Published online 11 March 2011

Nucleic Acids Research, 2011, Vol. 39, No. 12 ***5109–5118***
doi:10.1093/nar/gkr080

Escherichia coli σ^{70} senses sequence and conformation of the promoter spacer region

Shivani S. Singh[1], Athanasios Typas[2], Regine Hengge[2] and David C. Grainger[1,*]

[1]School of Life Sciences, The University of Warwick, Coventry CV4 7AL, UK and [2]Institut für Biologie, Mikrobiologie, Freie Universität Berlin, Königin-Luise-Strasse, 12-16, 14195 Berlin, Germany

Received December 21, 2010; Revised January 31, 2011; Accepted February 2, 2011

ABSTRACT

In bacteria, promoter identification by RNA polymerase is mediated by a dissociable σ factor. The housekeeping σ^{70} factor of *Escherichia coli* recognizes two well characterized DNA sequence elements, known as the '−10' and '−35' hexamers. These elements are separated by 'spacer' DNA, the sequence of which is generally considered unimportant. Here, we use a combination of bioinformatics, genetics and biochemistry to show that σ^{70} can sense the sequence and conformation of the promoter spacer region. Our data illustrate how alterations in spacer region sequence can increase promoter activity. This stimulatory effect requires σ^{70} side chain R451, which is located in close proximity to the non-template strand at promoter position −18. Conversely, R451 is not required to mediate transcriptional stimulation by improvement of the −10 element. Mutation of σ^{70} residue R451, which is highly conserved, results in reduced growth rate, consistent with a central role in promoter recognition.

INTRODUCTION

RNA polymerase requires specific DNA sequences known as promoters in order to recognize DNA and initiate the transcription of a gene. In bacteria, recognition of promoters is mediated by a dissociable subunit of RNA polymerase known as the σ factor (1). Most bacteria encode multiple σ factors, with different DNA binding specificities, and σ factor switching therefore represents a simple mechanism via which RNA polymerase can be directed to different sets of genes (2). The major σ factor in *Escherichia coli* is σ^{70}, which recognizes two DNA sequences known as the −10 and −35 hexamers (3). The −10 hexamer (5′-TATAAT-3′) is located ∼7-bp upstream of the transcription start site and becomes single stranded during transcription initiation. The −35 sequence (5′-TTGACA-3′) is usually located 17-bp upstream of the −10 element and is not unwound during open-complex formation (4). A subclass of 'extended −10' promoters have a 5′–TG-3′ motif at promoter positions −14 and −15, 1 bp upstream of the −10 hexamer. This serves to further stabilize RNAP–DNA interactions (5). The major promoter elements were originally identified on the basis of similarity between small numbers of very efficient promoters (6–9). Recent large-scale cataloguing of promoters in *E. coli* now permits hundreds of promoter sequences, many of which are sub-optimal, to be aligned. The overall conservation of different promoter elements can then be ascertained. Figure 1A shows a DNA sequence logo generated from the alignment of 554 *E. coli* promoters produced by Mitchell *et al.* (10). The −10 sequence is clearly the best conserved promoter element. Comparatively, the −35 and extended −10 elements are poorly conserved. Indeed, some sequences in the 17 bp spacer region, generally considered to be unimportant, are better conserved than parts of the −35 and extended −10 elements.

Interactions between the promoter elements described above and RNA polymerase are mediated by specific σ factor determinants. Primary σ factors, such as σ^{70}, consist of two or more conserved domains connected by flexible linkers. Sub-regions of domains 2, 3 and 4 mediate sequence specific RNA polymerase–DNA contacts. Region 2.4 of σ^{70} domain 2 contacts the −10 element, with side chains Q437 and T440 interacting with the DNA at promoter position −12 (11,12). The 5′-TG-3′ motif of extended −10 promoters is recognized by side chain E458 in σ^{70} domain 3 (5). Contact with the −35 element is mediated by domain 4 with amino acid side chains R584, E585 and Q589 making the critical interactions (13). Contacts between RNA polymerase and the spacer DNA were detected biochemically over 30 years

*To whom correspondence should be addressed. Tel: +44 (0)24 765 22509; Fax: +44 (0)24 765 23701; Email: d.c.grainger@warwick.ac.uk
Present address:
Athanasios Typas, Department of Microbiology and Immunology, University of California at San Francisco, 600 16th Street, San Francisco, CA 94158, USA.

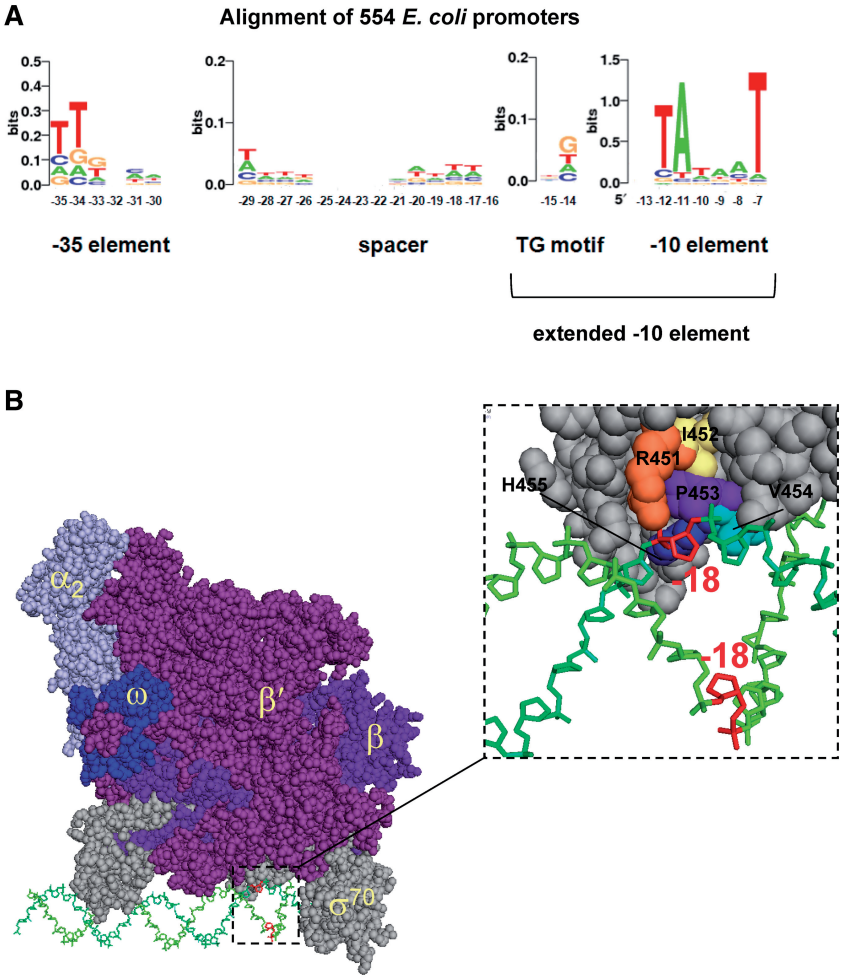

Figure 1. Promoter sequence, organization and recognition. (**A**) The panel shows a DNA sequence logo generated from the alignment of 554 *E. coli* promoters produced by Mitchell *et al.* (10). Different promoter elements are labelled. (**B**) A structural model of the RNA polymerase holoenzyme–DNA complex is shown (12). With the exception of the two α-subunits, each RNA polymerase component protein is shown in a different colour and is labelled. The DNA is shown in green with the base pair at position −18 highlighted in red. The expansion shows the close proximity of the loop between σ^{70} domains 2 and 3 and the promoter non-template strand. Residues in σ^{70} mutated during the course of this work are highlighted and labelled in the expansion.

ago (14–16). However, the nature or role of these contacts has not been probed further.

It is well established that the length of the 'spacer' DNA between the −10 and −35 elements is critical (17). Recent work has suggested that the sequence of the spacer may also be important (18,19). Structural modelling of the RNA polymerase holoenzyme–DNA complex places the linker between σ^{70} domains 2 and 3 within 2 Å of the non-template strand at promoter position −18, just upstream of the extended −10 element (12) (Figure 1B). Interestingly, as the data in Figure 1A show, the DNA sequence immediately upstream of the extended −10 element is not random. For example, T is the preferred base at both positions −17 and −18. In this work, we have investigated the role of the spacer region, and the linker between σ^{70} domains 2 and 3, in controlling promoter activity. We show that altering the base sequence at promoter position −18 modulates transcription initiation at many promoters. Moreover, mutational analysis reveals that σ^{70} side chain R451, located in the linker between σ^{70} domains 2 and 3, is required to respond to changes in DNA sequence at promoter position −18. Substitution of side chain R451 with alanine results in decreased growth rate consistent with R451 playing an important role.

MATERIALS AND METHODS

Strains plasmids and oligonucleotides

Bacterial strains and plasmids are listed in Table 1. Standard techniques for recombinant DNA manipulations were used throughout. Table 2 lists primers used to amplify sections of the *cbpA* regulatory region in such a way that it was flanked by EcoRI and HindIII restriction sites. After digestion, fragments carrying *cbpA* regulatory were cloned into pSR, sequenced and then sub-cloned into pRW50. The exception to this was the screen for 'up' mutations in the spacer region where fragments were cloned directly into pRW50. We have

Table 1. Bacterial Strains and plasmids

Name	Description	Source
Bacterial strains		
JCB387	Δ*nir* Δ*lac*	(20)
MC4100	F$^-$ *araD*139 Δ(*argF-lac*)U169 *rpsL*150 *relA*1 *deoC*1 *ptsF*25 *rbsR flbB*5301	(21)
MC4100*rpoS*::kan	MC4100*rpoS*::kan	(22)
T7 express (Invitrogen)	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10–*Tet^S*)2 [dcm] R(zgb-210::Tn10–*Tet^S*) endA1 Δ(mcrC-mrr)114::IS10*	
Plasmids		
pSR	pBR322-derived plasmid containing an EcoRI–HindIII fragment upstream of the λ*oop* transcription terminator	(23)
pRW50	Low-copy number broad-host-range *lac* fusion vector for cloning promoters on EcoRI–HindIII fragments: contains the RK2 origin of replication and encodes Tc^R	(24)
pRW224	A derivative of pRW50	(25)
pVRσ	pBR322 derivative encoding *rpoD* and mutant derivatives	(26)
pET21b (Novagen)	T7 Expression vector containing 6xHis tag	

Table 2. Oligonucleotides

Name	Sequence	Source
Primers to introduce random single base substitutions into the *cbpA* regulatory region spacer DNA element		
-22N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAA**N**TTTGAGGATTACCC 3′	This work
-21N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAAT**N**TTGAGGATTACCCT 3′	This work
-20N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATT**N**TGAGGATTACCCTA 3′	This work
-19N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTT**N**GAGGATTACCCTAC 3′	This work
-18N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTT**N**AGGATTACCCTACA 3′	This work
-17N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTG**N**GGATTACCCTACAC 3′	This work
-16N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTGA**N**GATTACCCTACACT 3′	This work
-15N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTGAG**N**ATTACCCTACACTT 3′	This work
-14N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTGAGG**N**TTACCCTACACTTA 3′	This work
-13N *cbpA*Δ45	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTGAGGA**N**TACCCTACACTTAT 3′	This work
Primers used for site-directed mutagenesis of the *cbpA* regulatory region		
*cbpA*Δ45 -18C	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTT**C**AGGATTACCC 3′	This work
*cbpA*Δ45-9A -10T	5′GGCTGC***GAATTC***ATATTCTGTGTTGGCATATGAAATTTTGAGGATTA**TA**CTACACTTATAGG 3′	This work
Primers used to amplify *cbpA* promoter inserts cloned in plasmid pSR		
pSR up	5′GCATTTATCAGGGTTATTGTCTC 3′	This work
pSR down	5′ CATCACCGAAACGCGCGAGG 3′	This work
Primers used to introduce alanine codons into *rpoD*		
HindIII oligo	5′GGGG***AAGCTT***TTAATCGTCCAGGAAGCTACGCAGCACTTCAGAACGGCTCGGGTGACGC AGTTTGCGCAGCGCCTTCGCTTC 3′	This work
R451A	5′CCTGGT***GGATCC***GTCAGGCGATCACCCGCTCTATCGCGGATCAGGCGCGCACCATC**GCT** ATTCCGGTGCATATGATTGAGACC 3′	This work
I452A	5′CCTGGT***GGATCC***GTCAGGCGATCACCCGCTCTATCGCGGATCAGGCGCGCACCATCCGT**GCT** CCGGTGCATATGATTGAGACC 3′	This work
P453A	5′CCTGGT***GGATCC***GTCAGGCGATCACCCGCTCTATCGCGGATCAGGCGCGCACCATCCGTATT**GCG** GTGCATATGATTGAGACC 3′	This work
V454A	5′CCTGGT***GGATCC***GTCAGGCGATCACCCGCTCTATCGCGGATCAGGCGCGCACCATCCGTATTCCG**GCG** CATATGATTGAGACC 3′	This work
H455A	5′CCTGGT***GGATCC***GTCAGGCGATCACCCGCTCTATCGCGGATCAGGCGCGCACCATCCGT ATTCCGGTG**GCT**ATGATTGAGACC 3′	This work
Primer used to amplify the *rpoD* RA451 allele for cloning into pET21b		
RpoD pET21b Up	5′AGCTCA***GCTAGC***GAGCAAAACCCGCAGTCACAGCTGAAAC 3′	This work
Primers used to amplify the LEE1 promoter and derivatives		
LEE up	5′GAATTCTTGACATTTAATGATAATGTATTTTACACATTAGAAAAAAG 3′	This work
LEE up -18A	5′GAATTCTTGACATTTAATGATAAT**A**TATTTTACACATTAGAAAAAAG 3′	This work
LEE up -18C	5′GAATTCTTGACATTTAATGATAAT**C**TATTTTACACATTAGAAAAAAG 3′	This work
LEE up -18T	5′GAATTCTTGACATTTAATGATAAT**T**TATTTTACACATTAGAAAAAAG 3′	This work
LEE down	5′AAGCTTATTCTCTTTTTTCTAATGTGTAAAA 3′	This work

Restriction sites used for cloning are shown in bold typeface and are italicized. Mutations introduced during the PCR are underlined.

numbered mutations in the *cbpA* regulatory region with respect to the $E\sigma^{70}$ transcription start point (+1) and with upstream and downstream locations denoted by '−' and '+' prefixes, respectively.

Proteins

Core *E. coli* RNA polymerase was purchased from Epicenter (Madison). To overproduce the σ^{70} RA451 protein the *rpoD* RA451 allele was amplified by PCR from plasmid pVRσ and cloned into pET21b. Preparations of all σfactors and derivatives were made by overexpression of the cloned *rpoD* and *rpoS* alelles in T7 express cells (NEB). Inclusion bodies were then solublized in 6 M Guanidine HCl, before being dialysed into buffer containing 20 mM Tris, 100 mM NaCl and 10% glycerol. Proteins were bound to a HiTrap QFF anion exchange column (Pharmacia) and eluted with a linear gradient to 1 M NaCl. RNA polymerase holoenzyme was reconstituted by incubating core RNA polymerase with equimolar amounts of σ^{70} and σ^{38} at room temperature for 20 min.

$KMnO_4$ footprinting

Purified AatII–HindIII DNAfragments were derived from maxi preparations (using a Qiagen maxiprep kit) of plasmid pSR carrying the *cbpA* regulatory region or derivatives. Fragments were labelled at the HindIII-end using [γ-^{32}P]-ATP and polynucleotide kinase. Footprints were performed at 37°C as in our previous work (27). DNA fragments were used at a final concentration of 10–40 nM in buffer containing 20 mM Tris pH 7, 10 mM $MgCl_2$, 100 mM EDTA and 120 mM KCl. Footprints were analysed on a 6% DNA sequencing gel (Molecular Dynamics). The results of all footprints and EMSA experiments were visualized by exposing the dried gel against a Fuji phosphor screen and analysed using a phosphorimager and Quantity One software.

In vitro transcription assays

The *in vitro* transcription experiments were performed as described previously (28) using the system of Kolb *et al.* (23). A Qiagen maxiprep kit was used to purify supercoiled pSR plasmid carrying the *cbpA* regulatory region or derivatives. This template (16 μg/ml) was incubated in buffer containing 20 mM Tris pH 7.9, 5 mM $MgCl_2$, 500 μM DTT, 50 mM KCl, 100 μg/ml BSA, 200 μM ATP, 200 μM GTP, 200 μM CTP, 10 μM UTP with 5 μCi [(α-^{32}P]-UTP. The reaction was started by adding purified *E. coli* $E\sigma^{70}$ and/or $E\sigma^{38}$. Labelled RNA products were analysed on a denaturing polyacrylamide gel.

β-Galactosidase assays

DNA fragments containing the *cbpA* regulatory region or the locus for enterocyte effacement 1 (LEE1) promoter were cloned into pRW50 or pRW224 respectively to generate promoter::*lacZ* fusions. β-Galactosidase levels in cells carrying these recombinants were measured by the Miller (29) method. Activities are the average of three or more independent experiments. Cells were grown aerobically in LB media as described.

DNA bending assays

DNA fragments generated by PCR were separated by electrophoresis on a 7.5% polyacrylamide non-denaturing gel. Electrophoresis was performed at 4°C in TBE buffer. After electrophoresis, the gels were stained with ethidium bromide and DNA was visualized by UV illumination. The DNA fragments for this analysis were generated by PCR. Thus, the *cbpA* regulatory region cloned in plasmid pSR was amplified using pSR up and pSR down oligos (Table 2). The LEE1 promoter and derivatives were generated using the oligonucleotides shown in Table 2. Note that the upstream and downstream LEE oligos overlap and can thus be used to generate a double-stranded DNA product without template DNA.

Modelling of DNA fragments *in silico*

Changes in DNA bending were modelled computationally using the 'model.it' web server (http://hydra.icgeb.trieste.it/dna/model_it.html) using the default parameters (30). Predicted DNA structures were downloaded in pdb format and PyMOL was used to prepare figures.

RESULTS

An *E. coli* gene regulatory region with overlapping promoters

The aim of this study was to investigate the role of the RNA polymerase σ^{70} subunit in sensing the sequence of the promoter spacer. To facilitate this, we chose to work with a regulatory DNA region containing overlapping promoters, with a shared spacer region, but different σ factor specificity. Our logic was that DNA sequences in the spacer region that make selective interactions with σ^{70} should affect only one of the two promoters. Conversely, sequences that non-selectively stimulate transcription should affect both promoters. Thus, the *E. coli cbpA* regulatory region contains overlapping promoters for RNA polymerase associated with σ^{70} ($E\sigma^{70}$) or σ^{38} ($E\sigma^{38}$) (Figure 2A) (31,32). Note that, throughout this work, we use the $E\sigma^{70}$ transcription start site as the point of reference for numbering mutations in the *cbpA* regulatory region.

To confirm that the two *cbpA* promoters were truly specific for their cognate σ factor we performed $KMnO_4$ footprinting, which detects open complex formation by RNA polymerase. As expected, $E\sigma^{70}$ and $E\sigma^{38}$ produce different patterns of DNA opening that are offset (Figure 2B). We also found that $E\sigma^{70}$ had to be added at higher concentrations than $E\sigma^{38}$, suggesting that the $E\sigma^{70}$ dependent promoter is less efficient. To further confirm the σ factor specificity of the two promoters, we utilized an *in vitro* transcription assay. Thus, an EcoRI–HindIII fragment carrying the *cbpA* regulatory region (illustrated in Figure 1A) was cloned into plasmid pSR. This places the two promoters upstream of the factor-independent λ*oop* transcription terminator.

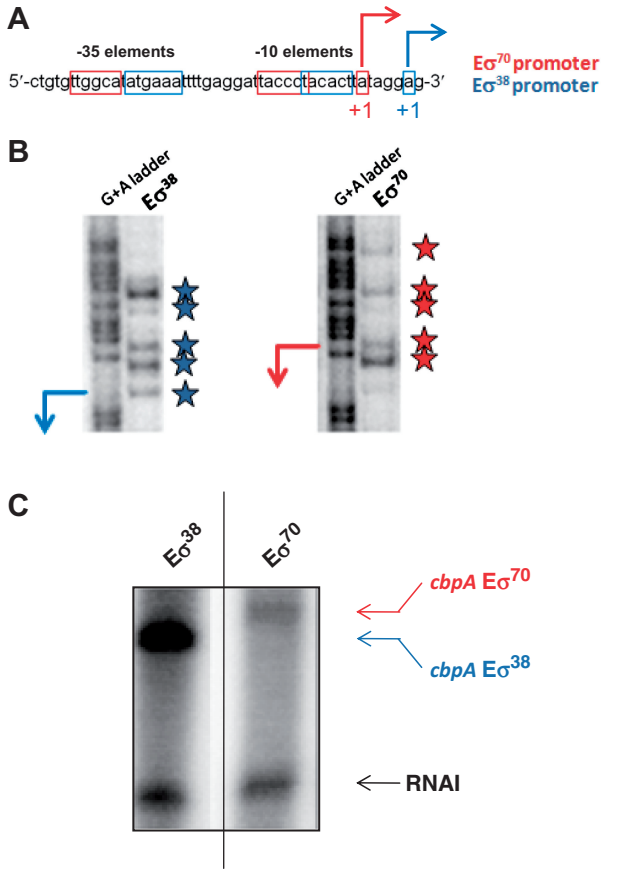

Figure 2. Overlapping promoters in the *cbpA* regulatory region. (**A**) Sequence of the *cbpA* regulatory region. Promoters for $E\sigma^{70}$ (red) and $E\sigma^{38}$ (blue) are highlighted. The transcription start sites are labelled with arrows and designated as '+1'. Note that, throughout this work, we have numbered all mutations in the *cbpA* regulatory region with respect to the $E\sigma^{70}$ '+1'. (**B**) Open complex formation by $E\sigma^{38}$ and $E\sigma^{70}$. The panel shows the results of a $KMnO_4$ footprinting experiment designed to detect open complex formation by RNA polymerase at the *cbpA* regulatory region. The positions of DNA opening by $E\sigma^{38}$ and $E\sigma^{70}$ are highlighted in blue and red, respectively and the transcription start sites are shown by arrows. $E\sigma^{70}$ was added in 3-fold excess of $E\sigma^{38}$. (**C**) *In vitro* transcription by $E\sigma^{38}$ and $E\sigma^{70}$. The results of an *in vitro* transcription assay are shown. The lower band corresponds to the RNAI transcript, which acts as an internal control. The upper bands are transcripts that initiate from the *cbpA* regulatory region and are due to transcription by either $E\sigma^{38}$ (blue) or $E\sigma^{70}$ (red). $E\sigma^{70}$ and $E\sigma^{38}$ were added in equal amounts.

Transcription *in vitro* with either $E\sigma^{38}$ or $E\sigma^{70}$ produces an RNA product, which can be detected after electrophoresis. As expected, while $E\sigma^{70}$ functions less efficiently than $E\sigma^{38}$, it produces a longer transcript (Figure 2C). Note that, the plasmid pSR replication origin encodes the 108 base RNAI transcript that acts as an internal control.

Isolation and analysis of spacer region mutations that stimulate transcription

To investigate the role of the promoter spacer, we created a library of DNA fragments carrying the *cbpA* regulatory region. The library was prepared so that only random single base mutations, introduced into the portion of the *cbpA* spacer region shared by $E\sigma^{70}$ and $E\sigma^{38}$, were selected. Multiple base changes were not permitted in order to exclude spacer regions with completely altered properties (e.g. very A:T rich spacer sequences). The library of DNA fragments that we prepared was then cloned upstream of *lacZ* in the plasmid pRW50 to create a library of *lacZ* fusion plasmids. This library was used to transform Lac^- JCB387 cells and transformants were plated on MacConkey agar medium. Note that, when fused to *lacZ* in plasmid pRW50, the wild type *cbpA* regulatory region stimulates only low levels of *lacZ* expression. Thus, JCB387 cells transformed with this plasmid construct have a Lac^- phenotype (i.e. appear white) on MacConkey indicator plates. This was also true for the majority of the 216 regulatory region derivatives screened from our library. However, 12 regulatory region derivatives resulted in a Lac^+ (red) phenotype. Colonies with a Lac^+ phenotype were purified by restreaking, levels of *lacZ* expression were quantified using β-galactosidase assays, and finally the sequence of the *cbpA* regulatory region insert was determined. The results of this analysis are displayed in Table 3. The data show that introduction of a T at promoter positions −18 and −17 (with respect to the $E\sigma^{70}$ transcription start) had the biggest stimulatory effect on transcription and were most frequently isolated. The −18A and −17C mutations also stimulated transcription, but to a lesser extent. Interestingly, mutations that optimized the extended −10 element (−14G and −15T) had the smallest stimulatory effect. Note that, our findings are largely consistent with the alignments presented in Figure 1A.

Table 3. Spacer DNA mutations that increase the activity of the *cbpA* regulatory region

Mutation	No. of isolates	β-Galactosidase activity
WT	N/A	95
−14G	1	127
−15T	2	126
−17C	1	158
−17T	4	199
−18T	3	249
−18A	1	221

The table shows β-galactosidase activities obtained from overnight cultures of JCB387 cells carrying different *cbpA::lacZ* fusions in plasmid pRW50. Mutations are numbered with respect to the σ^{70} dependent *cbpA* transcription start site (Figure 1A).

Mutations at position −18 stimulate transcription by $E\sigma^{70}$

Of all the spacer region derivatives that we identified, the substitutions at position −18 had the biggest stimulatory effect on transcription. Thus, we next sought to determine whether the base sequence at position −18 affected both of the overlapping *cbpA* promoters or specifically one of the two promoters. To do this, the EcoRI–HindIII fragments containing the wild type, −18T, or −18A derivatives of the *cbpA* regulatory region were cloned into plasmid pSR. We also used site directed mutagenesis to make a −18C derivative of this construct. Transcription was then measured *in vitro* with either purified $E\sigma^{70}$, or

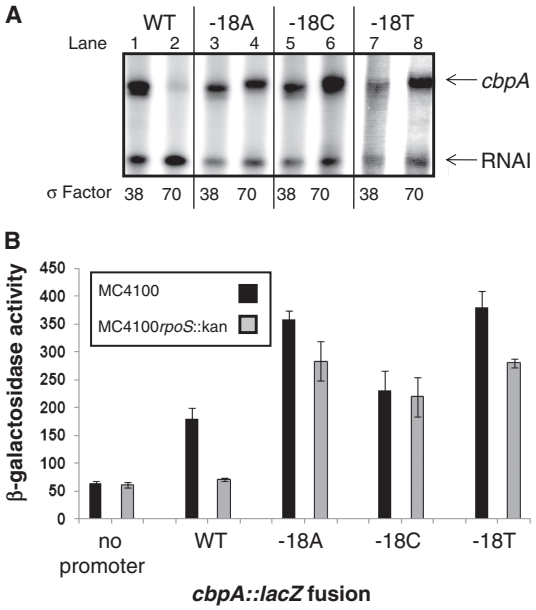

Figure 3. Mutations in the *cbpA* spacer region have differential effects on transcription by $E\sigma^{70}$ and $E\sigma^{38}$. (**A**) Effect of mutations at position −18 on transcription by $E\sigma^{38}$ and $E\sigma^{70}$ *in vitro*. The gel shows transcripts produced *in vitro* by $E\sigma^{38}$ and $E\sigma^{70}$ from the *cbpA* regulatory region and derivatives. (**B**) Activity of *cbpA* regulatory region derivatives *in vivo*. The bar chart shows β-galactosidase expression driven by different *cbpA* regulatory region derivatives, cloned in plasmid pRW50, in MC4100 and the *rpoS*::kan derivative. Values for 'no promoter' were obtained using pRW50 carrying no promoter insert. Assays were done using overnight cultures.

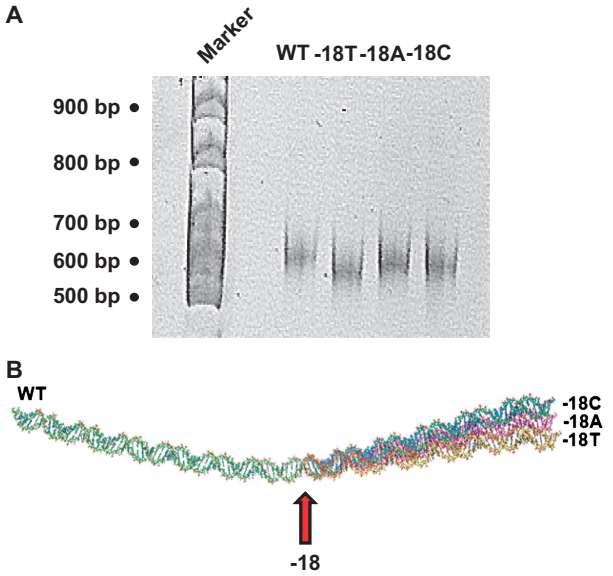

Figure 4. Changes in DNA conformation induced by mutations in the *cbpA* spacer region. (**A**) Derivatives of the *cbpA* regulatory region, with different sequences at position −18, have different mobility on a 7.5% native acrylamide gel. (**B**) Predicted topology of the wild-type (green), −18C (blue), −18A (pink) and −18T (yellow) derivatives of the *cbpA* regulatory region. Position −18 is highlighted by an arrow.

$E\sigma^{38}$, to test the specificity of the mutations. The results (Figure 3A) show that all of the mutations at position −18 stimulate transcription by $E\sigma^{70}$ (compare lanes 2,4,6 and 8) while not stimulating transcription by $E\sigma^{38}$ (compare lanes 1, 3, 5 and 7). To confirm our observations, the various promoter derivatives, cloned into the *lacZ* expression vector pRW50, were used to transform MC4100 or MC4100*rpoS::kan* cells. As expected, β-galactosidase expression driven by the wild-type *cbpA* regulatory region is almost totally dependent on $E\sigma^{38}$. Conversely, all of the regulatory region derivatives with mutations at position −18 showed an increased dependence on $E\sigma^{70}$ (Figure 3B).

To confirm that base changes at position −18 did not create an artificial promoter, we checked the positioning of open complexes formed by RNA polymerase using $KMnO_4$ footprinting. The data show that the −18T substitution stimulates open complex formation (Supplementary Figure S1) and that the open complex is not repositioned (compare Figure 2B and Supplementary Figure S1).

Effect of spacer region mutations on DNA bending

DNA fragments of equivalent length, but with different base sequences, can have different migration properties in native PAGE. These differences are due to changes in intrinsic DNA curvature. Thus, DNA fragments carrying the wild type, −18T, −18A and −18C, derivatives of the *cbpA* regulatory region were subjected to native PAGE analysis. The results confirm that these fragments have different mobility, consistent with altered DNA bending (Figure 4A). The wild-type DNA fragment was least mobile while the −18T derivative was most mobile. The −18A and −18C fragments had an intermediate mobility. Consistent with this, computational modelling of DNA topology for the different sequences predicted changes in conformation of the double helix centred around the −18 position (Figure 4B).

A σ^{70} side chain R451 is required to sense changes in spacer region sequence and conformation

Our data show that altering the sequence of the σ^{70}-dependent *cbpA* promoter at position −18 alters intrinsic promoter activity. This is intriguing since the linker between domains 2 and 3 of σ^{70} is within 2 Å of the DNA at promoter position −18 (12) (Figure 1B). Moreover, Fenton *et al.* (33) previously suggested that σ^{70} side chain R451, located in this linker, contacts the DNA upstream of the −10 element in the spacer region. Thus, we reasoned that changes in promoter conformation at position −18 might affect this σ^{70}–DNA contact. To test this, we introduced alanine substitutions into the *rpoD* gene, encoded by plasmid pVRσ, at the positions highlighted in Figure 1B. The starting pVRσ plasmid and derivatives were then used to transform MC4100*rpoS*::kan cells carrying the −18T version of the *cbpA* regulatory region in plasmid pRW50. Most of the substitutions in *rpoD* had no effect (data not shown) but the RA451 substitution drastically reduced the activity of the −18T promoter [Figure 5A(i)]. We could not test the effect of the RA451 mutation on transcription from the wild type *cbpA* regulatory region; $E\sigma^{70}$ drives hardly any LacZ

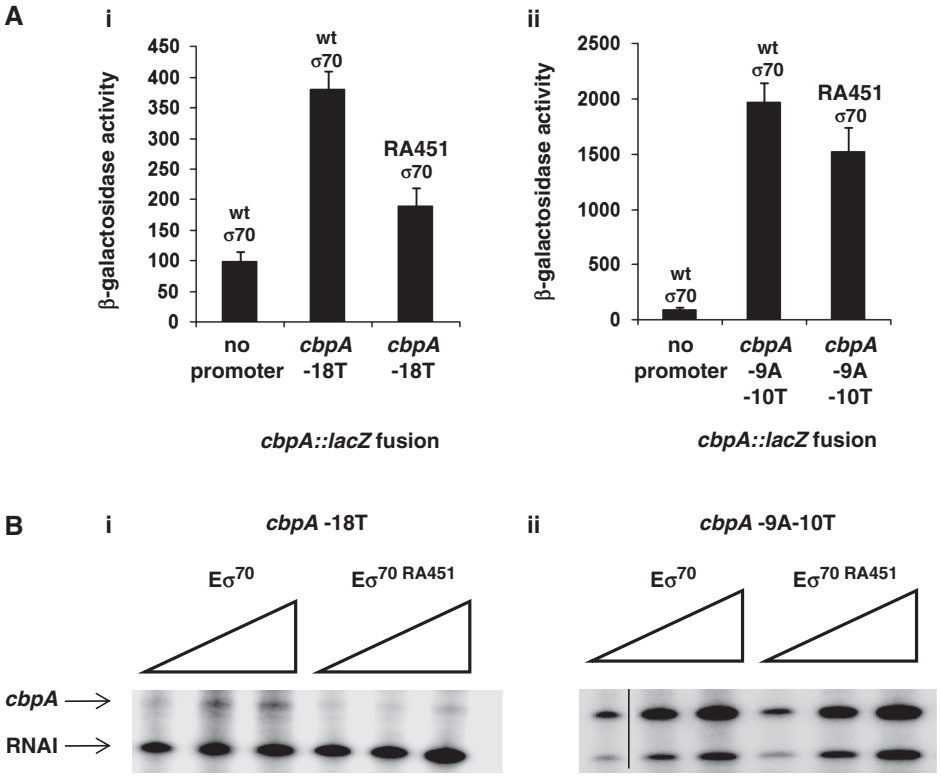

Figure 5. Stimulatory effects of the −18T substitution requires σ^{70} side chain R451. (**A**) The bar chart shows β-galactosidase expression driven by the −18T (i) and −9A−10T (ii) derivatives of the *cbpA* regulatory region, cloned in plasmid pRW50, in MC4100*rpoS*::kan cells carrying either pVRσ or pVRσ RA451. Values for 'no promoter' were obtained using pRW50 carrying no promoter insert. Assays were done using overnight cultures. (**B**) The gels show transcripts produced *in vitro* from the −18T (i) and −9A −10T (ii) derivatives of the *cbpA* regulatory region by purified $E\sigma^{70}$ and $E\sigma^{70}$ RA451. RNA polymerase was added at a concentration of 80, 160 or 240 nM.

expression from this promoter (compare 'no promoter' and 'wt' in Figure 3B). Thus, as a control, we instead used a derivative of the *cbpA* regulatory region with an improved −10 element for $E\sigma^{70}$ (due to the substitutions −9A and −10T). As expected, the data show that RA451 $E\sigma^{70}$ is functional at the −9A−10T derivative of the *cbpA* regulatory region [(Figure 5A(ii)]. To confirm our observations we also measured transcription *in vitro* with purified $E\sigma^{70}$ or the RA451 derivative. The data confirm that RA451 $E\sigma^{70}$ is defective at the −18T promoter but not the −9A−10T promoter (Figure 5B).

Role of promoter position −18 and σ^{70} side chain R451 at the LEE1 promoter

We next sought to determine whether similar phenomena could be reproduced at different promoters. The *E. coli* O157 LEE1 promoter drives expression of genes in the locus for enterocyte effacement (LEE). A maximal level of transcription from the LEE1 promoter requires the GrlA transcriptional activator but substantial basal levels of transcription are observed in *E. coli* K-12 cells that do not encode *grlA*. Recently, Islam *et al.* (25) isolated a LEE1 promoter derivative with increased GrlA-independent activity due to a G to A substitution at promoter position −18 (Figure 6A). Thus, we investigated the possibility that σ^{70} side chain R451 might be important for mediating this effect. The data show that, with the starting LEE1 promoter, RA451 $E\sigma^{70}$ was not defective. In contrast, while the −18A LEE1 promoter had increased activity, this increase was lost with RA451 $E\sigma^{70}$ (Figure 6B). Computational (Figure 6Ci) and native PAGE analysis (Figure 6Cii) confirm that base substitutions at position −18 of the LEE1 promoter alter DNA bending.

A σ^{70} side chain R451 is conserved and required for optimal rates of cell division

Alignment of primary σ factors from different bacteria reveals that R451 is highly conserved, consistent with it playing a key role in RNA polymerase function (13). We found that *E. coli* cells transformed with pVRσ, carrying the RA451 *rpoD* allele, displayed a growth defect (Figure 7). Thus, σ^{70} RA451 must be able to compete with wt σ^{70} (chromosomally encoded) for binding to core RNA polymerase. Once bound to RNA polymerase defects in DNA binding, presumably result in disrupted patterns of gene expression and thus impede growth.

DISCUSSION

Bacterial promoters have been the subject of intense investigation for decades. Recently, several studies have focused on the role of the promoter spacer sequence in controlling transcription. Thus, Liu *et al.* (18) showed that an 8 bp sequence upstream of the −10 element

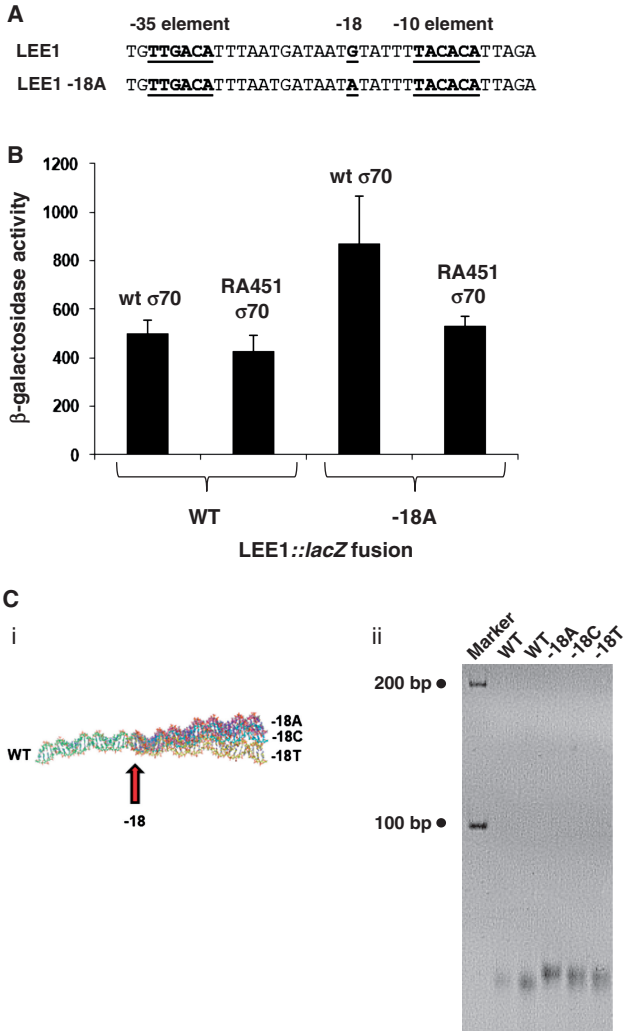

Figure 6. The LEE1 promoter responds to changes in the spacer region at position −18. (**A**) Sequence of the LEE1 promoter and −18A derivative. The −10 hexamer, −35 element and promoter position −18 are highlighted. (**B**) The bar chart shows β-galactosidase expression driven by the wild type and −18A derivatives of the LEE1 regulatory region in JCB387 cells transformed with either pVRσ or pVRσ RA451. Measurements were taken in mid-log phase using the LEE20-203 promoter::*lac* fusion described by Islam *et al.* (25). (**C**) (i) Predicted topology of the wild type (green), −18C (blue), −18A (pink) and −18T (yellow) derivatives of the LEE1 regulatory region. Position −18 is highlighted by an arrow. (ii) Derivatives of the LEE1 regulatory region, with different sequences at position −18, have different mobility on a 7.5% native acrylamide gel.

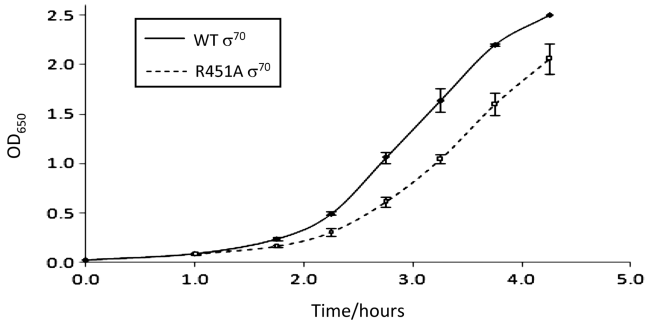

Figure 7. Mutation of σ^{70} side chain R451 induces growth defects. The graph shows growth curves for cultures of JCB387 cells transformed with either pVRσ (solid line) or pVRσ RA451 (dashed line). Cells were grown in LB medium with vigorous aeration at 37°C. The experiment was repeated three times and error bars show the standard deviation of the recorded OD_{650} values.

could stimulate transcription from the *lac* promoter. Similarly, Hook-Barnard and Hinton (19) showed that region 1.1 of σ^{70} was important for mediating the effects of changes in spacer region sequence. Here, we show that there is a clear T>A>C>G preference at promoter position −18, on the basis of promoter sequence alignments (Figure 1A) and biochemical/genetic experiments with three different *E. coli* promoters (Table 3, Figure 6, Supplementary Figure S2). Scrutiny of the literature reveals many other consistent observations. For example, Busby *et al.* (34) showed that a G>A substitution at position −18 stimulated transcription from a modified *gal* P2 promoter. At the *gap*A1 promoter a T>G substitution at position −18 reduces promoter activity 3-fold both *in vivo* and *in vitro* (35). The 8 bp motif identified by Liu *et al.* (18) has a T at position −18 that is essential for activity. Furthermore, Mitchell *et al.* (10) showed a preference for 5′-TTT-3′ trinucleotide sequences to be centred at position −18. Similar observations have also been made at the *rrn*B P1 promoter (36) and the *dps* promoter (Supplementary Figure S2). We note that the contribution of the base sequence at promoter position −18 may be dependent on overall promoter strength, conformation, and could also be influenced by transcriptional regulatory proteins. Thus, while it is unlikely that the base sequence at promoter position −18 plays a central role at all promoters, there are clearly numerous instances where it is important.

Structural and biochemical studies show that the loop between domains 2 and 3 of σ^{70} is correctly positioned to contact position −18 of the promoter non-template strand (12,37). Siebenlist and Gilbert (15) reported 'extensive' RNA polymerase contacts with the DNA in this region of the phage T7 A3 promoter. Similar observations have been made at the *lac*UV5 and the λc17 promoters (14,16). Side chain R451 of σ^{70} is proximal to the DNA backbone at position −18 and Fenton *et al.* (33) showed that an RS451 substitution abolished the promoter DNA binding activity of RNA polymerase *in vitro*. It was concluded that the R451–DNA interaction was non-specific. Our data do not exclude the possibility of a non-specific interaction with the DNA backbone. Indeed, we suggest that changes in DNA conformation centred at position −18 (Figure 4) can modulate the R451–DNA backbone contact, giving rise to pseudo sequence specificity. The extreme deleterious effect of the RS451 substitution constructed by Fenton *et al.* (33) may result from introduction of a polar serine side chain adjacent to the DNA backbone. Consistent with this, σ^{70} subunits with the RA451 substitution are still functional (Figures 5 and 6). In summary, this work describes a mechanism via which RNA polymerase can sense changes in the sequence and conformation of the

promoter spacer region. We note that our data also show how the sequence of the spacer region can play a key role in allowing RNA polymerase to differentiate between overlapping promoters (Figure 3). Intriguingly, the mutations that we have characterized at position −18 are only one-half a helical turn away from sites of DNaseI and singlet oxygen hypersensitivity that occur at −23 and −24 in open complexes (38).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Steve Busby, Nikolay Zenkin, Joseph Wade and Lars Westblade for helpful discussions. We thank Shahid Islam for providing materials and data prior to publication.

FUNDING

Wellcome Trust Research Career Development Fellowship (to D.C.G.); Deutsche Forschungsgemeinschaft (He1556/12-1to R.H.). Funding for open access charge: The Wellcome Trust.

Conflict of interest statement. None declared.

REFERENCES

1. Burgess,R.R., Travers,A.A., Dunn,J.J. and Bautz,E.K. (1969) Factor stimulating transcription by RNA polymerase. *Nature*, **221**, 43–46.
2. Gruber,T.M. and Gross,C.A. (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.*, **57**, 441–466.
3. Paget,M.S. and Helmann,J.D. (2003) The sigma70 family of sigma factors. *Genome Biol.*, **4**, 203.
4. Hsu,L.M. (2002) Promoter clearance and escape in prokaryotes. *Biochim. Biophys. Acta.*, **1577**, 191–207.
5. Barne,K.A., Bown,J.A., Busby,S.J. and Minchin,S.D. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO J.*, **16**, 4034–4040.
6. Pribnow,D. (1975) Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proc. Natl Acad. Sci. USA*, **72**, 784–788.
7. Pribnow,D. (1975) Bacteriophage T7 early promoters: nucleotide sequences of two RNA polymerase binding sites. *J. Mol. Biol.*, **99**, 419–443.
8. Maniatis,T., Ptashne,M., Backman,K., Kield,D., Flashman,S., Jeffrey,A. and Maurer,R. (1975) Recognition sequences of repressor and polymerase in the operators of bacteriophage lambda. *Cell*, **5**, 109–113.
9. Hawley,D.K. and McClure,W.R. (1983) Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res.*, **11**, 2237–2255.
10. 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.
11. Gross,C.A., Chan,C., Dombroski,A., Gruber,T., Sharp,M., Tupy,J. and Young,B. (1998) The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor Symp. Quant. Biol.*, **63**, 141–155.
12. 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*, **17**, 1285–1290.
13. Campbell,E.A., Muzzin,O., Chlenov,M., Sun,J.L., Olson,C.A., Weinman,O., Trester-Zedlitz,M.L. and Darst,S.A. (2002) Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Mol. Cell*, **9**, 527–539.
14. Siebenlist,U., Simpson,R.B. and Gilbert,W. (1980) E. coli RNA polymerase interacts homologously with two different promoters. *Cell*, **20**, 269–281.
15. Siebenlist,U. and Gilbert,W. (1980) Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc. Natl Acad. Sci. USA*, **77**, 122–126.
16. Rosenberg,M. and Court,D. (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.*, **13**, 319–353.
17. Mulligan,M.E., Brosius,J. and McClure,W.R. (1985) Characterization *in vitro* of the effect of spacer length on the activity of *Escherichia coli* RNA polymerase at the TAC promoter. *J. Biol. Chem.*, **260**, 3529–3538.
18. Liu,M., Tolstorukov,M., Zhurkin,V., Garges,S. and Adhya,S. (2002) A mutant spacer sequence between -35 and -10 elements makes the Plac promoter hyperactive and cAMP receptor protein-independent. *Proc. Natl Acad. Sci. USA*, **101**, 6911–6916.
19. Hook-Barnard,I.G. and Hinton,D.M. The promoter spacer influences transcription initiation via sigma70 region 1.1 of *Escherichia coli* RNA polymerase. *Proc. Natl Acad. Sci. USA*, **106**, 737–742.
20. Page,L., Griffiths,L. and Cole,J.A. (1990) Different physiological roles of two independent pathways for nitrite reduction to ammonia by enteric bacteria. *Arch. Microbiol.*, **154**, 349–354.
21. Casadaban,M.J. (1976) Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J. Mol. Biol.*, **5**, 541–555.
22. Conter,A., Menchon,C. and Gutierrez,C. (1997) Role of DNA supercoiling and *rpoS* sigma factor in the osmotic and growth phase-dependent induction of the gene *osmE* of *Escherichia coli* K12. *J. Mol. Biol.*, **273**, 75–83.
23. Kolb,A., Kotlarz,D., Kusano,S. and Ishihama,A. (1995) Selectivity of the *Escherichia coli* RNA polymerase Esigma38 for overlapping promoters and ability to support CRP activation. *Nucleic Acids Res.*, **23**, 819–826.
24. Lodge,J., Fear,J., Busby,S., Gunasekaran,P. and Kamini,N.R. (1992) Broad host range plasmids carrying the *Escherichia coli* lactose and galactose operons. *FEMS Microbiol. Lett.*, **74**, 271–276.
25. Islam,M.S., Bingle,L.E.H., Pallen,M.J. and Busby,S.J.W. (2011) Organisation of the LEE1 operon regulatory region of enterohaemorrhagic *Escherichia coli* O157:H7 and activation by GrlA. *Mol. Microbiol.*, **79**, 468–483.
26. Rhodius,V.A. and Busby,S.J. (2000) Interactions between activating region 3 of the Escherichia coli cyclic AMP receptor protein and region 4 of the RNA polymerase sigma(70) subunit: application of suppression genetics. *J. Mol. Biol.*, **299**, 311–324.
27. Grainger,D.C., Belyaeva,T.A., Lee,D.J., Hyde,E.I. and Busby,S.J. (2004) Transcription activation at the *Escherichia coli melAB* promoter: interactions of MelR with the C-terminal domain of the RNA polymerase alpha subunit. *Mol. Microbiol.*, **51**, 1311–1320.
28. Savery,N., Belyaeva,T. and Busby,S. (1996) Introduction to protein: DNA interactions, DNase I footprinting, hydroxyl radical footprinting, permanganate footprinting and supplementary protocols. In Docherty,K. (ed.), *Essential Techniques: Gene Transcription*. BIOS Scientific Publishers, Oxford, pp. 1–33.
29. Miller,J. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
30. Vlahovicek,K., Kaján,L. and Pongor,S. (2003) DNA analysis servers: plot.it, bend.it, model.it and IS. *Nucleic Acids Res.*, **31**, 3686–3687.
31. Yamashino,T., Kakeda,M., Ueguchi,C. and Mizuno,T. (1994) An analogue of the DnaJ molecular chaperone whose expression is controlled by sigma s during the stationary phase and phosphate starvation in *Escherichia coli*. *Mol. Microbiol.*, **13**, 475–483.

32. 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.
33. Fenton,M.S., Lee,S.J. and Gralla,J.D. (2000) *Escherichia coli* promoter opening and -10 recognition: mutational analysis of sigma70. *EMBO J.*, **19**, 1130–1137.
34. Busby,S., Truelle,N., Spassky,A., Dreyfus,M. and Buc,H. (1984) The selection and characterisation of two novel mutations in the overlapping promoters of the *Escherichia coli* galactose operon. *Gene*, **28**, 201–209.
35. Thouvenot,B., Charpentier,B. and Branlant,C. (2004) The strong efficiency of the *Escherichia coli gapA* P1 promoter depends on a complex combination of functional determinants. *Biochem. J.*, **383**, 371–382.
36. Gaal,T., Barkei,J., Dickson,R.R., deBoer,H.A., deHaseth,P.L., Alavi,H. and Gourse,R.L. (1989) Saturation mutagenesis of an *Escherichia coli* rRNA promoter and initial characterization of promoter variants. *J. Bacteriol.*, **171**, 4852–4861.
37. Colland,F., Fujita,N., Kotlarz,D., Bown,J.A., Meares,C.F., Ishihama,A. and Kolb,A. (1999) Positioning of sigma(S), the stationary phase sigma factor, in *Escherichia coli* RNA polymerase-promoter open complexes. *EMBO J.*, **18**, 4049–4059.
38. Buckle,M., Buc,H. and Travers,A. (1992) DNA deformation in nucleoprotein complexes between RNA polymerase, cAMP receptor protein and the *lac* UV5 promoter probed by singlet oxygen. *EMBO J.*, **11**, 2619–2652.