Sequence Determinants Spanning –10 Motif and Spacer Region Implicated in Unique *Ehrlichia chaffeensis* Sigma 32-Dependent Promoter Activity of dnaK Gene

Huitao Liu and Roman R. Ganta*

Center of Excellence for Vector-Borne Diseases, Department of Diagnostic Medicine/Pathobiology, College of Veterinary Medicine, Kansas State University, Manhattan, KS, United States

*Correspondence: Roman R. Ganta rganta@vet.k-state.edu

**Specialty section:** This article was submitted to Microbial Physiology and Metabolism, a section of the journal Frontiers in Microbiology

**Received:** 10 May 2019  
**Accepted:** 18 July 2019  
**Published:** 02 August 2019

**Citation:** Liu H and Ganta RR (2019) Sequence Determinants Spanning –10 Motif and Spacer Region Implicated in Unique Ehrlichia chaffeensis Sigma 32-Dependent Promoter Activity of dnaK Gene. Front. Microbiol. 10:1772. doi: 10.3389/fmicb.2019.01772

**Keywords:** gene regulation, intracellular bacteria, Anaplasmataceae, sigma factor, Ehrlichia chaffeensis, RNA polymerase

**INTRODUCTION**

*Ehrlichia chaffeensis*, a Gram-negative, obligate intracellular tick-borne rickettsial pathogen, causes human monocytic ehrlichiosis (HME) (Dunning Hotopp et al., 2006). HME as an emerging infectious disease, first reported in the United States in 1987, has become one of the most prevalent tick borne diseases in the United States and is also described from several other parts of the world...
Further, *E. chaffeensis* infects several other vertebrates, including dogs, goats, coyotes and white-tailed deer (Dawson et al., 1996; Lockhart et al., 1997; Breitschwerdt et al., 1998; Dugan et al., 2000; Kocan et al., 2000; Davidson et al., 2001). The pathogen infection in people may result in an acute flu-like illness with symptoms ranging from persistent fever, headache, myalgia, anorexia and chills (Walker et al., 2008). HME patients may exhibit leukopenia, thrombocytopenia, anemia, and upgraded levels of serum hepatic white-tailed deer (Dawson et al., 1996; Lockhart et al., 1997; several other vertebrates, including dogs, goats, coyotes and chills (Walker et al., 2008). Though some progress is made in establishing genetics in *E. chaffeensis* (Cheng et al., 2013; Wang et al., 2017), and similarly in other related *Ehrlichia* and *Anaplasma* (Long et al., 2005; Felsheim et al., 2006; Crosby et al., 2014; Wood et al., 2014; McClure et al., 2017), the genetic tool kit and its application is still limited. For example, it is not possible to investigate regulation of gene expression by transforming this group of important pathogens, possibly also because the pathogens lack naturally existing extrachromosomal plasmids. This major impediment limits the understanding of molecular mechanisms used by the pathogens in regulating gene expression in support of their continued survival in vertebrate and tick hosts and in causing pathogenesis (Dumler et al., 1993; Davidson et al., 2001; Unver et al., 2002).

Several prior studies reported differences in gene expression of *E. chaffeensis* impacted by different host environments (Seo et al., 2008; Kuriakose et al., 2011). However, it is unclear how the organism regulates its gene expression in support of its adaptation to the hosts.

Regulation of gene expression in bacteria is primarily controlled at the transcription. An RNA polymerase (RNAP) core enzyme with a sigma (σ) factor offers a simple and valid mechanism for bacteria to rapidly accommodate to diverse environmental changes by suitably modifying the transcriptional profiles (Gruber and Gross, 2003; Gunesekere et al., 2006; Browning and Busby, 2016). Typically, an RNAP holoenzyme is a multi-subunit complex consists of a core enzyme containing two alpha (α), a beta (β), a beta’ (β’), and a omega (ω) subunits and then the inclusion of a σ factor (Chamberlin et al., 1983). A σ factor enables a core enzyme in specific binding to the promoter region of a gene for initiating transcription. Numbers of σ factors differ depending on the genome size variations and the environmental diversification of a bacterium (Kill et al., 2005). For example, *Escherichia coli* possesses seven σ factors, while 109 σ factors are identified in *Sorangium cellulosum* (Han et al., 2013; Tripathi et al., 2014). Obligate intracellular bacteria generally tend to have reduced genomes and consequently also have fewer σ factors (Darby et al., 2007). For example, *E. chaffeensis* genome of 1,176 kb has only two σ factor genes; rpoD (ECH_0760) (the primary housekeeping σ70 gene) and rpoH (ECH_0655) (the alternate σ32 gene) (Dunning Hotopp et al., 2006) (GenBank # NC_007799.1).

To study gene regulation in *E. chaffeensis*, we previously described *in vitro* transcription system and an *E. coli* surrogate system that is valuable in investigating gene regulation driven from its σ70 (Faburay et al., 2011; Liu et al., 2013, 2016). Earlier, we also defined the promoters of several pathogen genes by utilizing *in vitro* transcription assays where *E. coli* RNAP core enzyme is reconstituted with the recombinant *E. chaffeensis* σ factors (Faburay et al., 2011; Liu et al., 2013). Our studies demonstrated that the RNAP binding motifs (−10 and −35 regions) of *E. chaffeensis* gene promoters share extensive homology and that they are recognizable by RNAP with either one of its only two sigma factors; σ32 or σ70, although affinities vary for different gene promoters (Liu et al., 2013). In *E. coli*, gene expression of heat shock proteins; DnaK-DnaJ-GrpE and GroES-GroEL chaperone complexes, are controlled by its σ32 (Nonaka et al., 2006). Similarly, we discovered that the *E. chaffeensis* chaperon protein gene (Ech_0471) encoding for DnaK protein is transcribed primarily by σ32 (Liu et al., 2013). Genes regulated by σ32 are known to induce cellular responses under varieties of stresses confronted during bacterial growth and are likely important for *E. chaffeensis* survival in its hostile host environments and that they may contribute to pathogenicity similar to other Gram negative bacteria (Du et al., 2005; Delory et al., 2006; Slamti et al., 2007; Matsui et al., 2008; Spector and Kenyon, 2012). For example, DnaK in *Vibrio cholerae*, the gene expression regulated by its σ32, is involved in causing virulence in a host (Sahu et al., 1994; Chakrabarti et al., 1999; Slamti et al., 2007).

Prior research on *E. coli* (a γ-proteobacteria) provides abundant knowledge regarding gene regulation from its σ32- and σ70-bound RNAP, while such knowledge for both σ32- and σ70 bound RNAP in other Gram-negative bacteria, particularly for α-proteobacteria, including for pathogenic organisms is very limited. Importantly, it is unclear how intracellular pathogens, such as *E. chaffeensis*, regulate gene expression to overcome the host stress and adapt to host environmental changes within its arthropod (tick) and vertebrate hosts. To extend knowledge on how *E. chaffeensis* regulates its gene expression, we continue investigations in defining the functions of its RNAP holoenzyme comprising σ32 or σ70. Transcription derived by an RNAP typically implicates in recognizing and binding to DNA sequence motifs of a promoter; −10 and −35 regions, and the spacer sequences located between the two motifs of a gene promoter (Gross et al., 1998; Paget and Helmann, 2003).

In the current study, we described the mapping of *dnaK* gene promoter recognized primarily by the *E. chaffeensis* RNAP containing σ32 using the previously developed *E. coli* surrogate system in the strain, CAG57101 (Koo et al., 2009a; Liu et al., 2013). In *E. coli* CAG57101, its endogenous rpoH gene (encoding for σ32) is inactivated (Koo et al., 2009a) and in its place, we expressed the *E. chaffeensis* σ32 from a plasmid in defining *dnaK* promoter mapping, as direct gene mapping studies are not possible in this and other related intracellular rickettsials. We also investigated the functional domains of *E. chaffeensis* σ32 likely important for the RNAP function and
in its interactions with the −10 motif and the spacer sequence of dnaK.

**MATERIALS AND METHODS**

**E. coli Strains and Plasmids**

*Escherichia coli* strains used in this study were TOP10 (Invitrogen Technologies, Carlsbad, CA, United States), BL21(DE3) (Novagen, San Diego, CA, United States), and CAG57101 (Koo et al., 2009a). Several plasmid constructs used in this study were obtained from commercial sources or modified from one or more of the existing plasmids. They include the derivatives of pSAKT-Eco_rpoH (previously known as pSAKT32) (Wang and deHaseth, 2003; Koo et al., 2009a), pQF50K-Ech_dnaK (Liu et al., 2013) and pMT504 (Tan and Engel, 1996). Genetic makeup of plasmids described in this study were included in Supplementary Table S1, except those obtained from a commercial source. The plasmid pSAKT-Eco_rpoH containing a p15A origin of replication and an ampicillin resistance gene has *E. coli* rpoH gene under the control of IPTG inducible P_{lac} promoter (Wang and deHaseth, 2003; Koo et al., 2009a). The *E. coli* rpoH from this plasmid was removed by using Q5 Site-Directed Mutagenesis Kit (New England Biolab, Inc., Ipswich, MA, United States) and the modified plasmid was referred as the pSAKT. The pSAKT-Ech_rpoH (previously known as pSAKT32-Ech_rpoH) contained *E. chaffeensis* rpoH (Liu et al., 2013). *E. chaffeensis* rpoH variants with substitutions within the 2.3, 2.4, and 3.0 regions of σ^{32} were constructed by mutagenesis using Q5 Site-Directed Mutagenesis Kit (New England Biolab, Inc., Ipswich, MA, United States). The names of the modified pSAKT-Ech_rpoH are provided in Supplementary Table S1.

The pQF50K-Ech_dnaK plasmid and the pQF50K-Ech_dnaK with deletion of −35 motif, which contains the β-galactosidase coding sequence (lacZ) with a P_{lac} origin of replication and with a kanamycin resistance gene, was reported earlier (Liu et al., 2013). The modified plasmid with deletion of −10 motif was generated from pQF50K-Ech_dnaK plasmid using Q5 Site-Directed Mutagenesis Kit (New England Biolab, Inc., Ipswich, MA, United States). Site directed mutagenesis constructs with mutations at every nucleotide of the −10 motif were also generated from the pQF50K-Ech_dnaK plasmid using Q5 Site-Directed Mutagenesis Kit. The mutants to modify the AT rich spacer sequence of the dnaK promoter were generated similarly by modifying the pQF50K-Ech_dnaK plasmid. The expression plasmids of *E. chaffeensis* σ^{32} was constructed and used for preparing purified recombinant proteins of σ^{32} as in early reports (Liu et al., 2013).

For *in vitro* transcription analysis, pMT504-Ech_dnaK as transcription template was prepared and reported earlier (Liu et al., 2013). Constructs with various mutations at −10 motif for the dnaK promoter for *in vitro* transcription assays were similarly prepared from this plasmid using Q5 Site-Directed Mutagenesis Kit. The lengths of transcripts for the various promoter segments of dnaK gene are 162 nucleotides. Integrity of all cloned segments in the plasmid constructs was confirmed by DNA sequence analysis using CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, CA, United States). The names of all engineered plasmids were listed in Supplementary Table S1. Mutagenic oligonucleotides were described in the Supplementary Table S2.

**E. coli Growth Conditions and β-Galactosidase Assays**

The *E. coli* strain CAG57101 transformed with the recombinant plasmids were grown as in early reports (Koo et al., 2009a; Liu et al., 2013). Briefly, cultures were grown at 30°C in Luria–Bertani (LB) medium with chloramphenicol (30 µg/ml) and spectinomycin (50 µg/ml) in support of the strain’s growth, and by ampicillin (100 µg/ml) for maintaining the pSAKT-derived plasmids. To assess the functions and impact of various mutations within the promoter regions of genes encoding dnaK, pQF50K-derived plasmid containing the promoter segments were also maintained by growing *E. coli* cultures with the addition of kanamycin (50 µg/ml). *E. coli* cultures of CAG57101 in LB medium were grown overnight with appropriate antibiotic supplements which were diluted 1:10 into a fresh medium containing appropriate antibiotics and the growth was continued for 2 h. Subsequently, cultures were then induced with 1 mM IPTG for 3 h before harvesting, when OD at 600 nm reached between 0.6 and 0.8. Lysates were prepared and used to measure β-galactosidase enzyme activity using a β-Gal Assay Kit (Invitrogen Technologies, Carlsbad, CA, United States). All experiments were performed three independent times with independently grown cultures; specific activity of β-galactosidase was calculated as outlined in the kit protocol.

**In vitro Transcription Assays**

*In vitro* transcription reactions were performed in 10 µl reaction mixture containing 0.1 picomoles each of the supercoiled plasmid DNA as the template and using RNAP holoenzyme containing recombinant *E. chaffeensis* σ^{32} (Liu et al., 2013). The holoenzyme was prepared by mixing 0.5 µl of 1:10 diluted stock of *E. coli* core enzyme (Epigen, Madison, WI, United States) mixed with 10-fold molar excess of purified recombinant *E. chaffeensis* σ^{32} and kept in ice for 30 min prior to using for the reactions. The transcription reactions were performed at 37°C for 20 min, and the reactions were terminated by adding 7 µl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol). Six microliters each of the samples were resolved on a 6% polyacrylamide sequencing gel with 7 M urea, then gels were transferred to a Whatman paper, dried and 162 nucleotide transcripts were visualized by exposing an X-ray film to the gels. The transcripts were quantified using ImageJ software^1^.

**Bioinformatics**

Multiple DNA alignments were done using the programs Clustal X version 2.0 with default parameters (Larkin et al., 2007).

---

1http://rsb.info.nih.gov/ij
Statistical Analysis
Statistical analyses were performed using Student’s t-test, and a P-value < 0.05 was considered significant with a single asterisk.

RESULTS
The −10 Motif Is Needed for the E. chaffeensis dnaK Gene Transcription by RNAP Holoenzyme Constituting of Its σ32
Our prior studies demonstrated that −35 motif, but not −10 motif, is required for the σ32-bound RNAP holoenzyme transcription from σ70-dependent promoters in E. chaffeensis (Liu et al., 2016). We also reported earlier that −35 motif is similarly required for the dnaK gene transcription by E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013).

To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013). To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013). To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013). To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013). To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013). To test whether or not the −10 motif of dnaK promoter is required for the σ32-dependent gene regulation, plasmid constructs lacking −10 or −35 motifs of the promoter were cloned upstream to the β-galactosidase coding sequence in an E. coli mutant deficient for its σ32 expression that is functionally complemented with the E. chaffeensis σ32-bound RNAP holoenzyme (Liu et al., 2013).

Identifying the Critical Sequence Determinants of −10 Motif of E. chaffeensis dnaK
The consensus sequence of −35 motif for E. chaffeensis promoters recognized by its σ32 (TTGTAT) is similar to the consensus −35 motif of σ32-dependent promoters in E. coli (TTGAAA) and similarly it shares extensive homology to −35 motif for the genes recognized by σ70 (TTGNTT) (Nonaka et al., 2006; Liu et al., 2013). The consensus −10 motif of E. chaffeensis promoters recognized by its σ32 (TATATN) is also similar to its consensus −10 motif recognized by its σ70 (TATTNT), however, it differs significantly from the consensus −10 motif of E. coli σ32-dependent promoters (CCCCATWT) (Nonaka et al., 2006; Koo et al., 2009a; Liu et al., 2013). While deletion of −10 motif from σ70-dependent promoters has no impact on promoter activities in E. chaffeensis (Liu et al., 2016), such deletion from the σ32-dependent dnaK resulted in significant and 90% reduction in the promoter activity (Figure 1). These novel data suggest that, contrary to σ70-bound RNAP (Liu et al., 2016), the −10 motif plays a critical role for σ32-bound RNAP in E. chaffeensis. We therefore performed detailed point mutation experiments to define the critical sequence determinants of the −10 motif for the dnaK promoter activity (Figure 2). We have made substitution mutations at each base of the six-nucleotide motif (TATATN) and evaluated the impact of each mutation by measuring changes in β-galactosidase expression in CAG57101 E. coli functionally complemented with E. chaffeensis σ32. A specific substitution mutation was indicated by combination of letters and numbers. For example, T1A indicates a change from T to A transversion at the first position in the −10 motif. One or more substitutions at all six nucleotide positions resulted in significant decline in the promoter activity of dnaK. Substitutions at the first five nucleotides to any other nucleotide resulted in significant promoter activity drop. Substitutions in the first position from T1G to T1C resulted in the near complete loss of promoter activity, while T1A resulted in the reduction of promoter activity to 39%. Similarly, in the second position, A2C and A2T mutations caused significant loss of promoter activity (reduced to 17% and 7% compared to the wildtype), whereas A2G mutation caused significant decline of two thirds of activity similar to T1A substitution. In the third and fourth positions, changes to any other nucleotide had the greatest impact in promoter activity decline (78–98% reduction). Mutations in the fifth position to T5G or T5C had a greater impact (near 90% decline), while substitution was less apparent for the T5A that is similar to the mutations in the first position T. In the sixth position, only G6G mutation resulted in significant decline in the promoter activity to 35%.

In vitro Transcription Assays to Verify Sequence Determinants of −10 Motif Mapped in E. coli CAG57101
To validate the results in defining the −10 motif in the E. coli surrogate system, we performed in vitro transcription assays using several randomly selected promoter mutation constructs; the assays were performed using RNAP holoenzyme reconstituted with the recombinant E. chaffeensis σ32. We selected five dnaK −10 motif mutants for this experiment and compared the results with the wildtype promoter. Both wildtype and mutated versions of dnaK promoter segments were cloned into the G-less cassette and used as templates in the in vitro
transcription assays (Figures 3A,B). Consistent with the results recorded with the E. coli CAG57101 system, mutants T1A, T1G and A2G produced lesser transcripts compared to the wildtype dnaK. There was no significant difference for the T5A mutation compared to the wildtype, which is also consistent with the results observed in the E. coli system. In the sixth position, C6T mutation caused an increase of in vitro transcript level, which is also similar to the enhanced promoter activity observed in the E. coli CAG57101 experiments.

The Spacer Sequences Affect Promoter Activity

Previous studies in E. coli demonstrate that spacer sequences located between −10 and −35 motifs contribute to promoter activities (Aoyama et al., 1983; Mulligan et al., 1985; Hook-Barnard and Hinton, 2007, 2009; Singh et al., 2011). In particular, nucleotides present in the spacer sequence and its length play critical role for a promoter activity. Further, a short C-rich region upstream to −10 motif in E. coli and in other γ-proteobacteria is identified as important extended −10 motif required for efficient transcription by σ32-containing RNA polymerase containing E. chaffeensis recombinant σ32. The mutants of promoter segments cloned upstream to the G-less cassette in the pMT504 plasmid were used in the assays. The abundance of transcripts was captured as 32P incorporation. Intensity of a band signal in a gel for in vitro transcripts made for the wild-type and mutant dnaK were assessed using the software ImageJ (https://imagej.nih.gov/ij/). Panel (A) has the image data and panel (B) included the quantitative data collected from the image signals. The bars show the relative transcription products of mutant promoters as the percentage of transcripts compared to the wild-type promoter for E. chaffeensis recombinant σ32. (WT and various mutant promoter constructs are identified as in Figure 2).

while keeping the spacer length constant (GC), or by increasing the spacer sequence lengths from 17 bp to 18, 19, or 20 bps or by decreasing it to 16, 15, or 14 bp and finally by deleting the spacer sequence completely. These different spacer mutant constructs
were depicted in Figure 4A. The β-galactosidase expression was then assessed for all these modified spacer promoter segments and compared to wildtype (WT) promoter construct in the E. coli surrogate system (CAG57101) (Figure 4B). The CP mutant caused a minor, non-significant increase in the promoter activity (30% increase). The spacer substitution with GC resulted in a significant 50% decline in the β-galactosidase expression. Increasing the spacer length to 18 bp or decreasing to 14 or 15 bp caused a major decline in the promoter activity, although the greatest decline was observed with the 14 bp spacer (96% drop), while reducing the length to 16 bp had no impact. Increasing the spacer lengths to 19 or 20 bp resulted in much higher enhancement of the promoter activity (586% and 249%, respectively). Deletion of the entire spacer sequence had no impact on the promoter activity compared to the WT promoter. We detected the presence of another 15 bp spacer like sequence and an alternative −35 motif sequence in the complete deletion spacer construct; thus, it is highly likely that these sequences served as alternate spacer and −35 motif for the RNAP (Supplementary Figure S1).

Substitution Mutations in E. chaffeensis σ32

Prior studies in E. coli revealed that the −10 motif of σ32-dependent promoters is recognized by few amino acids within the regions 2.3, 2.4, and 3.0 of σ32 protein (Kourennaia et al., 2005; Koo et al., 2009a). To determine what amino acids spanning these regions of E. chaffeensis σ32 would contribute to the promoter activity, we have made substitution mutations at 6 different amino acid positions likely alter the functional domains of the pathogen σ32. The amino acids for substitution mutations were selected based on their homology with the E. coli σ32 (Supplementary Figure S2). Tryptophan (W) 108 of E. coli σ32

![FIGURE 4 | AT-rich spacer sequence located between −10 and −35 motif contributes to altering the promoter activity of Ehrlichia chaffeensis dnaK gene. Promoter fragments used in the assays are as in Supplementary Figure S1 for wild-type dnaK. (A) Sequence spanning from +1 to −35 motif and the AT-rich spacer sequence is presented for the wild-type construct (WT) and for the constructs with modified spacer sequences which included replacing the AT-rich spacer with complementary sequence (CP), with GC rich spacer sequence (GC), changing the size of spacer from 17 bp (WT) to 20 bp (20bp), 19 bp (19bp), 18 bp (18bp), 16 bp (16bp), 15 bp (15bp), and 14 bp (14bp) and deleting entire 17 bp spacer (0bp). Lower case with underline indicated inserted nucleotides in spacer sequence and the break line refers to the deletion nucleotides from WT dnaK promoter. (B) The β-galactosidase activity was driven by promoters of WT, CP, GC, 20bp, 19bp, 18bp, 16bp, 15bp, 14bp, and 0bp in E. coli (CAG57101) with expressing E. chaffeensis σ32. All values are averages of at least three independent experiments; error bars indicated one standard deviation. *p-value < 0.05.](#)
within the region 2.4 is identified as important for recognition of −13C in a promoter recognized by it (Kourennaia et al., 2005; Koo et al., 2009a). In E. chaffeensis σ32, W is located at 106. Glutamic acid (E) at position 112 in E. coli σ32 is also implicated for its contribution to −13C recognition and interaction with its σ32-dependent promoters (Koo et al., 2009a). Phenylalanine (F) at position 110 of E. chaffeensis σ32 is in the homologous position to E112 in E. coli σ32. As E. chaffeensis dnaK lacks C-rich region upstream to its −10 motif, we expected that amino acid substitution mutations at these two positions in E. chaffeensis σ32 would not have any impact. Lysine (K) at position 130 in region 3.0 of E. coli σ32 is similarly implicated for the recognition of the C-rich region for the σ32-promoters; groE and grpE (Koo et al., 2009a). Glutamin (Q) at position 128 in region 3.0 of E. chaffeensis σ32 is the amino acid at the homologous position to K130 of E. coli σ32. Substitution of W106 to A caused significant reduction of the β-galactosidase expression (80% decline), while F110 substitution to A or E (alanine or glutamic acid) in E. chaffeensis σ32 did not significantly alter the promoter function (Figure 5). Similarly, substitution of Q128 to A did not significantly impact E. chaffeensis dnaK promoter activity. The mutational data with the exception of W106 are consistent with the lack of C-rich region in E. chaffeensis. However, W106 may be critical for the promoter activity independent of the C-rich sequence, at least in E. chaffeensis. Phenylalanine (F) at position 104 within the 2.3 region of E. coli σ32 is identified as critical for its structural integrity and activity of σ32 (Kourennaia et al., 2005). In E. chaffeensis σ32, a polar amino acid (at position 102) [tyrosine (Y)] is present at the position homologous to F104. Similarly, A111 of E. coli σ32 is implicated for its binding to core RNAP (Kourennaia et al., 2005). Y102 to A caused a significant decrease of E. chaffeensis σ32 activity (88% reduction). Likewise, substitution of A109 to glutamine (Q) in E. chaffeensis σ32 that is homologous to A111of E. coli σ32 resulted in a significant decline of activity (about 31% of WT level). Previous studies in E. coli demonstrate that substitution of F136A within the region 3.0 reduces the interaction between core RNAP and σ32 thus leading to 80% decline in promoter activity for its groE gene (Joo et al., 1997; Kourennaia et al., 2005; Koo et al., 2009a). A similar substitution in E. chaffeensis σ32; F134A also caused similar decline of its activity for the dnaK promoter (reduced to 28% activity compared to WT E. chaffeensis σ32). Together, E. chaffeensis σ32 substitution mutation experiments allowed the identification of critical functional domains engaged in σ32-bound RNAP interactions with the dnaK promoter, including in confirming that the C-rich region is not critical for its function.

**Differences in the −10 Motifs Between E. coli and E. chaffeensis Are Sufficient in Having Differential σ32 Functions**

Unlike E. coli and other γ- proteobacterial gene promoters, −10 motif of E. chaffeensis σ32-dependent dnaK promoter (TATATN) is distinct in lacking a C-rich sequence upstream to −10 motif (Nonaka et al., 2006; Koo et al., 2009a; Liu et al., 2013). Results presented in the previous section suggest that the C-rich region is indeed not required for E. chaffeensis σ32-dependent dnaK promoter function. To further map how the variations in −10 motif and spacer sequence in E. chaffeensis dnaK make it unique for this intracellular pathogen promoter function, we prepared two modified constructs where two or four nucleotides spanning between the spacer sequence and −10 motif (TT or TATT, respectively) were replaced with either two Cs or four Cs (Figure 6A). Importantly, these sequence modifications change the E. chaffeensis −10 motif to be more similar to the E. coli consensus −10 motif (CCCCCATWT) (Figure 6A). The modified constructs having 2Cs and 4Cs were then assessed in the E. coli surrogate system expressing either E. chaffeensis σ32 (Ech-σ32) or E. coli σ32 (Eco-σ32). Compared to the wildtype dnaK promoter, both the 2C and 4C mutants caused drastic decline in the β-galactosidase expression by as much as 15-fold for E. chaffeensis σ32 (Figure 6B). On the contrary, while wildtype dnaK promoter had a minimal β-galactosidase expression with the E. coli σ32, the 2Cs and 4Cs substitutions caused a significant and a step-wise increase of β-galactosidase expression to 2.8-fold and 14.5-fold, respectively. Together, these results suggest that the inclusion of C-rich sequences is sufficient in altering the promoter specificities of E. chaffeensis to be similar to E. coli σ32-dependent RNAP.

**DISCUSSION**

Because obligate intracellular bacteria reside within the cytosol or a phagosome of an infected host cell, they encounter minimal environmental changes, possibly compared to free-living bacteria, such as E. coli. Obligate intracellular bacteria typically have reduced genomes and consequently their genomes also have limited numbers of sigma factors. For example, E. chaffeensis genome has only two sigma factor genes encoding for a constitutive and an alternative sigma factor, σ20 and σ32, respectively (Dunning Hotopp et al., 2006). On the contrary, a...
non-pathogenic bacterium, such as \textit{E. coli}, contains seven sigma factors, including $\sigma^{70}$ and $\sigma^{32}$ (Gruber and Gross, 2003; Tripathi et al., 2014). Alternate sigma factor; $\sigma^{32}$ is known to regulate genes involved in overcoming environmental stresses (Zhao et al., 2005; Martinez-Salazar et al., 2009). Prior studies in \textit{E. coli}, \textit{V. cholerae}, \textit{Neisseria gonorhoeae}, and \textit{Francisella tularensis} suggest that $\sigma^{32}$ regulon involves several heat-shock proteins, also known as chaperon proteins; Hsp40, HspG, Dnak, Dnaj, GroES, and GroEL. These proteins are involved in temperature-induced stress control of the organisms (Zhao et al., 2005; Gunasekere et al., 2006; Nonaka et al., 2006; Slamti et al., 2007). These proteins play critical roles in both preventing denaturation of proteins and also to promote renaturation of aggregated proteins (Gragerov et al., 1992; Nishihara et al., 1998; Mogk et al., 1999). Some of the proteins involved in regulating stress response, such as ClpB, HspG, and Dnak, are also considered important for the intracellular survival of a pathogenic bacterium, \textit{F. tularensis} (Tempel et al., 2006; Su et al., 2007; Weiss et al., 2007; Meibom et al., 2008). The sigma factor, $\sigma^{32}$, is identified as critical for promoting transcription of genes engaged in overcoming stressful environments for bacteria to promote the reduction of accumulation of misfolded and aggregated proteins (Gragerov et al., 1992; Slamti et al., 2007; Guisbert et al., 2008).

\textit{Ehrlichia chaffeensis} and related \textit{Anaplasmataceae} family pathogens have the ability adapt to vertebrate and tick hosts. However, it is unclear how tick transmitted pathogens regulate gene expression in tick and vertebrate hosts. Studying gene regulation will aid in understanding how rickettsial pathogens adapt to dual hosts and sense nutrient, starving, temperature, and other stressful environments within an infected host cell. Previous studies revealed that \textit{E. chaffeensis} has two morphological forms, dense core cell as the pathogen’s infectious form and reticulate cell, which replicates within a phagosome of an infected host cell (Zhang et al., 2007; Dedonder et al., 2012). We recently reported that a stress response protein, ClpB, transcripts are higher during the replicative stage of the pathogen, while gene expression of another heat shock protein, DnaK, and the stress response sigma factor, RpoH, remained as constitutively expressed throughout the replicating stage (Liu et al., 2013; Zhang et al., 2013). Further, we also reported that \textit{E. chaffeensis} sigma factors, $\sigma^{32}$ and $\sigma^{70}$, function cooperatively in transcribing pathogen genes, but with varying affinities (Liu et al., 2013). The \textit{dnaK} has higher affinity for RNAP containing \textit{E. chaffeensis} $\sigma^{32}$ compared to $\sigma^{70}$ (Liu et al., 2013). The current study is the first in defining \textit{E. chaffeensis} $\sigma^{32}$-dependent gene promoter region of \textit{dnaK}. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6}
\caption{Sequence variations in –10 motif and its immediate upstream sequence define specificities for \textit{E. coli-} and \textit{E. chaffeensis} $\sigma^{32}$. The –10 motif for wild-type \textit{dnaK} promoter (WT) and for the modified sequences where TT and TATT near –10 motif are modified to CC (2C) and CCCC (4C), respectively (panel A). The inclusion of 2C and 4C make the –10 motif of \textit{dnaK} promoter to be similar to consensus sequence of –10 motif for \textit{E. coli} $\sigma^{32}$ (Ech-$\sigma^{32}$) and \textit{E. coli} $\sigma^{32}$ (Eco-$\sigma^{32}$) recognized the WT, 2C, and 4C promoters differently (panel B). $\beta$-galactosidase expression as specific activity values were in parenthesis. All values are averages of at least three independent experiments; error bars indicated one standard deviation.}
\end{figure}
Despite our recent advances in establishing tractable genetics for *E. chaffeensis* (Wang et al., 2017), studying gene regulation remains a challenge due to lack of a well-established methods to maintain extracellular plasmids, as the pathogen and other related *Anaplasmataceae* pathogens lack natural plasmids. To overcome this challenge, in the current study, we utilized the *E. coli* surrogate system to map promoter-binding domains required for gene expression in *E. chaffeensis* for a $\sigma^{32}$-dependent gene promoter. We validated the results from the *E. coli* surrogate system using *in vitro* transcription assays which we developed earlier (Liu et al., 2013, 2016). *In vitro* transcription system is applied well in understanding intracellular bacterial gene regulation for pathogens, such as *Chlamydia trachomatis*, for which a tractable genetic system is equally well not developed, which can support gene regulation studies (Mathews et al., 1993; Tan and Engel, 1996; Tan et al., 1998; Yu and Tan, 2003; Shen et al., 2004; Akers and Tan, 2006; Rao et al., 2009; Bao et al., 2011, 2012). Hence, in view of the technical challenges, methods described in the current study are innovative in mapping *E. chaffeensis* gene regulation, and that the study will lead the way for similar investigations in other related *Anaplasmataceae* family pathogens.

RNA polymerase holoenzyme containing $\sigma^{70}$ of *E. coli* transcribes housekeeping genes by recognizing two highly conserved motifs of a gene promoter; referred as $-10$ and $-35$ motifs (Gross et al., 1998). We reported earlier that the consensus $-10$ and $-35$ motifs for *E. chaffeensis* $\sigma^{70}$-dependent promoters (TATTN and TTGNTT, respectively) (Liu et al., 2013) are similar to *E. coli* consensus $-10$ and $-35$ (TATAAT and TTGACA, respectively) (Singh et al., 2011; Shimada et al., 2014). We reported that the $-35$ motif and the AT-rich spacer sequences are important for a gene-specific regulation by $\sigma^{32}$-dependent promoters (Liu et al., 2016). The predicted *E. chaffeensis* consensus $-35$ motif for the alternative sigma factor ($\sigma^{32}$) is also similar to *E. coli* consensus $-35$ motif (Nonaka et al., 2006; Koo et al., 2009a,b; Liu et al., 2013). The $-10$ motif of *E. chaffeensis* $\sigma^{32}$-dependent promoters (TATATN), however, differs substantially from the consensus *E. coli* $-10$ motif (CCCCATNT) (Nonaka et al., 2006; Koo et al., 2009a,b; Liu et al., 2013). The consensus $-35$ motif is also extensively conserved among $\sigma^{32}$-dependent promoters in proteobacteria (Supplementary Table S3), while $-10$ motif of $\sigma^{32}$-dependent promoters vary considerably among various classes of proteobacteria. For example, many $\gamma$-proteobacteria contain a C-rich sequence upstream to the $-10$ motif and is implicated in regulating $\sigma^{32}$-dependent gene regulation (Nonaka et al., 2006; Slamti et al., 2007; Grall et al., 2009; Stoll et al., 2009). Such sequence, however, is absent in *E. chaffeensis* $\sigma^{32}$-dependent gene promoters (Liu et al., 2013). Thus, we hypothesized that the pathogen is evolved to possess a distinctive $-10$ motif and that it plays a critical role in $\sigma^{32}$-dependent promoter regulation. Previous studies suggest that the length of a spacer sequence between the $-10$ and $-35$ motifs and the composition of specific nucleotides present within a spacer affect promoter activity (Aoyama et al., 1983; Mulligan et al., 1985; Hook-Barnard and Hinton, 2009; Singh et al., 2011; Liu et al., 2016). In particular, spacer sequences influence the three-dimensional structure of a promoter and any modifications to a spacer sequence, therefore, affect the efficiency of interactions between the RNAP holoenzyme with a promoter sequence either positively or negatively (Rud et al., 2006; Hook-Barnard and Hinton, 2009; Singh et al., 2011). Consistent with prior studies, our current investigation similarly revealed the importance of spacer sequence in contributing to the promoter activity for the *E. chaffeensis dnaK* gene transcribed by its $\sigma^{32}$. We previously reported that modifications to spacer sequences similarly impact *E. chaffeensis* genes transcribed by $\sigma^{70}$ genes and that changes to spacer sequences alter the curvature of a promoter region (Liu et al., 2016).

In summary, we mapped *E. chaffeensis* dnaK $-10$ and $-35$ motifs and the spacer sequence upstream to it by performing detailed mutational analysis. Furthermore, this study confirmed that the C-rich region–specific interactions between a gene promoter and $\sigma^{32}$ protein domains, known to be critical for *E. coli*, are absent in *E. chaffeensis*, which makes the pathogen gene regulation distinct, as judged from characterizing the *dnaK* gene promoter.

**DATA AVAILABILITY**

The datasets generated for this study are available on request to the corresponding author.

**AUTHOR CONTRIBUTIONS**

RG conceived and directed the research design. HL contributed to the research plan and performed the experiments. HL and RG evaluated the data and prepared the manuscript.

**FUNDING**

This work was supported by the PHS grant # AI070908 from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, United States. This manuscript is a contribution from the Kansas Agricultural Experiment Station (Contribution Number 19-291-J). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

**ACKNOWLEDGMENTS**

We thank Ms. Mal Rooks Hoover for her help in preparing the figures.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01772/full#supplementary-material
cells. *Cell Microbiol.* 9, 610–618. doi: 10.1111/j.1462-5822.2006.00812.x

Zhang, T., Kedzierska-Mieszkowska, S., Liu, H., Cheng, C., Ganta, R. R., and Zolkiewski, M. (2013). Aggregate-reactivation activity of the molecular chaperone ClpB from *Ehrlichia chaffeensis*. *PLoS One* 8:e62454. doi: 10.1371/journal.pone.0062454

Zhao, K., Liu, M., and Burgess, R. R. (2005). The global transcriptional response of *Escherichia coli* to induced sigma 32 protein involves sigma 32 regulon activation followed by inactivation and degradation of sigma 32 in vivo. *J. Biol. Chem.* 280, 17758–17768.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Liu and Ganta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.