5338–5355 Nucleic Acids Research, 2011, Vol. 39, No. 13 Published online 11 March 2011
doi:10.1093/nar/gkr129

In vitro transcription profiling of the σ^S subunit of bacterial RNA polymerase: re-definition of the σ^S regulon and identification of σ^S-specific promoter sequence elements

Anna Maciąg[1], Clelia Peano[2], Alessandro Pietrelli[2], Thomas Egli[3], Gianluca De Bellis[2] and Paolo Landini[1,*]

[1]Department of Biomolecular Sciences and Biotechnology, Università degli Studi di Milano, Milan, [2]Institute of Biomedical Technologies, National Research Council (CNR), Segrate (MI), Italy and [3]EAWAG, Swiss Institute for Environmental Technology, Dübendorf, Switzerland

Received July 23, 2010; Revised February 7, 2011; Accepted February 22, 2011

ABSTRACT

Specific promoter recognition by bacterial RNA polymerase is mediated by σ subunits, which assemble with RNA polymerase core enzyme (E) during transcription initiation. However, σ^{70} (the housekeeping σ subunit) and σ^S (an alternative σ subunit mostly active during slow growth) recognize almost identical promoter sequences, thus raising the question of how promoter selectivity is achieved in the bacterial cell. To identify novel sequence determinants for selective promoter recognition, we performed run-off/microarray (ROMA) experiments with RNA polymerase saturated either with σ^{70} ($E\sigma^{70}$) or with σ^S ($E\sigma^S$) using the whole *Escherichia coli* genome as DNA template. We found that $E\sigma^{70}$, in the absence of any additional transcription factor, preferentially transcribes genes associated with fast growth (e.g. ribosomal operons). In contrast, $E\sigma^S$ efficiently transcribes genes involved in stress responses, secondary metabolism as well as RNAs from intergenic regions with yet-unknown function. Promoter sequence comparison suggests that, in addition to different conservation of the −35 sequence and of the UP element, selective promoter recognition by either form of RNA polymerase can be affected by the A/T content in the −10/+1 region. Indeed, site-directed mutagenesis experiments confirmed that an A/T bias in the −10/+1 region could improve promoter recognition by $E\sigma^S$.

INTRODUCTION

Bacteria must cope with drastic changes in their environment, such as nutritional up- and downshifts, and variations in pH, osmolarity and temperature. Bacterial cells can quickly adapt to such environmental changes by modulating gene expression, at both transcriptional and post-transcriptional levels. At the transcription initiation level, gene expression can be regulated either through accessory transcription factors (activators and repressors), or via assembly of different forms of RNA polymerase. The latter mechanism of gene regulation involves the assembly of RNA polymerase core enzyme (indicated as E) with one of several σ factors that can direct RNA polymerase to specific promoter sequences (1). Typically, in the bacterial cell, one σ factor is devoted to transcription of a large part of the genome, including the essential cellular functions (housekeeping σ factor), while the so-called 'alternative σ factors' direct transcription of smaller sets of genes, often linked to specific functions (e.g. response to cellular stresses).

In *Escherichia coli*, seven σ factors have been identified: σ^{70} or σ^D (the housekeeping σ) and six alternativeσ factors: σ^E, σ^F, σ^H, σ^I, σ^N and σ^S (2). Most alternative σ factors recognize promoter sequences that strongly diverge from the consensus sequence for σ^{70}; in contrast, genes under the control of σ^S are characterized by promoter sequences very similar to σ^{70}-dependent genes (3,4). In line with this observation, *in vitro* selection of DNA sequences bound with high affinity by RNA polymerase associated with σ^S ($E\sigma^S$) led to the identification of a consensus sequence very similar to the one recognized by σ^{70} (5). Some level of overlapping in promoter recognition

*To whom correspondence should be addressed. Tel: +39 02 50315028; Fax: +39 02 50315044; Email: paolo.landini@unimi.it

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

by σ^{70} and σ^S might be consistent with σ^S function: indeed, in conditions leading to slow metabolic activity, such as nutrient starvation or oxidative stress, σ^S might take over in transcription of genes important for cell survival that are under σ^{70} control during faster growth (6). However, in order to switch from the fully active to the slow metabolic state, specific gene expression, and thus specific recognition of σ^{70}- versus σ^S-dependent promoters, must take place in the bacterial cell. Some promoter sequence determinants can favour recognition by either σ^{70} or σ^S (3): for instance, a C nucleotide upstream of the −10 promoter element (−13C) enhances transcription by $E\sigma^S$ (7). However, some sequence features favouring promoter recognition by σ^S seem to be dependent on specific promoter contexts: for instance, at the $E\sigma^S$-dependent *aidB* promoter, $E\sigma^S$, but not $E\sigma^{70}$, can recognize with equal efficiency either C or T as the first nucleotide in the −10 promoter element (8). However, the percentage of σ^S-dependent promoters carrying a −12C element is not significantly higher than in σ^{70}-dependent promoters (9), suggesting that the presence of a −12C might only contribute to specific promoter recognition by σ^S at selected promoters.

In addition to sequence determinants, it has been proposed that transcription factors such as CRP, IHF and Lrp can selectively block (or promote) promoter recognition by either $E\sigma^{70}$ or $E\sigma^S$ (10). A transcription regulator important for the modulation of promoter accessibility to different RNA polymerase holoenzymes is the H-NS protein, which can repress transcription by $E\sigma^{70}$, but not by $E\sigma^S$, at various promoters (11), a phenomenon known as transcriptional silencing (12). Specific promoter recognition by $E\sigma^S$ is also affected by the degree of DNA supercoiling (13). In addition, σ^S activity and intracellular concentrations are affected by various factors, such as the presence of an anti-sigma factor for σ^{70} (14), and by the accumulation of the signal molecules ppGpp (15) and polyphosphate (3).

Work aimed to the identification of σ^S-specific promoter elements has mostly been carried out *in vivo*, comparing relative gene expression in a wild type versus an *rpoS* mutant derivative unable to produce the σ^S protein (9,16–20). Although this approach has proven very useful for the identification of *rpoS*-dependent genes and in the characterization of σ^S-specific promoter elements such as the −13C, it cannot distinguish between promoters directly recognized by the $E\sigma^S$ form of RNA polymerase and promoters under the indirect control of the *rpoS* gene. In contrast, dependence on $E\sigma^S$, as determined by biochemical experiments with purified RNA polymerase, has only been determined for a limited number of promoters [e.g. *fic* (21), *csiD* (22) and *aidB* (23)]. In this work, we have performed *in vitro* transcription experiments with either $E\sigma^{70}$ or $E\sigma^S$, using the whole *E. coli* genome as template, to identify promoter regions selectively recognized by two forms of RNA polymerase. Our results support previous observations that $E\sigma^{70}$- and $E\sigma^S$-dependent promoters differ in conservation of the UP element and of the −35 sequence, and in the sequence immediately upstream of the −10 promoter element. In addition, we show that differences in the A/T content in the −10/+1 promoter region can favour transcription by either form of RNA polymerase. Finally, our work has led to the identification of novel σ^S-dependent genes, thus providing further insight on the physiological role of σ^S.

MATERIALS AND METHODS

Protein purification and RNA polymerase reconstitution

Escherichia coli RNA polymerase core enzyme was purchased from Epicentre (Madison, WI, USA); histidine-tagged σ factors were produced and purified as described (8,24). The σ^S protein appeared totally pure from contaminants as determined by denaturing protein gel electrophoresis; in contrast, in the σ^{70} preparations the presence of faint additional bands, corresponding to the molecular weight of the core RNA polymerase subunits α, β and β′, could be detected (data not shown). Weak contamination of σ^{70} preparations by core RNA polymerase subunits is consistent with the high affinity of σ^{70} for the core enzyme. For reconstitution of RNA polymerase holoenzymes, the core enzyme was incubated for 10 min at 37°C with either σ^S or σ^{70} at a 1:10 ratio. For calculation of RNA polymerase concentrations in transcription assays, it was assumed that, after reconstitution, core enzyme would be 100% active and fully saturated by either σ factor.

In vitro transcription on supercoiled plasmids

Promoter regions of interest were amplified from the genome of *E. coli* MG1655 (25) and cloned into the pJCD01 plasmid (22) using the BamHI and EcoRI sites, with the exception of the *ilvY* promoter, which was cloned using the BamHI and SphI sites, due to the presence of an EcoRI site in the *ilvY* promoter region. Single-round *in vitro* transcription experiments were carried out on supercoiled templates (3 nM) in the presence of 10 nM reconstituted RNA polymerase holoenzyme. Plasmid DNA and reconstituted holoenzyme were incubated for 10 min at 37°C in 18 µl of transcription buffer (40 mM HEPES pH 8.0, 10 mM magnesium chloride, 150 mM potassium glutamate, 2 mM dithiothreitol, 100 µg/ml bovine serum albumin) prior to the addition of 2 µl of a mixture of ribonucleotide triphosphates and heparin to a final concentration of 500 µM each for ATP, CTP, GTP and UTP and 250 µg/ml heparin. Transcription reactions were allowed to proceed for 10 min at 37°C and were stopped by addition of NaCl to a final concentration of 0.5 M followed by incubation at 70°C for 5 min. Samples were extracted with a 1:1 phenol–chlorophorm mixture, precipitated with ethanol and resuspended in 12 µl TE buffer. A volume of 2.5 µl of resuspended samples were treated with DNaseI (1 U in 10 mM Tris–HCl pH 7.6, 2.5 mM $MgCl_2$, 0.5 mM $CaCl_2$, in a final volume of 20 µl) for 1 h at 37°C, and DNaseI was heat-inactivated at 65°C for 10 min. Transcript amounts were determined by quantitative real-time polymerase chain reaction (PCR): 1 µl of DNaseI-treated samples was retrotranscribed and the cDNAs were amplified in quantitative real-time PCR using the RNA-I transcript as

reference gene. The sequences of the primers used in real-time PCR are available upon request. DNaseI-treated transcription reactions not incubated with reverse transcriptase were added as negative controls for the presence of undigested plasmid DNA. To verify that transcripts would indeed originate from the promoters of interest, and not from non-specific transcription initiation events, *in vitro* transcription start sites were determined by Rapid Amplification of cDNA Ends (RACE) analysis as described (26): start sites observed in the *in vitro* transcription experiments corresponded to those described in the literature (data not shown). Site-directed mutagenesis of the *ssrS* P1 promoter was performed using the three-step PCR method (27).

Run-off microarray experiments on *E. coli* genome

Genomic DNA from *E. coli* MG1655 was isolated with GenElute™ Bacterial Genomic DNA Kit (Sigma) and digested with EcoRI overnight. The size of digested genomic DNA ranged from 0.6 to 30 kb as judged from visualization on agarose gels. After purification with Wizard SV Gel and PCR Clean-Up System (Promega), 1 μg of genomic DNA was used in single transcription assay for run-off micro array (ROMA) experiments. RNA polymerase holoenzyme concentrations were 100 nM in 50-μl reaction mixture. Digested genomic DNA and reconstituted holoenzyme were incubated in transcription buffer (40 mM HEPES pH 8.0, 10 mM magnesium chloride, 150 mM potassium glutamate, 2 mM dithiothreitol, 100 μg/ml bovine serum albumin) for 15 min at 37°C to allow binding of RNA polymerase to DNA. Transcription reactions were started with the addition of 2.5 μl of ATP, CTP, GTP and UTP to a final concentration of 500 μM each, were allowed to proceed for 30 min at 37°C and were stopped by incubation at 65°C for 5 min. The reaction mixture was incubated with 10 U of DNaseI at 37°C for 30 min to degrade genomic DNA. RNA was purified from reaction with miRNeasy Mini Kit (Qiagen), which allows recovering of total RNA, including small RNA molecules (<200 nt). The quantity of RNA produced by *in vitro* transcription was determined using Nanodrop Spectrophotometer, while the average length of the transcripts, as well as of the corresponding cDNA from reverse transcription (see below), was evaluated by capillary electrophoresis on an Agilent Bioanalyzer 2100 using an RNA pico and nano assay, respectively. Capillary electrophoresis analysis of transcription reactions performed in the presence of $E\sigma^S$, $E\sigma^{70}$ or core RNA polymerase clearly showed that, while the average transcript size generated by any form of RNA polymerase was comparable, a significantly lower amount of RNA was produced in transcription reactions performed with core RNA polymerase alone compared to either form of RNA polymerase holoenzyme (Supplementary Figure S1). Indeed, *in vitro* transcription reactions performed using RNA polymerase core enzyme generated 4 ng/μl of transcripts, versus 31 ng/μl and 23 ng/μl total RNA produced in the presence of $E\sigma^{70}$ and $E\sigma^S$, respectively. Dependence on the presence of a σ factor for efficient *in vitro* transcription strongly suggests that, in our experimental conditions, transcription initiation does not originate from non-specific initiation sites such as 5′ DNA termini, but requires interaction of RNA polymerase holoenzyme with promoter sequences. Ten independent run-off transcription assays with either $E\sigma^S$ or $E\sigma^{70}$ were performed, and transcripts were pooled together (~1 μg total RNA). From the pool of the 10 independent multi-round transcription reactions two distinct hybridizations on microarrays were performed.

For hybridization onto microarrays, we used the Affymetrix Genechip *E. coli* Genome 2.0 array, which includes 10 144 probe sets covering all the predicted transcripts from four strains of *E. coli*: the laboratory strain MG1655, the uropathogenic CFT073 and the enteropathogenic strain OH157:N7, subtypes EDL933 and SAKAI. Due to the high degree of similarity between the *E. coli* strains, typically a single probe set is tiled to represent the equivalent orthologue in all four strains, and strain-specific probes are only used for genes displaying low levels of conservation, or present solely in one strain. Probe sets match every open reading frame (ORF) in *E. coli*; in addition, 1427 probe sets targeting 714 *E. coli* MG1655 intergenic regions, probes for various antibiotic resistance markers, and additional control and reporter genes from the previous generation *E. coli* arrays are also represented in the Affymetrix Genechip *E. coli* Genome 2.0 array.

The RNA samples were processed for microarray hybridization, following the instructions of the GeneChip® Expression Analysis Technical Manual (Chapter 5, 6 and 7, Prokaryotic target Preparation, Hybridization, Washing, Staining and Scanning), except that cDNA synthesis by reverse transcription was performed with 1 μg RNA. cDNA was fragmented with DNase I treatment and labelled with biotin using Terminal Deoxynucleotidyl Transferase. The fragmented and labelled cDNAs were then hybridized for 16 h at 45°C on individual *E. coli* Genome 2.0 arrays. After hybridization, GeneChips were washed and stained with streptavidin-conjugated phycoerythrin by using the Fluidic Station FS450 (Affymetrix) according to the FS450_0006 Protocol. Fluorescent images of the arrays were acquired using a GeneChip Scanner 3000 7G (Affymetrix). All Chip images and files have been deposited in the GEO (Gene Expression Omnibus) (http://www.ncbi.nlm.nih.gov/geo/) repository (accession number:GSE22207). After quality control of data distribution, the raw data (CEL files) were used to perform normalization and probe set summarization through Robust Multiarray Analysis (RMA) algorithm (by using the Affymetrix Gene Expression Console Software (www.affymetrix.com). Normalized data were also analyzed with OneChannelGUI Software (http://www.bioinformatica.unito.it/oneChannelGUI/) in order to perform an additional expression analysis based on different parameters. OneChannelGUI is an add-on Bioconductor package extending the capability of the affylmGUI package (28); it is a library providing a graphical interface (GUI) for Bioconductor libraries to be used for the complete single-channel microarray analysis.

To perform statistic analysis on the two hybridizations, we decided to apply a simple non-parametric statistical method based on ranks of fold changes to perform a two-class paired differential analysis. To select genes with significantly different transcription levels, we set the following parameters for the Rank Product analysis: 100 permutations and 0.1 cut-off percentage of false positives (pfp) which corresponds to a *P*-value <0.01. Fold differences higher than 1.5-fold were considered indicative of significantly different expression, similar to previous ROMA experiments (29): genes more efficiently transcribed by either $E\sigma^S$ or $E\sigma^{70}$ are listed in Tables 1 and 2, respectively. Determination of start sites on transcripts generated in the *in vitro* transcription assays was carried out by RACE analysis (26).

In order to manage and retrieve microarray data, a genome browser was set up (http://155.253.6.64/cgi-bin/gbrowse/provakappa/#search). The genome browser is based on the Generic Genome Browser (Gbrowse) which is a combination of database and interactive webpage for manipulating and displaying annotations on genomes. This bioinformatic tool allows users to view and navigate through the MG1655 genome (GenBank Accession Number:U00096) with information about the gene annotation, coming from the website http://www.genome.wisc.edu/tools/asap.htm (downloaded as GFF3 file), and all the probe sets contained in the Affymetrix *E. coli* Genome 2.0 Array manually remapped on the genome.

Determination of *rpoS*-dependent gene expression in bacterial cells

Bacterial strains used were MG1655 (wild type) and its *rpoS* mutant derivative EB1.3 (30). Strains were grown at 30°C in three different media: the complex Luria Bertani (LB) broth, the glucose-based M9Glu/sup medium (31) and LB medium diluted 1:4 (LB1/4). The LB1/4 medium was utilized since it was shown to stimulate expression of *rpoS*-dependent genes such as the *csg* operons (32). Thus, growth in LB1/4 medium might positively affect either σ^S concentration or σ^S activity. Samples for RNA extraction were taken both in late logarithmic (OD_{600} = 0.6–0.7) and in late stationary phase (overnight cultures, $OD_{600} \geq 1.5$). RNA was extracted using the small RNA miRNeasy Mini Kit (Qiagen), and further reverse transcription and cDNA amplification in quantitative real-time-PCR were performed as described (33). Primer sequences are available upon request. All reactions were performed twice, each time in duplicate, and always showed very similar results. The relative amounts of the transcripts were determined using 16S rRNA as the reference gene ($[Ct_{\text{Gene of interest}} - Ct_{16S}] = \Delta Ct$).

RESULTS

ROMA analysis of σ^S- and σ^{70}-dependent promoters

To identify sequence determinants that can direct selective promoter recognition by either the σ^{70}- or the σ^S- associated form of RNA polymerase, we compared $E\sigma^S$ and $E\sigma^{70}$ in ROMA experiments (29,34). Run-off transcription assays were performed using as DNA template the whole genome of *E. coli* MG1655 after digestion with EcoRI; cDNAs generated from the transcripts were hybridized on microarrays to determine their relative amounts. Unlike genome expression studies performed in living cells, which do not distinguish between direct and indirect effects, such as transcription factor-dependent promoter recognition, the ROMA analysis solely detects genes whose promoter are recognized by either $E\sigma^S$ or $E\sigma^{70}$ (or both) in the absence of any additional factor. Out of ~10 000 probe sets on the microarray, only 173 (~1.7%) showed differences in transcription levels higher than 1.5-fold, considered to be significant (pfp<0.1, *P*-value <0.01; see 'Materials and Methods' section). For several operons, only a portion of the transcriptional unit (or even a single gene) showed significant difference in transcription levels by either $E\sigma^S$ or $E\sigma^{70}$: in the case of reduced transcription of distal genes within a given operon, this effect might depend either on premature arrest of transcription by RNA polymerase or on the presence of an EcoRI restriction site within the operon, since the EcoRI restriction nuclease was used to digest chromosomal DNA used in ROMA experiments (see 'Materials and Methods' section). However, the effects of EcoRI digestion were not so severe as it could be expected, possibly due to the fact that chromosomal DNA was only partially digested by the enzyme in our conditions. For instance, despite the presence of an EcoRI site in the *ydhY* gene, transcription extending into the downstream genes was still detectable (Table 1 and Supplementary Table S1). In the case of distal genes of an operon showing more significant differences in $E\sigma^S$- versus $E\sigma^{70}$-dependent transcription than genes proximal to the promoter, this might be due to either lack of detection of proximal genes in the microarray experiment or the presence of unidentified promoters internal to the operons. We verified that genes within the same operon would not show dependence on different forms of RNA polymerase: out of the several hundred operons in the *E. coli* genome, only in the weakly $E\sigma^{70}$-dependent *atpBEFHAGDC* and *flgBCDEFGHIJ* operons could we observe the presence of a single gene showing dependence on $E\sigma^S$ in our ROMA experiments (data not shown). This result might suggest the presence of $E\sigma^S$-dependent promoters, or promoter-like sequences, within these operons; however, we decided to focus our investigation on operons consistently showing preferential recognition by either form of RNA polymerase, and the *atpBEFHAGDC* and *flgBCDEFGHIJ operons* were not considered further in our study. In Supplementary Table S1, we show the ratios of $E\sigma^S$- versus $E\sigma^{70}$-dependent transcription for operons featuring genes preferentially transcribed in the presence of $E\sigma^S$.

Complete results of ROMA experiments are summarized in Tables 1 and 2, listing genes and intergenic regions transcribed more efficiently either by $E\sigma^S$ (Table 1) or by $E\sigma^{70}$ (Table 2). Out of the 54 genes preferentially transcribed by $E\sigma^S$ in ROMA experiments that are annotated in *E. coli* MG1655, 21 (39%) had already been described as *rpoS*-dependent genes, either from genetic characterization or from microarray experiments comparing wild-type and *rpoS* mutant strains (underlined in Table 1). In

Table 1. Genes transcribed more efficiently in the presence of $E\sigma^S$

Genes and promoter regions[a]	Gene product[b]	b number[c]	Transcription levels ($E\sigma^S/E\sigma^{70}$ ratio)[d]	Known regulatory factors[e]
Stress response				
yieF	Chromate reductase	b3713	1.50	σ^E (59); σ^S (19)
IG1341353_1341620-r	Intergenic region including *osmB* promoter	N.A.	1.51	σ^S, RcsBA (60)
osmB	Osmolarity-inducible lipoprotein	b1283	1.51	σ^S, RcsBA (60)
gshA	γ-glutamate–cysteine ligase (glutathione biosynthesis)	b2688	2.21	
ahpF	Alkyl-peroxidase reductase	b0606	1.52	OxyR (61)
pphA	Protein phosphatase	b1838	1.54	σ^H (62,63)
macA	Macrolide resistance efflux pump	b0878	1.54	
DNA and RNA metabolism and modification				
ppk	Polyphosphate kinase, part of RNA degradosome,	b2501	1.50	Regulator of intracellular σ^S concentrations (3)
IG2522899_2523146-r	Intergenic region including *xapA* promoter	N.A.	1.62	XapR (64)
xapA	Xanthosine phosphorylase, nucleotide synthesis/degradation	b2407	1.56	XapR (64)
gyrB	B subunit of DNA gyrase	b3699	1.65	Fis (65)
recT	Rac prophage, recombinase, in *recET-lar-ydaCQ* operon	b1349	1.93	σ^S-Dependent in biofilm-growing cells (18)
lar	Rac prophage,restriction alleviation and modification enhancement, in *recET-lar-ydaCQ* operon	b1348	1.93	
ydaC	Rac prophage, in *recET-lar-ydaCQ* operon	b1347	1.74	as *recT*
ydaQ	Rac prophage, possible recombinase, in *recET-lar-ydaCQ* operon	b1346	1.66	as *recT*
Polyamine metabolism				
puuB	γ-glutamyl–putrescine oxidase	b1301	1.61	σ^S (19)
gabD	Succinate semialdehyde dehydrogenase	b2661	1.65	σ^S, Nac (9,66)
speB	Agmatinase (in arginine/putrescine degradation pathway)	b2937	1.66	
Transcription regulation				
IG582284_582903-f	Intergenic region including *appY* promoter	N.A.	1.69	H-NS (67)
appY	Regulator of *hyaABCDEF* operon	b0564	1.51	H-NS (67)
yciT	Putative *deoR*-type transcription regulator	b1284	1.56	
ilvY	Regulator of *ilvC* operon	b3773	1.75	
rhaS	Regulator of rhamnose transport	b3905	1.81	CRP (29); RhaS (68)
Sugar metabolism				
treF	Trehalase	b3519	1.53	σ^S (9,19)
ytfT	Putative galactose ABC transporter	b4230	1.55	σ^S (19)
otsB	Trehalose-6-phosphate phosphatise	b1897	1.56	σ^S (7)
araF	Arabinose transporter, component of an ABC transport system	b1901	1.57	CRP, AraC (69); σ^S (19)
glgP	Glycogen phosphorylase, part of *glgCAP* operon	b3428	1.60	CRP (70); σ^S controls *glgCAP* operon (19)
gpmM	Putative 2,3-bisphosphoglycerate mutase	b3612	1.61	
glvC	Sugar transport phosphotransferase	b3683	1.79	
Other metabolic functions				
IG1030936_1031361-f	Intergenic region including *hyaABCDEF* promoter	N.A.	2.02	
hyaA	Hydrogenase small subunit; in *hyaABCDEF* operon	b0972	1.68	σ^S, AppY, anaerobic regulation by ArcA and NarP/NarL (20,71,72)
hyaB	Hydrogenase large subunit; in *hyaABCDEF* operon	b0973	1.66	as *hyaA*
hyaF	Hydrogenase subunit (nickel-binding protein); in *hyaABCDEF* operon	b0977	1.50	as *hyaA*
syd	SecY-interacting protein	b2793	1.59	
murP	Acetyl-muramic acid permease	b2429	1.59	CRP, MurR (73)
murR	Transcriptional repressor of *murQP*	b2427	1.62	
tam	Trans-aconitate methylatransferase	b1519	1.60	σ^S (74)
ydhY	Predicted oxidoreductase, Fe-S protein; in *ydhYVWXUT* operon	b1674	1.80	FNR, NarL (75)
ydhV	Predicted oxidoreductase, Fe-S protein; in *ydhYVWXUT* operon	b1673	1.62	FNR, NarL (75). σ^S-dependent in biofilm-growing cells (18)

(continued)

Table 1. Continued

Genes and promoter regions[a]	Gene product[b]	b number[c]	Transcription levels ($E\sigma^S/E\sigma^{70}$ ratio)[d]	Known regulatory factors[e]
Unknown and miscellaneous functions				
yccT	Unknown	b0964	1.51	Induced in stationary phase (76)
yffP	Predicted protein, prophage	b2447	1.52	
yidK	Putative membrane transporter	b3679	1.53	
ynfD	Predicted protein	b1586	1.54	
yfiL	Putative lipoprotein	b2602	1.55	
yi91a	Unknown, in CP4-6 prophage sequence	b0255	1.56	
yfjL	Unknown, possible prophage gene	b2625	1.57	σ^S (9)
yagL	Unknown, prophage protein	b0278	1.59	σ^S-Dependent in biofilm-growing cells (18)
eutA	Reactivating factor for ethanolamine ammonia lyase	b2451	1.59	*eutH*, in *eutHA* operon, is σ^S-dependent (9)
yhjG	Predicted outer membrane protein	b3524	1.61	σ^S (9)
ybeH	Hypothetical protein	b0625	1.66	
yedS	Unknown	b1964	1.69	
yfgJ	Unknown, mutant affecting swarming motility	b2510	1.74	
G7353	Phantom gene	b2596	1.79	Upstream of a ribosome modulation factor induced in stationary phase (*yfiA*)
yqiG	Unknown, interrupted by IS element	b3046	1.81	
IG1006824_1007066-r	Intergenic region including *ycbX* promoter	N.A.	2.26	
ycbX	Unknown	b0947	2.05	
yfjH	Unknown	b2623	2.30	
IG2755422_2755664-r	Intergenic region downstream of *yfjH*	N.A.	2.01	
ychS	Unknown	b1228	3.21	
Non-coding RNAs				
ryeE (cyaR)	Small RNA, promotes degradation of *ompX* and *nadE* RNA	b4438	1.50	CRP, σ^E (77,78)
sgrS|sgrT	Small RNA, inhibits *ptsG* translation/SgrT protein	b4577	1.51	SgrR protein
ECs3934	*sibD/sibE* non coding RNA- *ibsD/ibsE* toxic peptides	b4447 b4664	1.62	Complex locus including two non coding RNAs overlapped by two small ORFs encoding a putative toxin/antitoxin system
micA	*micA* small RNA (downregulates *ompA* expression)	b4442	2.00	
Intergenic regions, genes not annotated in MG1655				
c3878	Unknown, annotated in CFT073	N.A.	1.50	
IG2922538_2922756-f	Intergenic region between *syd* (predicted protein) and *csrB* (ncRNA), antisense	N.A.	1.50	
c5008	Unknown, annotated in CFT073, downstream of *malM*	N.A.	1.50	
IG2438141_2438404-f	Between *fabB* and *ycfJ*, antisense	N.A.	1.50	
ECs5537	Unknown, annotated in O157:H7 SAKAI	N.A.	1.51	
IG2885243_2885600_r	Intergenic region, upstream of predicted helicase *ygcB*	N.A.	1.52	
c4656	Unknown, annotated in CFT073 (antisense of *atpC*)	N.A.	1.52	
IG3665211_3665420-f	Intergenic region upstream of *gadA*, antisense	N.A.	1.52	
c1010	Unknown, annotated in CFT073, upstream of *aqpZ*	N.A.	1.52	*aqpZ* (aquaporin) is σ^S-dependent (18,79)
c0723	Unknown, annotated in CFT073, upstream of *dacA*, transcribed in opposite direction	N.A.	1.54	
IG2481360_2481774-r	Intergenic region upstream *emrKY* (TolC-multidrug efflux pump)	N.A.	1.54	
IG2228406_2228643-f	Intergenic region upstream of *yohJK* (inner membrane proteins)	N.A.	1.54	
IG2201932_2202549-r	Intergenic region between *yehI-yehK*, antisense	N.A.	1.54	
c2806	Unknown, annotated in CFT073 (antisense of *menB*)	N.A.	1.55	
c3010	Unknown, annotated in CFT073, upstream of *perM* gene, transcribed in opposite direction	N.A.	1.55	
IG3669525_3669971-r	Intergenic region, upstream of *yhjB*	N.A.	1.58	
IG2898371_2898613-f	Intergenic region upstream of *yqcE*	N.A.	1.60	
c5221	Unknown, annotated in CFT073 (antisense of *aspA*)	N.A.	1.63	
IG223409_223770-r	Intergenic region upstream of *rrsH* ribosomal operon (transcribed in opposite direction)	N.A.	1.64	

(continued)

Table 1. Continued

Genes and promoter regions[a]	Gene product[b]	b number[c]	Transcription levels ($E\sigma^S/E\sigma^{70}$ ratio)[d]	Known regulatory factors[e]
IG3420831_3421058-r	Intergenic region, downstream of *rrfF* ribosomal operon	N.A.	1.66	
c4942	Unknown, annotated in CFT073 (antisense of *rplL*)	N.A.	1.67	
IG2428784_2429041-f	Intergenic region, upstream of *cvpA* gene, transcribed in opposite direction	N.A.	1.71	
c1908	Unknown, annotated in CFT073 (antisense of *yddM*)	N.A.	1.71	
c0703	Unknown, annotated in CFT073 (antisense of *citF*)	N.A.	1.74	
ECs5165	*yjfO* (biofilm-related protein) in O157:H7 SAKAI	N.A.	1.86	
IG1903284_1903567-r	Intergenic region downstream of *yobD*, transcribed in opposite direction	N.A.	1.87	
IG3358642_3358810-f	Intergenic region downstream of *gltD*	N.A.	1.89	
IG2190243_2190534-r	Intergenic region downstream of *yehE*	N.A.	1.94	
c3113	Unknown, annotated in CFT073 (upstream of *rrsG* ribosomal operon, transcribed in opposite direction)	N.A.	1.95	
c2317	Unknown, annotated in CFT073 (antisense of *azuC*)	N.A.	1.99	Complex locus including small RNA *isrB*
IG2755422_2755664-r	Intergenic region downstream *yfjH*	N.A.	2.01	
IG2519349_2519612-f	Intergenic region downstream of *xapR*	N.A.	2.05	
IG330721_331594-r	Intergenic region upstream of *yahA*, transcribed in opposite direction	N.A.	2.07	

[a]For known genes, we used the nomenclature reported in the NCBI database (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi); intergenic region are indicated using the nomenclature by Affymetrix, which indicates the start and the end of the intergenic region (IG) covered by the various probe sets; –f or –r indicate if the probe sets are in forward or reverse orientation relative to the (+) strand of the *E. coli* chromosome; genes already described as *rpoS*-dependent *in vivo* are underlined.
[b]Gene product (or predicted product).
[c]Relative location of the various ORFs on the *E. coli* MG1655 chromosome.
[d]Determined by microarray analysis as described in 'Materials and Methods' section.
[e]When not otherwise stated, the information is taken from http://ecocyc.org/.

contrast, no known *rpoS*-dependent gene was found to be preferentially transcribed by $E\sigma^{70}$ (Table 2). However, several genes clearly identified by previous works as *rpoS*-dependent (e.g. *katE*, *dps* and *gadA*) only showed slight (<1.5-fold) preferential recognition by $E\sigma^S$ in our experimental conditions (Supplementary Table S2). This would suggest that, for these *rpoS*-dependent genes, promoter recognition by $E\sigma^S$ might be mediated by regulatory proteins, or facilitated by additional factors such as DNA supercoiling or effector molecules (e.g. ppGpp), missing in ROMA experiments. Alternatively, in ROMA experiments, $E\sigma^{70}$ might recognize promoters which might not be accessible to this form of RNA polymerase *in vivo*, due to selective negative recognition by regulatory proteins such as H-NS (11,12).

A significant fraction of genes more efficiently transcribed by $E\sigma^S$ is involved in sugar and in polyamine metabolism, in response to cellular stresses and in nucleic acid synthesis, modification and turnover (Table 1), consistent with the role of σ^S as a starvation- and stress-related protein; no essential gene was found. Interestingly, several $E\sigma^S$-dependent transcripts correspond to non-coding RNAs, intergenic regions and ORFs of unknown function only annotated in pathogenic *E. coli* strains. This last result was surprising, since the DNA template used in the *in vitro* transcription experiments came from *E. coli* MG1655; however, probe set analysis and sequence comparison performed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) revealed that these ORFs are also present in MG1655, although they are not annotated in this strain. Several of these unlisted ORFs overlap known genes in an antisense direction, as indicated in Table 1. Interestingly, one of such ORFs (*c3113*, annotated in the uropathogenic strain CFT073), as well as one intergenic region (*IG223409_223770-r*), is located in the promoter regions of ribosomal operons *rrsG*, and *rrsH* in an antisense direction, possibly suggesting that they may play a role in $E\sigma^S$-dependent control of ribosomal operon transcription.

In contrast, $E\sigma^{70}$ appeared to transcribe with higher-efficiency genes encoding ribosomal proteins and other protein synthesis-related genes, such as rRNA- and tRNA-encoding genes and *prfB*, encoding for release factor 2 (Table 2). Several of the protein synthesis-related genes listed in Table 2 are essential, as are the *yjeE* and the *amiB* genes, part of a multifunctional operon also preferentially transcribed by $E\sigma^{70}$. Interestingly, regulatory genes directly affecting σ^{70} activity, such as *relA*, responsible for biosynthesis of the ppGpp alarmone in response to amino acid starvation (35), and *ssrS*, encoding a 6S

Table 2. Genes transcribed more efficiently in the presence of $E\sigma^{70}$

Genes and promoter regions[a]	Gene product[b]	b number[c]	Transcription levels (Es^{70}/Es^{s} ratio)[d]	Known regulatory factors[e]
Ribosomal genes and protein synthesis				
IG1286552_1286760-r	Intergenic region including *tyrT* promoter	N.A.	2.97	FIS (80)
rttR	Non-coding RNA, part of *tyrT* transcript	b4425	1.97	FIS (80)
tpr	Small protamine-like protein part of *tyrT* transcript	b1229	2.03	
serT	Serine tRNAs gene	b0971	1.89	FIS (79); upregulated in *rpoS* mutant of MG1655 (17)
rpmI	Ribosomal protein L35	b1717	1.91	
prfB	Release Factor RF2	b2891	1.91	Upregulated in a biofilm-growing *rpoS* mutant derivative of MG1655 (18)
rimP	Ribosomal maturation protein	b3170	2.01	
rpsF	Ribosomal protein S6	b4200	2.04	
rpmG	Ribosomal protein L33	b3636	2.10	
relA	ppGpp alarmone biosynthetic enzyme	b2784	2.35	
rplU	Ribosomal protein L21, in *rplU-rpmA* operon	b3186	2.50	
rpmA	Ribosomal protein L27, in *rplU-rpmA* operon	b3185	2.58	
rpsU	Ribosomal protein S21	b3065	2.78	Upregulated in a biofilm-growing *rpoS* mutant derivative of MG1655 (18)
IG3426400_3426656-r	Intergenic region including *rrnH* promoter	N.A.	2.87	
DNA repair				
sulA	SOS response inhibitor of cell division	b0958	2.04	LexA (81)
dinI	AP endonuclease, SOS response	b1061	2.18	LexA (82); upregulated in an *rpoS* mutant derivative of OH157:H7 EDL 933 (19)
Multifunctional operons				
cvpA	Colicin V production; in *cvpA-purF-ubiX* operon	b2313	1.85	PurR (83); upregulated in an *rpoS* mutant derivative of OH157:H7 EDL 933 (19)
purF	Amidophosphoribosyl transferase (ribonucleotide metabolism); in *cvpA-purF-ubiX* operon	b2312	2.12	as *cvpA*
ubiX	3-octaprenyl-4-hydroxybenzoate decarboxylase (ubiquinone biosynthesis); in *cvpA-purF-ubiX* operon	b2311	1.90	as *cvpA*
c2854	Unknown, annotated in CFT073 as part of the *cvpA-purF-ubiX* operon	N.A.	2.00	
yjeF	Putative carbohydrate kinase; in *yjeFE-amiB-mutL* operon	b4167	1.88	
yjeE	Essential protein with weak ATPase activity; in *yjeEF-amiB-mutL* operon	b4168	1.95	
amiB	*N*-acetylmuramyl-L-alanine amidase needed for septum formation during cell division; in *yjeEF-amiB-mutL* operon.	b4169	2.02	
mutL	Methyl-directed mismatch repair, subunit in MutHLS complex; in *yjeEF-amiB-mutL* operon.	b4170	2.06	
yhbE	Inner membrane protein; in *yhbE-obgE* operon	b3184	2.36	
obgE	GTP-binding protein, involved in ppGpp turnover; in *yhbE-obgE* operon	b3183	1.91	
Transcription regulation				
IG3717398_3717677-f	Intergenic region including *cspA* promoter	N.A.	1.85	
cspA	Cold shock protein A	b3556	2.34	Upregulated in an *rpoS* mutant derivative of OH157:H7 EDL 933 (19)
alpA	CP4-57 prophage gene, regulator of tmRNAs	b2624	2.31	
Metabolic functions				
fhuF	Iron reductase	b4367	1.83	Fur, OxyR (84)
ydhR	Putative monooxygenase	b1667	1.91	
artJ	Arginine transporter	b0860	2.13	Upregulated in a biofilm-growing *rpoS* mutant derivative of MG1655 (18)
Unknown and miscellaneous functions				
ycgY	Unknown	b1196	1.83	

(continued)

Table 2. Continued

Genes and promoter regions[a]	Gene product[b]	b number[c]	Transcription levels (Es^{70}/Es^{s} ratio)[d]	Known regulatory factors[e]
yefM	Antitoxin in *yefM-yoeB* toxin-antitoxin system	b2017	1.85	
yehL	Unknown, possible component of ABC transport system	b2119	1.92	
yebN	Unknown, putative membrane protein	b1821	2.06	
ydiE	Putative lipoprotein	b1705	2.10	
Non-coding RNAs				
isrB	Small RNA	b4434	2.25	
spf	Small RNA, regulates DNA polymerase I activity	b3864	2.72	
ssrS	6S RNA, modulates σ^{70} activity	b2911	3.74	Upregulated in a *rpoS* mutant derivative of MG1655 (20)
Intergenic regions, genes not annotated in MG1655				
c1714	Unknown, annotated in CFT073, upstream of *cls*, transcribed in opposite direction	N.A.	1.83	
IG330721_331594-f	Intergenic region including *yahA* promoter region	N.A.	1.85	
ECs1613	Unknown, annotated in O157:H7 SAKAI, possible prophage gene (*renD*)	N.A.	1.86	
IG2424809_2425028-r	Intergenic region between *argT* and *hisJ*	N.A.	1.92	
IG1120179_1120464-r	Intergenic region upstream *bssS* (regulator of biofilm formation)	N.A.	1.92	
c2481	Unknown, annotated in CFT073, upstream of *cobU*	N.A.	1.94	
IG127588_127911-f	Intergenic region upstream of *lpd*	N.A.	1.96	
IG2404662_2405580-r	Intergenic region upstream of *lrhA*	N.A.	1.98	
Z0043	Annotated in O157:H7 OH157:H7 EDL 933, *caiC*	N.A.	2.05	
c2568	Annotated in CFT073, *wcaM*	N.A.	2.13	
Z5055	Annotated in O157:H7 OH157:H7 EDL 933, *rfaG*	N.A.	2.16	
c4052	Unknown, annotated in CFT073	N.A.	2.27	
IG2815526_2815805-r	Intergenic region upstream *yqaBA*	N.A.	2.31	
IG583654_583902-r	Intergenic region downstream of *ompT*	N.A.	2.35	
c2230	Unknown, annotated in CFT073 (antisense of *cspC*)	N.A.	2.42	
IG1905616_1906284-f	Intergenic region, between *yobF* and *yebO*	N.A.	2.44	
Z3239	Annotated in O157:H7 OH157:H7 EDL 933 (antisense of *yegI*)	N.A.	2.49	
c4719	Annotated in CFT073 homologous to *aslA*	N.A.	2.54	
c4352	Annotated in CFT073 as *dppC*	N.A.	2.99	
IG2428784_2429041-f	Intergenic region, between *yobD* and *yebN*	N.A.	3.10	
c1434	Unknown, annotated in CFT073, *ydfR*	N.A.	3.45	
Z5945	Annotated in O157:H7 OH157:H7 EDL 933 (toxic peptide)	N.A.	4.02	
Z5868	Annotated in O157:H7 OH157:H7 EDL 933; *yjgM*, putative acetyltransferase	N.A.	4.44	

[a]For known genes, we used the nomenclature reported in the NCBI database (http://www.ncbi.nlm.nih.gov/genomes/lproks.cgi); intergenic region are indicated using the nomenclature by Affymetrix, which indicates the start and the end of the intergenic region (IG) covered by the various probe sets; –f or –r indicate if the probe sets are in forward or reverse orientation relative to the (+) strand of the *E. coli* chromosome; genes previously described as being upregulated in an *rpoS* mutant strain are double underlined.
[b]Gene product (or predicted product).
[c]Relative location of the various ORFs on the *E. coli* MG1655 chromosome.
[d]Determined by microarray analysis as described in 'Materials and Methods' section.
[e]When not otherwise stated, the information is taken from http://ecocyc.org/.

RNA able to modulate $E\sigma^{70}$-dependent at several promoters (36), showed dependence on $E\sigma^{70}$ in ROMA experiments (Table 2). For *ssrS*, our data confirm literature data showing that the main *ssrS* promoter (*ssrS* P1) is strictly $E\sigma^{70}$-dependent *in vitro* (37). Interestingly, several genes preferentially transcribed by $E\sigma^{70}$ in ROMA experiments are upregulated in *rpoS* mutant strains (underlined in Table 2); it has been proposed that negative regulation by σ^{S} can depend on competition with σ^{70} for a limiting amount of RNA polymerase core enzyme, which results in lower intracellular $E\sigma^{70}$ concentrations and, in turn, in impaired transcription initiation at strictly σ^{70}-dependent promoters (38).

In vivo validation of $E\sigma^{S}$-dependent genes identified in ROMA experiments

In order to validate the results of ROMA experiments, we tested *in vivo* expression of 18 genes preferentially

transcribed by $E\sigma^S$, comparing their transcript levels in MG1655 and in its *rpoS* mutant derivative EB1.3. Transcript levels were determined by quantitative real-time PCR. Different genetic backgrounds (MG1655, MC4100 and OH157:H7 EDL933 strains) and different growth conditions (LB medium, glucose-based medium and biofilm growth) can strongly affect *rpoS*-dependent gene expression (9,17–20), suggesting that most genes belonging to the σ^S regulon are also subject to additional forms of gene expression regulation. Thus, *in vivo* gene expression was determined in three different growth media: the peptone-based LB medium, either full strength or diluted 1:4, and the glucose-based M9Glu/sup medium (Table 3). In addition, we took samples from cultures both in late exponential phase and in late stationary phase of growth; indeed, although σ^S-mediated gene expression is typically associated with stationary phase, several *rpoS*-dependent genes are maximally expressed at the transition between exponential and stationary phase or even in mid-exponential phase (19,39; Landini,P., unpublished data).

For five genes (*recT*, *hyaA*, *gabD*, *puuB* and *treF*), dependence on a functional *rpoS* gene has been described (Table 1); however, for the *recT* gene, part of the *recET-lar-ydaCQ* operon, both positive and negative control by *rpoS* was reported (17,18). As a positive control, we tested the expression of *dps* that, although showing only weak dependence on $E\sigma^S$ in our ROMA experiments (Supplementary Table S2), has consistently been described as *rpoS*-dependent *in vivo* in several reports (9,16,17); in agreement with the literature data, *dps* expression showed strong *rpoS*-dependence in all conditions tested (Table 3). We also tested the expression of *cspA*, which was transcribed with higher efficiency by $E\sigma^{70}$ in ROMA experiments (Table 2); consistent with this result, *in vivo* gene expression studies show that a functional *rpoS* allele is not required for *cspA* expression, which is in fact upregulated in the *rpoS* mutant when grown in peptone-based media (Table 3), in agreement with previous observations (20). For *in vivo* gene expression experiments, we considered as significant a fold difference $\geq$2.5 in WT versus *rpoS* mutant relative expression ratio.

Table 3. Gene expression in bacterial cells

Gene	LB		M9Glu/sup		LB 1/4	
	Exp	Stat	Exp	Stat	Exp	Stat
ychS	2.4	0.1	1.1	**6.7**	1.7	2.0
gshA	1.0	1.5	1.0	1.1	1.9	**3.3**
ycbX	1.1	0.5	0.7	2.0	1.5	**2.6**
recT	0.9	0.4	0.7	**2.5**	1.8	2.0
rhaS	1.8	0.3	1.5	**2.7**	1.5	**2.7**
cyaR	1.0	1.7	1.5	**10.0**	0.8	0.6
ilvY	1.7	0.3	1.0	**10.2**	1.9	2.2
hyaA	1.2	**5.6**	2.1	**2.9**	2.3	**17.5**
speB	1.5	1.3	1.4	**2.7**	1.2	**86.3**
gyrB	1.4	0.8	0.9	**2.8**	1.0	**3.1**
gabD	**4.8**	0.7	**12.1**	1.2	**5.1**	**69.0**
ydhY	1.5	0.3	0.9	**19.2**	1.1	2.1
appY	1.2	0.7	0.9	**3.4**	2.0	**11.4**
puuB	1.7	0.2	0.6	**3.5**	**3.0**	1.8
treF	1.5	**3.0**	**7.2**	0.6	2.1	**52.6**
bsmA	1.1	1.0	2.3	0.9	1.0	2.1
xapA	2.4	0.2	0.8	**4.3**	1.1	**2.7**
c3113	1.0	1.4	1.3	**8.1**	**2.7**	1.3
dps	0.9	**10.7**	**10.7**	0.3	**14.2**	**30.3**
cspA	1.3	0.2	1.1	0.8	0.7	0.2

Relative expression is indicated as WT/*rpoS* ratio.
Values higher than 2.5-fold were considered significant and are shown in boldface type.
Values are the average of two independent experiments performed in duplicate.

As shown in Table 3, growth conditions strongly affected gene expression: in LB medium, expression of only three genes (*hyaA*, *gabD* and *treF*) was dependent on a functional *rpoS* allele, while seven genes were in fact upregulated in the *rpoS* mutant, suggesting negative control by σ^S. In contrast, 12 genes were expressed in an *rpoS*-dependent fashion when bacteria were grown in 1:4 diluted LB, while 15 genes showed dependence on *rpoS* in M9Glu/sup medium (Table 3). Of all genes tested, only the *bsmA* gene did not show relative expression values $\geq$2.5 in any growth condition, while *recT* and *ycbX* only displayed weak dependence on a functional *rpoS* allele (2.5- and 2.6-fold increase in one growth condition). All other genes showed either strong dependence on the *rpoS* gene (up to 86.3-fold for *speB* in LB1/4 medium) or were affected by the *rpoS* mutation in more than one growth condition (e.g. *rhaS,* induced 2.7-fold in the wild-type strain both in M9Glu/sup and in LB1/4 growth media). Finally, *in vivo* gene expression experiments demonstrated that the *c3113* gene, tested as a representative of ORFs only annotated in the uropathogenic *E. coli* CFT073, is indeed expressed in MG1655, and it shows dependence on the *rpoS* gene when bacteria are grown in M9Glu/sup (Table 3).

In vitro transcription assays on single promoters

To further confirm the results of ROMA experiments, we performed *in vitro* transcription assays comparing $E\sigma^S$ and $E\sigma^{70}$ on single promoters. To this aim, promoter regions of the *ilvY*, *speB*, *xapA*, *ydhY* and *ssrS* genes were cloned into the pJCD01 vector (22) and the obtained plasmids were used as DNA template for *in vitro* transcription. The *ilvY*, *speB*, *xapA* and *ydhY* genes were selected since their promoter region has been experimentally identified and they show *rpoS*-dependence *in vivo* (Table 3). In contrast, *ssrS* shows the highest dependence on $E\sigma^{70}$ in ROMA experiments among genes whose promoter has been characterized (Table 2). Strong dependence on $E\sigma^{70}$ is in agreement with previous results showing that *ssrS* P1, the main *ssrS* promoter, is $E\sigma^{70}$-dependent *in vitro* (37), and consistent with upregulation of *ssrS* in an *rpoS* mutant of MG1655 (20). Transcription of the different genes of interest was normalized to the RNA-I transcript, as previously described (22), and transcript quantitation was performed by real-time PCR as described in 'Materials and Methods' section.

As shown in Figure 1, *in vitro* transcription experiments performed on plasmids showed that the *ilvY*, *speB*, *xapA* and *ydhY* are transcribed more efficiently in the presence of $E\sigma^S$. Extent of dependence on $E\sigma^S$ was higher than in ROMA experiments, ranging from a 3.4-fold difference for *speB* to a 10.2-fold difference for *xapA*, thus suggesting that, at least at these promoters, DNA supercoiling does not negatively affect promoter recognition by $E\sigma^S$. In contrast, the *ssrS* P1 promoter showed clear dependence on $E\sigma^{70}$, in agreement with previous observations (37). Thus, results of *in vitro* transcription experiments performed on single promoters were fully consistent with ROMA experiments.

Sequence elements involved in specific $E\sigma^S$- versus $E\sigma^{70}$ recognition of promoter regions

The genes identified in ROMA experiments define, at least partially, what can be considered as the 'core σ^S regulon', i.e. a set of genes whose transcription is directly controlled by $E\sigma^S$. The core σ^S regulon can be opposed to the 'expanded σ^S regulon', i.e. genes dependent on a functional *rpoS* allele *in vivo*, whose promoters are, however, not necessarily recognized by $E\sigma^S$. In order to identify sequence features important for specific promoter recognition by $E\sigma^S$, it can be very informative to compare the sequences of $E\sigma^S$-dependent promoters; however, it is important to limit this comparison to the promoters that are exclusively, or at least preferentially, under $E\sigma^S$ control. Thus, we performed a sequence alignment on the promoter sequences of genes preferentially transcribed by either $E\sigma^S$ or $E\sigma^{70}$. Known promoter sequences were retrieved from the Ecocyc database (http://ecocyc.org/): only promoters whose transcription initiation start site had been experimentally determined were considered for sequence alignment. However, since many *rpoS*-dependent genes are controlled by multiple promoters *in vivo*, suggesting complex regulation that might involve different sigma factors, we verified that their transcription start sites in the *in vitro* transcription reactions did indeed correspond to the transcription sites reported in the literature. Thus, using RACE analysis, we determined transcription start sites on *in vitro* transcription assays carried out with $E\sigma^S$ and performed as in the ROMA experiment (data not shown): we were able to identify precisely the *in vitro* transcription start sites at 31 promoter regions controlling 29 different genes, as listed in Supplementary File S1. For most genes with already known promoter regions, we could confirm the transcription start observed *in vivo*, although for several genes reported to be controlled by multiple promoters, only one promoter was found to be recognized by $E\sigma^S$ *in vitro* (e.g. *osmB*, *otsB*, *glgC* and *murQ*). For the *gabD* gene, two of the three promoters described as functional in the bacterial cell were also recognized by $E\sigma^S$ in the *in vitro* transcription assays. In contrast, for the *araF* gene, we identified a second promoter additional to the one already described in the literature. Finally, we determined the transcription start sites for eight genes (or operons) preferentially transcribed by $E\sigma^S$ in ROMA experiments with yet-unknown transcription start sites, namely *puuCBE*, *tam*, *treF*, *yciT*, *yffOP*, *yhjG*, *yi91a* and *yqiG*: their transcription start site and putative promoter elements are listed in Supplementary File S1. For $E\sigma^{70}$-dependent promoters, we verified the transcription start sites for the *ssrS* transcript obtained in ROMA experiments performed with $E\sigma^{70}$: the transcription start site matched the known transcription site for the *ssrS* P1 promoter previously identified (37).

For promoter sequence comparison, we considered DNA sequences extending from −100 to +2 bp relative to the transcription start site, i.e. an area that includes all promoter elements described for $E\sigma^{70}$. For sequence analysis, we divided the −100 to +2 promoter sequence in

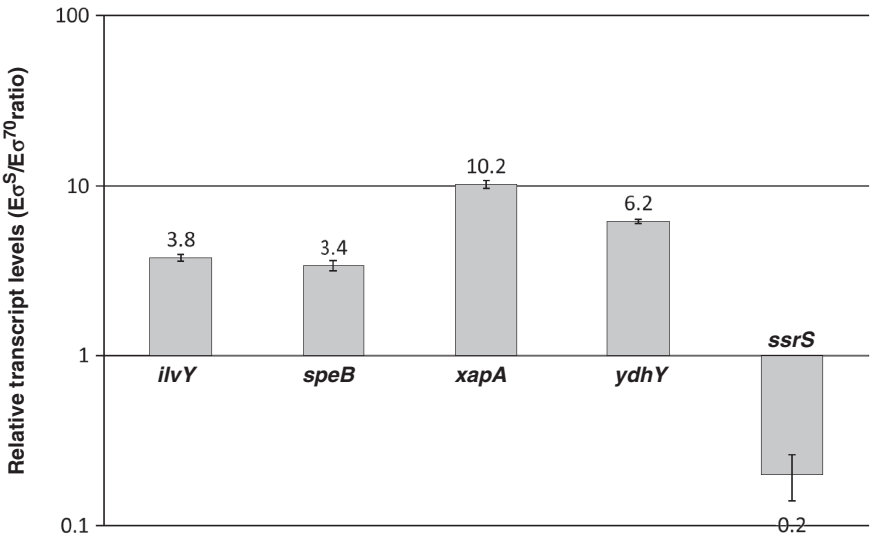

Figure 1. *In vitro* transcription experiments on supercoiled plasmids. Transcription from the *ilvY*, *speB*, *xapA*, *ydhY* and *ssrS* P1 promoter regions cloned into pJCD01 plasmid was performed in the presence of either $E\sigma^S$ or $E\sigma^{70}$; transcript amounts were determined by quantitative real-time PCR as described in 'Materials and Methods' section, using the RNA-I transcript as reference as previously described (22). Relative transcript levels are shown as $E\sigma^S/E\sigma^{70}$ ratio. Experiments were performed three times in duplicate, and standard errors are shown.

three parts: the −17 to +2 region, carrying the −10 element and the transcription start site; the −60 to −18 region, containing the UP element and the −35 sequence; and the −100 to −61 region. Alignment of the −17 to +2 regions was centred on the first nucleotide of the −10 element (conventionally referred to as the −12 position, Figure 2B). For the −60 to −18 regions, the alignment was centred either on the first nucleotide of the −35 element (when present) or on the nucleotide located 22 bp upstream of the −10 element, which would correspond to the first nucleotide of a hypothetical −35 element placed at the optimal 17 bp distance from the −10 sequence (Figure 2A). In total,

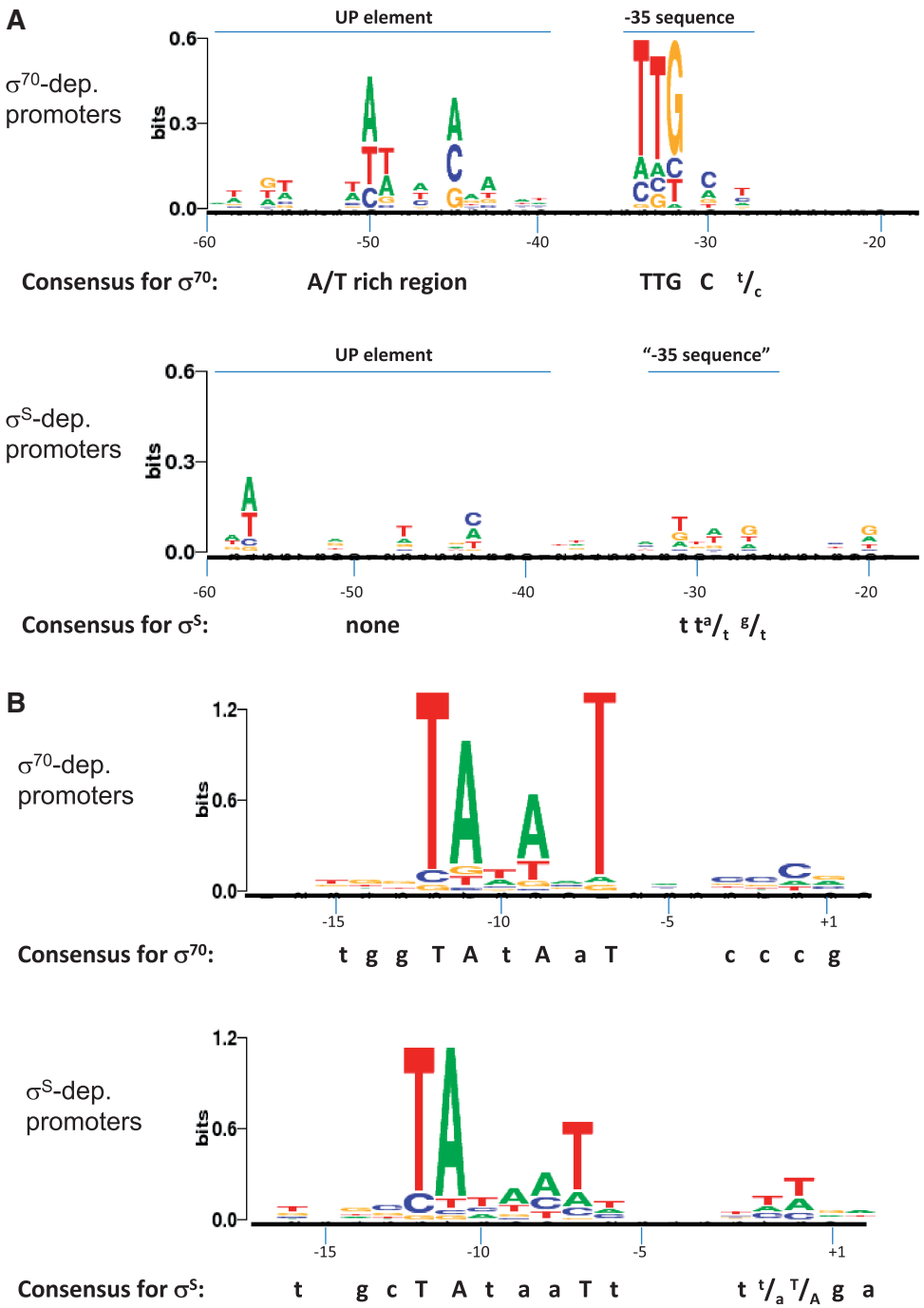

Figure 2. Conserved sequence features in promoters of genes showing preferential transcription by either Eσ^{70} or Eσ^{S}, shown as a sequence logo derived from multiple sequence alignments. (**A**) Comparison of alignments in the −60 to −18 promoter regions. (**B**) Comparison of alignments in the −17 to +2 promoter regions. Note that different *y*-axis scales were used in the two panels to account for the different levels of sequence conservation. Multiple alignment included 31 promoters controlling genes preferentially transcribed by Eσ^{S} in ROMA experiments (listed in Supplementary File S1) and 29 promoters preferentially recognized by Eσ^{70} (listed in Supplementary File S2).

31 promoters showing preferential recognition by $E\sigma^S$ (listed in Supplementary File S1) and 30 promoters preferentially recognized by $E\sigma^{70}$ (Supplementary File S2) were selected for sequence analysis. Consensus sequence conservation within promoters preferentially recognized by either $E\sigma^S$ or $E\sigma^{70}$ was displayed as sequence logos using the Weblogo 2.8.2 application (40) (http://weblogo.berkeley.edu/). Results of sequence analysis showed that, as expected, the most conserved sequences corresponded to the known promoter elements and were comprised between −60 and +2 (Figure 2). Comparison of the sequences located further upstream (−100 to −61) is shown in Supplementary Figure S2.

For promoters more efficiently recognized by $E\sigma^{70}$, the −35 and −10 sequences, i.e. the most important $E\sigma^{70}$-dependent promoter elements, are clearly recognizable in the sequence logos (Figure 2). In addition, the −60/−40 region is characterized by a high occurrence of A and T residues (Figure 2A), consistent with strong conservation of the UP element, i.e. a binding site for the α subunit of RNA polymerase (41). Other than the already known promoter features for σ^{70}, only the transcription start site (−3/+1) showed a moderately conserved sequence (CCCG, Figure 2B).

Several differences in conserved regions were detectable in the promoter set for genes more efficiently transcribed by $E\sigma^S$ in ROMA experiments. The −10 sequence is clearly the main conserved promoter element in this set (Figure 2B). No conserved sequences were detectable in the −35 region (Figure 2A), in agreement with previous works reporting that the −35 element does not play an important role in promoter recognition by $E\sigma^S$ (42,43), while a weakly conserved sequence similar to a −35 element seems to be located at around −30 (Figure 2A). Another difference between the two promoter sets resides in the lack of an A/T rich region between −60 and −40 (Figure 2A). In contrast, however, promoters more efficiently transcribed by $E\sigma^S$ seem to possess an increased occurrence of T residues in the −90 to −70 region (Supplementary Figure S2). Finally, the discriminator, i.e. the sequence located between the −10 element and the transcription start site, appears to be biased towards a high A/T content in promoters transcribed more efficiently by $E\sigma^S$. In particular, T residues appear to be conserved at positions −6, −2 and −1. Rather than being associated to specific nucleotide position, however, the bias towards an A/T-rich discriminator seems to be a common feature in promoters of genes preferentially transcribed by $E\sigma^S$. Indeed, while the average length of the discriminator is identical for both $E\sigma^{70}$- and $E\sigma^S$-dependent promoters (6.1 bp, Supplementary Table S3), the GC content is significantly higher in $E\sigma^{70}$-dependent promoters (0.58 versus 0.41, Supplementary Table S3).

Finally, an additional deviation between $E\sigma^{70}$- and $E\sigma^S$-specific promoter elements could be observed immediately upstream of the −10 hexamer: the $E\sigma^{70}$-dependent promoters showed some conservation of a TGG motif upstream of the −10 [the 'extended −10′ (44)], which was replaced by a TNGC motif, which includes the characteristic −13C element, in $E\sigma^S$-dependent promoters.

Mutations in the discriminator region affect selective promoter recognition at the $E\sigma^{70}$-dependent promoter *ssrS* P1

It might be inferred that differences in the specific elements of the promoter regions might account for different transcription efficiency by either $E\sigma^S$ or $E\sigma^{70}$ in ROMA experiments. In particular, we investigated whether the presence of an A/T rich region immediately downstream of the −10 element, conserved among promoters preferentially recognized by $E\sigma^S$ in ROMA experiment (Figure 2B), could indeed be a determinant for selective promoter recognition by $E\sigma^S$. To test this possibility, we targeted for mutagenesis the *ssrS* P1 promoter, which shows preferential recognition by $E\sigma^{70}$ in our *in vitro* assays (Table 1 and Figure 1) and whose dependence on $E\sigma^{70}$ had already been reported in the literature (37). We changed the CG nucleotides at positions −3/−2 of the *ssrS* P1 promoter to TA (Figure 3), thus making the −6/+1 region of the promoter A/T-rich. *In vitro* transcription experiments on supercoiled templates in the presence of either $E\sigma^S$ or $E\sigma^{70}$ showed that CG to TA substitutions at positions −3/−2 of the *ssrS* P1 promoter resulted in loss of specific recognition by $E\sigma^{70}$ from 3.5- to 1.6-fold (Figure 3), suggesting that the A/T content in the −6/+1 region can indeed play a role in modulating transcription efficiency by either $E\sigma^S$ or $E\sigma^{70}$.

DISCUSSION

In this work, we have attempted to identify *bona fide* $E\sigma^S$-dependent promoters through *in vitro* transcription experiments using the whole *E. coli* genome as template, followed by identification of $E\sigma^S$-dependent transcripts

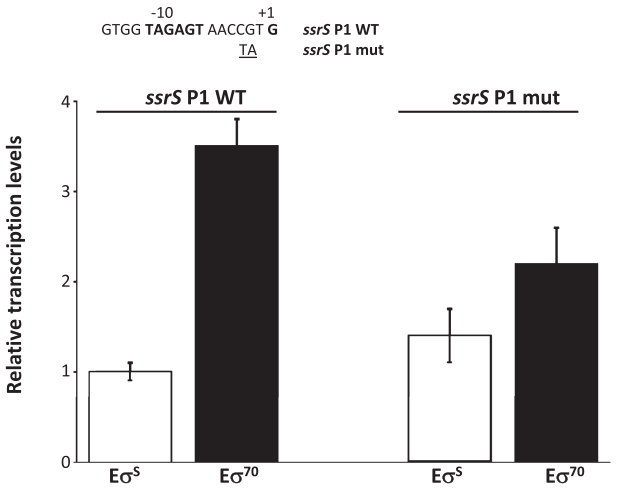

Figure 3. *In vitro* transcription experiments on pJCD01 plasmid derivatives in which either the *ssrS* P1 promoter (*ssrS* P1 WT) or a mutated derivative (*ssrS* P1 mut) had been cloned. The *ssrS* P1 mut carries a double substitution (CG to TA) at positions −3/−2 (shown in the figure). *In vitro* transcription was performed in the presence of either $E\sigma^S$ or $E\sigma^{70}$; transcript amounts were determined by quantitative real-time PCR as described in 'Materials and Methods' section, using the RNA-I transcript as reference as previously described (22). Experiments were performed three times in duplicate, and standard errors are shown.

by microarray analysis (ROMA experiments). The *in vitro* transcription experiments have been performed on linear DNA, in the absence of any additional factor or molecule able to affect transcription initiation by either $E\sigma^S$ or $E\sigma^{70}$ (i.e. DNA supercoiling, transcription regulators, histone-like proteins, ppGpp), and thus they represent a direct measurement of sequence-specific interactions between promoters and RNA polymerase. Our results indicate that selective promoter utilization by $E\sigma^{70}$ and $E\sigma^S$ is mediated by specific sequence features, and confirm the role of σ^S-specific promoter elements previously identified. Promoters more efficiently recognized by $E\sigma^{70}$ in ROMA experiments are characterized by a high occurrence of the UP-like element, i.e. an A/T-rich region located immediately upstream of the −35 sequence and acting as a binding site for the α subunit of RNA polymerase (41), suggesting that, at least *in vitro*, the UP element might favour promoter recognition by $E\sigma^{70}$. It is worth mentioning that, in *Bacillus subtilis*, UP elements are highly conserved among σ^A-dependent promoters (45), thus suggesting that UP elements might favour promoter recognition by the housekeeping σ factor in different bacteria. In contrast, UP-like elements flanking the −35 sequence are less conserved in promoters better recognized by $E\sigma^S$, where, however, A/T-rich elements seem to be scattered in the region spanning 70−90 nt upstream of the transcription start (Supplementary Figure S2). It might be speculated that, similar to the UP element for $E\sigma^{70}$-dependent promoters, AT-rich sequences located in the −70 to −90 region might also be involved in interaction with RNA polymerase α subunit at σ^S-dependent promoters. It is conceivable that the α subunit might contact alternative upstream promoter elements when assembled in different forms of RNA polymerase holoenzyme; indeed, differential ability to interact with UP elements has already been described for $E\sigma^{70}$ and $E\sigma^S$ (46).

A sequence element showing strong differences between the two promoter sets is the discriminator, i.e. the region spanning between −10 and +1. This region shows a high T/A content for promoters more efficiently recognized by $E\sigma^S$; in contrast, $E\sigma^{70}$-dependent promoters identified in ROMA experiments are biased towards a CG-rich region in the −3/+1 residues (CCCG, Figure 2). Mutations in the discriminator region of the $E\sigma^{70}$-dependent *ssrS* P1 promoter increasing its A/T content result in partial loss of preferential recognition by $E\sigma^{70}$ (Figure 3), providing further confirmation for a role of the −10/+1 region in specific recognition by either form of RNA polymerase. GC versus AT content in the discriminator region can directly affect promoter melting, and GC-rich discriminators are a common feature among promoters subject to negative regulation by ppGpp (47,48), which, indeed, inhibits $E\sigma^{70}$-dependent transcription while promoting transcription by alternative forms of RNA polymerase (49).

The sequence features indicated by our promoter analysis are consistent with previous observations based on sequence analysis of *rpoS*-dependent promoters in the bacterial cell (4); this also includes the fairly strong conservation of a C nucleotide immediately upstream of the −10 hexamer (the −13 C element) among σ^S-dependent promoters. Indeed, the −13 C element occurs in almost half of the promoters transcribed more efficiently by $E\sigma^S$ in ROMA experiments (14 out of 31 promoters; Supplementary File S1), opposed to a much lower frequency in $E\sigma^{70}$-dependent promoters (4 out of 29, Supplementary File S2). This result is not particularly surprising: indeed, a large amount of data has clearly shown that the presence of a C residue immediately upstream of the −10 nt favours promoter recognition by $E\sigma^S$ (7,8,50). The −13C element might play a similar role as the TG motif at $E\sigma^{70}$-dependent promoters lacking a −35 region (44). It is noteworthy that an *in vivo* analysis suggests that the −13C element occurs in more than 70% of putative *rpoS*-dependent promoters (4): this observation would suggest that the −13C element might be needed to improve $E\sigma^S$–promoter interaction in the bacterial cell, possibly to overcome the negative effects of DNA-binding proteins, such as H-NS, that can modulate RNA polymerase–promoter interaction.

The results of the ROMA approach have also expanded our knowledge of the σ^S regulon: the σ^S protein is considered a central element of the so-called 'general stress response' (51). Intracellular σ^S concentration and expression of σ^S-dependent genes respond to reduction in growth rate (52); thus, any cellular stress affecting growth rate is likely to induce σ^S accumulation, which in turn plays a direct role in oxidative, acid and osmotic stress through activation of specific genes. In addition, σ^S activates metabolic genes associated to stationary-phase metabolism, in particular carbon storage genes involved in glycogen (19,20) and trehalose metabolism (9,20); consistent with these observations, stress response and carbon metabolism genes are highly represented among genes preferentially transcribed by $E\sigma^S$ in ROMA experiments (Table 1). In addition, our results underline the importance of σ^S for the expression of genes involved in polyamine metabolism. Polyamines, in particular putrescine, play an important role in various cellular processes and can affect intracellular concentrations of the regulatory proteins σ^N, Cra and H-NS, thus impacting global gene expression (53). As shown in Figure 4, putrescine, the most abundant polyamine in bacterial cells (54), is a product of arginine degradation. Putrescine accumulates at the transition between exponential and stationary phase (55), and it can be subsequently converted into other polyamines or degraded to succinate, which can be shunted into the tricarboxylic acid (TCA) cycle, in a NADPH-generating process (Figure 4). Interestingly, mutants unable to synthesize polyamines are more sensitive to oxidative stress, since they cannot induce *ahpC*, *katE* and *katG* genes, encoding three different peroxidase, suggesting that polyamine accumulation might control expression of *rpoS*-dependent genes involved in the response to oxidative stress (54). In turn, our results indicate that genes encoding enzymes involved in both accumulation and degradation of putrescine, one of the main polyamines found in the bacterial cell, are *rpoS*-dependent. Indeed, the *speB* gene, encoding the putrescine-biosynthetic enzyme agmatinase, is preferentially transcribed by $E\sigma^S$ *in vitro* (Table 1 and Figure 1) and is

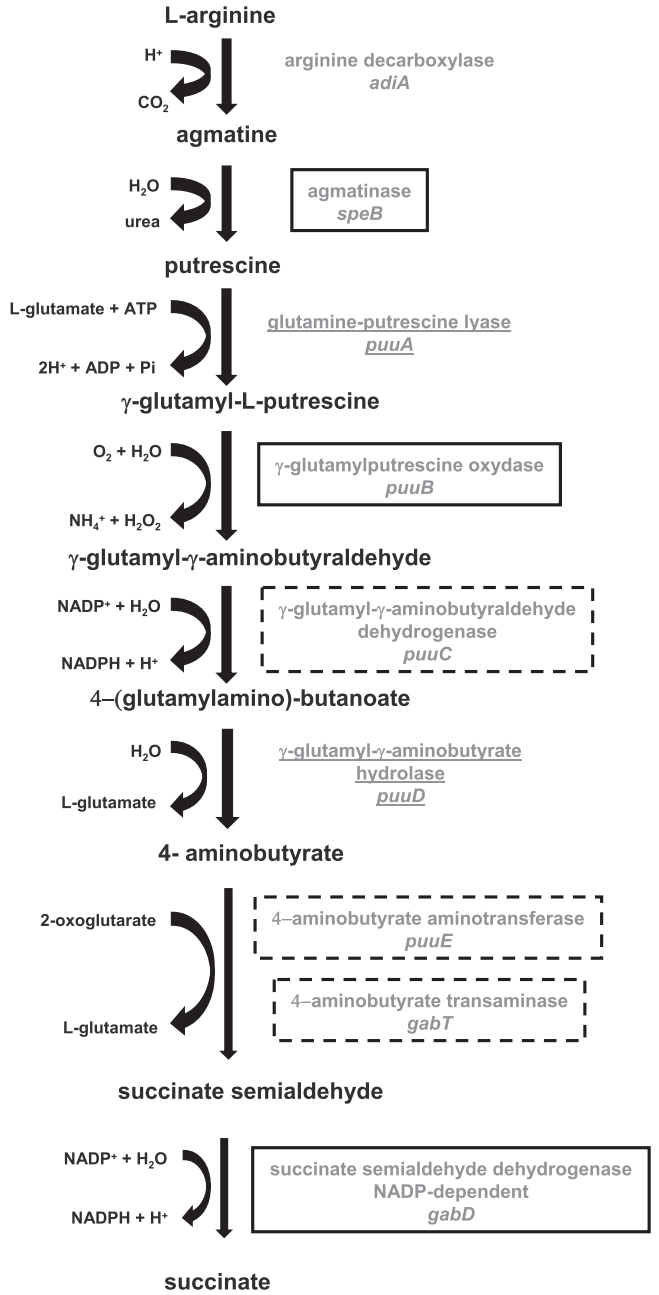

Figure 4. Schematic representation of a biochemical pathway under the control of the *rpoS* gene: arginine degradation to succinate via putrescine (adapted from EcoCyc; http://ecocyc.org/). Genes belonging to this pathway and found to be preferentially transcribed by $E\sigma^S$ in ROMA experiments are boxed with solid lines (*speB*, *puuB* and *gabD*); genes co-transcribed with either *puuB* or *gabD* (i.e. belonging to the *puuCBE* and *gabDTP* operons; see also Supplementary Table 1) are boxed with dashed lines. Other genes of the pathway known to be dependent on the *rpoS* gene *in vivo* (*puuA* and *puuD*) (20,55) are underlined.

rpoS-dependent *in vivo*, in particular in LB1/4 medium (Table 3). Likewise, the genes involved in putrescine degradation, belonging to the *puuCBE* and *gabDTP* operons, are efficiently transcribed by $E\sigma^S$ in ROMA experiments (Table 1 and Supplementary Table S1) and show dependence on a functional *rpoS* gene *in vivo* (Table 3). The *gabDTP* operon, as well as the *puuA* and *puuD* genes, also involved in putrescine degradation, have already been reported to be *rpoS*-dependent *in vivo* (20,55,56). Thus, σ^S could control every step in putrescine metabolism by regulating genes responsible for both its biosynthesis and its degradation, as shown in Figure 4. Activation of putrescine biosynthesis rather than degradation might respond to different growth conditions and environmental cues, as also suggested by the very different levels of *rpoS*-dependent regulation of putrescine-related genes (i.e. *speB*, *gabD* and *puuB*) in different growth media (Table 3).

rpoS-dependent control of intracellular polyamines concentrations represents an important mechanism of indirect gene regulation by the σ^S protein, since polyamines act as signal molecules able to impact global transcription pattern in the bacterial cell (53). Indirect control of gene expression by σ^S can also occur through activation of regulatory proteins and of regulatory RNAs. Indeed, ROMA experiments would suggest that $E\sigma^S$ directly controls at least four genes encoding regulatory proteins (*appY*, *ilvY*, *rhaS* and *yciT*) and four genetic loci encoding regulatory RNAs (*cyaR*, *micA*, *sgrS*/*sgrT* and *ECs3934*) (Table 1). Interestingly, the *sgrS*/*sgrT* locus, encoding both a non-coding RNA (SgrS) and a small regulatory protein (SgrT), negatively affects expression of the *ptsG* glucose uptake system (57), which is overexpressed in *rpoS* mutant strains of *E. coli* (20,58). However, the extent of σ^S-dependent indirect regulation of gene expression through non-coding RNA might not be limited to these four loci. Indeed, a significant number of intergenic regions, often located immediately upstream of known ORFs in antisense direction, was efficiently transcribed by $E\sigma^S$ (Table 1). It is possible that at least a part of these intergenic regions might indeed be transcribed in the bacterial cell and function as *cis*-acting regulatory RNAs. Similarly, several ORFs only annotated in pathogenic *E. coli* strains were detectable as transcripts in the ROMA experiment, although MG1655 genomic DNA had been used in the experiments. Sequence comparison allowed us to determine that these ORFs are indeed present in MG1655, but they are not accounted for in the available databases. As observed for intergenic regions efficiently transcribed by $E\sigma^S$, several of these ORFs also overlap known genes in an antisense direction, suggesting that at least some of them might be involved in modulating gene expression. One ORF, *c3113*, partially overlaps the promoter region of the *rrsG* ribosomal operon, and quantitative real-time PCR experiments suggest that it is transcribed in a stationary-phase-dependent and *rpoS*-dependent manner in bacterial cells grown in glucose-based medium (Table 3). Similar to *c3113*, an intergenic region showing preferential transcription by $E\sigma^S$ (*IG223409_223770-r*) also partially overlaps the ribosomal operon *rrsH*, but in an antisense direction. Thus, our observations would suggest that small proteins or non-coding RNAs might play an important role in *rpoS*-dependent negative regulation of genes associated with fast growth, such as ribosomal operons and the *ptsG* glucose uptake system.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

The Swiss National Fund for Scientific Research (Project 3100A0-109433 to T.E. and P.L.); Italian Ministry for University and Research (CHEM-PROFARMA-NET Research Program, Project RBPR05NWWC_004 to P.L., and FIRB-MIUR NG-LAB Research Program, Project RBLA03ER38_004 to G.D.B.). Funding for open access charge: Italian Ministry for University and Research (MIUR).

Conflict of interest statement. None declared.

REFERENCES

1. Lloyd,G., Landini,P. and Busby,S. (2001) Activation and repression of transcription initiation in bacteria. *Essays Biochem.*, **37**, 17–31.
2. Gourse,R.L., Ross,W. and Rutherford,S.T. (2006) General pathway for turning on promoters transcribed by RNA polymerases containing alternative sigma factors. *J. Bacteriol.*, **188**, 4589–4591.
3. Hengge-Aronis,R. (2002) Stationary phase gene regulation: what makes an *Escherichia coli* promoter sigmaS-selective? *Curr. Opin. Microbiol.*, **5**, 591–595.
4. 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.
5. Gaal,T., Ross,W., Estrem,S.T., Nguyen,L.H., Burgess,R.R. and Gourse,R.L. (2001) Promoter recognition and discrimination by EsigmaS RNA polymerase. *Mol. Microbiol.*, **42**, 939–954.
6. Tanaka,K., Takayanagi,Y., Fujita,N., Ishihama,A. and Takahashi,H. (1993) Heterogeneity of the principal sigma factor in *Escherichia coli*: the *rpoS* gene product, sigma38, is a second principal sigma factor of RNA polymerase in stationary-phase *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **90**, 8303.
7. Becker,G. and Hengge-Aronis,R. (2001) What makes an *Escherichia coli* promoter sigma(S) dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of sigma(S). *Mol. Microbiol.*, **39**, 1153–1165.
8. Lacour,S., Kolb,A. and Landini,P. (2003) Nucleotides from -16 to -12 determine specific promoter recognition by bacterial sigmaS-RNA polymerase. *J. Biol. Chem.*, **278**, 37160–37168.
9. 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*: sigmaS-dependent genes, promoters, and sigma factor selectivity. *J. Bacteriol.*, **187**, 1591–1603.
10. Colland,F., Barth,M., Hengge-Aronis,R. and Kolb,A. (2000) sigma factor selectivity of *Escherichia coli* RNA polymerase: role for CRP, IHF and Lrp transcription factors. *EMBO J.*, **19**, 3028–3037.
11. Arnqvist,A., Olsen,A. and Normark,S. (1994) Sigma S-dependent growth-phase induction of the *csgBA* promoter in *Escherichia coli* can be achieved *in vivo* by sigma 70 in the absence of the nucleoid-associated protein H-NS. *Mol. Microbiol.*, **13**, 1021–1032.
12. Mujacic,M. and Baneyx,F. (2006) Regulation of *Escherichia coli hchA*, a stress-inducible gene encoding molecular chaperone Hsp31. *Mol. Microbiol.*, **60**, 1576–1589.
13. Kusano,S., Ding,Q., Fujita,N. and Ishihama,A. (1996) Promoter selectivity of *Escherichia coli* RNA polymerase Esigma70 and Esigma38 holoenzymes. Effect of DNA supercoiling. *J. Biol. Chem.*, **271**, 1998–2004.
14. Jishage,M. and Ishihama,A. (1998) A stationary phase protein in *Escherichia coli* with binding activity to the major sigma subunit of RNA polymerase. *Proc. Natl Acad. Sci. USA*, **95**, 4953–4958.
15. Gentry,D.R., Hernandez,V.J., Nguyen,L.H., Jensen,D.B. and Cashel,M. (1993) Synthesis of the stationary-phase sigma factor sigmaS is positively regulated by ppGpp. *J. Bacteriol.*, **175**, 7982–7989.
16. Lacour,S. and Landini,P. (2004) SigmaS-dependent gene expression at the onset of stationary phase in *Escherichia coli*: function of sigmaS-dependent genes and identification of their promoter sequences. *J. Bacteriol.*, **186**, 7186–7195.
17. Patten,C.L., Kirchhof,M.G., Schertzberg,M.R., Morton,R.A. and Schellhorn,H.E. (2004) Microarray analysis of RpoS-mediated gene expression in *Escherichia coli* K-12. *Mol. Genet. Genomics*, **272**, 580–591.
18. Ito,A., May,T., Kawata,K. and Okabe,S. (2008) Significance of *rpoS* during maturation of *Escherichia coli* biofilms. *Biotechnol. Bioeng.*, **99**, 1462–1471.
19. Dong,T. and Schellhorn,H.E. (2009) Control of RpoS in global gene expression of *Escherichia coli* in minimal media. *Mol. Genet. Genomics*, **281**, 19–33.
20. Dong,T. and Schellhorn,H.E. (2009) Global effect of RpoS on gene expression in pathogenic *Escherichia coli* O157:H7 strain EDL933. *BMC Genomics*, **10**, 349.
21. Hiratsu,K., Shinagawa,H. and Makino,K. (1995) Mode of promoter recognition by the *Escherichia coli* RNA polymerase holoenzyme containing the sigmaS subunit: identification of the recognition sequence of the *fic* promoter. *Mol. Microbiol.*, **18**, 841–850.
22. Marschall,C., Labrousse,V., Kreimer,M., Weichart,D., Kolb,A. and Hengge-Aronis,R. (1998) Molecular analysis of the regulation of *csiD*, a carbon starvation-inducible gene in *Escherichia coli* that is exclusively dependent on sigmaS and requires activation by cAMP-CRP. *J. Mol. Biol.*, **276**, 339–353.
23. Landini,P., Hajec,L.I., Nguyen,L.H., Burgess,R.R. and Volkert,M.R. (1996) The leucine-responsive regulatory protein (Lrp) acts as a specific repressor for sigmaS-dependent transcription of the *Escherichia coli aidB* gene. *Mol. Microbiol.*, **20**, 947–955.
24. Checroun,C., Bordes,P., Leroy,O., Kolb,A. and Gutierrez,C. (2004) Interactions between the 2.4 and 4.2 regions of sigmaS, the stress-specific sigma factor of *Escherichia coli*, and the -10 and -35 promoter elements. *Nucleic Acids Res.*, **32**, 45–53.
25. Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F. *et al*. (1997) The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1462.
26. Argaman,L., Hershberg,R., Vogel,J., Bejerano,G., Wagner,E.G., Margalit,H. and Altuvia,S. (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. *Curr. Biol.*, **11**, 941–950.
27. Li,X.M. and Shapiro,L.J. (1993) Three step PCR mutagenesis for "linker scanning". *Nucleic Acids Res.*, **21**, 3745–3478.
28. Wettenhall,J.M., Simpson,K.M., Satterley,K. and Smyth,G.K. (2006) affylmGUI: a graphical user interface for linear modeling of single channel microarray data. *Bioinformatics*, **22**, 897–899.
29. Zheng,D., Constantinidou,C., Hobman,J.L. and Minchin,S.D. (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res.*, **32**, 5874–5893.
30. Prigent-Combaret,C., Brombacher,E., Vidal,O., Ambert,A., Lejeune,P., Landini,P. and Dorel,C. (2001) Complex regulatory network controls initial adhesion and biofilm formation in *Escherichia coli* via regulation of the *csgD* gene. *J. Bacteriol.*, **183**, 7213–7223.
31. Brombacher,E., Dorel,C., Zehnder,A.J. and Landini,P. (2003) The curli biosynthesis regulator CsgD co-ordinates the expression of both positive and negative determinants for biofilm formation in *Escherichia coli*. *Microbiology*, **149**, 2847–2857.
32. Perrin,C., Briandet,R., Jubelin,G., Lejeune,P., Mandrand-Berthelot,M.A., Rodrigue,A. and Dorel,C. (2009) Nickel promotes biofilm formation by *Escherichia coli* K-12 strains that produce curli. *Appl. Environ. Microbiol.*, **75**, 1723–1733.
33. Gualdi,L., Tagliabue,L. and Landini,P. (2007) Biofilm formation-gene expression relay system in *Escherichia coli*: modulation of sigmaS-dependent gene expression by the CsgD regulatory protein via sigmaS protein stabilization. *J. Bacteriol.*, **189**, 8034–8043.

34. Cao,M., Kobel,P.A., Morshedi,M.M., Wu,M.F., Paddon,C. and Helmann,J.D. (2002) Defining the *Bacillus subtilis* sigmaW regulon: a comparative analysis of promoter consensus search, run-off transcription/macroarray analysis (ROMA), and transcriptional profiling approaches. *J. Mol. Biol.*, **316**, 443–457.
35. Cashel,M., Gentry,D.R., Hernandez,V.J. and Vinella,D. (1996) *The Stringent Response. In Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. American society for Microbiology, Washington DC.
36. Wassarman,K.M. and Storz,G. (2000) 6S RNA regulates *E. coli.* RNA polymerase activity. *Cell*, **101**, 613–623.
37. Kim,K.S. and Lee,Y. (2004) Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.*, **32**, 6057–6068.
38. Farewell,A., Kvint,K. and Nystrom,T. (1998) Negative regulation by RpoS: a case of sigma factor competition. *Mol. Microbiol.*, **29**, 1039–1051.
39. Dong,T., Kirchhof,M.G. and Schellhorn,H.E. (2008) RpoS regulation of gene expression during exponential growth of *Escherichia coli* K12. *Mol. Genet. Genomics*, **279**, 267–277.
40. Crooks,G.E., Hon,G., Chandonia,J.M. and Brenner,S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.*, **14**, 1188–1190.
41. 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.
42. Espinosa-Urgel,M., Chamizo,C. and Tormo,A. (1996) A consensus structure for sigma S-dependent promoters. *Mol. Microbiol.*, **21**, 657–659.
43. Lee,S.J. and Gralla,J.D. (2001) Sigma38 (*rpoS*) RNA polymerase promoter engagement via -10 region nucleotides. *J. Biol. Chem.*, **276**, 30064–30071.
44. 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.
45. Helmann,J.D. (1995) Compilation and analysis of *Bacillus subtilis* sigmaA-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. *Nucleic Acids Res.*, **23**, 2351–2360.
46. Typas,A. and Hengge,R. (2005) Differential ability of sigmaS and sigma70 of *Escherichia coli* to utilize promoters containing half or full UP-element sites. *Mol. Microbiol.*, **55**, 250–260.
47. Travers,A.A. (1980) Promoter sequence for stringent control of bacterial ribonucleic acid synthesis. *J. Bacteriol.*, **141**, 973–976.
48. Travers,A.A. (1980) A tRNATyr promoter with an altered *in vitro* response to ppGpp. *J. Mol. Biol.*, **141**, 91–97.
49. Condon,C., Squires,C. and Squires,C.L. (1995) Control of rRNA transcription in *Escherichia coli. Microbiol. Rev.*, **59**, 623–645.
50. Bordes,P., Repoila,F., Kolb,A. and Gutierrez,C. (2000) Involvement of differential efficiency of transcription by EsigmaS and Esigma70 RNA polymerase holoenzymes in growth phase regulation of the *Escherichia coli osmE* promoter. *Mol. Microbiol.*, **35**, 845–853.
51. Hengge-Aronis,R. (2000) A role for the sigmaS subunit of RNA polymerase in the regulation of bacterial virulence. *Adv. Exp. Med. Biol.*, **485**, 85–93.
52. Ihssen,J. and Egli,T. (2004) Specific growth rate and not cell density controls the general stress response in *Escherichia coli. Microbiology*, **150**, 1637–1648.
53. Terui,Y., Higashi,K., Taniguchi,S., Shigemasa,A., Nishimura,K., Yamamoto,K., Kashiwagi,K., Ishihama,A. and Igarashi,K. (2007) Enhancement of the synthesis of RpoN, Cra, and H-NS by polyamines at the level of translation in *Escherichia coli* cultured with glucose and glutamate. *J .Bacteriol.*, **189**, 2359–2368.
54. Jung,I.L. and Kim,I.G. (2003) Transcription of *ahpC*, *katG*, and *katE* genes in *Escherichia coli* is regulated by polyamines: polyamine-deficient mutant sensitive to H_2O_2-induced oxidative damage. *Biochem. Biophys. Res. Commun.*, **301**, 915–922.
55. Kurihara,S., Oda,S., Kumagai,H. and Suzuki,H. (2006) Gamma-glutamyl-gamma-aminobutyrate hydrolase in the putrescine utilization pathway of *Escherichia coli* K-12. *FEMS Microbiol. Lett.*, **256**, 318–323.
56. Mendoza-Vargas,A., Olvera,L., Olvera,M., Grande,R., Vega-Alvarado,L., Taboada,B., Jimenez-Jacinto,V., Salgado,H., Juarez,K., Contreras-Moreira,B. *et al.* (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli. PLoS ONE*, **4**, e7526.
57. Vanderpool,C.K. and Gottesman,S. (2004) Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol. Microbiol.*, **54**, 1076–1089.
58. Seeto,S., Notley-McRobb,L. and Ferenci,T. (2004) The multifactorial influences of RpoS, Mlc and cAMP on *ptsG* expression under glucose-limited and anaerobic conditions. *Res. Microbiol.*, **155**, 211–215.
59. Rhodius,V.A., Suh,W.C., Nonaka,G., West,J. and Gross,C.A. (2006) Conserved and variable functions of the sigmaE stress response in related genomes. *PLoS Biol.*, **4**, e2.
60. Boulanger,A., Francez-Charlot,A., Conter,A., Castanie-Cornet,M.P., Cam,K. and Gutierrez,C. (2005) Multistress regulation in *Escherichia coli*: expression of *osmB* involves two independent promoters responding either to sigmaS or to the RcsCDB His-Asp phosphorelay. *J. Bacteriol.*, **187**, 3282–3286.
61. Storz,G., Tartaglia,L.A. and Ames,B.N. (1990) Transcriptional regulator of oxidative stress-inducible genes: direct activation by oxidation. *Science*, **248**, 189–194.
62. Missiakas,D. and Raina,S. (1997) Signal transduction pathways in response to protein misfolding in the extracytoplasmic compartments of *E. coli:* role of two new phosphoprotein phosphatases PrpA and PrpB. *EMBO J.*, **16**, 1670–1685.
63. Missiakas,D. and Raina,S. (1997) Protein misfolding in the cell envelope of *Escherichia coli*: new signaling pathways. *Trends Biochem. Sci.*, **22**, 59–63.
64. Seeger,C., Poulsen,C. and Dandanell,G. (1995) Identification and characterization of genes (*xapA*, *xapB*, and *xapR*) involved in xanthosine catabolism in *Escherichia coli. J. Bacteriol.*, **177**, 5506–5516.
65. Schneider,R., Travers,A., Kutateladze,T. and Muskhelishvili,G. (1999) A DNA architectural protein couples cellular physiology and DNA topology in *Escherichia coli. Mol. Microbiol.*, **34**, 953–964.
66. Metzner,M., Germer,J. and Hengge,R. (2004) Multiple stress signal integration in the regulation of the complex sigmaS-dependent *csiD-ygaF-gabDTP* operon in *Escherichia coli. Mol. Microbiol.*, **51**, 799–811.
67. Atlung,T., Sund,S., Olesen,K. and Brondsted,L. (1996) The histone-like protein H-NS acts as a transcriptional repressor for expression of the anaerobic and growth phase activator AppY of *Escherichia coli. J. Bacteriol.*, **178**, 3418–3425.
68. Wickstrum,J.R., Skredenske,J.M., Kolin,A., Jin,D.J., Fang,J. and Egan,S.M. (2007) Transcription activation by the DNA-binding domain of the AraC family protein RhaS in the absence of its effector-binding domain. *J. Bacteriol.*, **189**, 4984–4993.
69. Hendrickson,W., Stoner,C. and Schleif,R. (1990) Characterization of the *Escherichia coli araFGH* and *araJ* promoters. *J. Mol. Biol.*, **215**, 497–510.
70. Romeo,T. and Preiss,J. (1989) Genetic regulation of glycogen biosynthesis in *Escherichia coli*: *in vitro* effects of cyclic AMP and guanosine 5'-diphosphate 3'-diphosphate and analysis of *in vivo* transcripts. *J. Bacteriol.*, **171**, 2773–2782.
71. Atlung,T., Knudsen,K., Heerfordt,L. and Brondsted,L. (1997) Effects of sigmaS and the transcriptional activator AppY on induction of the *Escherichia coli hya* and *cbdAB-appA* operons in response to carbon and phosphate starvation. *J. Bacteriol.*, **179**, 2141–2146.
72. Richard,D.J., Sawers,G., Sargent,F., McWalter,L. and Boxer,D.H. (1999) Transcriptional regulation in response to oxygen and nitrate of the operons encoding the [NiFe] hydrogenases 1 and 2 of *Escherichia coli. Microbiology*, **145**, 2903–2912.
73. Jaeger,T. and Mayer,C. (2008) The transcriptional factors MurR and catabolite activator protein regulate N-acetylmuramic acid catabolism in *Escherichia coli. J. Bacteriol.*, **190**, 6598–6608.
74. Cai,H. and Clarke,S. (1999) A novel methyltransferase catalyzes the methyl esterification of trans-aconitate in *Escherichia coli. J. Biol. Chem.*, **274**, 13470–13479.

75. Partridge,J.D., Browning,D.F., Xu,M., Newnham,L.J., Scott,C., Roberts,R.E., Poole,R.K. and Green,J. (2008) Characterization of the *Escherichia coli* K-12 *ydhYVWXUT* operon: regulation by FNR, NarL and NarP. *Microbiology*, **154**, 608–618.
76. Bergholz,T.M., Wick,L.M., Qi,W., Riordan,J.T., Ouellette,L.M. and Whittam,T.S. (2007) Global transcriptional response of *Escherichia coli* O157:H7 to growth transitions in glucose minimal medium. *BMC Microbiol.*, **7**, 97.
77. Johansen,J., Eriksen,M., Kallipolitis,B. and Valentin-Hansen,P. (2008) Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and sigmaE-dependent CyaR-*ompX* regulatory case. *J. Mol. Biol.*, **383**, 1–9.
78. De Lay,N. and Gottesman,S. (2009) The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. *J. Bacteriol.*, **191**, 461–476.
79. Soupene,E., King,N., Lee,H. and Kustu,S. (2002) Aquaporin Z of *Escherichia coli*: reassessment of its regulation and physiological role. *J. Bacteriol.*, **184**, 4304–4307.
80. Neidhardt,F.C., Curtiss,R. III, Ingraham,J.L., Lin,E.C.C., Low,K.B. Jr, Magasanik,B., Reznikoff,W.S., Riley,M., Schaechter,M. and Umbarger,H.E. (1996) *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn. American Society for Microbiology, Washington DC.
81. Mizusawa,S., Court,D. and Gottesman,S. (1983) Transcription of the *sulA* gene and repression by LexA. *J. Mol. Biol.*, **171**, 337–343.
82. Yasuda,T., Nagata,T. and Ohmori,H. (1996) Multicopy suppressors of the cold-sensitive phenotype of the *pcsA68* (*dinD68*) mutation in *Escherichia coli*. *J. Bacteriol.*, **178**, 3854–3859.
83. Rolfes,R.J. and Zalkin,H. (1988) *Escherichia coli* gene *purR* encoding a repressor protein for purine nucleotide synthesis. Cloning, nucleotide sequence, and interaction with the *purF* operator. *J. Biol. Chem.*, **263**, 19653–19661.
84. Zheng,M., Wang,X., Doan,B., Lewis,K.A., Schneider,T.D. and Storz,G. (2001) Computation-directed identification of OxyR DNA binding sites in *Escherichia coli*. *J. Bacteriol.*, **183**, 4571–4579.