

Modulation of extracytoplasmic function (ECF) sigma factor promoter selectivity by spacer region sequence

Ahmed Gaballa[†], Veronica Guariglia-Oropeza[†], Franziska Dürr, Bronwyn G. Butcher, Albert Y. Chen, Pete Chandrangsu and John D. Helmann^{*}

Department of Microbiology, Cornell University, Ithaca, NY 14853-8101, USA

Received February 28, 2017; Revised September 26, 2017; Editorial Decision October 04, 2017; Accepted October 05, 2017

ABSTRACT

The ability of bacteria to adapt to stress depends on the conditional expression of specific sets of genes. *Bacillus subtilis* encodes seven extracytoplasmic function (ECF) sigma (σ) factors that regulate functions important for survival under conditions eliciting cell envelope stress. Of these, four have been studied in detail: σ^M , σ^W , σ^X and σ^V . These four σ factors recognize overlapping sets of promoters, although the sequences that determine this overlapping recognition are incompletely understood. A major role in promoter selectivity has been ascribed to the core -10 and -35 promoter elements. Here, we demonstrate that a homopolymeric T-tract motif, proximal to the -35 element, functions in combination with the core promoter sequences to determine selectivity for ECF sigma factors. This motif is most critical for promoter activation by σ^V , and contributes variably to activation by σ^M , σ^X and σ^W . We propose that this motif, which is a feature of the deduced promoter consensus for a subset of ECF σ factors from many species, imparts intrinsic DNA curvature to influence promoter activity. The differential effect of this region among ECF σ factors thereby provides a mechanism to modulate the nature and extent of regulon overlap.

INTRODUCTION

Transcription is the first step in gene expression and can be regulated at multiple levels (1,2). The global transcriptional profile of the cell results from the activity of RNA

polymerase (RNAP) as determined by many transcription factors that affect initiation, elongation, and termination of RNA synthesis. In bacteria, a key initial step is the recognition of the promoter by formation of sequence specific DNA–protein interactions between RNAP, and in particular the σ subunit, and promoter DNA (3–5). The widespread σ^{70} family of initiation factors recognize conserved DNA sequences centered near -35 and -10 bp upstream of the transcription start site that are recognized by σ regions 4 and 2, respectively (3,5,6). Bacteria often contain multiple σ factors, each recognizing distinct promoter sites, thereby allowing the cell to express specific genes in response to changing growth conditions (7,8). The genes controlled by each σ factor are defined as its regulon.

The expression profile of individual genes is often determined by the combinatorial action of multiple regulators (2,9). For many operons, transcription can initiate from more than one promoter element, depending on environmental conditions. Multiple promoter sites are often revealed by the presence of multiple transcription start sites, including some that maybe within genes or within operons, and the resulting transcripts may differ in their stability or the efficiency or site of translation initiation (10). In addition, it is possible for two or more σ factors to recognize the same or largely overlapping -35 and -10 core promoter elements and this may lead to initiation from the same position (11). We refer to this latter phenomenon, where multiple holoenzymes can initiate from the same promoter and from the same start site(s), as overlapping promoter recognition.

One of the best studied cases of overlapping promoter recognition by σ factors occurs in *Escherichia coli* between the primary housekeeping σ factor, σ^{70} and the alternative σ factor, σ^S . These two closely related σ factors recognize very similar core promoter elements, although σ^S is able to

^{*}To whom correspondence should be addressed. Tel: +1 607 255 3086; Fax: +1 607 255 3904; Email: jd9@cornell.edu

[†]These authors contributed equally to the paper as first authors.

Present addresses:

Ahmed Gaballa, Department of Food Science, 350 Stocking Hall, Cornell University, Ithaca, NY 14853, USA.

Veronica Guariglia-Oropeza, Department of Food Science, 350 Stocking Hall, Cornell University, Ithaca, NY 14853, USA.

Franziska Dürr, Technische Universität (TU) Dresden, Institute of Microbiology, 01062 Dresden, Germany.

Bronwyn G. Butcher, Cornell Lab of Ornithology, Cornell Univ., 159 Sapsucker Woods Rd, Ithaca, NY 14853, USA.

Albert Y. Chen, Department of Bacteriology, University of Wisconsin-Madison, 1550 Linden Drive, Madison, WI 53706, USA.

Pete Chandrangsu, Department of Microbiology, 370 Wing Hall, Cornell University, Ithaca, NY 14853, USA.

recognize less conserved -35 elements (12), and its activity is stimulated by A/T stretches at positions -29 to -25 and -22 to -18 (11). Moreover, σ^S is able to better tolerate sub-optimal spacer lengths, whereas σ^{70} strongly prefers promoters with 17 bp spacing (12). A cytosine at position -13 directly upstream of the -10 element is one characteristic of σ^S -dependent promoters and, conversely, counter-selected in σ^{70} -dependent promoters (13–16). In addition, these two σ factors differ in their response to the presence and location of upstream stimulatory elements (UP elements) recognized by the RNAP α subunit: a distal UP-element site is beneficial for σ^S recognition, whereas a proximal UP-element favors σ^{70} selectivity. Most likely, a combination of several of these factors ultimately determines the σ specificity of each promoter, with many promoter elements being well recognized by both σ factors. Overlapping promoter recognition has also been reported for σ^{70} and the heat shock σ factor, RpoH (17). In this case, the two σ factors recognize different consensus sequences, but these can be interdigitated in such a manner as to allow initiation from the same start site.

Here, we focus on the phenomenon of overlapping promoter recognition among promoters recognized by extracytoplasmic function (ECF) family σ factors. The ECF family of σ factors is characterized by their generally small size relative to the primary σ factor, corresponding to the presence of only conserved σ regions 2 and 4 (18,19). Multiple ECF σ factor paralogs are present in many organisms with, in extreme cases, 50 or more paralogs in a single species (20). The activation of these alternative σ factors is often associated with signal sensing pathways initiated at the cell surface that inactivate the corresponding membrane-associated anti- σ factor leading to release of active σ (19,21). The resulting stress responses may be largely distinct, or may overlap extensively with those controlled by other ECF σ factors, depending on the extent of overlapping promoter recognition.

The genome of *Bacillus subtilis* encodes 7 ECF σ factors: σ^M , σ^W , σ^X , σ^V , σ^Y , σ^Z and σ^{ylaC} (22,23). The regulons for five of these have been defined in detail (reviewed in (23)) and are thought to be determined largely by the sequence determinants in the -35 and -10 elements. σ^M regulates a large set of genes that include essential functions of cell division and envelope synthesis (24). The σ^W regulon comprises of at least 60 genes that are expressed in response to membrane active agents and involved in inactivation, sequestering, or eliminating toxic compounds from the cell (25). The σ^X regulon includes genes which serve to alter cell surface properties to provide protection against antimicrobial peptides (26), whereas σ^V regulates an overlapping set of ~ 30 genes and plays a primary role in lysozyme resistance (27,28). Finally, σ^Y appears to control a small regulon of less than a dozen genes with poorly defined functions (29,30). The regulons and functions of σ^Z and σ^{ylaC} have not been clearly defined (31,32), but they are induced in response to some stress conditions (33,34).

Functional redundancy and overlapping promoter recognition among the ECF σ factors of *B. subtilis* has been thoroughly documented (27,35–38). Indeed, some phenotypes associated with lack of expression of specific genes can only be obtained by mutating two or more of the activating σ factors (31,32). In many cases, this has been ascribed to over-

lapping promoter recognition, as supported by *in vitro* transcription experiments (26,35,39) and high resolution start-site mapping (40). Indeed, analysis of the global, condition-dependent transcriptome of *B. subtilis* revealed a computationally inferred cluster of promoters with the general characteristics of ECF-class promoters (σ^{ECF}), but the individual regulons could not be discerned under the set of conditions tested (41). In general, these σ^{ECF} promoters often share a conserved ‘AAC’ motif in the -35 region, and promoters recognized by two or more of σ^X , σ^W and σ^M (collectively defined as σ^{XWM}) have a -10 element consensus of ‘CGT’ (32). As shown previously, a subset of σ^{XWM} promoters are also recognized σ^V , thereby defining a σ^{XWMV} sub-group (27).

Even though overlapping promoter recognition has been frequently observed among ECF σ factors, the mechanisms that allow some promoters, but not others, to be recognized by more than one σ factor have not been fully resolved. By analogy with the example of *E. coli* σ^{70} and σ^S , overlapping promoter recognition among ECF σ factors is primarily due to similarities in the preferred consensus sequences in the -35 and -10 elements (23), but is strongly modulated by other features within the promoter region. There are clearly some promoters that are highly specific for a single ECF σ factor (there is little if any cross-talk), and others that can be recognized by three or more. For example, sequence comparisons indicate that promoters recognized by σ^X and σ^W are distinguished by two key positions in the -10 element (39), and changes to the -10 element were sufficient to switch a promoter from the σ^X to the σ^W regulon and *vice versa* (42). With this particular pair of σ factors, a -10 sequence of CGAC is found in the σ^X -specific *sigX* autoregulatory promoter and CGTA confers highly selective recognition of the *sigW* autoregulatory promoter. However, promoters with a -10 element of CGTC may be recognized by both σ factors (overlapping promoter recognition) (42), and indeed may also be recognized by σ^M (36) and/or σ^V (27,28).

Here we identify a -35 proximal homopolymeric T-tract as a novel promoter element involved in promoter discrimination by ECF σ factors in *B. subtilis*, likely by altering the trajectory of promoter DNA during engagement with RNAP. Based on its conservation in a significant subset of deduced consensus sequences for ECF σ factors, we surmise that this element likely plays a general role in modulating the extent and impact of overlapping promoter recognition among ECF σ factors.

MATERIALS AND METHODS

Strain construction and growth conditions

The *B. subtilis* 168 wild-type (WT) strain (168) was grown in LB medium. Unless otherwise indicated, liquid media were inoculated from an overnight preculture and incubated at 37°C with shaking at 200 rpm. For selection in *B. subtilis*, antibiotics were added at the following concentrations: erythromycin ($1\ \mu\text{g ml}^{-1}$) and lincomycin ($25\ \mu\text{g ml}^{-1}$) [for selecting for macrolide-lincosamide-streptogramin B (MLS) resistance], spectinomycin ($100\ \mu\text{g ml}^{-1}$), chloramphenicol ($10\ \mu\text{g ml}^{-1}$), kanamycin ($10\ \mu\text{g ml}^{-1}$) and neomycin (10

$\mu\text{g ml}^{-1}$). Routine molecular biology procedures were carried out using *E. coli* DH5 α as described (43). Isolation of *B. subtilis* chromosomal DNA, transformation and specialized SP β transduction were performed as described (44). Restriction enzymes, DNA ligase and DNA polymerases were used according to the manufacturer's instructions (New England Biolabs).

To construct *lacZ* transcriptional fusions, the regulatory regions were amplified from *B. subtilis* 168 genomic DNA by PCR and cloned as a HindIII–BamHI fragment into the integrational plasmid pJPM122 (45). The resulting constructs were linearized with ScaI and transformed into *B. subtilis* strain ZB307A (46), a strain containing a temperature-sensitive SP β phage carrying a Tn917-element with a *lacZ* gene, selecting for neomycin resistance. After plasmid integration adjacent to the *lacZ* gene, the promoter of choice (P) generates a P-*cat-lacZ* operon fusion expressing both chloramphenicol acetyltransferase (*cat*) and β -galactosidase (*lacZ*), as described (45). An SP β transducing lysate was prepared by heat induction and used to transduce different strain backgrounds as described (45). β -galactosidase activity was assayed in biological triplicate using a modification of the procedure of Miller (47) as described in (48).

Microarray analysis

The WT *B. subtilis* 168 parental strain and an isogenic strain lacking all seven ECF σ factor genes (designated $\Delta 7$) were used for microarray studies (31,49). The WT and $\Delta 7$ strains with P_{*spac-sigV*} were grown in LB to OD₆₀₀ ~0.4. Cultures were divided into two parts, and 2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to one half of the culture and further incubated for 20 min. RNA isolation, cDNA synthesis, microarray hybridization and analysis was done according to (50) using a microarray platform containing 65-mer oligonucleotides (one per annotated open reading frame; Sigma-Genosys). The resulting datasets have been deposited to the NCBI GEO database under accession number GSE95393.

Promoter site directed mutagenesis

The sequence of each promoter region was mutated using overlap extension polymerase chain reaction (PCR) (51). Briefly, for each promoter a set of primers were designed, two universal and flanking, and two pairs of mutagenic overlapping primers to introduce the TTTTT (T₅) \rightarrow AAAAA (A₅), T₅ \rightarrow TAAAT or T₅ \rightarrow TTATT changes. Each mutagenic primer was used in pair with a flanking primer to generate two fragments that will have overlapping ends. These two fragments were used in a second PCR joining reaction using only the flanking primers, and the resulting full length product was amplified by PCR.

Protein purification

RNAP was purified from *B. subtilis* expressing His-tagged β' subunit as described (52). The *sigV* gene was PCR amplified from *B. subtilis* chromosomal DNA and cloned into the pMCSG19 vector (53), which allows the expression of

SigV as a His-tagged-maltose binding protein fusion with a site for *in vivo* cleavage to remove the MBP domain. The resultant plasmid was transformed into BL21/DE3(pLysS) cells harboring the plasmid pRK1037, which expresses the tobacco vein mottling virus protease to cleave the MBP domain. Cells were grown in 1 L of LB medium with ampicillin (100 $\mu\text{g ml}^{-1}$) and kanamycin (1 $\mu\text{g ml}^{-1}$) at 37°C to OD₆₀₀ 0.4. IPTG was added to 0.3 mM final concentration and the culture was incubated with shaking overnight at 14°C. Cells were collected by centrifugation and His-tagged σ^V was purified from the soluble fraction using Prepease Histidine-tagged purification resin (Affymetrix) according to the manufacturer's recommendations. The σ^M , σ^W and σ^X proteins were purified as described (24).

In vitro transcription

Run-off *in vitro* transcription was performed as previously described (54). Briefly, purified RNAP was incubated with different purified σ factors in a 1:5 molar ratio in transcription buffer (10 mM Tris–HCl pH 8.0, 10 mM MgCl₂, 0.5 mM dithiothreitol (DTT), 50 mM KCl, 500 $\mu\text{g/ml}$ acetylated bovine serum albumin, 5% glycerol) and incubated on ice for 15 min. The *in vitro* transcription reactions contained 10 nM of linear promoter fragment in transcription buffer and 40 nM of RNAP. After 10 min of incubation 37°C, transcription was initiated by adding 0.25 mM (final concentration) of adenosine triphosphate (ATP), guanosine triphosphate (GTP) and cytidine triphosphate (CTP) and 0.025 mM uridine triphosphate (UTP) and 25 μCi of [α -³²P]-UTP. After 10 min of incubation, the reaction products were ethanol precipitated in the presence of 2 mM ethylenediaminetetraacetic acid, 0.3 M sodium acetate pH 5.2 and 3 μg glycogen. The RNA pellet was washed with 70% cold ethanol, dried and dissolved in formamide-containing loading buffer and separated on a 6% denaturing polyacrylamide sequencing gel. The gel was then dried and exposed to a phosphorimager screen. The resulting phosphorimage was visualized using a Typhoon FLA 7000 imaging system (GE Healthcare) and analyzed using ImageJ software.

Bioinformatics analysis

Promoter sequences used for alignment and WebLogo analyses for ECF families were obtained from the supplementary data of Staron *et al.* (20). Promoter sequences for the auto-regulatory region of different sigma factor orthologs from different bacilli were selected based on similarity and Gene context (55). Promoter regions from *B. subtilis* ECF σ controlled genes were obtained from experimentally determined start sites (23,24,27,35,41,56). Promoter sequence alignment was done using BioEdit and WebLogo 3 (57).

P_{*dltA*} DNA was modeled onto the crystal structure of the *Thermus aquaticus* RNAP-fork junction complex (PDB ID:1L9Z) (58) by aligning the non-template strand –35 element of the fork junction DNA and P_{*dltA*} DNA using the 'align' command in PyMol (www.pymol.org). Figure preparation was performed using the UCSF Chimera package (59).

Native PAGE analysis of DNA bending

Annealed oligos encoding the T₅ *dltA* promoter (endpoints: –50 to +10) or the G₅ variant were electrophoresed on a 15% native polyacrylamide gel at 100 V for 10 h at 4°C in 0.5× tris-acetic acid-ethylenediaminetetraacetic acid (TAE) buffer. The gel was stained for 10 min in EtBr and visualized using a BioRad ChemiDoc XRS.

RESULTS

The extent of the σ^V regulon is increased in the absence of other ECF σ factors

Here, we compared the transcriptional response to induction of σ^V (using an IPTG-based induction system) in a WT strain (168) and in a strain deleted for all seven ECF σ factor-encoding genes (designated $\Delta 7$; (31,49)) (Figure 1). The observed transcriptional response extends and refines that described previously as resulting from induction of σ^V (27,60).

The σ^V regulon includes a single, autoregulated operon (*sigVrsiVoatAyrhK*) transcribed from a characteristic ECF-class promoter with a –35 element (GCAAAC) and –10 element (CGTC) as defined by primer extension start site mapping (56). This promoter is, for reasons not entirely clear, strongly activated only by σ^V *in vivo* (27,28,60). The majority of the σ^V -induced genes are associated with promoter sites known to be activated by σ^X , σ^W and/or σ^M and associated with promoter sequences previously defined as the σ^{XWMV} sub-group (27). Since these ECF σ factors recognize very similar promoter consensus sequences (23), we interpret this as indicative of overlapping promoter recognition. Since these studies were conducted in the $\Delta 7$ strain that lacks the other ECF σ factors (including σ^M and σ^W), it was unclear whether the ability of σ^V to activate this full set of promoters is a characteristic of WT cells.

When we compared the transcriptional response to induction of σ^V in the *B. subtilis* 168 parental strain and the $\Delta 7$ strain we noted a significantly broader and more robust response in the $\Delta 7$ background. This effect was most pronounced for a subset of those promoters previously assigned to the σ^W regulon, which were generally unresponsive to σ^V induction in the 168 parental strain (Figure 1A), but were strongly induced in the $\Delta 7$ background (Figure 1B). This difference is clear when induction under the two conditions is compared directly (Figure 1C). We note that this is the opposite of what we would predict if σ^V normally functions to activate expression of one or more other ECF σ encoding genes in a cascade fashion, as occurs for example with the σ factors controlling sporulation (61) and with many of the 13 σ factors of *Mycobacterium tuberculosis* (62).

To gain further insights into the increased breadth and amplitude of gene induction in the $\Delta 7$ strain we sorted all of the genes based on the sequence of the –10 region of the promoter, and in particular whether the last base was an A (CGTA) or a C (CGTC) (Supplementary Table S1). Previously, we have shown that the sequence CGTA is associated with recognition by σ^W , whereas promoters with CGTC are recognized by σ^M or σ^X and, in some cases, also by σ^W (24,42). Promoters with a –10 region CGTA

sequence were generally not induced upon σ^V induction in the WT strain, but many were highly responsive in the $\Delta 7$ strain (Figure 1D). In contrast, for those promoters with a –10 sequence of CGTC the average induction in WT was higher, but was generally increased less in the $\Delta 7$ strain (an average increase in induction of ~2-fold versus >10-fold for the CGTA promoters). These observations support the notion that promoters belonging to the σ^W regulon (and with CGTA in the –10 region) are somehow less responsive to σ^V induction in WT, perhaps due to promoter occlusion by basal levels of σ^W holoenzyme. Indeed, the σ^W regulon has a significant level of basal activity under a wide variety of growth conditions (63).

The homopolymeric T-stretch is correlated with σ^V responsiveness

We previously noted that T residues in the –35 proximal region (–26 to –29) seemed to be associated with responsiveness to σ^V (27). Here, we have extended this analysis by generating sequence alignments and WebLogos of known ECF σ -regulated genes based on their induction in response to σ^V (Supplementary Figure S1). Overall, those promoters known to be activated by σ^M , σ^W and/or σ^X and also induced (at least 2-fold) upon induction of σ^V displayed extended conservation in the –35 region (tgaAACntttt; lower case bases are less conserved as shown in Supplementary Figure S1A), whereas those relatively non-responsive to σ^V induction (<2-fold increase) lacked conservation of these T residues (tgaAAC; Supplementary Figure S1B). Among the σ^V responsive promoters, conservation of this T-tract is most pronounced for those promoters with a CGTA –10 element (and therefore generally associated with activation by σ^W), but is also apparent for those with a CGTC –10 element (Supplementary Figure S1C and D). For example, among σ^W -dependent promoters with a –10 region consensus of CGTA, the –35 region consensus for those induced at least 2-fold by σ^V (tGAAACnTTTT) is quite distinct from that for those promoter non-responsive to σ^V induction (TGAAAC) (Supplementary Figure S2).

Effects of mutations affecting the homopolymeric T-stretch *in vivo*

We used the *dltA* promoter (P_{dltA}), which can be activated by σ^X (26), σ^M (24) and σ^V (27), as a model system to assess the importance of this homopolymeric T-stretch for expression by holoenzymes bearing different σ factors. The *dltABCDE* operon encodes proteins involved in D-alanylation of teichoic acids, a cell wall modification pathway that changes the overall net charge (64) and is involved in resistance to cationic antimicrobial peptides and lysozyme (26,27).

We constructed *lacZ* transcriptional fusions with P_{dltA} and mutant derivatives that altered the –35 proximal homopolymeric T-stretch (Figure 2). In the *B. subtilis* 168 parental strain disruption of this T-stretch decreased activity substantially, with a 2-fold reduction noted for the single base change converting TTTTT (T₅) to TTATT and a 10-fold reduction when T₅ was replaced by A₅ (Figure 2B). This indicates that this promoter has significant basal activ-

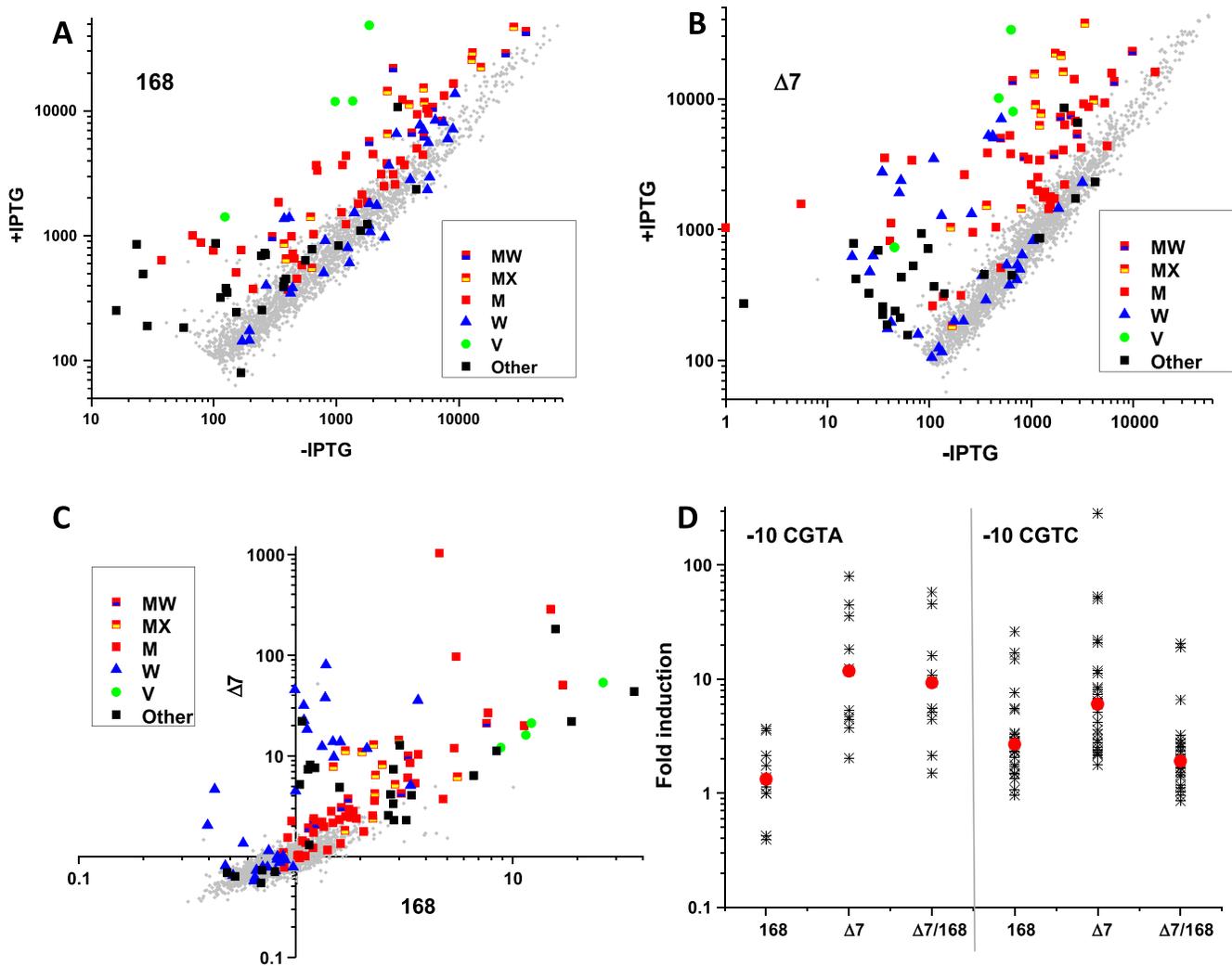


Figure 1. Overlap of ECF σ factor regulons. (A) Gene expression as measured by cDNA microarray analysis in WT *Bacillus subtilis* 168 cells expressing $P_{\text{spac-sigV}}$ showing normalized fluorescence values corresponding to uninduced cells (x-axis) versus those harvested 20 min after the addition of 2 mM IPTG (y-axis). Each gene associated with an ECF σ factor regulon is represented by a symbol as indicated in the inset (e.g. MW indicates genes known to be activated by both σ^M and σ^W), with other genes indicated in small circles. (B) Gene expression as measured by cDNA microarray analysis in the $\Delta 7$ background expressing $P_{\text{spac-sigV}}$ showing normalized fluorescence values corresponding to uninduced cells (x-axis) versus those harvested 20 min after the addition of 2 mM IPTG (y-axis). (C) Fold induction of genes in the *B. subtilis* 168 parental (168) versus $\Delta 7$ backgrounds over-expressing *sigV*. (D) Fold induction of genes in the 168 WT versus $\Delta 7$ backgrounds after over-expression of σ^V sorted by the last position of the -10 region (either CGTA or CGTC) of the cognate promoter. The mean value is shown in circles for each dataset.

ity in *B. subtilis* 168, and this activity is strongly modulated by this homopolymeric T-stretch.

To determine how these same mutations might affect recognition by each of the ECF σ factors reported to recognize P_{dltA} , we repeated this analysis in the $\Delta 7$ mutant background after induction of σ^V , σ^M or σ^X (Figure 2C–E). Even a single base change within this T-tract (from T_5 to TTATT) led to a >3 -fold reduction in activity when driven by σ^X or σ^V . In contrast, this same change actually improved promoter activity (by nearly 2-fold) when driven by σ^M (Figure 2D). Thus, mutations that affect the integrity of this homopolymeric T-stretch can differentially affect the activity of ECF σ factors, although in all cases when T_5 was replaced by A_5 there was a drastic loss of activity.

Generally consistent findings were noted for mutations affecting T-tracts in three other ECF σ factor dependent promoters. Mutation of a -35 proximal T_4 tract in the *bcrC* promoter decreased activity in WT cells and in the $\Delta 7$ strain when promoter activity was driven by induction of σ^V or σ^M (Supplementary Figure S3). A T_4 to TTAT change reduced σ^V activity, but not that of σ^M .

Note that this promoter was not strongly activated by induction of σ^X . Similarly, mutations affecting a -35 proximal T_5 -tract in the *abh* and *pbpX* promoters confirm the general stimulatory effect of this sequence *in vivo*, and support the notion that this sequence is particularly important when transcription is activated by the σ^V holoenzyme (Supplementary Figure S4).

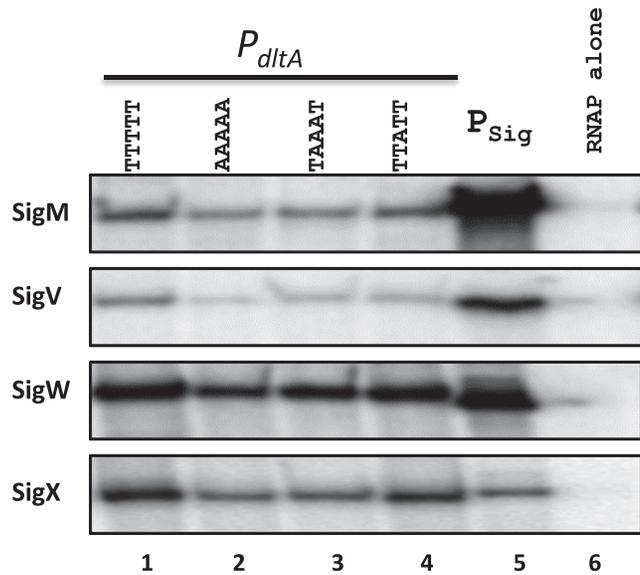


Figure 3. Effect of T₅-tract on *in vitro* transcription of *dltA* promoter. RNAP without addition of any σ factors was unable to transcribe the cognate autoregulatory promoter region for each ECF σ (lane 6), indicating the absence of significant amounts of the cognate σ factor in the purified RNAP. Addition of σ^M , σ^V , σ^W or σ^X to RNAP at 5:1 molar ratio (in lanes 1–5) enables the resulting RNAP- σ complex to transcribe the autoregulatory promoter of each σ , indicating that all purified proteins are functional (lane 5). Transcription activity of WT (lane 1) and mutant (lanes 2–4) *dltA* promoter regions was determined using different RNAP- σ holoenzymes.

When the P_{murG} -30 to -26 region was mutated from CCGAG to T₅ the level of activity in WT cells increased 6-fold (Figure 4B). This is consistent with the general stimulatory effect of the T₅ sequence, as noted above. To determine how individual ECF σ factors might respond to the introduction of the T₅ sequence in this promoter context, we used the $\Delta 7$ strain expressing σ^M , σ^X or σ^V under xylose induction. For both σ^M and σ^X , induction led to activation of P_{murG} , but the level of expression was not strongly impacted (~2-fold) by the introduction of the T₅ sequence. In contrast, the ability of σ^V to initiate transcription from this promoter was increased ~10-fold by addition of the T₅ sequence (Figure 4B).

These results were corroborated using *in vitro* transcription assays with P_{murG} as template. With the WT promoter sequence, transcription was most active with the σ^M holoenzyme, as expected, with little to no transcription detected in the presence of σ^V , σ^W and σ^X (Figure 4C). Introduction of the T₅-tract results in a promoter that can be transcribed by all four ECF σ factors, although activity with σ^M is still the highest (Figure 4C). Consistent with the *in vivo* results, the stimulatory effect of the T₅-tract was highest with the σ^V holoenzyme (~8-fold). These results support the notion that a homopolymeric T-tract in an ECF σ factor regulated promoter is, in general, stimulatory and can extend the range of σ factors that can potentially recognize the promoter. Conversely, the absence of a stimulatory T-tract may be beneficial when seeking to restrict the activation of a promoter to a single ECF σ factor. Thus, -35

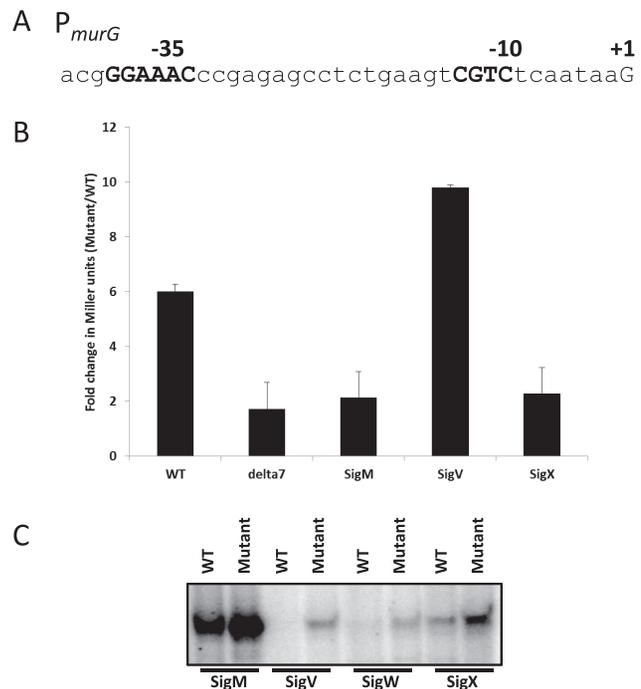


Figure 4. Addition of a T₅-tract enhances promoter recognition by ECF σ factors. (A) Sequence of the P_{murG} promoter that lacks a T₅-tract showing the -35 and -10 (bold), +1 transcription start site. (B) Fold change in P_{murG} -*lacZ* expression of mutant (CCGAG to T₅) to WT promoter region in WT, $\Delta 7$ or $\Delta 7$ over-expressing σ^M , σ^V or σ^X (values are mean \pm SD; $n = 3$). Actual mean values of Miller units for the WT promoter were 4.95 (168), 0.34 ($\Delta 7$), 4.45 (induction of σ^M), 0.4 (induction of σ^V) and 0.32 (induction of σ^X). (C) *In vitro* transcription of the WT P_{murG} promoter and its mutant (CCGAG to T₅) by RNAP associated with σ^M , σ^V , σ^W or σ^X .

region proximal T-tracts provide a tool for modulating the extent of regulon overlap.

Homopolymeric T-tract sequences in the -35 proximal region of ECF σ -dependent promoters

ECF σ factors are the smallest alternative σ factors, as well as the most abundant and diverse group (20,23). A previous bioinformatic analysis of 2708 ECF σ factors (those containing conserved sigma regions 2 and 4, while lacking region 3) revealed that 2/3 of these could be organized into 43 phylogenetically distinct clusters with at least 10 members each (designated ECF01–ECF43), with the remainder forming smaller clusters (20). Since most ECF σ factors regulate their own synthesis, inferences regarding the likely promoter selectivity of each ECF σ group were developed from a comparison of known and conserved promoter elements upstream of each σ factor operon. Promoters recognized by ECF σ factors typically have a conserved AAC motif in their -35 regions with group-specific -10 region consensus sequences (20).

In *B. subtilis*, σ^W , σ^V and σ^Y belong to groups ECF01, ECF30 and ECF31, respectively, whereas σ^M and σ^X did not fall into any of the major groups (20). As reported previously (20), comparison of the deduced consensus sequences for the ECF01, ECF30, and ECF31 families (Figure 5A)

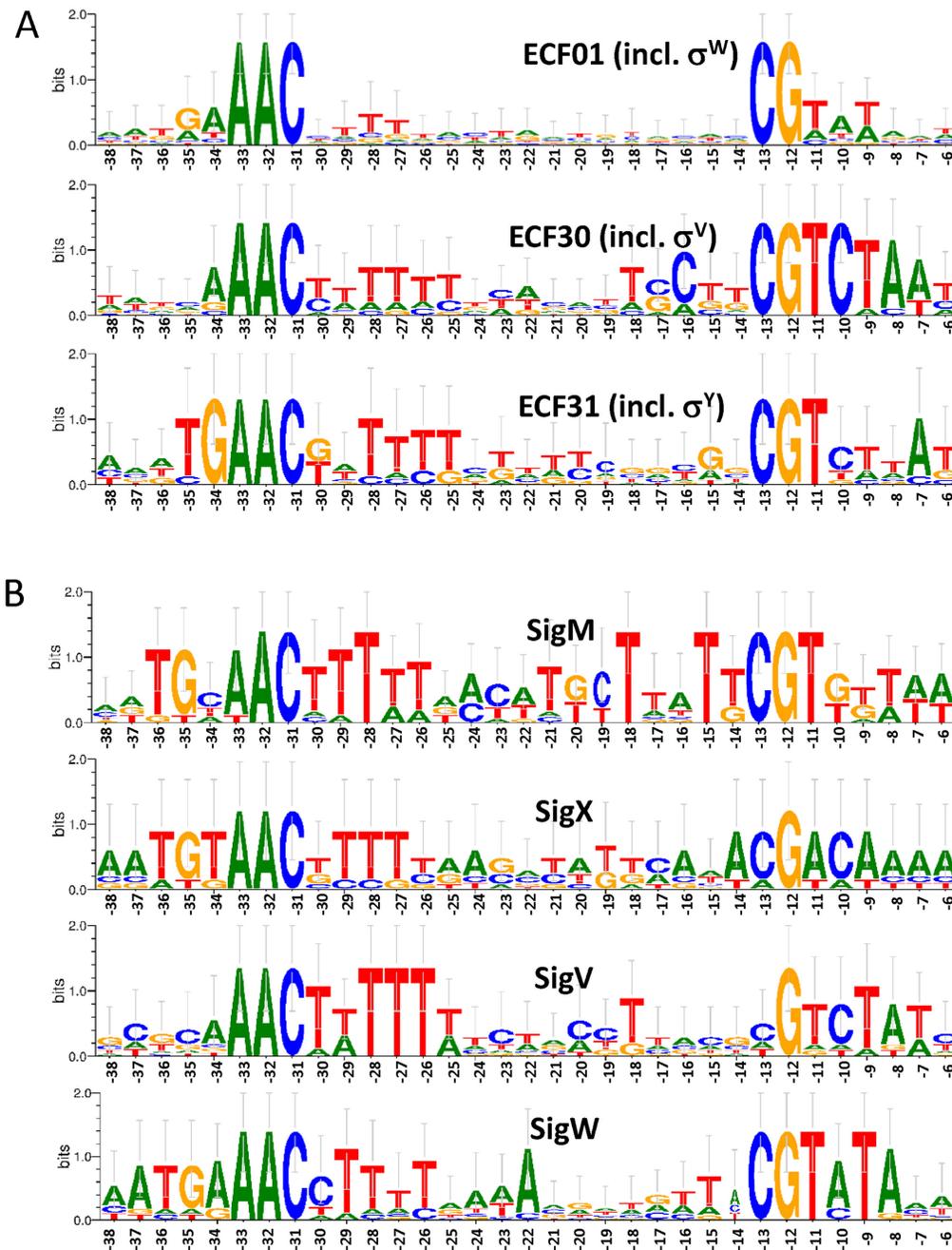


Figure 5. Conservation of a T₅-tract in ECF σ factor families. Alignment of ECF σ factor-dependent promoter regions was analyzed using WebLogo 3 (57). (A) WebLogo analysis of ECF01, ECF30 and ECF31 families. Promoter sequences data were extracted from the Supplementary Tables in (20). (B) WebLogo analysis of the auto-regulatory region of σ^M , σ^X , σ^V and σ^W orthologs from different bacilli. ECF σ factors were selected based on similarity and Gene context (55).

indicates a weak conservation of T residues proximal to the -35 element for ECF01(σ^W) and a more pronounced conservation for ECF30(σ^V) and ECF31(σ^Y). It is notable that the consensus for σ^W derived from the regulon defined in *B. subtilis* includes a T-rich sequence (23,35), whereas that derived from only autoregulatory promoters does not. One interpretation of this observation is that some σ^W targets belong to more than one regulon, whereas autoregulatory

promoters, at least in *B. subtilis*, tend to be highly selective for their cognate ECF σ factor (23,39,60,65).

To gain insights into whether conservation of this T-rich region extends to σ^M and σ^X family proteins (which were not part of the 43 major groups defined in (20)), we generated sequence alignments using an analogous process in which the auto-regulatory regions were recovered from orthologous σ factor genes (restricting the search however to *Bacillus* spp.) and used to generate promoter consen-

sequences. In this case, orthology was judged by conservation of both σ factor sequences and genomic context, rather than relying on σ factor sequence alone. For comparison with (20), we included in this analysis those ECF σ factors orthologous to σ^W (ECF01) and σ^V (ECF30). For each σ , the -35 proximal T-tract was conserved (Figure 5B). The σ^V consensus derived from *Bacillus* strains (Figure 5B), is generally consistent with that derived for the broader ECF30 cluster (Figure 5A), and also with the previously reported σ^V consensus derived from comparison of promoters activated by overproduction of σ^V in *B. subtilis* (27). The σ^X autoregulatory consensus differs in the -10 region (CGAC) from that of the regulon as a whole (CGwC; w = A or T). This is consistent with the previous observations that the autoregulatory promoter is highly specific for σ^X (due to the CGAC -10 element), whereas many promoters in the broader σ^X regulon are also recognized by other ECF σ factors (and have a more widely recognized CGTC -10 element) (42).

Collectively, the results above indicate that the ability of a -35 proximal T-tract sequence to stimulate transcription varies significantly among the ECF σ factors tested. The largest stimulatory effect was noted for σ^V , with the least effect for σ^W . The presence of short oligo-dT (alternatively known as oligo-dA) tracts are known to induce intrinsic DNA bends in B-form DNA, with maximal bending noted for phased T₅₋₆ sequences (66). We therefore postulated that the stimulatory effect of this -35 proximal oligo-dT sequence was likely due to changes in the trajectory of DNA during binding of the promoter region with RNAP.

To assess the effect of the T₅-tract in the *dltA* promoter we first modeled the promoter region as B-form DNA with and without a T-tract sequence. As expected, the T₅-tract is predicted to impart an intrinsic DNA bend (Supplementary Figure S5A). We also used native polyacrylamide gel electrophoresis to compare the mobilities of WT P_{*dltA*} DNA and P_{*dltA*} DNA in which the T₅-tract was replaced with a G₅-tract (Supplementary Figure S5B). Even though the two DNA fragments are the same length (60 bp), WT P_{*dltA*} migrated more slowly than the G₅ variant, consistent with the increased bend angle of WT P_{*dltA*} observed *in silico*. The effect of this change in DNA trajectory on the interaction with RNAP was modeled by docking both forms of the P_{*dltA*} promoter (with and without the T-tract) onto the crystal structure of the *T. aquaticus* RNAP holoenzyme in complex with fork junction DNA (PDB ID: 1L9Z) (58). While crystal structures of complete RNAP open complexes are available (67), the fork junction complex represents an earlier intermediate (68,69) and may more closely emulate initial binding of RNAP to promoter DNA. Modeling indicates that this motif would likely bend the DNA toward RNAP, leading to a near superposition on the intermediate state represented by the fork-junction DNA (Figure 6). If RNAP binds initially at the -35 consensus element, such bending is predicted to facilitate the subsequent engagement of σ region 2 with the -10 consensus element. We hypothesize that this facilitating mechanism might be most important for those ECF σ factors that are relatively constrained in their ability to engage simultaneously with both the -35 and -10 consensus element due to small linker region between the corresponding protein recognition domains, do-

main 4 and domain 2, respectively. One factor that can restrict RNA interactions with promoter DNA is the limited flexibility of σ domain 4, which is constrained by interactions with the β flap tip helix (Figure 6) (67). In general, we note that there is a correlation between short interdomain linker regions (which potentially restrict σ factor flexibility) and the stimulatory activity of the -35 proximal T-rich sequence (Supplementary Figure S6).

DISCUSSION

RNAP holoenzyme contacts a large region of DNA during the processes of promoter engagement, melting, and clearance (5). It is therefore not surprising that sequences throughout this extended region can impact activity. These include the -35 and -10 (core) promoter elements (18), upstream AT-rich UP elements (70), and the region surrounding the transcription start site (including the discriminator element) (71). Here, we postulate that the presence of a spacer-region T-tract, proximal to the -35 recognition element, changes the trajectory of the DNA (Figure 6) and facilitates productive engagement of the holoenzyme with the promoter region.

Previous studies have begun to provide insights into the basis of promoter recognition by ECF σ factors (18,72,73). On average, promoters recognized by ECF σ factors adhere closely to consensus, suggestive of strong initial binding (22,23). Promoters activated by ECF σ factors may be specific for only one σ , or may display overlapping promoter recognition. Promoter selectivity is likely due to differences in the -35 and -10 consensus elements and in optimal spacer lengths, which range from 14 to 17 bp (74). Sub-optimal spacer lengths alter both the separation and relative rotational positioning of the -35 and -10 elements (75–78), and this presumably requires conformational adaptation on the part of the σ factor.

Here, we describe the novel role of a T-tract in the spacer region between -30 and -26 as an additional element affecting promoter selectivity by ECF σ factors in *B. subtilis*. Previous studies with *E. coli* σ^{70} have demonstrated that spacer region sequences can influence promoter strength by affecting either RNAP–DNA contacts or DNA conformation. For example, *E. coli* σ^{70} region 3 contacts the extended -10 region (consensus TGTGn) (3), the linker between σ^{70} region 2 and 3 interacts with the DNA backbone at -18 (79,80), and the β' subunit zipper region interacts with the ‘Z-element’ near position -21 (81). The spacer region tolerates large scale substitutions of DNA sequences and these modulate, but generally do not eliminate, promoter activity (75,82–84). Previously, T-tract sequences were found to reduce activity of an already consensus σ^{70} promoter (83,84), but the role of sequence-directed DNA-bending in the -35 proximal region described here has not been previously reported.

Bioinformatic studies suggest that T-tract sequences may be a common feature in promoter recognition by ECF σ factors. Among 29 of the major ECF σ groups (20), three showed a relatively high conservation of T residues in this part of the spacer region. These groups include ECF02 (including *E. coli* RpoE), ECF30 (including *B. subtilis* σ^V) and ECF31 (74). However, these consensus sequences were all

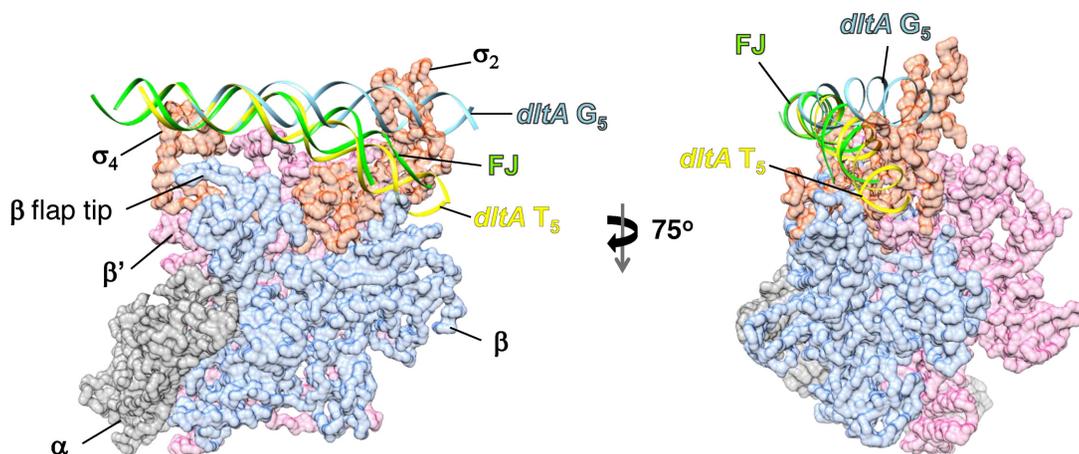


Figure 6. Effect of T₅ tract on DNA trajectory in during RNAP–DNA complex formation. The RNAP holoenzyme (PDB ID:1L9Z) (58) is shown as a surface representation (α_2 , β , β' , σ^A are indicated) and the fork junction DNA is shown as a ribbon (FJ). B-form WT (*dltA T₅*) and variant (*dltA G₅*) *dltA* promoter DNA was generated and modeled on to the *Thermus aquaticus* RNAP–fork junction DNA crystal structure. The ‘Align’ command in PyMol (www.pymol.org) was used to align the resulting B-form models to the non-template strand of the fork junction –35 element and the figure was generated using UCSF Chimera (59).

derived from presumed autoregulatory promoters, which by analogy with *B. subtilis* (23,60), may have been selected to have a restricted level of crosstalk. The prevalence of T-tracts in the overall regulons for most ECF σ factors is unknown, and the contribution of the T-tract to promoter recognition may have been systematically underestimated in this prior analysis.

The role of this –35 proximal T-tract in promoter recognition is not yet clear, but we speculate that the key parameter is introduction of an intrinsic DNA-bend (Figure 6). The ability of short 4 to 5 nt T-tracts to bend DNA is well documented (85), and the observation that disrupting this sequence (e.g. changing T₅ to TTATT) reduces the stimulatory activity significantly argues in favor of this interpretation. The possible role of intrinsic DNA-bending within the spacer region has not been systematically investigated, but bending is known to accompany promoter engagement (5). Perhaps the closest analogy to the results reported here is the introduction of DNA-bending and twisting by MerR-family regulatory proteins. For example, *E. coli* CueR binds to spacer region DNA and, in a Cu(I)-responsive manner, twists and bends the DNA to enable RNAP to effectively engage with both the –35 and –10 elements (86). In this case, protein-induced changes in DNA conformation are required to compensate for the non-optimal 19 bp spacer region (87). By analogy, the presence of a T-tract directed DNA-bend in some ECF σ factor dependent promoters may compensate for conformational restrictions imposed by a relatively short amino acid linker between the –10 and –35 recognition domains of σ (Supplementary Figure S5B). Indeed, there is an inverse correlation between the stimulatory activity of a T-tract and the length of the amino acid linker separating the promoter recognition domains in σ .

Our results provide further insights into the mechanisms that contribute to the nature and extent of functional overlap among the ECF σ factor regulons in *B. subtilis*. Spacer-region T-tracts modulate promoter respon-

siveness and thereby fine-tune the extent of activation of each promoter by the various σ factors, and therefore in response to various stresses. A set of ECF σ factors (selected to represent diverse groups) was shown previously to recognize their cognate promoters with little if any apparent cross-talk (orthogonality) (74). Since the –35 and –10 promoter elements are recognized by two distinct protein domains, hybrid promoters can be constructed by matching –35 and –10 elements that can then be recognized by domain-swapped σ factors (74). Consideration of the possible effects of –35 proximal T-tract sequences will likely be important in efforts to use ECF σ factors in systems biology, and consideration of linker length when constructing chimeric ECF σ factors may also be advisable. Ultimately, a fuller understanding of all of these components will be essential for efforts to predict the regulons controlled by ECF σ factors and for efforts to co-opt these regulatory systems for use as tools for bioengineering of new pathways and regulatory circuits.

AVAILABILITY

Microarray datasets have been deposited to the NCBI GEO database under accession number GSE95393.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

National Institutes of Health [GM047446, GM122461 to J.D.H.]. Funding for open access charge: National Institutes of Health [GM122461].

Conflict of interest statement. None declared.

REFERENCES

- Hellmann, J.D. (2009) RNA polymerase: a nexus of gene regulation. *Methods*, **47**, 1–5.

2. Browning, D.F. and Busby, S.J. (2016) Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.*, **14**, 638–650.
3. Feklistov, A., Sharon, B.D., Darst, S.A. and Gross, C.A. (2014) Bacterial sigma factors: a historical, structural, and genomic perspective. *Annu. Rev. Microbiol.*, **68**, 357–376.
4. Lee, J. and Borukhov, S. (2016) Bacterial RNA polymerase-DNA interaction—the driving force of gene expression and the target for drug action. *Front. Mol. Biosci.*, **3**, 73.
5. Ruff, E.F., Record, M.T. Jr and Artsmovitch, I. (2015) Initial events in bacterial transcription initiation. *Biomolecules*, **5**, 1035–1062.
6. Hook-Barnard, I.G. and Hinton, D.M. (2007) Transcription initiation by mix and match elements: flexibility for polymerase binding to bacterial promoters. *Gene Regul. Syst. Bio.*, **1**, 275–293.
7. Gruber, T.M. and Gross, C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.*, **57**, 441–466.
8. Paget, M.S. and Helmann, J.D. (2003) The σ^{70} family of sigma factors. *Genome Biol.*, **4**, 203.
9. Imam, S., Noguera, D.R. and Donohue, T.J. (2015) An integrated approach to reconstructing genome-scale transcriptional regulatory networks. *PLoS Comput. Biol.*, **11**, e1004103.
10. Wade, J.T. and Grainger, D.C. (2014) Pervasive transcription: illuminating the dark matter of bacterial transcriptomes. *Nat. Rev. Microbiol.*, **12**, 647–653.
11. Typas, A., Becker, G. and Hengge, R. (2007) The molecular basis of selective promoter activation by the sigmaS subunit of RNA polymerase. *Mol. Microbiol.*, **63**, 1296–1306.
12. Typas, A. and Hengge, R. (2006) Role of the spacer between the -35 and -10 regions in sigmaS promoter selectivity in *Escherichia coli*. *Mol. Microbiol.*, **59**, 1037–1051.
13. Weber, H., Polen, T., Heuveling, J., Wendisch, V.F. and Hengge, R. (2005) Genome-wide analysis of the general stress response network in *Escherichia coli*: σ^S -dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.*, **187**, 1591–1603.
14. Gaal, T., Ross, W., Estrem, S.T., Nguyen, L.H., Burgess, R.R. and Gourse, R.L. (2001) Promoter recognition and discrimination by E sigmaS RNA polymerase. *Mol. Microbiol.*, **42**, 939–954.
15. Lee, S.J. and Gralla, J.D. (2001) σ^{38} (RpoS) RNA polymerase promoter engagement via -10 region nucleotides. *J. Biol. Chem.*, **276**, 30064–30071.
16. Mitchell, J.E., Zheng, D., Busby, S.J. and Minchin, S.D. (2003) Identification and analysis of 'extended -10' promoters in *Escherichia coli*. *Nucleic Acids Res.*, **31**, 4689–4695.
17. Wade, J.T., Castro Roa, D., Grainger, D.C., Hurd, D., Busby, S.J., Struhl, K. and Nudler, E. (2006) Extensive functional overlap between sigma factors in *Escherichia coli*. *Nat. Struct. Mol. Biol.*, **13**, 806–814.
18. Campagne, S., Allain, F.H. and Vorholt, J.A. (2015) Extra cytoplasmic function sigma factors, recent structural insights into promoter recognition and regulation. *Curr. Opin. Struct. Biol.*, **30**, 71–78.
19. Mascher, T. (2013) Signaling diversity and evolution of extracytoplasmic function (ECF) sigma factors. *Curr. Opin. Microbiol.*, **16**, 148–155.
20. Staron, A., Sofia, H.J., Dietrich, S., Ulrich, L.E., Liesegang, H. and Mascher, T. (2009) The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) sigma factor protein family. *Mol. Microbiol.*, **74**, 557–581.
21. Ho, T.D. and Ellermeier, C.D. (2012) Extra cytoplasmic function sigma factor activation. *Curr. Opin. Microbiol.*, **15**, 182–188.
22. Helmann, J.D. (2002) The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.*, **46**, 47–110.
23. Helmann, J.D. (2016) *Bacillus subtilis* extracytoplasmic function (ECF) sigma factors and defense of the cell envelope. *Curr. Opin. Microbiol.*, **30**, 122–132.
24. Eiamphungporn, W. and Helmann, J.D. (2008) The *Bacillus subtilis* σ^M regulon and its contribution to cell envelope stress responses. *Mol. Microbiol.*, **67**, 830–848.
25. Helmann, J.D. (2006) Deciphering a complex genetic regulatory network: the *Bacillus subtilis* sigmaW protein and intrinsic resistance to antimicrobial compounds. *Sci. Prog.*, **89**, 243–266.
26. Cao, M. and Helmann, J.D. (2004) The *Bacillus subtilis* extracytoplasmic-function σ^X factor regulates modification of the cell envelope and resistance to cationic antimicrobial peptides. *J. Bacteriol.*, **186**, 1136–1146.
27. Guariglia-Oropeza, V. and Helmann, J.D. (2011) *Bacillus subtilis* σ^V confers lysozyme resistance by activation of two cell wall modification pathways, peptidoglycan O-acetylation and D-alanylation of teichoic acids. *J. Bacteriol.*, **193**, 6223–6232.
28. Ho, T.D., Hastie, J.L., Intile, P.J. and Ellermeier, C.D. (2011) The *Bacillus subtilis* extracytoplasmic function sigma factor σ^V is induced by lysozyme and provides resistance to lysozyme. *J. Bacteriol.*, **193**, 6215–6222.
29. Cao, M., Salzberg, L., Tsai, C.S., Mascher, T., Bonilla, C., Wang, T., Ye, R.W., Marquez-Magana, L. and Helmann, J.D. (2003) Regulation of the *Bacillus subtilis* extracytoplasmic function protein σ^Y and its target promoters. *J. Bacteriol.*, **185**, 4883–4890.
30. Mendez, R., Gutierrez, A., Reyes, J. and Marquez-Magana, L. (2012) The extracytoplasmic function sigma factor SigY is important for efficient maintenance of the Spbeta prophage that encodes subblancin in *Bacillus subtilis*. *DNA Cell Biol.*, **31**, 946–955.
31. Luo, Y., Asai, K., Sadaie, Y. and Helmann, J.D. (2010) Transcriptomic and phenotypic characterization of a *Bacillus subtilis* strain without extracytoplasmic function sigma factors. *J. Bacteriol.*, **192**, 5736–5745.
32. Mascher, T., Hachmann, A.B. and Helmann, J.D. (2007) Regulatory overlap and functional redundancy among *Bacillus subtilis* extracytoplasmic function sigma factors. *J. Bacteriol.*, **189**, 6919–6927.
33. Hashimoto, M., Seki, T., Matsuoka, S., Hara, H., Asai, K., Sadaie, Y. and Matsumoto, K. (2013) Induction of extracytoplasmic function sigma factors in *Bacillus subtilis* cells with defects in lipoteichoic acid synthesis. *Microbiology*, **159**, 23–35.
34. Ryu, H.B., Shin, I., Yim, H.S. and Kang, S.O. (2006) YlaC is an extracytoplasmic function (ECF) sigma factor contributing to hydrogen peroxide resistance in *Bacillus subtilis*. *J. Microbiol.*, **44**, 206–216.
35. Cao, M., Kobel, P.A., Morshedi, M.M., Wu, M.F., Paddon, C. and Helmann, J.D. (2002) Defining the *Bacillus subtilis* σ^W regulon: a comparative analysis of promoter consensus search, run-off transcription/microarray analysis (ROMA), and transcriptional profiling approaches. *J. Mol. Biol.*, **316**, 443–457.
36. Eiamphungporn, W. and Helmann, J.D. (2009) Extracytoplasmic function sigma factors regulate expression of the *Bacillus subtilis* *yabE* gene via a cis-acting antisense RNA. *J. Bacteriol.*, **191**, 1101–1105.
37. Luo, Y. and Helmann, J.D. (2009) Extracytoplasmic function sigma factors with overlapping promoter specificity regulate subblancin production in *Bacillus subtilis*. *J. Bacteriol.*, **191**, 4951–4958.
38. Pietiainen, M., Gardemeister, M., Mecklin, M., Leskela, S., Sarvas, M. and Kontinen, V.P. (2005) Cationic antimicrobial peptides elicit a complex stress response in *Bacillus subtilis* that involves ECF-type sigma factors and two-component signal transduction systems. *Microbiology*, **151**, 1577–1592.
39. Huang, X., Fredrick, K.L. and Helmann, J.D. (1998) Promoter recognition by *Bacillus subtilis* sigmaW: autoregulation and partial overlap with the σ^X regulon. *J. Bacteriol.*, **180**, 3765–3770.
40. Cao, M. and Helmann, J.D. (2002) Regulation of the *Bacillus subtilis* *berC* bacitracin resistance gene by two extracytoplasmic function sigma factors. *J. Bacteriol.*, **184**, 6123–6129.
41. Nicolas, P., Mader, U., Dervyn, E., Rochat, T., Leduc, A., Pigeonneau, N., Bidnenko, E., Marchadier, E., Hoebeke, M., Aymerich, S. et al. (2012) Condition-dependent transcriptome reveals high-level regulatory architecture in *Bacillus subtilis*. *Science*, **335**, 1103–1106.
42. Qiu, J. and Helmann, J.D. (2001) The -10 region is a key promoter specificity determinant for the *Bacillus subtilis* extracytoplasmic-function sigma factors σ^X and σ^W . *J. Bacteriol.*, **183**, 1921–1927.
43. Sambrook, J. and Russell, D.W. (2001) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory, NY.
44. Cutting, S.M. and VanderHorn, P.B. (1990) Genetic analysis. In: Harwood, C.R. and Cutting, S.M. (eds). *Molecular Biological Methods for Bacillus*. John Wiley and Sons, Ltd., Chichester, pp. 27–74.
45. Slack, F.J., Mueller, J.P. and Sonenshein, A.L. (1993) Mutations that relieve nutritional repression of the *Bacillus subtilis* dipeptide permease operon. *J. Bacteriol.*, **175**, 4605–4614.
46. Zuber, P. and Losick, R. (1987) Role of AbrB in Spo0A- and Spo0B-dependent utilization of a sporulation promoter in *Bacillus subtilis*. *J. Bacteriol.*, **169**, 2223–2230.

47. Miller, J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory, NY.
48. Bsat, N., Herbig, A., Casillas-Martinez, L., Setlow, P. and Helmann, J.D. (1998) *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. *Mol. Microbiol.*, **29**, 189–198.
49. Asai, K., Ishiwata, K., Matsuzaki, K. and Sadaie, Y. (2008) A viable *Bacillus subtilis* strain without functional extracytoplasmic function sigma genes. *J. Bacteriol.*, **190**, 2633–2636.
50. Zhao, H., Sun, Y., Peters, J.M., Gross, C.A., Garner, E.C. and Helmann, J.D. (2016) Depletion of undecaprenyl pyrophosphate phosphatases disrupts cell envelope biogenesis in *Bacillus subtilis*. *J. Bacteriol.*, **198**, 2925–2935.
51. Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K. and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
52. Qi, Y. and Hulett, F.M. (1998) PhoP-P and RNA polymerase sigmaA holoenzyme are sufficient for transcription of Pho regulon promoters in *Bacillus subtilis*: PhoP-P activator sites within the coding region stimulate transcription in vitro. *Mol. Microbiol.*, **28**, 1187–1197.
53. Donnelly, M.I., Zhou, M., Millard, C.S., Clancy, S., Stols, L., Eschenfeldt, W.H., Collart, F.R. and Joachimiak, A. (2006) An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. *Protein Expr. Purif.*, **47**, 446–454.
54. Gaballa, A., MacLellan, S. and Helmann, J.D. (2012) Transcription activation by the siderophore sensor Btr is mediated by ligand-dependent stimulation of promoter clearance. *Nucleic Acids Res.*, **40**, 3585–3595.
55. Martinez-Guerrero, C.E., Ciria, R., Abreu-Goodger, C., Moreno-Hagelsieb, G. and Merino, E. (2008) GeConT 2: gene context analysis for orthologous proteins, conserved domains and metabolic pathways. *Nucleic Acids Res.*, **36**, W176–W180.
56. Zellmeier, S., Hofmann, C., Thomas, S., Wiegert, T. and Schumann, W. (2005) Identification of sigma(V)-dependent genes of *Bacillus subtilis*. *FEMS Microbiol. Lett.*, **253**, 221–229.
57. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
58. Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O. and Darst, S.A. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme-DNA complex. *Science*, **296**, 1285–1290.
59. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
60. Asai, K., Yamaguchi, H., Kang, C.M., Yoshida, K., Fujita, Y. and Sadaie, Y. (2003) DNA microarray analysis of *Bacillus subtilis* sigma factors of extracytoplasmic function family. *FEMS Microbiol. Lett.*, **220**, 155–160.
61. Fimlaid, K.A. and Shen, A. (2015) Diverse mechanisms regulate sporulation sigma factor activity in the Firmicutes. *Curr. Opin. Microbiol.*, **24**, 88–95.
62. Chauhan, R., Ravi, J., Datta, P., Chen, T., Schnappinger, D., Bassler, K.E., Balazsi, G. and Gennaro, M.L. (2016) Reconstruction and topological characterization of the sigma factor regulatory network of *Mycobacterium tuberculosis*. *Nat. Commun.*, **7**, 11062.
63. Zweers, J.C., Nicolas, P., Wiegert, T., van Dijk, J.M. and Denham, E.L. (2012) Definition of the sigma(W) regulon of *Bacillus subtilis* in the absence of stress. *PLoS One*, **7**, e48471.
64. Perego, M., Glaser, P., Minutello, A., Strauch, M.A., Leopold, K. and Fischer, W. (1995) Incorporation of D-alanine into lipoteichoic acid and wall teichoic acid in *Bacillus subtilis*. Identification of genes and regulation. *J. Biol. Chem.*, **270**, 15598–15606.
65. Huang, X. and Helmann, J.D. (1998) Identification of target promoters for the *Bacillus subtilis* σ^X factor using a consensus-directed search. *J. Mol. Biol.*, **279**, 165–173.
66. Crothers, D.M., Haran, T.E. and Nadeau, J.G. (1990) Intrinsically bent DNA. *J. Biol. Chem.*, **265**, 7093–7096.
67. Zuo, Y. and Steitz, T.A. (2015) Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell*, **58**, 534–540.
68. Guo, Y. and Gralla, J.D. (1998) Promoter opening via a DNA fork junction binding activity. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 11655–11660.
69. Tsujikawa, L., Tsodikov, O.V. and deHaseth, P.L. (2002) Interaction of RNA polymerase with forked DNA: evidence for two kinetically significant intermediates on the pathway to the final complex. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 3493–3498.
70. Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K. and Gourse, R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science*, **262**, 1407–1413.
71. Travers, A.A. (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *J. Bacteriol.*, **141**, 973–976.
72. Guzina, J. and Djordjevic, M. (2016) Promoter recognition by extracytoplasmic function sigma factors: analyzing DNA and protein interaction motifs. *J. Bacteriol.*, **198**, 1927–1938.
73. Guzina, J. and Djordjevic, M. (2017) Mix-and-matching as a promoter recognition mechanism by ECF sigma factors. *BMC Evol. Biol.*, **17**, 12.
74. Rhodius, V.A., Segall-Shapiro, T.H., Sharon, B.D., Ghodasara, A., Orlova, E., Tabakh, H., Burkhardt, D.H., Clancy, K., Peterson, T.C., Gross, C.A. et al. (2013) Design of orthogonal genetic switches based on a crosstalk map of sigmas, anti-sigmas, and promoters. *Mol. Syst. Biol.*, **9**, 702.
75. Auble, D.T., Allen, T.L. and deHaseth, P.L. (1986) Promoter recognition by *Escherichia coli* RNA polymerase. Effects of substitutions in the spacer DNA separating the -10 and -35 regions. *J. Biol. Chem.*, **261**, 11202–11206.
76. Ayers, D.G., Auble, D.T. and deHaseth, P.L. (1989) Promoter recognition by *Escherichia coli* RNA polymerase. Role of the spacer DNA in functional complex formation. *J. Mol. Biol.*, **207**, 749–756.
77. Helmann, J.D. and deHaseth, P.L. (1999) Protein-nucleic acid interactions during open complex formation investigated by systematic alteration of the protein and DNA binding partners. *Biochemistry*, **38**, 5959–5967.
78. Warne, S.E. and deHaseth, P.L. (1993) Promoter recognition by *Escherichia coli* RNA polymerase. Effects of single base pair deletions and insertions in the spacer DNA separating the -10 and -35 regions are dependent on spacer DNA sequence. *Biochemistry*, **32**, 6134–6140.
79. Singh, S.S., Typas, A., Hengge, R. and Grainger, D.C. (2011) *Escherichia coli* σ^{70} senses sequence and conformation of the promoter spacer region. *Nucleic Acids Res.*, **39**, 5109–5118.
80. Fenton, M.S., Lee, S.J. and Gralla, J.D. (2000) *Escherichia coli* promoter opening and -10 recognition: mutational analysis of σ^{70} . *EMBO J.*, **19**, 1130–1137.
81. Yuzenkova, Y., Tadigotla, V.R., Severinov, K. and Zenkin, N. (2011) A new basal promoter element recognized by RNA polymerase core enzyme. *EMBO J.*, **30**, 3766–3775.
82. Hook-Barnard, I.G. and Hinton, D.M. (2009) The promoter spacer influences transcription initiation via σ^{70} region 1.1 of *Escherichia coli* RNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 737–742.
83. Lozinski, T., Adrych-Rozek, K., Markiewicz, W.T. and Wierchowski, K. (1991) Effect of DNA bending in various regions of a consensus-like *Escherichia coli* promoter on its strength in vivo and structure of the open complex in vitro. *Nucleic Acids Res.*, **19**, 2947–2953.
84. Lozinski, T. and Wierchowski, K.L. (1996) Effect of reversed orientation and length of An.Tn DNA bending sequences in the -35 and spacer domains of a consensus-like *Escherichia coli* promoter on its strength in vivo and gross structure of the open complex in vitro. *Acta Biochim. Pol.*, **43**, 265–279.
85. Haran, T.E. and Mohanty, U. (2009) The unique structure of A-tracts and intrinsic DNA bending. *Q. Rev. Biophys.*, **42**, 41–81.
86. Philips, S.J., Canalizo-Hernandez, M., Yildirim, I., Schatz, G.C., Mondragon, A. and O'Halloran, T.V. (2015) Allosteric transcriptional regulation via changes in the overall topology of the core promoter. *Science*, **349**, 877–881.
87. Martell, D.J., Joshi, C.P., Gaballa, A., Santiago, A.G., Chen, T.Y., Jung, W., Helmann, J.D. and Chen, P. (2015) Metalloregulator CueR biases RNA polymerase's kinetic sampling of dead-end or open complex to repress or activate transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 13467–13472.