proP* based on the annealing position of the primer used to create cDNA for 5′ RACE (see Materials and Methods). This σ^{54}-dependent transcript may be processed to an sRNA, as suggested by enrichment of a *proP* transcript in Hfq immunoprecipitation (IP) (4) and evidence for the prevalence of sRNAs generated from the 3′ untranscribed regions (UTRs) in *Salmonella*, *E. coli*, *Vibrio cholerae*, and *Streptomyces coelicolor* (56). In the analysis of the *S.* Typhimurium transcriptome by Kröger et al. (1), an sRNA (STnc630) maps to the intergenic region downstream of *proP*, but a TSS for the sRNA mapped to an intergenic promoter by dRNA-seq.

Although the microarray analysis did not reveal an upregulated transcript in association with the intragenic Eσ^{54} binding site in *rumA* (STM14_3565/STM2957 [Table 1]), we previously demonstrated that this intragenic site, which is immediately upstream of the ppGpp synthase gene (*relA*), is an active σ^{54}-dependent promoter in the presence of DctD250 (25). The physiological relevance of the orthologous σ^{54}-dependent promoter in *E. coli* was recently reported by Brown et al. (57). Under nitrogen stress, the bEBP NtrC activates the intragenic σ^{54}-dependent promoter to express *relA*, thereby linking nitrogen stress response to the stringent response (57).

(iii) Indirect σ^{54}-dependent downregulation of gene transcripts. There were four ORF transcripts that exhibited >2-fold downregulation in the presence of σ^{54} and DctD250: STM14_1795, *lysA* (STM14_3638), *malE* (STM14_5085), and *lamB* (STM24_5087) (Table 2). The downregulation of *malE* was confirmed by qRT-PCR; the relative transcript level (WT DctD250/Δ*rpoN* DctD250) for *malE* was 0.49 ± 0.18 ($P = 0.038$). Based on the ChIP-chip analysis, there is an Eσ^{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 S2), and there are no Eσ^{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 resulting from an indirect mechanism.

STM14_1795 is annotated in NCBI as encoding an acid shock protein precursor that is required for growth under moderately acid conditions, but the mechanism for regulation of this gene has not been 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 (48). The σ^{54}-dependent gene *argT* (STM14_2901), which codes for 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 2) and relief of GcvB negative regulation by SroC (31), which is also expressed from a σ^{54}-dependent promoter (Fig. 3B). The increased levels of lysine, due to ArgT lysine transporter activity in the lysine-rich nutrient broth medium, are likely to result in repression of *lysA* expression. It should be noted here that there is another σ^{54}-dependent gene that is annotated as a diaminopimelate decarboxylase gene, STM14_2907 (Table 2), but the product of this gene has not been confirmed. The remaining two genes that exhibited σ^{54}-dependent downregulation are *malE*, which encodes a periplasmic protein involved in maltose transport, and *lamB*, which encodes a porin involved in transport of maltose and maltodextrins (58). These genes are contained 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 cyclic AMP (cAMP) and cAMP receptor protein (CRP) (58). Three σ^{54}-dependent operons encode different mannose family phosphotransferase systems (PTSs) and associated enzymes. Substrates have been identified for two of these mannose family PTSs: D-glucosaminic acid (*dgaABCDEF* [32]) and fructose lysine/glucose lysine (*gfrABCDEF* [33]). Expression of one or more of these PTS operons may result in decreased expression of *malE* and *lamB* through a catabolite repression mechanism (58).

(iv) σ^{54}-dependent downregulation of operon transcripts. Several downregulated operon transcripts (*speF-potE*-STM14_0816, *ybjX*, *cbiABCDEFGHIJKLMNOPQ-cobUT*, *glmY*, and *hypO-hybABCDE*) that are associated with Eσ^{54} binding sites did not meet the 2-fold cutoff for differential expression, but transcript levels for genes in these five operons were significantly reduced (up to 1.8-fold; $P \leq 0.01$) in the WT DctD250 strain compared to the Δ*rpoN* DctD250 strain (Table S3). 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 (50, 59). Potential transcriptional interference mechanisms for the Eσ^{54} DNA binding sites associated with the *speF-potE*-STM14_0816 and *hypO-hybABCDE* operons are illustrated in Fig. 3B. 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^{70}-dependent transcripts for *speF-potE*-STM14_0816 (0.68 ± 0.086; $P = 0.038$) and for *hypO* (0.75 ± 0.17; $P = 0.041$) (Fig. 3B). The Eσ^{54} binding sites associated with these downregulated σ^{70}-dependent transcripts both exhibit σ^{54}-dependent promoter activity: the Eσ^{54} binding site in the sense orientation within the last 200 bp of STM14_0816 (STM0669) was shown to have σ^{54}-dependent promoter activity in a *lacZ* transcriptional fusion assay (25), and the intergenic Eσ^{54} binding site that is antisense to and overlaps the *hypO* σ^{70}-dependent promoter (and is in the sense orientation relative to *yghW*) has σ^{54}-dependent promoter activity (3.2 ± 1.1 relative transcript level in WT DctD250 versus the Δ*rpoN* DctD250 mutant; $P = 0.0025$), as measured by qRT-PCR.

For the Eσ^{54} DNA binding sites with potential regulatory roles in gene transcription in the *S.* Typhimurium strains expressing DctD250, the growth conditions and bEBPs that may activate the σ^{54}-dependent promoters are generally not evident. However, since the hydrogenase 2 (*hypO-hybABCDE*) operon is differentially expressed under conditions that favor fermentation (60, 61), we hypothesized that FhlA, a bEBP that is activated by the fermentation product formate (62), stimulates transcription from the σ^{54}-dependent promoter in the intergenic region between *hypO* and *yghW* ($P_{\text{inter-}hypO\text{-}yghW}$) and thereby modulates expression of the hydrogenase 2 operon (Fig. 3B). Transcript levels associated with $P_{inter\text{-}hypO\text{-}yghW}$, the first gene of the hydrogenase 2 operon (*hypO*), and *fdhF* (one of the four σ^{54}-dependent operons normally activated by FhlA [25, 62]) were assayed by qRT-PCR for the *S.* Typhimurium 14028s WT, Δ*rpoN*, and Δ*fhlA* strains grown under anaerobic conditions in LB and LB plus 30 mM formate. Transcription of *hypO* and from $P_{inter\text{-}hypO\text{-}yghW}$ was unaltered in the WT strain compared to the Δ*rpoN* and Δ*fhlA* strains (see Table S4 in the supplemental material); transcript levels for the positive control gene, *fdhF*, were significantly reduced in both the Δ*rpoN* and Δ*fhlA* strains compared to the WT, as expected (Table S4). Thus, the condition under which $P_{inter\text{-}hypO\text{-}yghW}$ is expressed and whether the hydrogenase 2 operon is regulated by $P_{inter\text{-}hypO\text{-}yghW}$ are still unknown.

Consistent with the σ^{54}-dependent downregulation of operon transcripts reported here, a recent study of σ^{54}-dependent regulation of gene expression in *E. coli* strain BW25113 by Schaefer et al. (63) identified 26 genes with associated intergenic or intragenic $E\sigma^{54}$ binding sites that had reduced expression in the presence of σ^{54}. Promoter binding competition was indicated as a mechanism for transcriptional interference for six of these genes; the remaining downregulated genes have $E\sigma^{54}$ binding sites positioned within the coding sequence or downstream antisense, suggesting interference with transcription through roadblock or collision mechanisms (63). Only two of the 26 *E. coli* $E\sigma^{54}$ binding sites that were reported to negatively regulate gene expression were identified in *S.* Typhimurium: the class A $E\sigma^{54}$ site associated with *argT* and the class C $E\sigma^{54}$ site in *glgA* (Table 1). In our transcriptome analysis of WT and Δ*rpoN S.* Typhimurium strains expressing DctD250, *argT* expression was upregulated 3.6-fold in the presence of σ^{54} (Table 2) and *glgA* expression was not significantly changed. The different regulatory activities of the *argT*-associated $E\sigma^{54}$ binding sites in *E. coli* (63) and *S.* Typhimimurium (this study) illustrate how $E\sigma^{54}$ binding sites can have different regulatory roles depending on the potential promoter activity of the $E\sigma^{54}$ binding site. Because the *E. coli* study did not utilize a promiscuous, constitutively active bEBP or growth conditions that activate NtrC, which is the cognate bEBP for the *argT* σ^{54}-dependent promoter (64), $E\sigma^{54}$ is likely to remain in stable closed complex at its promoter and block binding of $E\sigma^{70}$ to the overlapping primary σ^{70}-dependent promoter for *argT*. In our study, transcription from the *argT* σ^{54}-dependent promoter was stimulated by the promiscuous, constitutively active DctD250. For the different classes of $E\sigma^{54}$ binding sites identified in this study, the potential and mechanism for regulatory activity depend on whether the $E\sigma^{54}$-bound site is simply a protein-DNA complex or is a promoter that can respond to an activated bEBP under particular growth conditions.

Many identified $E\sigma^{54}$ binding sites are not associated with σ^{54}-dependent transcripts. Of the 186 identified $E\sigma^{54}$ genomic binding sites, only 24 are associated with transcripts that were considered differentially expressed in WT DctD250 versus Δ*rpoN* DctD250 (Table 2). As detailed in Materials and Methods, the designation of differentially expressed genes from intergenic $E\sigma^{54}$ binding sites (class A sites) required a >2-fold change (with $P < 0.01$) in the averaged expression levels for all the probes that span the gene (assessed from Data Set S1 in the supplemental material); and for potential differentially expressed transcripts from $E\sigma^{54}$ binding sites in classes B to F, including potential regulatory RNAs, a cluster of individual probes immediately downstream must have exceeded 2-fold change in expression with $P < 0.05$ (assessed from Data Set S2 in the supplemental material). The constraints of these designations may have excluded some actual σ^{54}-dependent transcripts. For example, several probes downstream of the intergenic sites associated with *clpP* and *fdhF* (Table 1) exhibited >2-fold increased expression (Data Set S2), but the average for probes across the associated gene (Data Set S1) did not meet the 2-fold cutoff; *fdhF* was differentially expressed in our previous study in *S.* Typhimurium LT2 (25).

It is likely that additional $E\sigma^{54}$ binding sites would have σ^{54}-dependent promoter activity under conditions in which the cognate bEBP would be activated, as suggested by the lack of transcription activity associated with four intragenic $E\sigma^{54}$ binding sites that were previously shown to be functional promoters in *lacZ* fusion assays (sites associated with STM0699, STM2430 [*cysK*], STM2939 [*ygcH*], and STM2957 [*rumA*]) (25). The intragenic Eσ^{54} binding site within *rumA* has a TSS mapped to it under nitric oxide shock and in "NonSPI2 medium" in the transcriptome study by Kröger et al. (2). Under the conditions of our assays, factors that influence the activity of σ^{54}-dependent promoters may be absent or affected, such as the indirect positive effects of DksA and ppGpp (65) and a promoter sequence-directed specificity for a particular bEBP even when the bEBP acts from solution, as demonstrated for the *Klebsiella pneumoniae nifH* promoter for activation by NtrC and NifA (66, 67). Unexpressed σ^{54}-dependent genes for transcription factors or regulatory RNAs would impact the extent of the σ^{54} transcriptome assessed in this study.

It is also possible that σ^{54}-dependent transcripts associated with some Eσ^{54} binding sites were expressed but not detected in our comparative microarray assays. Poor detection of transcripts may arise from transcript instability, overlap of non-σ^{54}-dependent promoters, or expression below the level of detection for the microarray assays. Low σ^{54}-dependent promoter activity may be due to titration of DctD250, which was expressed at a low level (25), or the need for higher concentrations of DctD250 to activate transcription because it is not tethered to the promoter through binding to an enhancer sequence, thus reducing the likelihood of proper protein-protein interactions for activation (68).

Our simulations with random sequences (Fig. 2) suggest that many of the Eσ^{54} DNA binding sites may have arisen in the genome by chance in the absence of selective constraints and thus may not necessarily provide a direct benefit to the organism. These sites are unlikely to be functional promoters, but may play a role in the evolution of new regulatory networks, as discussed in "Concluding remarks" below.

Sequence determinants for Eσ^{54} and σ^{54} binding in a stable closed complex. 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 (11). However, σ^{54} in its native state has been shown to bind specifically to the *Sinorhizobium meliloti nifH* promoter, albeit at an ~100-fold-lower affinity than the holoenzyme (40, 69). To address whether DNA binding of σ^{54} in the absence of RNAP contributed to the genomic binding sites identified in our ChIP-chip assays, 11 sites of various PSSM scores and binding signal ratios from the six classes of binding sites were assessed for binding by σ^{54} and Eσ^{54} in electrophoretic mobility shift assays (EMSAs) (Table 3). A derivative of the *Klebsiella pneumoniae nifH* σ^{54}-dependent promoter, designated *nifH049* (40), was used as a positive control for both Eσ^{54} and σ^{54} *in vitro* binding; *nifH049* 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 (40, 70). EMSAs for binding of Eσ^{54} and σ^{54} were performed with 50-bp heteroduplex probes containing two unpaired bases immediately 3′ of the conserved −12 GC motif, which mimics the DNA distortion that is a definitive feature of the stable Eσ^{54} closed complex (reviewed in reference 71). Previous *in vitro* DNA binding studies in the Buck laboratory (15, 17, 70, 72, 73) showed Eσ^{54} interacts with bases on the bottom strand of the DNA distortion at positions −11 and −10 in a stable closed complex, and σ^{54} (in the absence of core RNAP) has a 6-fold-higher affinity for *S. meliloti nifH* promoter sequence containing the DNA distortion than for *nifH* homoduplex promoter sequence.

Table 3 summarizes the EMSA results for all 11 sites, the *nifH049* positive control and the −24TT, −12TT *proP* negative control, which is mutated in the −24GG and −12GC promoter elements of the *proP* intragenic site. Two new sites that bind σ^{54} in its native state were identified: the intragenic sites in *proP* and STM14_0816 (Fig. 5). The probes for these two sites were shifted by 100 to 500 nM σ^{54}, which is in the range for the predicted σ^{54} intracellular concentration (~140 nM [see Materials and Methods]); Eσ^{54} bound the same probes with ~10-fold greater affinity (Fig. 5; Table 3). The specificity of Eσ^{54} and σ^{54} binding in the EMSAs was assessed for the *proP* site by competition with 10- to 500-fold molar excess of nonspecific or specific competitor DNA (see

TABLE 3 Summary of EMSA results for Eσ^{54} and σ^{54} binding to selected sites

Binding site-associated ORF(s)	Binding site class	Binding signal ratio	PSSM score for WT binding site	Eσ^{54} binding[a]	σ^{54} binding[a]	Altered Eσ^{54} core binding sequence on top strand of heteroduplex probe[b]
Selected sites with binding signal ratios of >3						
STM14_0530 (*clpP*)	A	12.5	17.6	+++	−	TGTCACGTATTTTGC**CG**G
STM14_0816	E	15.4	18.0	+++	+	TGGCATCGATATTGC**CC**A
STM14_1057 (*ybjX*)	C	13.3	17.5	+	−	TGGCCTGAATCTTGC**GC**A
STM14_2345 (*otsA*)	D	5.84	5.04	+++	−	GGGAATGGAATATGC**CT**G
STM14_2985 (*cysK*)	C	18.8	16.8	+++	−	TGGCATCACTGTTGC**CT**T
STM14_3816 (*yghW*)	B	11.0	14.6	+++	−	TGGCTTTTATTTTGC**CA**T
STM14_4295 (*rpoH*)	A	14.1	19.4	+	−	TGGCACGGTTGTTGC**GA**G
STM14_4722 (*wecD*)	C	15.6	14.3	+++	−	TGGCGCGGAAATTGC**CA**A
STM14_4820 (*glnA*)	A	14.4	18.3	−	−	TGGCACAGATTTCGC**GG**T
STM14_5080 (*yjbA*)	F	6.0	11.5	++	−	AGGCGCGAATAATGC**CG**C
STM14_5161 (*proP*)	C	16.1	17.6	+++	+	TGGCCTGATTTTTGC**CA**G
Control binding sites						
nifH049				+++	+	TGGTATGTTTTTTGC**CA**T
−24TT, −12TT *proP*				−	−	TTTCCTGATTTTTTT**CA**G

[a]Eσ^{54} binding affinity for DNA site based on estimated protein concentration required for 50% binding of probe from 3 replicate assays (see Materials and Methods): +++, ≤0.3 μM; ++, >0.3 μM and ≤0.8 μM; +, >0.8 μM and ≤2 μM; −, >2 μM. Values for 50% bound heteroduplex probes are not reported as equilibrium dissociation constants to avoid confusion with the affinity of Eσ^{54} or σ^{54} for homoduplex sites.

[b]Core DNA-binding sequence for Eσ^{54} or σ^{54} in the 50-bp double-stranded oligonucleotide probe; the two boldface bases in the top strand sequence differ from the wild-type sequence of the binding site, causing a DNA distortion adjacent to the −12GC promoter element in the double-stranded oligonucleotide, in which the bottom strand is the wild-type sequence.

Materials and Methods); the Eσ^{54}-*proP* and σ^{54}-*proP* complexes were resistant to 500-fold molar excess of nonspecific DNA but were reduced by 10 to 50% at 10-fold molar excess specific DNA competitor and by >95% at 500-fold molar excess specific competitor. As expected, binding of core RNAP alone to the assayed DNA binding sites was weak or not detectable at all concentrations used in the binding reactions, and the low levels of shifted complex were reduced to undetectable by 10-fold molar excess of nonspecific competitor DNA.

The lower DNA binding activity of σ^{54} in the absence of RNA polymerase and the previously demonstrated high affinity of σ^{54} for RNAP (13–15) (Table 3) suggest that σ^{54} is unlikely to occupy the genomic binding sites in the absence of RNAP. This is consistent with the estimated colocalization of RNAP with most *E. coli* Eσ^{54} genomic binding sites, based on ChIP-seq analyses (52). However, it has been suggested that σ^{54} may regulate promoter activity by binding to some promoters after open complex formation and affecting promoter escape or abortive recycling (71). DNA sequence features that determine binding of σ^{54} in the absence of RNAP are not clear; sequence alignment of the two newly identified binding sites with the previously defined σ^{54} binding sites indicates a consensus sequence, but comparison with a strong Eσ^{54} binding site that did not bind σ^{54} in the EMSA analysis reveals no conserved sequence that is unique to the σ^{54} binding sites (Fig. 6). This result suggests that a composite of sequence determinants is required for σ^{54} binding.

An unexpected result from the EMSA analysis of holoenzyme (Eσ^{54}) binding to the heteroduplex DNA substrates was the effect on binding affinity of specific nucleotide substitutions in the top strand that created the DNA distortion 3′ of the −12 GC promoter element (Table 3), since previous reports on *in vitro* binding of Eσ^{54} to the heteroduplex probes indicated Eσ^{54} interacts with bases on the bottom strand of the DNA distortion at positions −11 and −10 in the stable closed complex (15, 17, 70, 72, 73). The heteroduplex DNA substrate corresponding to the known σ^{54}-dependent *glnA* promoter, which had GG substituted for TT at positions −11 and −10 in the top strand, did not exhibit a shift in the EMSA even at an Eσ^{54} concentration >1,000-fold higher than the molar concentration that gave 50% binding of a 43-bp homoduplex probe containing the *E. coli glnA* σ^{54}-dependent promoter sequence (74), which has the identical 18-bp core binding sequence to the *S.* Typhimurium promoter. Comparison of

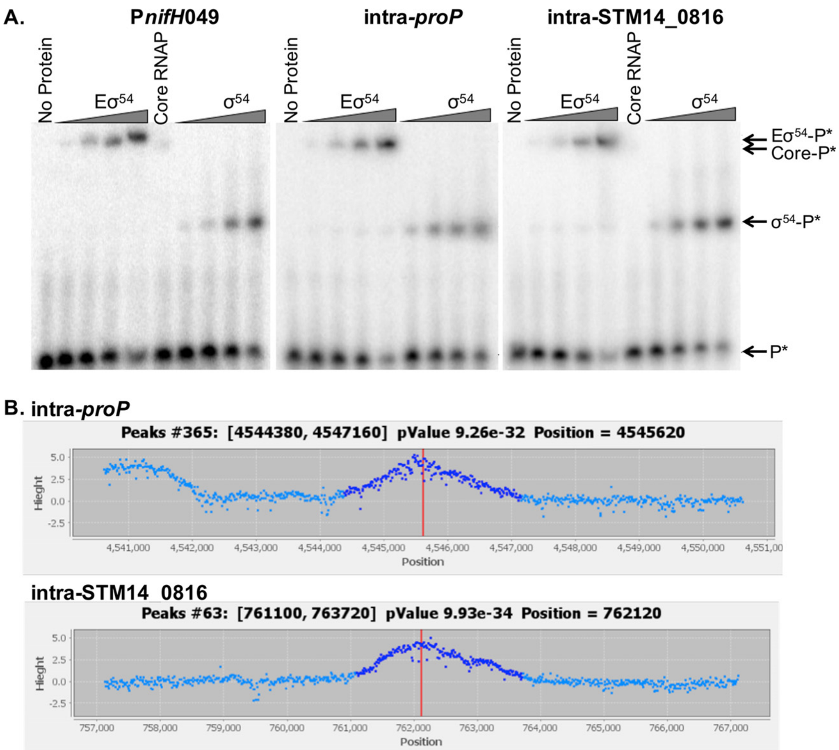

FIG 5 *In vitro* assays of Eσ^{54} and σ^{54} binding to DNA sequences identified in ChIP-chip analysis of Eσ^{54} genomic binding sites. (A) Representative EMSAs for binding reaction mixtures containing 16 nM ^{32}P-labeled 50-bp heteroduplex oligonucleotide probes (*P) for the positive control (*nifH049* promoter) and the intragenic Eσ^{54} binding sites in *proP* and STM14_0816 with 0, 10, 50, 100, or 200 nM Eσ^{54}, 100 nM core RNAP, or 100 nM, 500 nM, 1 μM, or 2 μM σ^{54} protein, unbound probe, and the protein-DNA complexes, Eσ^{54}-*P, σ^{54}-*P, or core-*P (marked by arrows) separated by native PAGE (6.5% acrylamide). Images are from Typhoon scans of gel-exposed phosphorimager screens. (B) Examples of ChIPeak output from analysis of Eσ^{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*.

the heteroduplex top strand sequences for all of the assayed sites revealed that the sites with G substituting for T at the −11 position (*ybjX*, *rpoH*, and *glnA* [Table 3]) exhibited weak or no binding of Eσ^{54} regardless of the binding signal ratio of the WT sites from ChIP-chip (Table 3). Sequence comparison of the ChIP-chip-identified 186 genomic Eσ^{54} DNA binding sites revealed only 12 sites (6.5%) with a G at the −11 position and no sites with G at both

K. p. nifH049	TAAACAGGCACGGCT**GG**TATGTTTTTT**GC**ACTTCTCTGCTGGCAAACACT
S. m. nifH	TTATTTCAGACGGCT**GG**CACGACTTTT**GC**ACGATCAGCCCTGGGCGCGCA
S. T. *proP*	AACAGTAACGTTATT**GG**CCTGATTTTT**GC**ACGTTTGTTGATGCTGGCGGT
STM14_0816	TTTCGCCACCGGACT**GG**CATCGATATT**GC**AAACGCGCGAGGAGATGCGCT
Consensus	WWH~~HBHVR~~S~~VB~~KR~~YT~~~~**GG**~~~~Y~~M~~Y~~S~~DHTWTT~~~~**GC**~~~~AMD~~H~~BYVBBVN~~K~~R~~S~~DN~~RC~~R~~S~~W~~
*S.*T. *hypO-yghW*	CCGTTACGAAGACCT**GG**CTTTTATTTT**GC**ACTGTTCGCGAAGAAGTTATT

FIG 6 Consensus sequence for σ^{54} binding in the absence of RNAP core. Alignment of the sequence from −40 to +10 (relative to the +1 transcription start site) of the σ^{54}-dependent promoters *K. pneumoniae* (*K. p.*) *nifH049* and *S. meliloti* (*S. m.*) *nifH*, which were previously shown to bind σ^{54} in the absence of RNAP (40, 70), with the newly identified *S.* Typhimurium (*S.*T.) *proP* and STM14_0816 σ^{54} binding sites (Table 3). The −24 GG and −12 GC promoter elements are in boldface. The extent of the DNase I footprint for σ^{54} in the closed complex with *nifH049* and *S. meliloti nifH* (40, 70) is indicated by the black bar underneath the sequence. The consensus sequence was generated for the four σ^{54} binding 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 that does not bind σ^{54} but has the −14 to −17 T-tract previously proposed to be associated with σ^{54} binding (40) and the same bases at the DNA distortion in the probes used for EMSA as *nifH049* and *proP* (Table 3). Nucleic acid residues shared between the consensus sequence and nonbinding sequence are struck through.

the −10 and −11 positions. This result provides further insight into the sequence determinants for Eσ^{54} binding to DNA to form a stable closed complex.

Comparison of the multiple-sequence alignment of the 33 promoter sequences (Fig. 1A) to the multiple-sequence alignment of all 186 *S.* Typhimurium Eσ^{54} binding sites (Fig. 1C) and the Eσ^{54} binding sites with PSSM scores of >14 (Fig. 1D) suggests that the active promoters have less variance in the CA nucleotide pair at −23 and −22, adjacent to the highly conserved GG promoter element at −25 and −24, and more A/T-rich sequence at the positions from −11 to −9 (relative to the +1 TSS), which is consistent with the compilation analysis of σ^{54}-dependent promoters from 47 different bacterial species (75).

Concluding remarks. We have characterized the σ^{54} regulon of *S.* Typhimurium strain 14028s that is expressing the promiscuous, constitutively active bEBP (DctD250). One hundred eighty-six Eσ^{54} genomic binding sites were identified, most of which are located within genes (77.4%), and 24 σ^{54}-dependent transcripts were defined, 22 of which are associated with intergenic Eσ^{54} binding sites. These results, together with our previous microarray and promoter-*lacZ* fusion assays (25) and the characterization of *glmY* and *glmZ* transcription in *S.* Typhimurium by Gopel et al. (50), confirm 33 σ^{54}-dependent promoters in *S.* Typhimurium. In addition, nine transcripts appear to be downregulated in a σ^{54}-dependent manner.

The position and context of the Eσ^{54} genomic binding sites suggest potential roles in transcription regulation, ranging from the promoter for expressing mRNA and regulatory RNAs to directing transcription interference through promoter competition, collision, or roadblock mechanisms (49), which are consistent with regulatory activities described herein for some of the novel and confirmed Eσ^{54} binding sites from this study. These regulatory mechanisms allow the σ^{54} regulon to intersect and impact the regulons of σ^{70} and other alternative sigma factors under changing growth conditions.

For the many Eσ^{54} binding sites that are likely to have arisen by chance in the genome in the absence of selective constraints, we speculate that they could play a generic role in facilitating the target search by Eσ^{54} for true promoters (76) or simply be tolerated if they cause no harm; selection against sites that do cause harm is reflected in the reduced frequency of intragenic antisense Eσ^{54} binding sites with high PSSM scores. In addition, the dynamic emergence of (weak) Eσ^{54} binding sites throughout the genome could play an important role in adaptations of the organism to novel environmental conditions. Along these lines, analysis of evolution of regulatory networks suggested that the networks evolve rapidly by emergence of new regulatory sites and regulatory proteins (77, 78), and frequent random emergence of transcription factor binding sites in a genome could be one of the mechanisms that facilitate the rapid evolution of regulatory networks.

MATERIALS AND METHODS

Oligonucleotides, enzymes, media, and chemicals. All oligonucleotides used in this work were synthesized by Integrated DNA Technologies and are described in Table S1 in the supplemental material. All enzymes were purchased from New England BioLabs, unless otherwise indicated, and used according to manufacturer's recommendations. Cells were grown in Lennox broth (LB [Fisher]), nutrient broth (NB [Difco]), InSPI2 medium [inducing *Salmonella* pathogenicity island 2 medium [2]), or MOPS (morpholinepropanesulfonic acid) medium with ammonium chloride as the nitrogen source and glucose as the carbon source (32, 79). Medium supplements where noted are as follows: 5 mM glutamine (Sigma-Aldrich), 100 μg/ml ampicillin (Amp [Fisher]), 50 μg/ml kanamycin (Kan [Roche Life Science]), 100 μg/ml rifampin (Rif [Fisher]), and 30 mM potassium formate (Sigma-Aldrich).

Bacterial strains and plasmids. *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) strain ATCC 14028s is the wild-type strain (WT) in these studies. pPBHP192, a derivative of pTrcHisC (Invitrogen) that expresses the *Sinorhizbium meliloti* DctD AAA+ domain (E141-S390, designated DctD250) with an N-terminal His_6 tag (described in reference 25), was introduced by electroporation into wild-type and Δ*rpoN S.* Typhimurium 14028s strains. *S.* Typhimurium strains 14028s Δ*rpoN* (ACB01), 14028s Δ*fhlA* (ACB03), and 14028s Δ*ntrC* (CEH01) were created via the Lambda Red recombination system (80), using the method described by Miller et al. (32). Oligonucleotides 1 and 2 (for Δ*rpoN*), 3 and 4 (for Δ*fhlA*), and 73 and 74 (for Δ*ntrC*) were used to amplify the kanamycin resistance cassette (*kan*) from pKD4 (80), and the Δ*rpoN*::*kan*, Δ*fhlA*::*kan*, and Δ*ntrC*::*kan* mutant strains were confirmed by PCR with primer pairs 55 and 56, 3 and 4, and 75 and 76, respectively. The substitution mutations were transduced by P22 HT *int* into a clean genetic background, and *kan* was excised as described in reference 32. Excision of *kan* in each of the mutants ACB01 (Δ*rpoN*), ACB03 (Δ*fhlA*), and CEH01 (Δ*ntrC*) was confirmed by kanamycin sensitivity and DNA sequencing (Genewiz, Inc.) of PCR products from amplification of chromosomal DNA with the same primers used to confirm the insertion of the cassette. These mutant strains have a cassette scar that encodes a ribosome binding site and translation start codon, which minimizes polar effects on downstream genes (80).

All plasmids were maintained in *Escherichia coli* strain DH5α (81). pJES937, which is a pET28a(+)-derived plasmid expressing σ^{54} with an N-terminal His$_6$ tag (82), was introduced into One Shot chemically competent *E. coli* BL21(DE3) (Life Technologies). Plasmids that were introduced into *S.* Typhimurium 14028s by electroporation were first passed through the *S. enterica* restriction-negative modification-positive strain MS1868 (83).

ChIP-chip and transcriptome profiling on tiling microarrays. Wild-type and Δ*rpoN* (ACB01) *S.* Typhimurium 14028s cells containing pPBHP192 (WT DctD250 and Δ*rpoN* DctD250 mutant, respectively), which expresses DctD250 at a low level without induction (84), were grown overnight with aeration at 37°C from single colonies in NB-Amp. Three biological replicates were prepared for each strain. Overnight cultures were used to inoculate 100 ml NB-Amp, and the cultures were aerobically grown to the mid-exponential growth phase (optical density at 600 nm [OD$_{600}$] of ≈0.5). For the transcriptome profiling, 40 ml was removed from each culture and centrifuged; the cell pellets were stored on dry ice for subsequent RNA isolation (see below). For chromatin immunoprecipitation (ChIP), 50 ml of the remaining culture was treated with rifampin (100 μg/ml) for 10 min at 37°C before treatment with formaldehyde (1.1%) for 10 min and then used for preparation of the ChIP samples as previously described in Samuels et al. (25). Primers 61 and 62 were used in the ligation-mediated PCR (LM-PCR). The size range of the LM-PCR products was 0.15 to 1.0 kb, with most products present in the 0.3- to 0.5-kb range; the products were labeled using Cy3- and Cy5-dCTP (GE Healthcare), and efficiency of incorporation was determined as described in reference 85. Dye switches were performed between the biological replicates (R01-03).

RNA for transcriptome profiling was extracted from the frozen cell pellets, described above, using the RNAsnap method (86). RNA was treated with RNase-free DNase I (Ambion) and ethanol precipitated. Absence of residual DNA was confirmed by PCR using primers to amplify the *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 as described in reference 87. As for the ChIP samples, the dyes used to label cDNA from the WT DctD250 and Δ*rpoN* DctD250 strains were flipped in different biological replicates, to assess whether labeling efficiency for Cy3 or Cy5 impacts the microarray results; no difference was observed.

The labeled ChIP-enriched DNA samples and the labeled strand-specific, single-stranded cDNAs were hybridized, as described in reference 38, to a NimbleGen tiling microarray of ~387,000 50-mer oligonucleotides tiling the *S.* Typhimurium strain 14028s genome at overlapping 24-base steps on both strands (88). Arrays were scanned with a GenePix 4000B laser scanner (Molecular Devices) at a 5-μm resolution. Signals were quantified by NimbleScan software v2.4 (Roche NimbleGen).

Analysis of ChIP-chip assay data. Background calculations for each array and normalization of intensity values within and between arrays were performed as described in reference 88, using WebarrayDB (89). Differential analyses for biological replicates R01 and R02 for the Δ*rpoN* DctD250 mutant versus R01, R02, and R03 (dye-switch) for WT 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 Δ*rpoN* DctD250 and WT DctD250) and *P* values (from LIMMA analysis) (89).

Analysis of the ChIP-chip data was accomplished with the original software ChIPeak, designed for this study (see the supplemental material). ChIPeak estimates statistical significance of signal peaks (*P* value), maps peak positions in the genome, and predicts associated binding sites using the standard position-specific scoring matrix (PSSM) method, as previously described for the Motif Locator software (90). 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 were the start and end positions for each probe (NCBI GenBank accession no. CP001363.1) and the associated *M* values from the WebarrayDB LIMMA differential analysis for the combined biological replicates. *M* values were obtained for the Δ*rpoN* DctD250 mutant relative to WT DctD250, so the option in ChIPeak to inverse input data was selected so that resulting peak data are for WT DctD250 versus the Δ*rpoN* DctD250 mutant (peaks point upward). ChIPeak determines significant peaks for the enriched ChIP DNA using a sliding window average; the window size was set to 500 bp (average size of the ChIP DNA fragments), and the window was moved along the genomic sequence in 10-bp steps. In the analysis of ChIP-chip data from this study, a conservative *P* value cutoff of 10^{-19} was selected based on visual analysis of the peaks with different *P* values provided by the ChIPeak program. The output from the ChIPeak analysis included 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 a 500-bp window that yields the local maximum), and a list of up to 3 predicted binding sites with the highest PSSM scores within a 350-bp (user-selected) distance from the peak. The peak intensity, which is a log$_2$ value, is converted to a binding signal ratio for WT DctD250 to Δ*rpoN* DctD250 mutant as follows: $2^{\text{peak intensity}}$ = binding signal ratio. The PSSM that was utilized by the program to predict the Eσ^{54} sites was generated from core 18-bp sequences of 52 Eσ^{54} binding sites; the core 18-bp sequences include the highly conserved −12 and −24 sequence elements of σ^{54}-dependent promoters and span positions −9 to −26 (relative to a +1 transcription start site). These 52 sites comprise 33 sequences that have been shown to be associated with σ^{54}-dependent transcripts in *Salmonella* (25, 50; this study) and 19 additional Eσ^{54} binding sites that were associated with peaks enriched >10-fold and predicted based on a PSSM from the 33 known sequences (Fig. 1A and B). In this study, the highest-scoring predicted site is reported for each peak unless the top two predicted binding sites have PSSM scores that differ by less than 1 point, in which case the site closest to the peak maximum is reported.

Analysis of transcriptome profiling data. As in the analysis of the ChIP-chip data, the background calculations for each array and normalization of intensity values within and between arrays were performed as described in reference 88, using WebarrayDB (89). Differential analyses for biological replicates R01, R02, and R03 for WT DctD250 versus R01, R02, and R03 for the Δ*rpoN* DctD250 mutant, utilizing normalized and averaged intensity $\log_2$ values corresponding to each probe on the arrays, were performed in WebarrayDB to obtain the median *M* values ($\log_2$ ratio of Δ*rpoN* DctD250 mutant and WT DctD250) and *P* values (from LIMMA analysis) (89). Data Set S1 lists the *S*. Typhiumurium 14028s annotated genes that have statistically significant median *M* values for all tiling array probes corresponding to the gene ($P < 0.01$, LIMMA analysis for the three biological replicates). Genes were considered differentially expressed if the *M* value was <-1.0 or >1.0, reflecting at least a 2-fold change in gene expression in the Δ*rpoN* DctD250 mutant versus WT DctD250, and the *P* value was <0.01. Operons immediately downstream of ChIP-chip-identified intergenic $E\sigma^{54}$ binding sites were considered differentially expressed if one gene within the operon met the criterion for a differentially expressed gene. To identify potential novel transcripts associated with ChIP-chip-identified $E\sigma^{54}$ binding sites, the median *M* values for individual probes on the high-density tiled microarray ($P < 0.05$, LIMMA analysis of three biological replicates [Data Set S2]) were assessed for clustered probes with *M* values of <-1.0 (reflecting at least 2-fold decreased expression in the Δ*rpoN* DctD250 mutant relative to WT DctD250) immediately downstream of the identified $E\sigma^{54}$ binding site.

The *M* values for the transcriptome profiling of the Δ*rpoN* DctD250 mutant and WT DctD250 were converted to expression signal ratios for WT DctD250 to Δ*rpoN* DctD250 mutant as follows: $1/(2^{M\ \mathrm{value}})$ = expression signal ratio.

Generation of randomized genome and gene sequences. One thousand randomized genome sequences were generated from the *S*. Typhimurium 14028s genome using the "m1c1" model implemented in Genome Randomizer (91, 92) (available for download at http://www.cmbl.uga.edu/software.html). This program first divides the actual genome DNA sequence into segments corresponding to individual genes and intergenic regions. Subsequently, a random sequence is generated of the same length as each gene and intergenic region as a Markov chain trained to mimic the nucleotide and dinucleotide composition of each particular intergenic region and the nucleotide, dinucleotide, and codon usage of each gene. Finally, the randomized gene and intergenic region sequences are reassembled into a randomized genome sequence. For the randomized gene sequences, only the protein-coding genes of the *S*. Typhimurium genome were used in the program to generate the 1,000 sets of randomized genes.

qRT-PCR with selected 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 (Bio-Rad) using oligonucleotides 29-54 and 69-70 as primer pairs (Table S1). The 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), to determine transcript levels. Three technical replicates were performed for each biological replicate. Averages of the technical replicates for target genes were normalized to the averaged *rpoD* transcript levels. (The *rpoD* transcript levels were constant in the WT and mutant strains.) Primers and annealing conditions were optimized to give 90 to 100% PCR efficiency (based on the slope of the standard curves). The averages of biological replicates were used for the reported ratios of WT DctD250 to the Δ*rpoN* DctD250 mutant and calculation of standard deviation and *P* values (Student's *t* test).

qRT-PCR assays for expression of *hypO* and the intergenic region between *hypO* and *yghW* under physiologically relevant conditions. Expression from σ^{54}-dependent $P_{\mathrm{inter}\text{-}hypO\text{-}yghW}$ and from the primary/secondary σ^{70}-dependent promoters for the *hypO-hybABCDEFG* operon was assessed in *S*. Typhimurium WT, AB01 (Δ*rpoN*), and AB03 (Δ*fhlA*) grown aerobically or anaerobically (as described in reference 60) in LB or LB containing 30 mM formate. Cultures were inoculated with overnight cultures (grown in LB containing 5 mM glutamine) to an OD_{600} of ~0.01 and then grown to an OD_{600} of ~0.5. RNA was isolated and reverse transcribed 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 the *hyb* operon]), 69/70 (*fdhF* [known FhlA- and σ^{54}-dependent gene]), and 43/44 (inter-*hypO-yghW*); the latter 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 biological replicates, with three technical replicates each, were utilized in the qRT-PCR; averages of the three biological replicates were used for the reported ratios and calculation of standard deviation and *P* values (Student's *t* test).

5′ RACE. Transcription start sites associated with selected $E\sigma^{54}$-binding sites were determined using the 5′ RACE system for rapid amplification of cDNA ends kit, version 2.0 (Life Technologies), per the manufacturer's instruction. A sample from the RNA that was prepared from biological replicate R02 for the microarray analysis was reverse transcribed with gene-specific primers 65 (*glnA*) and 66 (*proP*). After synthesis of a 3′-poly(C) tail, primers 63 and 67 were each used in combination with the Abridged Anchor primer to amplify the 5′ region. Nested PCR with primer 64 or 68 and the Abridged Anchor primer was performed, and the amplified RACE products were sequenced after TOPO TA cloning (Invitrogen). Sequences were mapped against the genomic sequence (NCBI GenBank accession no. CP001363.1) to identify the transcription start site.

Purification of His$_6$-σ^{54}. His$_6$-tagged σ^{54} was purified from *E. coli* BL21(DE3) containing pJES937, encoding His$_6$-RpoN. Purification was performed as described by Cannon et al. with the modifications described by Kelly and Hoover (93, 94). Protein was quantified via the Bradford assay (Bio-Rad), per the manufacturer's instructions. σ^{54} purity and quality were assessed by Coomassie stain and Western blot, as described previously (94).

EMSA. The affinity of Eσ^{54} or σ^{54} for selected predicted binding sites from the ChIP-chip analysis was assessed using heteroduplex DNA probes in EMSAs as described by Gallegos and Buck (95), with the exception that the 88-bp heteroduplex probe was replaced with a 50-bp heteroduplex probe (corresponding to sequence from −40 to +10 relative to the predicted +1 transcription start site), which did not appear to affect the affinity of either Eσ^{54} or σ^{54} for the *nifH049* probe (data not shown). Oligonucleotides 5-28 and 71-72 (Table S1) were used to form the heteroduplex probes, which were labeled with [γ-^{32}P]ATP (Perkin-Elmer Life Sciences) on the top or bottom strand depending on the 5′ nucleotide (96). Binding reactions were performed, as in reference 95, with Eσ^{54} (2:1 RpoN-RNA polymerase [Epicentre]), σ^{54}, or core RNAP. The range of Eσ^{54} and σ^{54} concentrations utilized, given in the Results, overlaps calculated estimates for the intracellular concentrations of core RNAP (~260 nM) and σ^{54} (~140 nM) in *E. coli* during exponential growth in LB; to calculate these estimated intracellular concentrations, we used the ratio of σ^{54} to σ^{70} in the cell as 0.16 (97), the level of σ^{70} in the cell as 60 to 170 fmol/μg total protein (19, 97), the level of RNAP core as 46 fmol/μg total protein (13), 450 μg total protein/10^9 cells (19), and a cell volume of 0.8×10^{-15} liter (98). Electrophoresis of binding reactions, imaging, and quantitation of protein-DNA complexes were performed as described previously (99).

Competition EMSAs, to determine the specificity of σ^{54} and Eσ^{54} binding to target DNA, were implemented as described above with the following modifications: nonspecific competitor DNA (sonicated calf thymus DNA, GE Healthcare) or specific, unlabeled target DNA was added to the binding reactions at 1-, 10-, 50-, 100- or 500-fold molar excess relative to the labeled target DNA before addition of core RNAP, Eσ^{54}, or σ^{54}, which were added at empirically determined concentrations that give ~50% bound labeled DNA target.

Accession number(s). The data for both the ChIP-chip and expression microarrays have been deposited in GEO; the superseries accession number is GSE75180.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/JB.00816-16.

SUPPLEMENTAL FILE 1, XLSX file, 0.2 MB.
SUPPLEMENTAL FILE 2, XLSX file, 0.4 MB.
SUPPLEMENTAL FILE 3, PDF file, 1.9 MB.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant MCB-1051175 (A.C.K.). M.M. and S.P. acknowledge support from the North American Meat Institute (NAMI) and from the Binational Agricultural Research and Development Fund (BARD). J.G.F. acknowledges the support of USDA project no. 6040-32000-006-00.

We appreciate the insightful ideas and suggestions provided by Timothy Hoover, Lawrence Shimkets, Claiborne Glover, David Garfinkel, and Russell Karls. We thank the following undergraduate students who assisted with analysis of the Eσ^{54} binding sites and σ^{54}-dependent transcription in *S.* Typhimurium: Caleb Gulledge and Deanna Walden, who were participants in the NSF REU site for Research in Prokaryotic Biology at the University of Georgia (DBI-1460671), and Selin Odman and Isabella Tondi-Resta, who received UGA Center for Undergraduate Research Opportunities Assistantships. We thank Nathan Hartman for technical assistance.

The mention of trade names or commercial products in the manuscript is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

REFERENCES

1. Kröger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp K, Chao Y, Sittka A, Hebrard M, Handler K, Colgan A, Leekitcharoenphon P, Langridge GC, Lohan AJ, Loftus B, Lucchini S, Ussery DW, Dorman CJ, Thomson NR, Vogel J, Hinton JC. 2012. The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. Proc Natl Acad Sci U S A 109:E1277–E1286. https://doi.org/10.1073/pnas.1201061109.
2. Kröger C, Colgan A, Srikumar S, Handler K, Sivasankaran SK, Hammarlof DL, Canals R, Grissom JE, Conway T, Hokamp K, Hinton JC. 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. Cell Host Microbe 14:683–695. https://doi.org/10.1016/j.chom.2013.11.010.
3. Hebrard M, Kröger C, Srikumar S, Colgan A, Handler K, Hinton JC. 2012. sRNAs and the virulence of *Salmonella enterica* serovar Typhimurium. RNA Biol 9:437–445. https://doi.org/10.4161/rna.20480.
4. Chao Y, Papenfort K, Reinhardt R, Sharma CM, 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. https://doi.org/10.1038/emboj.2012.229.

5. Adkins JN, Mottaz HM, Norbeck AD, Gustin JK, Rue J, Clauss TR, Purvine SO, Rodland KD, Heffron F, Smith RD. 2006. Analysis of the *Salmonella* Typhimurium proteome through environmental response toward infectious conditions. Mol Cell Proteomics 5:1450–1461. https://doi.org/10.1074/mcp.M600139-MCP200.
6. Shi L, Ansong C, Smallwood H, Rommereim L, McDermott JE, Brewer HM, Norbeck AD, Taylor RC, Gustin JK, Heffron F, Smith RD, Adkins JN. 2009. Proteome of *Salmonella enterica* serotype Typhimurium grown in a low Mg/pH medium. J Proteomics Bioinform 2:388–397. https://doi.org/10.4172/jpb.1000099.
7. Qiu Y, Nagarajan H, Embree M, Shieu W, Abate E, Juarez K, Cho BK, Elkins JG, Nevin KP, Barrett CL, Lovley DR, Palsson BO, Zengler K. 2013. Characterizing the interplay between multiple levels of organization within bacterial sigma factor regulatory networks. Nat Commun 4:1755. https://doi.org/10.1038/ncomms2743.
8. Bang IS, Frye JG, McClelland M, Velayudhan J, Fang FC. 2005. Alternative sigma factor interactions in *Salmonella*: sigma and sigma promote antioxidant defences by enhancing sigma levels. Mol Microbiol 56:811–823. https://doi.org/10.1111/j.1365-2958.2005.04580.x.
9. Mauri M, Klumpp S. 2014. A model for sigma factor competition in bacterial cells. PLoS Comput Biol 10:e1003845. https://doi.org/10.1371/journal.pcbi.1003845.
10. deHaseth P, Lohman T, Burgess R, Record MT, Jr. 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. https://doi.org/10.1021/bi00602a006.
11. Feklistov A, Sharon B, Darst S, Gross C. 2014. Bacterial sigma factors: a historical, structural, and genomic perspective. Annu Rev Microbiol 68:357–376. https://doi.org/10.1146/annurev-micro-092412-155737.
12. Merrick M. 1993. In a class of its own—the RNA polymerase sigma factor σ^{54} (σ^{N}). Mol Microbiol 10:903–909. https://doi.org/10.1111/j.1365-2958.1993.tb00961.x.
13. Maeda H, Fujita N, Ishihama A. 2000. Competition among seven *Escherichia coli* sigma subunits: relative binding affinities to the core RNA polymerase. Nucleic Acids Res 28:3497–3503. https://doi.org/10.1093/nar/28.18.3497.
14. Scott D, Ferguson A, Gallegos M, Pitt M, Buck M, Hoggett J. 2000. Interaction of sigma factor σ^{N} with *Escherichia coli* RNA polymerase core enzyme. Biochem J 352:539–547.
15. Gallegos M-T, Buck M. 1999. Sequences in σ^{N} determining holoenzyme formation and properties. J Mol Biol 288:539–553. https://doi.org/10.1006/jmbi.1999.2704.
16. Campbell EA, Tupy JL, Gruber TM, Wang S, Sharp MM, Gross CA, Darst SA. 2003. Crystal structure of *Escherichia coli* sigmaE with the cytoplasmic domain of its anti-sigma RseA. Mol Cell 11:1067–1078. https://doi.org/10.1016/S1097-2765(03)00148-5.
17. Bush M, Dixon R. 2012. The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. Microbiol Mol Biol Rev 76:497–529. https://doi.org/10.1128/MMBR.00006-12.
18. Wang W, Carey M, Gralla JD. 1992. Polymerase II promoter activation: closed complex formation and ATP-driven start site opening. Science 255:450–453. https://doi.org/10.1126/science.1310361.
19. Grigorova IL, Phleger NJ, Mutalik VK, Gross CA. 2006. Insights into transcriptional regulation and sigma competition from an equilibrium model of RNA polymerase binding to DNA. Proc Natl Acad Sci U S A 103:5332–5337. https://doi.org/10.1073/pnas.0600828103.
20. Gruber TM, Gross CA. 2003. Assay of *Escherichia coli* RNA polymerase: sigma-core interactions. Methods Enzymol 370:206–212. https://doi.org/10.1016/S0076-6879(03)70018-4.
21. Li J, Overall CC, Johnson RC, Jones MB, McDermott JE, Heffron F, Adkins JN, Cambronne ED. 2015. ChIP-Seq analysis of the sigmaE regulon of *Salmonella enterica* serovar Typhimurium reveals new genes implicated in heat shock and oxidative stress response. PLoS One 10:e0138466. https://doi.org/10.1371/journal.pone.0138466.
22. Skovierova H, Rowley G, Rezuchova B, Homerova D, Lewis C, Roberts M, Kormanec J. 2006. Identification of the sigmaE regulon of *Salmonella enterica* serovar Typhimurium. Microbiology 152:1347–1359. https://doi.org/10.1099/mic.0.28744-0.
23. Ibanez-Ruiz M, Robbe-Saule V, Hermant D, Labrude S, Norel F. 2000. Identification of RpoS (sigma(S))-regulated genes in *Salmonella enterica* serovar Typhimurium. J Bacteriol 182:5749–5756. https://doi.org/10.1128/JB.182.20.5749-5756.2000.
24. Mutalik VK, Nonaka G, Ades SE, Rhodius VA, Gross CA. 2009. Promoter strength properties of the complete sigma E regulon of *Escherichia coli* and *Salmonella enterica*. J Bacteriol 191:7279–7287. https://doi.org/10.1128/JB.01047-09.
25. Samuels DJ, Frye JG, Porwollik S, McClelland M, Mrazek J, Hoover TR, Karls AC. 2013. Use of a promiscuous, constitutively-active bacterial enhancer-binding protein to define the sigma54 (RpoN) regulon of *Salmonella* Typhimurium LT2. BMC Genomics 14:602. https://doi.org/10.1186/1471-2164-14-602.
26. Appia-Ayme C, Hall A, Patrick E, Rajadurai S, Clarke TA, Rowley G. 2012. ZraP is a periplasmic molecular chaperone and a repressor of the zinc-responsive two-component regulator ZraSR. Biochem J 442:85–93. https://doi.org/10.1042/BJ20111639.
27. Karlinsey JE, Maguire ME, Becker LA, Crouch ML, Fang FC. 2010. The phage shock protein PspA facilitates divalent metal transport and is required for virulence of *Salmonella enterica* sv. Typhimurium. Mol Microbiol 78:669–685. https://doi.org/10.1111/j.1365-2958.2010.07357.x.
28. Mills PC, Richardson DJ, Hinton JC, Spiro S. 2005. Detoxification of nitric oxide by the flavorubredoxin of *Salmonella enterica* serovar Typhimurium. Biochem Soc Trans 33:198–199. https://doi.org/10.1042/BST0330198.
29. Palacios S, Escalante-Semerena JC. 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. https://doi.org/10.1128/JB.182.4.905-910.2000.
30. Kustu S, Burton D, Garcia E, McCarter L, McFarland N. 1979. Nitrogen control in *Salmonella*: regulation by the *glnR* and *glnF* gene products. Proc Natl Acad Sci U S A 76:4576–4580. https://doi.org/10.1073/pnas.76.9.4576.
31. Miyakoshi M, Chao Y, 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. https://doi.org/10.15252/embj.201490546.
32. Miller KA, Phillips RS, Mrazek J, Hoover TR. 2013. *Salmonella* utilizes D-glucosaminate via a mannose family phosphotransferase system permease and associated enzymes. J Bacteriol 195:4057–4066. https://doi.org/10.1128/JB.00290-13.
33. Miller KA, Phillips RS, Kilgore PB, Smith GL, Hoover TR. 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. https://doi.org/10.1128/JB.00339-15.
34. Hartman CE, Samuels DJ, Karls AC. 2016. Modulating *Salmonella* Typhimurium's response to a changing environment through bacterial enhancer-binding proteins and the RpoN regulon. Front Mol Biosci 3:41. https://doi.org/10.3389/fmolb.2016.00041.
35. Swords W, Cannon B, Benjamin WH, Jr. 1997. Avirulence of LT2 strains of Salmonella typhimurium results from a defective *rpoS* gene. Infect Immun 65:2451–2453.
36. Wilmes-Riesenberg M, Foster J, Curtiss R, III. 1997. An altered *rpoS* allele contributes to the avirulence of *Salmonella typhimurium* LT2. Infect Immun 65:203–210.
37. Jarvik T, Smillie C, Groisman E, Ochman H. 2010. Short-term signatures of evolutionary change in the *Salmonella enterica* serovar Typhimurium 14028 genome. J Bacteriol 192:560–567. https://doi.org/10.1128/JB.01233-09.
38. Santiviago CA, Reynolds MM, Porwollik S, Choi SH, Long F, Andrews-Polymenis HL, McClelland M. 2009. Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. PLoS Pathog 5:e1000477. https://doi.org/10.1371/journal.ppat.1000477.
39. Zhao K, Mingzhu L, Burgess R. 2010. Promoter and regulon analysis of nitrogen assimilation factor, σ^{54}, reveal alternative strategy for *E. coli* MG1655 flagellar biosynthesis. Nucleic Acids Res 38:1273–1283. https://doi.org/10.1093/nar/gkp1123.
40. Buck M, Cannon W. 1992. Specific binding of the transcription factor sigma-54 to promoter DNA. Nature 358:422–424. https://doi.org/10.1038/358422a0.
41. Lutz S, Bohm R, Beier A, 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. https://doi.org/10.1111/j.1365-2958.1990.tb02010.x.

42. Leonhartsberger S, Huber A, Lottspeich F, 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. https://doi.org/10.1006/jmbi.2000.4451.
43. Huang Y, Mrázek J. 2014. Assessing diversity of DNA structure-related sequence features in prokaryotic genomes. DNA Res 21:285–297. https://doi.org/10.1093/dnares/dst057.
44. Aparicio O, Geisberg JV, Struhl K. 2004. Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences *in vivo*. Curr Protoc Cell Biol Chapter 17:Unit 17.7. https://doi.org/10.1002/0471143030.cb1707s23.
45. Myers KS, Yan H, Ong IM, Chung D, Liang K, Tran F, Keles S, Landick R, Kiley PJ. 2013. Genome-scale analysis of *Escherichia coli* FNR reveals complex features of transcription factor binding. PLoS Genet 9:e1003565. https://doi.org/10.1371/journal.pgen.1003565.
46. Wade JT, Reppas NB, Church GM, Struhl K. 2005. Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites. Genes Dev 19:2619–2630. https://doi.org/10.1101/gad.1355605.
47. Tanay A. 2006. Extensive low-affinity transcriptional interactions in the yeast genome. Genome Res 16:962–972. https://doi.org/10.1101/gr.5113606.
48. Stragier P, Borne F, Richaud F, Richaud C, Patte JC. 1983. Regulatory pattern of the *Escherichia coli lysA* gene: expression of chromosomal *lysA-lacZ* fusions. J Bacteriol 156:1198–1203.
49. Shearwin KE, Callen BP, Egan JB. 2005. Transcriptional interference—a crash course. Trends Genet 21:339–345. https://doi.org/10.1016/j.tig.2005.04.009.
50. Gopel Y, Luttmann D, Heroven AK, Reichenbach B, Dersch P, 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. https://doi.org/10.1093/nar/gkq986.
51. Georg J, Hess WR. 2011. cis-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev 75:286–300. https://doi.org/10.1128/MMBR.00032-10.
52. Bonocora RP, Smith C, Lapierre P, Wade JT. 2015. Genome-scale mapping of *Escherichia coli* sigma54 reveals widespread, conserved intragenic binding. PLoS Genet 11:e1005552. https://doi.org/10.1371/journal.pgen.1005552.
53. Studholme D. 2002. Enhancer-dependent transcription in *Salmonella enterica* Typhimurium: new members of the σ^N regulon inferred from protein sequence homology and predicted promoter sites. J Mol Microbiol Biotechnol 4:367–374.
54. Fuentes DN, Calderon PF, Acuna LG, Rodas PI, Paredes-Sabja D, Fuentes JA, Gil F, Calderon IL. 2015. Motility modulation by the small non-coding RNA SroC in *Salmonella* Typhimurium. FEMS Microbiol Lett 362:fnv135. https://doi.org/10.1093/femsle/fnv135.
55. Dong T, Yu R, Schellhorn H. 2011. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in *Escherichia coli*. Mol Microbiol 79:375–386. https://doi.org/10.1111/j.1365-2958.2010.07449.x.
56. Miyakoshi M, Chao Y, Vogel J. 2015. Regulatory small RNAs from the 3′ regions of bacterial mRNAs. Curr Opin Microbiol 24:132–139. https://doi.org/10.1016/j.mib.2015.01.013.
57. Brown DR, Barton G, Pan Z, Buck M, Wigneshweraraj S. 2014. Nitrogen stress response and stringent response are coupled in *Escherichia coli*. Nat Commun 5:4115. https://doi.org/10.1038/ncomms5115.
58. Boos W, Shuman H. 1998. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. Microbiol Mol Biol Rev 62:204–229.
59. Reichenbach B, Gopel Y, Gorke B. 2009. Dual control by perfectly overlapping sigma 54- and sigma 70-promoters adjusts small RNA GlmY expression to different environmental signals. Mol Microbiol 74:1054–1070. https://doi.org/10.1111/j.1365-2958.2009.06918.x.
60. Zbell A, Benoit S, Maier R. 2007. Differential expression of NiFe uptake-type hydrogenase genes in *Salmonella enterica* serovar Typhimurium. Microbiology 153:3508–3516. https://doi.org/10.1099/mic.0.2007/009027-0.
61. Sawers RG, Jamieson DJ, Higgins CF, Boxer DH. 1986. Characterization and physiological roles of membrane-bound hydrogenase isoenzymes from *Salmonella* Typhimurium. J Bacteriol 168:398–404. https://doi.org/10.1128/jb.168.1.398-404.1986.
62. Maupin JA, Shanmugam KT. 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. https://doi.org/10.1128/jb.172.9.4798-4806.1990.
63. Schaefer J, Engl C, Zhang N, Lawton E, Buck M. 2015. Genome wide interactions of wild-type and activator bypass forms of sigma54. Nucleic Acids Res 43:7280–7291. https://doi.org/10.1093/nar/gkv597.
64. Ferro-Luzzi Ames G, Nikaido K. 1985. Nitrogen regulation in *Salmonella* Typhimurium. Identification of an NtrC protein-binding site and definition of a consensus binding sequence. EMBO J 4:539–547.
65. Bernardo LM, Johansson LU, Skarfstad E, Shingler V. 2009. Sigma54-promoter discrimination and regulation by ppGpp and DksA. J Biol Chem 284:828–838. https://doi.org/10.1074/jbc.M807707200.
66. 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. https://doi.org/10.1128/jb.166.2.545-551.1986.
67. Buck M, Khan H, Dixon R. 1985. Site-directed mutagenesis of the *Klebsiella pneumoniae nifL* and *nifH* promoters and *in vivo* analysis of promoter activity. Nucleic Acids Res 13:7621–7638. https://doi.org/10.1093/nar/13.21.7621.
68. Beck L, Smith T, Hoover T. 2007. Look, no hands! Unconventional transcriptional activators in bacteria. Trends Microbiol 15:530–537. https://doi.org/10.1016/j.tim.2007.09.008.
69. Cannon W, Claverie-Martin F, Austin S, Buck M. 1993. Core RNA polymerase assists binding of the transcription factor σ^{54} to promoter DNA. Mol Microbiol 8:287–298. https://doi.org/10.1111/j.1365-2958.1993.tb01573.x.
70. Morris L, Cannon W, Claverie-Martin F, Austin S, 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.
71. Wigneshweraraj S, Bose D, Burrows PC, Joly N, Schumacher J, Rappas M, Pape T, Zhang X, Stockley P, Severinov K, Buck M. 2008. Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. Mol Microbiol 68:538–546. https://doi.org/10.1111/j.1365-2958.2008.06181.x.
72. Cannon W, Wigneshweraraj SR, Buck M. 2002. Interactions of regulated and deregulated forms of the sigma54 holoenzyme with heteroduplex promoter DNA. Nucleic Acids Res 30:886–893. https://doi.org/10.1093/nar/30.4.886.
73. Bose D, Pape T, Burrows PC, Rappas M, Wigneshweraraj SR, Buck M, Zhang X. 2008. Organization of an activator-bound RNA polymerase holoenzyme. Mol Cell 32:337–346. https://doi.org/10.1016/j.molcel.2008.09.015.
74. Vogel SK, Schulz A, Rippe K. 2002. Binding affinity of *Escherichia coli* RNA polymerase-sigma54 holoenzyme for the *glnAp2*, *nifH* and *nifL* promoters. Nucleic Acids Res 30:4094–4101. https://doi.org/10.1093/nar/gkf519.
75. Barrios H, Valderrama B, Morett E. 1999. Compilation and analysis of sigma(54)-dependent promoter sequences. Nucleic Acids Res 27:4305–4313. https://doi.org/10.1093/nar/27.22.4305.
76. Mahmutovic A, Berg OG, Elf J. 2015. What matters for Lac repressor search *in vivo*—sliding, hopping, intersegment transfer, crowding on DNA or recognition? Nucleic Acids Res 43:3454–3464. https://doi.org/10.1093/nar/gkv207.
77. Babu MM, Luscombe NM, Aravind L, Gerstein M, Teichmann SA. 2004. Structure and evolution of transcriptional regulatory networks. Curr Opin Struct Biol 14:283–291. https://doi.org/10.1016/j.sbi.2004.05.004.
78. Babu MM, Aravind L. 2006. Adaptive evolution by optimizing expression levels in different environments. Trends Microbiol 14:11–14. https://doi.org/10.1016/j.tim.2005.11.005.
79. Maloy S, Stewart V, Taylor R. 1996. Genetic analysis of pathogenic bacteria: a laboratory manual. Cold Spring Harbor Laboratory Press, Plainview, NY.
80. Datsenko K, Wanner B. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A 97:6640–6645. https://doi.org/10.1073/pnas.120163297.
81. Perkins-Balding D, Duval-Valentin G, Glasgow AC. 1999. Excision of IS*492* requires flanking target sequences and results in circle formation in *Pseudoalteromonas atlantica*. J Bacteriol 181:4937–4948.
82. Popham D, Keener J, Kustu S. 1991. Purification of the alternative sigma factor, sigma 54, from *Salmonella* Typhimurium and characterization of sigma 54-holoenzyme. J Biol Chem 266:19510–19518.

83. Grana D, Youderian P, Susskind MM. 1985. Mutations that improve the *ant* promoter of *Salmonella* phage P22. Genetics 110:1–16.
84. Xu H, Gu B, Nixon BT, Hoover TR. 2004. Purification and characterization of the AAA+ domain of *Sinorhizobium meliloti* DctD, a sigma54-dependent transcriptional activator. J Bacteriol 186:3499–3507. https://doi.org/10.1128/JB.186.11.3499-3507.2004.
85. Lieu PT, Jozsi P, Gilles P, Peterson T. 2005. Development of a DNA-labeling system for array-based comparative genomic hybridization. J Biomol Tech 16:104–111.
86. Stead MB, Agrawal A, Bowden KE, Nasir R, Mohanty BK, Meagher RB, Kushner SR. 2012. RNAsnap: a rapid, quantitative and inexpensive, method for isolating total RNA from bacteria. Nucleic Acids Res 40:e156. https://doi.org/10.1093/nar/gks680.
87. Wang Q, Frye JG, McClelland M, Harshey RM. 2004. Gene expression patterns during swarming in *Salmonella* Typhimurium: genes specific to surface growth and putative new motility and pathogenicity genes. Mol Microbiol 52:169–187. https://doi.org/10.1111/j.1365-2958.2003.03977.x.
88. Morales EH, Collao B, Desai PT, Calderon IL, Gil F, Luraschi R, Porwollik S, McClelland M, Saavedra CP. 2013. Probing the ArcA regulon under aerobic/ROS conditions in *Salmonella enterica* serovar Typhimurium. BMC Genomics 14:626. https://doi.org/10.1186/1471-2164-14-626.
89. Xia XQ, McClelland M, Porwollik S, Song W, Cong X, Wang Y. 2009. WebArrayDB: cross-platform microarray data analysis and public data repository. Bioinformatics 25:2425–2429. https://doi.org/10.1093/bioinformatics/btp430.
90. Mrazek J, Xie S, Guo X, Srivastava A. 2008. AIMIE: a web-based environment for detection and interpretation of significant sequence motifs in prokaryotic genomes. Bioinformatics 24:1041–1048. https://doi.org/10.1093/bioinformatics/btn077.
91. Mrazek J. 2006. Analysis of distribution indicates diverse functions of simple sequence repeats in *Mycoplasma* genomes. Mol Biol Evol 23:1370–1385. https://doi.org/10.1093/molbev/msk023.
92. Mrazek J, Guo X, Shah A. 2007. Simple sequence repeats in prokaryotic genomes. Proc Natl Acad Sci U S A 104:8472–8477. https://doi.org/10.1073/pnas.0702412104.
93. Cannon W, Chaney M, Wang X-Y, Buck M. 1997. Two domains within σ^{N} (σ^{54}) cooperate for DNA binding. Proc Natl Acad Sci U S A 94:5006–5011. https://doi.org/10.1073/pnas.94.10.5006.
94. Kelly M, Hoover T. 1999. Mutant forms of *Salmonella typhimurium* σ^{54} defective in transcription initiation but not promoter binding activity. J Bacteriol 181:3351–3357.
95. Gallegos M-T, Buck M. 2000. Sequences in σ^{54} region I required for binding to early melted DNA and their involvement in sigma-DNA isomerisation. J Mol Biol 297:849–859. https://doi.org/10.1006/jmbi.2000.3608.
96. van Houten V, Denkers F, van Dijk M, van den Brekel M, Brakenhoff T. 1998. Labeling efficiency of oligonucleotides by T4 polynucleotide kinase depends on 5′-nucleotide. Anal Biochem 265:386–389. https://doi.org/10.1006/abio.1998.2900.
97. Jishage M, Iwata A, Ueda S, 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. https://doi.org/10.1128/jb.178.18.5447-5451.1996.
98. Kubitschek HE, Friske JA. 1986. Determination of bacterial cell volume with the Coulter counter. J Bacteriol 168:1466–1467. https://doi.org/10.1128/jb.168.3.1466-1467.1986.
99. Tobiason DM, Lenich AG, Glasgow AC. 1999. Multiple DNA binding activities of the novel site-specific recombinase, Piv, from *Moraxella lacunata*. J Biol Chem 274:9698–9706. https://doi.org/10.1074/jbc.274.14.9698.
100. Schneider TD, Stephens RM. 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res 18:6097–6100. https://doi.org/10.1093/nar/18.20.6097.
101. Crooks GE, Hon G, Chandonia JM, Brenner SE. 2004. WebLogo: a sequence logo generator. Genome Res 14:1188–1190. https://doi.org/10.1101/gr.849004.
102. Klose KE, Mekalanos JJ. 1997. Simultaneous prevention of glutamine synthesis and high-affinity transport attenuates *Salmonella* Typhimurium virulence. Infect Immun 65:587–596.