Characterization of the Escherichia coli $\sigma^E$ Regulon*

Claire Dartigalongue‡, Dominique Missiakas‡, and Satish Raina‡

From the ‡Department of Biochimie Medicale, Centre Medicale Universitaire, Universite de Geneve, I Rue Michel Servet, 1211 Geneva 4, Switzerland, and the ¶Department of Microbiology, Immunology, and Molecular Genetics, UCLA, Los Angeles, California 90095

Received for publication, January 17, 2001, and in revised form, March 20, 2001
Published, JBC Papers in Press, March 23, 2001, DOI 10.1074/jbc.M100464200

Escherichia coli responds to the accumulation of misfolded proteins by inducing the transcription of heat shock genes. $\sigma^E$ RNA polymerase controls one of the two heat shock regulons of E. coli. This regulon is activated upon accumulation of misfolded polypeptides in the double membrane envelope of E. coli, $\sigma^E$ (RpoE) is a member of the extracytoplasmic function subfamily of sigma factors. Here we asked how many genes are activated by $\sigma^E$ RNA polymerase and what is the identity of these genes. Using two independent genetic approaches, 20 E. coli promoters were identified which activate reporter gene transcription in a $\sigma^E$-dependent manner. In all cases examined, a canonical $\sigma^E$ binding site could be revealed upon mapping transcriptional start sites. 10 identified promoters activated the transcription of previously identified genes with four genes acting directly on the folding of E. coli envelope proteins (dsbC, fkpA, skp, and surA). The remaining promoters transcribed genes that are presumed to encode hitherto unknown extracytoplasmic functions and were named ecf (ecfA–ecfM). Two of these ecf genes were found to be essential for E. coli growth.

Heat shock and other environmental stresses result in the misfolding of polypeptides in all cells. Escherichia coli responds to the accumulation of misfolded polypeptides by activating the transcription of heat shock genes. Heat is a drastic stress that leads to protein unfolding in general and triggers two heat shock responses controlled by two distinct RNA polymerase species in E. coli: $\sigma_32$, $\sigma^A$, $\sigma^B$, and $\sigma^E$, respectively (1, 2). The unfolding of proteins in the envelope of E. coli uniquely induces the $\sigma^E$ regulon but not $\sigma^A$ or $\sigma^B$. E. coli (RpoE) is a member of the extracytoplasmic function (ECF) subfamily of sigma factors which function as effector molecules responding to extracytoplasmic stimuli (3, 5). Some microorganisms such as Streptomyces coelicolor harbor multiple ECFs that seem specialized in responding to different extracytoplasmic stimuli (5, 6). The E. coli $\sigma^E$ regulon is induced specifically in response to imbalanced synthesis of outer membrane proteins (7) and to misfolding of polypeptides that have been translocated across the cytoplasmic membrane (8).

Previous work identified several genes that are transcribed by $\sigma^E$ (4). $\sigma^E$ directs its own expression. rpoE is the first gene of an operon that also contains rseA, rseB, and rseC (regulator of $\sigma^E$) (9, 10). RseA is a short hydrophobic polypeptide that integrates into the cytoplasmic membrane. The N-terminal cytoplasmic domain of RseA binds to $\sigma^E$, sequestering the sigma factor from core RNA polymerase (E) (9, 10). The C-terminal domain of RseA protrudes into the periplasm, a compartment located between the cytoplasmic and outer membranes of E. coli. The C-terminal domain of RseA interacts with RseB, a periplasmic soluble protein (9, 10). RseB is believed to sense the concentration of misfolded polypeptides, causing RseB dissociation from RseA and liberating cytoplasmic $\sigma^E$ for interaction with core RNA polymerase (11). Another model suggested proteolytic cleavage of RseA in response to the accumulation of outer membrane proteins (12). The function of RseC, encoded by the fourth gene of the rpoE operon, remains unknown. $\sigma^E$ also transcribes htrA and fkpA, encoding a periplasmic protease (HtrA/DegP) for the removal of misfolded polypeptides (13, 14) and a periplasmic peptidyl prolyl isomerase (FkpA) involved in folding envelope proteins (8, 15).

Earlier work described the isolation of rpoE knockout mutants (16, 17). E. coli appears to require rpoE for viability and growth under physiological conditions, as the mutant strains cope with loss of rpoE function by acquiring compensatory mutations (18). The nature of compensatory mutations as well as the number and identity of the affected genes are still unknown. Even though rpoE seems to be essential, none of the known $\sigma^E$-transcribed genes (rpoH, htrA, fkpA, rseA, rseB, rseC) is required for either growth or viability of E. coli. Taken together, all previous work suggests that $\sigma^E$ must transcribe additional genes that are involved in the folding of envelope proteins. To identify genes that are transcribed by $\sigma^E$ and to approximate the size of the $\sigma^E$ regulon, we have used two different experimental strategies. Small DNA segments, generated by fragmentation of the E. coli chromosome, were fused to a promoterless lacZ reporter gene carried on a single copy plasmid. Further, the λMu53-lacZ transposon was used to generate sets of random fusion between the promoterless lacZ reporter and regulatory sequences of the chromosome of E. coli. Screening of both libraries of reporter fusions in various genetic backgrounds identified 20 promoters that activated LacZ expression in a $\sigma^E$-dependent manner. A hypothesis is presented to account for the essential function of the $\sigma^E$ regulon and to describe the role of the identified genes in responding to misfolded polypeptides within the envelope of E. coli.

---

* This work was supported by Fond National Scientifique Suisse Grant FN3100-069131.99/1 (to S. R.) and United States Public Health Service Grant GM58266 (to D. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1794 solely to indicate this fact.

To whom correspondence should be addressed: Dept. of Microbiology, Immunology, and Molecular Genetics, UCLA, 609 Charles Young Dr., Los Angeles, CA 90095. E-mail: missiaka@microbio.ucla.edu.

The abbreviations used are: E and $\alpha_3\beta_3\rho$, core RNA polymerase; RpoE and $\sigma^E$, sigma E transcription factor; $\sigma^E$ and $\alpha_2\beta_2\rho^\sigma$, holoenzyme complexed to sigma E; Rse, regulator of $\sigma^E$; Δ28RseA, a variant of RseA lacking the first 28 amino acids; rpoER178G, an allele of rpoE encoding a mutant of $\sigma^E$ with severely impaired transcriptional activity; Ecf, extracytoplasmic function gene product.
ExperimenTal procedures
bacterial strains and growth conditions—Most strains used in this study are listed in Table I. Strains carrying promoter fusions of hitherto unknown genes are referred to as ecf-lacZ. Sequences of primers used in this study can be obtained from the authors upon request. Luria Bertani (LB), MacConkey, and M9 minimal media were prepared as described [19]. When necessary, media were supplemented with 100 μg/ml ampicillin, 50 μg/ml spectinomycin, 15 μg/ml tetracycline, 50 μg/ml kanamycin, or 20 μg/ml chloramphenicol. The indicator dye 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside was used at a final concentration of 40 μg/ml in the agar medium. Mutations were transduced into various backgrounds using P1 bacteriophage [19]. Labeling experiments using [35S]methionine in the M9 high-sulfur medium were performed as described previously [20].

Construction of Promoter Fusion Libraries—A library of chromosomal transcriptional fusions was constructed using λMu53-lacZ (Kan6) [21]. Briefly, strain MC4100 (LacZ+) was mutagenized at 30 °C with λMu53-lacZ, and colonies were formed on MacConkey plates at 30 °C and 43 °C. Colonies that developed red staining at 43 °C but not at 30 °C were isolated. The site of λMu53 insertions into the chromosome was determined by DNA sequencing with the oligonucleotide primer (5′-GTTCAGTGTTCTTCTGTC-3′). For this step, DNA regions carrying the λMu53-lacZ fusions were cloned into a cosmid, taking advantage of the λMu53-lacZ (Kan6) marker. A second library was constructed using the single copy F-based promoter probe vector pFZY [22] essentially as described earlier [23]. Putative Eo5-regulated promoters identified by this strategy were analyzed by DNA sequencing using the synthetic oligonucleotide described above.

Cloning Procedures and Gene Replacement—The DNA regions corresponding to Eo5-regulated promoters were amplified by polymerase chain reaction using appropriate primers and cloned into pRS535 using restriction sites BamHI and EcoRI [24]. Plasmids were characterized by DNA sequence analysis, and each fusion was transferred on the chromosome from plasmid pKO3 when the cells were grown in the presence of arabinose (0.2%). The presence of the ecf-lacZ fusions was verified by transducing linked markers with bacteriophage P1. In all cases, disruption of ecf genes was verified by polymerase chain reaction amplification using chromosomal template DNA and appropriate primers.

Primer Extension Analysis—Total RNA was isolated using the RNeasy kit from Qiagen. Cultures were grown at 30 °C, and aliquots were shifted to 50 °C for a period of either 5 or 10 min. Immediately after the heat shock, cultures were lysed with guanidinium isothiocyanate following the protocol of the RNeasy kit. To define the transcriptional start site(s) of each gene, ~10 ng of complementary oligonucleotide probe was annealed with 10 μg of total RNA. Strand extension from the annealed primer was achieved using the avian myeloblastosis virus reverse transcriptase. Primer extension products were separated on 8 % urea-containing gels, and their migration profile was compared by running on the same gel the deoxy sequencing reactions using the same oligonucleotide.

Biochemical Assays—β-Galactosidase activity was determined as described previously [19]. Bacterial cultures were grown overnight at 30 °C, diluted 1:100, and allowed to reach A600 nm between 0.5 and 0.7. Aliquots were maintained or shifted to 14, 37, or 43 °C for 20 min. Measurements were performed in duplicate, and the data represent the average of at least three independent experiments. Two-dimensional equilibration gel electrophoresis was performed as described previously [28].

results

Approximating the Number of Eo5-regulated Genes—To measure the size of the RpoE regulon, we analyzed pulse-labeled E. coli proteins by two-dimensional gel electrophoresis. This technique has been employed routinely for the analysis of heat shock regulation [29]. E. coli strain BL21 (DE3), carrying a plasmid overexpressing rpoE or an empty vector control, was grown in M9 minimal medium to A600 nm 0.6. Cells were pulse labeled with [35S]methionine for 2 min, and all further incorporation of radioactivity into polypeptide was quenched by the addition of excess unlabeled methionine. E. coli cells were lysed in buffer containing detergent and ampholytes. Proteins in the extracts were separated by charge electrophoresis of the ampholytes within the pH range 3.5–10 and then separated on SDS-polyacrylamide gel electrophoresis in the second dimension based on their molecular mass. Of the 4,500 polypeptides encoded by the E. coli genome, the two-dimensional gel elec-

<table>
<thead>
<tr>
<th>Strains</th>
<th>Relevant characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC4100</td>
<td>F araD139 (pargF-lac) U169</td>
<td>Our collection</td>
</tr>
<tr>
<td>LMG194</td>
<td>F lacZΔ74 galE thi rpsL phoA ara714 leu-19 orn-10</td>
<td>Our collection</td>
</tr>
<tr>
<td>SR1502</td>
<td>MC4100 rpoER178G</td>
<td>16</td>
</tr>
<tr>
<td>SR3323</td>
<td>MC4100 transformed with a derivative plasmid of pET-24d overproducing wild type RseA</td>
<td>10</td>
</tr>
<tr>
<td>SR3076</td>
<td>MC4100 transformed with a derivative plasmid of pET-24d overproducing RseAΔ28 (RseA)</td>
<td>10</td>
</tr>
<tr>
<td>DM1055</td>
<td>BL21 transformed with a derivative plasmid of pBADΔ1 overproducing gfp</td>
<td>Our collection</td>
</tr>
<tr>
<td>DM1695</td>
<td>BL21 transformed with a derivative plasmid of pBADΔ12 overproducing PrpA</td>
<td>28</td>
</tr>
<tr>
<td>SR2912</td>
<td>MC4100 rseA::Tn10</td>
<td>10</td>
</tr>
<tr>
<td>SR2306</td>
<td>MC4100 surA::Tn10 Tet</td>
<td>Our collection</td>
</tr>
<tr>
<td>MU1</td>
<td>Δlsp</td>
<td>36</td>
</tr>
<tr>
<td>PND325</td>
<td>M1400 cpxR::Spec</td>
<td>32</td>
</tr>
<tr>
<td>SR1795</td>
<td>MC4100 dsbC::Tn10</td>
<td>41</td>
</tr>
<tr>
<td>SR3316</td>
<td>MC4100 htrA::Tn10</td>
<td>Our collection</td>
</tr>
<tr>
<td>SR3676</td>
<td>MC4100 fhpA::TcM</td>
<td>Our collection</td>
</tr>
<tr>
<td>SR4425</td>
<td>LM9144 ecfA::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR4442</td>
<td>MC4100 ecfD::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR4177</td>
<td>LM9144 ecfE::TcTet carrying plasmid pBAD-ecfE</td>
<td>This study</td>
</tr>
<tr>
<td>SR4300</td>
<td>MC4100 ecfF::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR672</td>
<td>MC4100 ecfG::Tn10</td>
<td>This study</td>
</tr>
<tr>
<td>SR4292</td>
<td>MC4100 ecfH::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR650</td>
<td>MC4100 ecfI::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR4172</td>
<td>MC4100 ecfJ::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR4455</td>
<td>MC4100 ecfK::TcKan</td>
<td>This study</td>
</tr>
<tr>
<td>SR4418</td>
<td>LM9144 ecfL::TcKan carrying plasmid pBAD-ecfL</td>
<td>This study</td>
</tr>
</tbody>
</table>
when insoluble material was removed by centrifugation (13,000 × g, 3 min). Producing cells contained 13 polypeptide spots with significantly higher levels of expression in panel B as compared with panel A, with increased abundance when the pEAD vector alone (panel A) or pDM1055 (overexpressing rpoE) (panel B) were transformed into E. coli MC4100 carrying E. coli rpoE::Tn5. The letters K, EL, and E identify DnaK, GroEL, and σE, respectively, and bold arrows identify some proteins of equal abundance in both experiments.

Fig. 1. Global effect of rpoE overexpression as visualized using two-dimensional equilibrium gel electrophoresis. E. coli BL21(DE3) cultures were grown at 30 °C to an A435 nm of 0.6 in M9 medium supplemented with glucose, and the expression of T7 polymerase was induced by the addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside. After a 30-min incubation period, cells were labeled with [35S]methionine for 1 min. Cells were lysed and extracted with detergents (2% Nonidet P-40) and 8 M urea followed by several freeze-thawing cycles, producing cells contained 13 polypeptide spots with significantly increased intensity (empty arrows in Fig. 1B), suggesting that these polypeptides represent σE-regulated genes. Further, overexpression of σE appears to have a negative effect on the regulation of some E. coli genes because the abundance of nine polypeptides was severely diminished compared with Fig. 1B (circled peptide spots).

Transposon Mutagenesis to Search for σE-regulated Genes

50,000 MC4100 μMu53::lacZ transposon mutants

MacConkey agar, 43°C

1,000 red colonies pooled

μMu53::lacZ insertion transduced in MC4100 surA::Tn10

MacConkey agar, 30°C

200 red colonies transformed with p(serA+)

MacConkey agar, 30°C

78 white colonies

FIG. 2. Genetic strategy using transposon mutagenesis to search for σE-regulated genes. 50,000 μMu53::lacZ (KanR) E. coli MC4100 mutants were screened by growing colonies at 43 °C on MacConkey plates with kanamycin. 1,000 red colonies were pooled (filled circles on the agar plate drawing) and transduced into E. coli MC4100 surA::Tn10. 200 colonies that developed a red staining on MacConkey agar were transformed with a plasmid overexpressing rseA. 78 colonies that developed white staining on MacConkey agar were analyzed further by characterizing the transposon insertion.

Genes—We sought to identify σE-regulated genes by searching for promoter sequences that activate transcription under heat shock conditions (43 °C). The μMu53::lacZ (KanR) transposon inserts randomly into the chromosome of E. coli and generates a promoterless lacZ reporter with regulatory sequences flanking the insertion sites. 50,000 μMu53::lacZ (KanR) E. coli MC4100 mutants were screened by growing colonies at 43 °C on MacConkey plates with kanamycin (Fig. 2). 1,000 mutant red colonies were picked and pooled for further analysis. Our initial screen could not distinguish between RpoE-regulated promoters and those that are transcribed by other polymersases. 1,000 μMu53::lacZ (KanR) insertions were transduced into E. coli strain SR3206 (surA::Tn10) using bacteriophage P1. Transductants were selected for growth at 30 °C on MacConkey agar containing kanamycin. surA encodes a periplasmic chaperone required for folding of outer membrane proteins (8, 30, 31). E. coli surA mutants express the σE-regulated promoters htrA and rpoEP2 constitutively, even when cells are grown on agar medium at 30 °C (8). 200 transposon transductants of E. coli strain SR3206 formed red colonies at 30 °C. These μMu53::lacZ (KanR) insertions were analyzed further and transduced into E. coli strain SR3323 containing RseA encoded on a high copy number plasmid. Production of RseA reduces the ωR RNA polymerase transcription because the anti-sigma factor sequesters σR in the cytoplasmic membrane compartment. 78 E. coli μMu53::lacZ transductants formed white colonies on MacConkey agar. The site of transposon insertion in these strains was determined by DNA sequence analysis. μMu53::lacZ insertions identified nine σE-regulated genes: htrA, fhpA, cutC, nlpB, purA, mdoG, mdoH, yggN, and ytfJ. yggN and ytfJ were identified previously by genome sequencing; however, a physiological role of these genes has not yet been described. Henceforth, we refer to these genes as ecfF (yggN) and ecfJ (ytfJ), for extracytoplasmic function genes F and J.

Promoter Fusion to Search for σE-regulated Genes—Our transposon insertion mutagenesis cannot identify σE-regulated genes that are essential for E. coli growth. To identify all...
Role of the E. coli RpoE Regulon

E. coli MC4100 transformants of a promoter library in pFZY were isolated by growing colonies at 30 °C on lactose minimal agar. 200,000 colonies were replica plated and grown at 43 °C on MacConkey agar. 10,000 red colonies (filled circles) were selected by transformation of Recombinant plasmids containing transcriptionally active promoterless lacZ fusions were selected by transformation of MacConkey agar at 30 °C. 500 transformants were plated on lactose minimal agar. 200,000 colonies were replica plated and grown at 43 °C on MacConkey agar. 10,000 red colonies (filled circles) were picked and analyzed for promoter content using Southern hybridization (htrA, rpoE, and rpoH promoter probes) and DNA sequencing.

RpoE-regulated genes, even those that are essential, 0.8–1.2-kilobase pair DNA fragments were generated by Sau3A digestion of the E. coli chromosome and cloned into pFZY digested with BamHI. pFZY is a single copy F factor plasmid containing a promoterless lacZ gene downstream of the BamHI site (22). Recombinant plasmids containing transcriptionally active promoter fusions were selected by transformation of E. coli MC4100 (lacZ ). Transforms were plated on lactose minimal agar at 30 °C. 200,000 Lac colonies were replica plated on MacConkey agar and incubated at 43 °C (Fig. 3). 10,000 colonies displayed red staining (Lac+) under heat shock conditions, representing possible ρE-regulated promoter fusions. The red colonies were pooled and plasmids purified and transformed into E. coli MC4100 (rseA::Tn10). Transforms were plated on MacConkey agar at 30 °C and screened for a red colony phenotype, consistent with sra-dependent induction of ρE-transcribed promoters. 10,000 red colonies were pooled and made competent for transformation with pSR323 (28), a high copy number plasmid encoding rseA+. Transforms were plated on MacConkey agar at 30 °C. 350 white colonies were picked. Plasmids were extracted from the pool and used to transform strain SR1502. E. coli SR1502 is an MC4100 variant carrying a mutation in the rpoE gene (rpoER178G) which displays a temperature-sensitive growth phenotype because the mutant of E polymerase cannot adequately transcribe ρE-regulated genes (16). To determine whether promoter fusions were transcribed by E polymerase, SR1502 transformants were plated on MacConkey agar at 30 °C. 500 transformants were pooled and transformed into E. coli strain SR1502. 500 transformants were plated on lactose minimal agar. 200,000 colonies were replica plated and grown at 43 °C on MacConkey agar. 10,000 red colonies (filled circles) were selected by transformation of MacConkey agar at 30 °C. 500 transformants that failed to hybridize in the dot blot experiment were analyzed by restriction mapping and DNA sequence analysis, which identified 22 distinct promoters. Results described below revealed that 19 of the 22 promoters are transcribed in an E polymerase-dependent manner. The promoters activate transcription of cutC, dsbC, fkpA, htrM, mdoG, nlpB, ostA, rpoD, skp, ecfA (p288), ecfD (yfiO), ecfE (yaeL), ecfF (yggN), ecfG (htrG), ecfH (yraP), ecfI (yidQ), ecfJ (yfJ), ecfK (UP0), and ecfL (yqA). Genes without a previously assigned physiological function are referred to as ecf, for extracytoplasmic function genes.

EρE-transcribed Promoters Are Regulated by the rpoE rseA Operon.—To quantify transcriptional regulation of E polymerase-transcribed promoters, fusions were inserted into attB (bacteriophage λ attachment site) of E. coli MC4100 and SR3206 (sra::Tn10) using λRS45 as cloning vector (Fig. 4A). λRS45 lysogens were grown in Luria broth to Abio 0.5–0.7, and LacZ activity was measured in a spectrophotometer using o-nitrophenyl-β-D-galactopyranoside as a substrate. htrA promoter activity was monitored as a control for a known ρE-regulated gene. When λ lysogens of the surA mutant strain SR3206 were examined at 30 °C, htrA promoter activity was increased by 4-fold compared with wild-type E. coli. All 19 isolated promoters displayed a similar phenotype with a 2–4-fold increase of reporter transcription in the surA mutant strain (cutC, dsbC, fkpA, htrM, mdoG, nlpB, ostA, rpoD, skp, ecfA, ecfD, ecfE, ecfF, ecfG, ecfH, ecfI, ecfJ, ecfK, and ecfL fusions to lacZ) (Fig. 4A). When E. coli SR1502 (rpoER178G) was lysogenized with λRS45 derivatives, the promoter fusions expressed the LacZ reporter only with background activity (Fig. 4B). Transformation of SR1502 harboring a htrA-lacZ insertion with a plasmid encoding wild-type rpoE led to a 20-fold increase in expression of LacZ reporter. All 19 promoter fusions behaved similarly (Fig. 4B). Increased expression was also observed in the presence of wild-type RpoE (between 10- and 30-fold). As a final test to determine whether the isolated promoters are transcribed by E polymerase, E. coli MC4100 carrying insertions of λRS45 derivatives were transformed with three different plasmids: pRS3323 (overexpression of wild-type RseA), pRS3076 (overexpression of RseAD28, a mutant lacking the first 28 amino acids), and pET-24d (vector lacking RseA) (Fig. 4C). Compared with strains carrying the vector alone, all transformants expressing wild-type rseA transcribed between 10 and 60% of lacZ reporter gene. In contrast, expression of the mutant rseA allele (rseAA28) caused no reduction in reporter transcription. Together these data indicate that the 20 isolated gene promoters are transcribed by E polymerase RNA polymerase in a manner that is also subject to regulation by RseA. With the exception of rpoD, the promoters were isolated multiple times, suggesting that our search for RpoE-regulated genes has been nearly exhaustive. Comparison of the number of isolated gene promoters with the number of ρE-regulated protein spots identified by two-dimensional gel electrophoresis corroborates this view further. The search for EρE-transcribed promoters identified 24 new promoters. At least 12 of the genes transcribed by E polymerase encode lipoproteins or membrane proteins. These hydrophobic proteins will not be identified by two-dimensional gel electrophoresis.

Characterization of EρE-transcribed Promoters—Promoter
FIG. 4. Transcriptional activity of the selected $\sigma^E$-dependent promoter fused to lacZ. Bacterial cells containing a single copy lacZ promoter fusions (inserted with lambda vector at attB) were assayed for promoter activity by measuring $\beta$-galactosidase expression. Panel A, promoter activities as assayed in E. coli MC4100 wild-type (empty bars) and surA::Tn10 mutant cells (gray bars). Panel B, promoter activities as assayed in E. coli MC4100 carrying the chromosomal rpoER178G mutation (partial loss-of-function mutant of $\sigma^E$) and the pEAD vector with (gray bars) or without the wild-type rpoE gene (empty bars). Panel C, promoter activities as assayed in E. coli MC4100 transformed with pET-24d (empty bars) or pET-24d derivatives overexpressing wild-type rseA (black bars) or rseA$\Delta28$ (gray bars), encoding a truncated form of RseA which lacks a $\sigma^E$ binding domain.
sequences were subjected to a BLAST search against the E. coli genome data bank, thereby identifying the entire gene sequences (Table II). Primers were designed to map transcriptional start sites. Total mRNA was purified and used as a template for reverse transcriptase reactions primed with specific oligonucleotides, and the generated data are summarized in Table III. In all cases examined, the transcriptional start sites are spaced appropriately with respect to −10 and −35 recognition sites. Comparison of DNA sequences allowed the identification of a presumed canonical Eσ regulon recognition site. YCTGA is positioned 7–9 nucleotides upstream of the transcriptional start site (−35 site). Some Eσ promoters showed no effect when analyzed in vivo, the activity of all 20 isolated promoters was measured in wild-type and cpxR mutant strains. cpxP promoter activity was monitored as a control for a known CpxR-regulated gene (4). cpxP promoter activity was decreased by 30% in the cpxR mutant strain. The phosphtase PrpA modulates the activity of CpxAR (28), and overproduction of PrpA led to an 80% increase of cpxP promoter activity. Of the 20 promoters examined here, dsbC, skp, and ecfI behaved similarly to the cpxP promoter (Fig. 5), whereas all other promoters showed no effect when analyzed in cpxR mutant strains (data not shown). The CpxR binding site has been identified as tandem repeats of the nucleotide sequence GGTNAY. The dsbC, skp, and ecfI promoter sequences were found to harbor DNA repeat elements that matched the consensus sequence of CpxR binding sites (Table IV).

Some θσ-regulated Genes That Are Essential for E. coli Growth—Eσ represents a minor RNA polymerase species and transcribes only 20 of the 4,500 genes encoded by the genome of E. coli. We wondered why rpoE may be essential for E. coli growth. Two Eσ-transcribed genes encode sigma factors for major RNA polymerase species, rpoD (θσ,θσ) and rpoH (θσ,θσ′,θσ′). rpoH is essential for growth at all temperature. It seems unlikely that Eσ transcription of rpoD is essential for E. coli growth because rpoD is transcribed by multiple promoters and RNA polymerase species. The rpoH gene is also transcribed by several RNA polymerase species, and Eσ recognizes only one of the three known promoters. Deletion of rpoH is tolerated at elevated temperature upon overexpression of groEL and dnaK operons (33). Thus, if Eσ transcription of rpoD and rpoH is not essential, can RpoE-mediated tran-
The corresponding −10 and −35 regions of the promoters are depicted in bold. The +1 transcriptional start site is underlined. Some genes contain multiple promoters, only the σE-dependent promoter is shown here (e.g. htrMP4).

<table>
<thead>
<tr>
<th>Gene name</th>
<th>DNA sequences surrounding the CpxR-dependent boxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>sbpC</td>
<td>CATT GCTAAAAGGAGGAACTCAAGGCTTGTGGCTGTTGGTAAGAAGG</td>
</tr>
<tr>
<td>skp</td>
<td>CATGGTTAAAAGGCAGGACACGCTCCTGTCGGTTGCTGTTGG</td>
</tr>
<tr>
<td>ecfI</td>
<td>GCAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfJ</td>
<td>AAGCTTGATGATCGGATTGCTGTTGGTTGAAGGAGG</td>
</tr>
<tr>
<td>ecfK</td>
<td>GCTTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfL</td>
<td>GCTTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfM</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfN</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfP</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfQ</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfR</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfS</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfT</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfU</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfV</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfW</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfX</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfY</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
<tr>
<td>ecfZ</td>
<td>TCTAGGTTAAAGGAGGAACTCAAGGCTTGTGGCTGTTGG</td>
</tr>
</tbody>
</table>

"TABLE IV

σE-transcribed genes containing canonical CpxR-binding boxes"

DNA repeats shown in bold are presumed to be CpxR binding boxes.

To test this assumption, ecf gene sequences were disrupted by insertion of Ω elements and cloned on a plasmid carrying a temperature-sensitive replicon and the sacB marker. After transformation of plasmids into E. coli MC4100, single cross

formation of some other genes be required for E. coli growth? To test this assumption, ecf gene sequences were disrupted by insertion of Ω elements and cloned on a plasmid carrying a temperature-sensitive replicon and the sacB marker. After transformation of plasmids into E. coli MC4100, single cross-over recombination events with wild-type ecf sequences were isolated by plating bacteria at 43 °C, a condition that stalls plasmid replication, and by selecting for plasmid-encoded chloramphenicol resistance during sucrose fermentation. The resulting plasmid cointegrates into the E. coli chromosome are merodiploid and contain two copies of the ecf gene under study, a wild-type and a mutant allele. Growth of cointegrate strains on sucrose-containing media serves as a counterselection for plasmid-encoded sacB because as expression of the sacB gene product leads to the accumulation of toxic metabolites during sucrose fermentation. Thus, when cointegrated strains are streaked on agar medium containing sucrose as well as antibiotic selection for the Ω element, the resulting colonies represent ecf mutants arising from double crossover recombination. Using this experimental scheme, ecfA, ecfD, ecfF, ecfH, ecfI, ecfJ, and ecfK knockout mutants were obtained. A Tn10 insertional knockout mutation of ecfG had been isolated previously. Two of the isolated knockout mutant strains, ecfG and ecfJ, displayed a temperature-sensitive growth phenotype above 43 °C.

Cointegrates that formed after plasmid insertion into ecfE and ecfL could not be resolved, suggesting that these genes may be essential for viability and growth of E. coli. This hypothesis was tested by repeating the resolution of cointegrate strains after transformation with a second plasmid, containing the wild-type ecf gene under control of the arabinose-inducible araBAD promoter. Streaking cointegrates on arabinose-containing sucrose plates produced the desired knockout mutants, whereas streaking on sucrose media without arabinose failed to produce any colonies. Thus, the chromosomal copy of E. coli expressing plasmid-encoded ecfE or ecfL can be deleted by homologous recombination. The resulting strains henceforth require arabinose-containing media for growth, indicating that ecfE and ecfL are essential genes.

"Synthetic Lethality and Synthetic Conditional Phenotypes for σE-regulated Genes—"
ing factors, skp and fkpA. Double mutants dsbC/htrA, skp/htrA, and fkpA/htrA display a synthetic conditional lethal phenotype. Thus, at elevated temperatures E. coli cells require HtrA protease to remove misfolded polypeptides in the periplasm, a condition that is aggravated when specific folding catalysts are nonfunctional. Loss of both fkpA and surA also leads to a synthetic conditional lethal phenotype. When combined with htrA, the triple mutant fkpA/surA/htrA is nonviable (Table V).

Cold Shock Weakly Induces the σE Regulon—The σE regulon can be viewed as providing essential folding functions for proteins that are located in the bacterial envelope. An increase in temperature (38–45 °C) weakens the interactions that maintain the three-dimensional structure of polypeptides at physiological temperature (25–37 °C): hydrogen bonds, ion bonds as well as van der Waal’s forces. Other conditions that alter the above mentioned parameters of protein folding and stability should therefore also induce the σE regulon. We wondered whether a reduction in temperature (14–24 °C) could induce the σE regulon. E. coli cold shock appears to be a regulated response requiring many genes; however, a specific sensing or transcriptional regulatory mechanism has thus far not been established. Using MC4100 strains carrying single copy insertions of σE-regulated promoters fused to lacZ, we observed about a 20–30% increase in transcription after incubating cells at elevated temperatures (38–45 °C) weakens the interactions that maintain the three-dimensional structure of polypeptides at physiological temperature (25–37 °C): hydrogen bonds, ion bonds as well as van der Waal’s forces. Other conditions that alter the above mentioned parameters of protein folding and stability should therefore also induce the σE regulon. We wondered whether a reduction in temperature (14–24 °C) could induce the σE regulon. E. coli cold shock appears to be a regulated response requiring many genes; however, a specific sensing or transcriptional regulatory mechanism has thus far not been established. Using MC4100 strains carrying single copy insertions of σE-regulated promoters fused to lacZ, we observed about a 20–30% increase in transcription after incubating cells...
FIG. 6. The $\sigma^E$ regulon is cold shock-inducible. Bacterial cells containing a single copy lacZ promoter fusions (inserted with lambda vector at attB) were assayed for promoter activity by measuring $\beta$-galactosidase expression. Promoter activities were assayed in E. coli MC4100 at 37 °C (empty bars) or 14 °C (gray bars).

for 1 h at 14 °C (Fig. 6). Thus, rapid reduction of ambient temperature also stimulates $E\sigma^K$ polymerase, causing a small increase in the expression of folding catalysts.

**DISCUSSION**

$E\sigma^K$ RNA polymerase is thought to be dedicated to expressing folding catalysts that act on proteins in the bacterial envelope. Here we measured the size of the $E\sigma^K$ regulon with two methods: two-dimensional gel electrophoresis of RpoE-induced cells and cloning of RpoE-regulated promoters. Results from both experiments as well as previous work suggest that $E\sigma^K$ transcribes some 43 genes. We describe here 20 new promoters that are recognized by $E\sigma^K$ RNA polymerase. Some of the genes regulated by $E\sigma^K$ were hitherto unknown and have been designated $ecf$, for extracytoplasmic encoding function. Some of the encoded gene products are located in the periplasmic space and act directly on misfolded proteins: DsbC, FkpA, HtrA, Skp, and SurA. Some other gene products are located in the bacterial cytoplasm and serve regulatory functions that coordinate the expression of the $E\sigma^K$ regulon with environmental conditions. RpoE, RpoH, and RpoD represent components of various RNA polymerase species, whereas RseA, RseB, and RseC regulate the availability of $E\sigma^K$ for core RNA polymerase. Several $E\sigma^K$-regulated gene products are involved in the synthesis of lipopolysaccharide, a component of the outer membrane of Gram-negative bacteria. Lipopolysaccharide has been proposed to act as a cofactor for the membrane assembly of outer membrane proteins, a pathway that appears to require Skp activity (8). Skp has also been shown to play other roles in envelope assembly (36). It seems noteworthy however that skp mutant cells contain increased amounts of lipopolysaccharide within the periplasm (36). It is as if deletion of the presumed folding factor (Skp) may lead to the simultaneous accumulation of its cofactor (lipopolysaccharide). The rpoDFCL and lpxDA genes provide known components of the lipopolysaccharide biosynthetic pathway and are transcribed by $E\sigma^K$ polymerase. In fact, the lpxD lpxA fabZ genes are regulated by two $E\sigma^K$-dependent promoters: one placed in front of skp (the first gene of the operon) and a second one in front of lpxD. Our preliminary results suggest that the ecfABC gene products may also be involved in the lipopolysaccharide biosynthetic pathway.3

Two $E\sigma^K$-regulated genes encode proteins with sensory functions. MdoG is involved in coordinating cellular pressure with the biosynthesis of periplasmic membrane-derivated lipopolysaccharides (37), whereas CutC has been postulated to be involved in copper homeostasis (38). The requirement of these gene products for protein folding in the periplasmic space is not immediately apparent. In this and perhaps in other cases, the presence of a $E\sigma^K$ promoter may provide growth advantages for the E. coli host which are not related to protein folding. The largest group of $E\sigma^K$-regulated genes encodes proteins located in the inner (NlpB, EcfD, EcfG, EcfI, and EcfL) and outer membranes (EcfK and EcfM). The precise function of these proteins remains to be established; however, it is conceivable that the membrane proteins act directly on misfolded membrane proteins and promote either polypeptide degradation or insertion into the lipid bilayer. Alternatively, membrane proteins may be involved in the transport and assembly of lipopolysaccharide into the physiological bilayer structures.

Two new members of the RpoE regulon were observed to be essential: $ecfE$ and $ecfL$. Because these genes appear to be transcribed by several RNA polymerases (data not shown) and have no definitive function attributed, it is impossible to draw conclusions as to why the $E\sigma^K$ regulon is essential for E. coli growth. EcfE appears to be a member of a large group of proteases designated RIP (regulated intramembrane proteolysis). Proteases of the RIP family are needed for diverse functions such as lipid metabolism, cell differentiation, and response to unfolded proteins (39, 40). We are currently investigating the role of EcfE in signaling envelope stress in E. coli.

In summary, the $E\sigma^K$ regulon has evolved to control at least two cellular processes, folding of polypeptides in the bacterial envelope and biosynthesis/transport of lipopolysaccharide. Conditions that cause unfolding of polypeptides are signaled by the RseA and RseB proteins (11). It is conceivable that the $E\sigma^K$

3 C. Dartigalongue, D. Missiakas, and S. Raina, unpublished data.
regulon can sense and respond to changes in lipopolysaccharide metabolism. Our future work will address this possibility.

Acknowledgments—We thank O. Schneewind (UCLA) for a critical review of this manuscript.

REFERENCES


Characterization of the *Escherichia coli*\(^E\) Regulon
Claire Dartigalongue, Dominique Missiakas and Satish Raina

doi: 10.1074/jbc.M100464200 originally published online March 23, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M100464200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 39 references, 22 of which can be accessed free at http://www.jbc.org/content/276/24/20866.full.html#ref-list-1