Biochemical and Biophysical Research Communications **288,** 91–100 (2001)
doi:10.1006/bbrc.2001.5734, available online at http://www.idealibrary.com on IDEAL®

Influence of the *yihE* Gene of *Shigella flexneri* on Global Gene Expression: On Analysis Using DNA Arrays

Ming-Shi Li, J. Simon Kroll, and Jun Yu[1]
Molecular Infectious Diseases Group, Department of Paediatrics, Faculty of Medicine, Imperial College St. Mary's Campus, Norfolk Place, London W2 1PG, United Kingdom

Received September 10, 2001

Inactivation of *dsbA* (disulfide bond formation), either by an insertion (Sh4, *dsbA::kan*) or by alteration of the active site (Sh42, *dsbA33G*), renders *Shigella flexneri* avirulent. However, Sh4 and Sh42 behave differently in many ways *in vitro* and *in vivo*. A gene of unknown function, *yihE*, up-stream and cotranscribed with *dsbA*, is thought to differentiate Sh4 and Sh42 as the *kan* insertion may result in a truncated unstable *yihE-dsbA* mRNA in Sh4. To gain insight into the function of *yihE*, DNA array hybridization was performed to study the genomic expression in Sh4, Sh42, and a newly constructed *yihE* mutant (Sh54). Compared to the wild-type, M90TS, Sh4, and Sh54 demonstrated significantly changed transcription levels of about 100 genes, of which many involved in energy metabolism and stress response were down- and up-regulated, respectively. In contrast, Sh42 showed altered transcription levels of only 20 genes. The results argue that *yihE* is principally responsible for the changed genomic expression in Sh4 and Sh54. Given the fact that the transcription of *yihE-dsbA* is regulated by the CpxRA two-component signal transduction system, *yihE* is probably involved in the extracytoplasmic stress response in a manor deserving further studies. © 2001 Academic Press

***Key Words: Shigella flexneri; yihE; dsbA;* DNA array; stress response.**

Shigella flexneri is a facultative intracellular pathogen, causing bacillary dysentery in humans and primates. The major virulence determinants, including the invasion plasmid antigens (Ipa) and a dedicated *mxi-spa* (type III) secretion system, are encoded by a 220-kb plasmid (1). But factors encoded by chromosomal pathogenicity islands also contribute to virulence (2–4). Gene products not conventionally related to virulence, such as cytochrome bd and TonB, are necessary to support intracellular growth (5, 6).

We have previously constructed two mutants to investigate the role of the periplasmic thiol:disulfide oxidoreductase, DsbA, in *Shigella* pathogenecity. Sh4 expresses no DsbA as a result of a kanamicin insertion (*dsbA::kan*), while Sh42 expresses nonfunctional DsbA33G due to the substitution of glycine for cystine at the active site (7, 8). Although both mutants have lost virulence in *in vitro* and *in vivo* assays, they behave quite differently in many ways. Under normal aerobic growth conditions, Sh4 grows more slowly than Sh42 (doubling times of 45 and 35 min, respectively) compared to the wild-type doubling time of 30 min. In the guinea pig keratoconjunctivitis model of *Shigella* infection, immunization with Sh42 can fully protect animals from virulent *S. flexneri* challenge, whereas Sh4 offers little protection (Hartman, in preparation). *In vitro*, the two mutants possess different cytotoxicity to murine- and human-derived macrophages, and induce different levels of cytokines from host cells (9). An explanation for these differences must be sought elsewhere beyond DsbA function, as both Sh4 and Sh42 are defective in disulfide bond formation, resulting in a conditionally defective *mxi-spa* secretion system. This has been explained by the folding of Spa32, a constituent of the *mxi-spa* secretion system, being impaired in both Sh4 and Sh42, because Spa32 must form its single internal disulfide bond to be functional (10). As a result, Sh4 and Sh42 can invade epithelial cells and escape from the phagocytic vacuoles, but are trapped in the inter-epithelial protrusions because they do not secrete sufficient amount of Ipa, responsible for lysis of protrusion membranes, at the latter location (8).

A previous study has shown that *dsbA* can be transcribed alone or together with *yihE*, a gene of unknown function(s) lying upstream (11). We speculated that the difference between Sh4 and Sh42 might be related to a difference in *yihE* expression, through the *kan* cassette in Sh4 affecting the stability of the *yihE-dsbA* transcript, resulting in phenotypic changes in addition to

[1] To whom correspondence should be addressed. Fax: (+44) 20 7886 6284. E-mail: jun.yu@ic.ac.uk.

0006-291X/01 $35.00

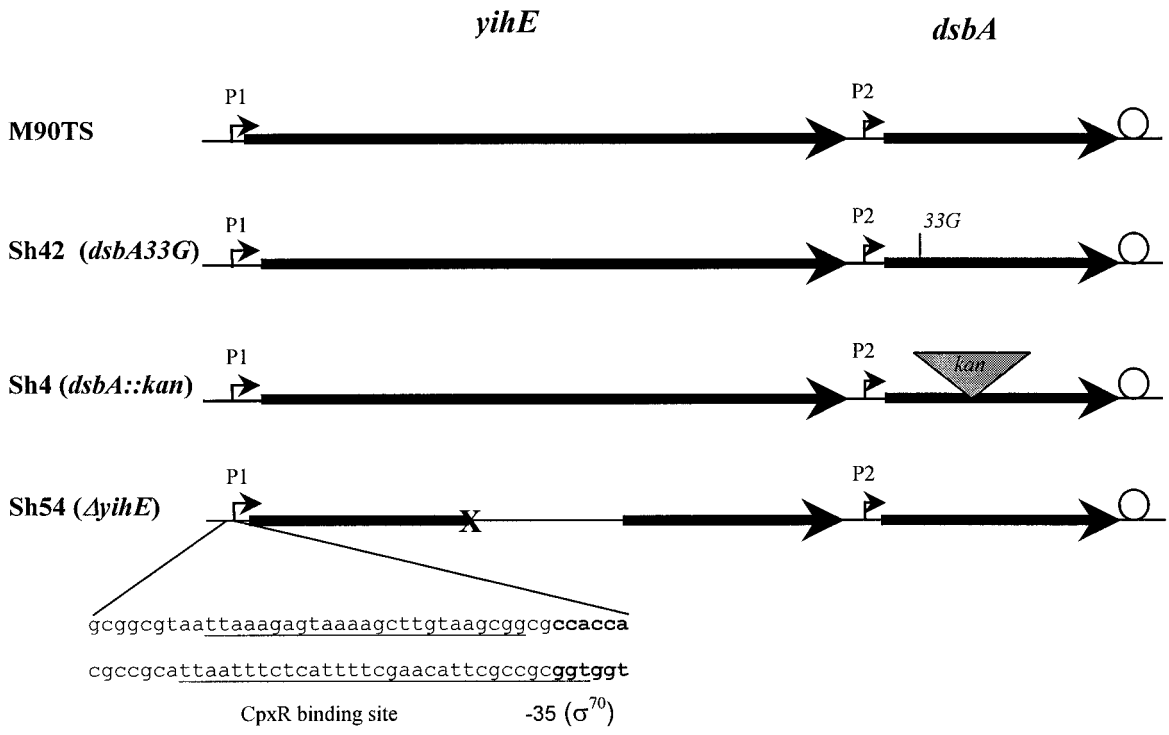

FIG. 1. Schematic illustration (not to scale) the *yihE-dsbA* locus. Solid arrows indicate the *yihE* and *dsbA* genes. The ρ independent transcriptional terminator after the stop coden of *dsbA* is represented as an open circle. The thin line illustrates the internal deletion and "x" indicates the stop coden introduced in the Sh54 *yihE* gene. The P1 promoter region is enlarged and the CpxR biding sequence is shown.

those caused by defective disulfide bond formation. Furthermore, transcription of *yihE-dsbA* is regulated by the two-component signal transduction system, CpxRA, which in conjunction with σ^E governs the extracytoplasmic stress response (12) implicating YihE may be involved in this process.

MATERIALS AND METHODS

Bacterial strains, media, and growth. *Escherichia coli* XL-1 Blue (*recA1, lac, endA1, gyrA96, thi, hsdR17, supE44, rclA1,* (*F′, proAB, lacIq, lacZΔM15, Tn10*). Δ(*ara-leu*),*araD,* Δ*lacX74, galE, galK, phoA20*) was used as host for general cloning, and CC118 λ*pir* (*thi-1, rpsE, rpoB, argE*(*Am*), *recA1,* λ*pir* phage lysogen) was used for maintenance of the suicide vector pCVD422 and its derivatives (13). Strains were routinely grown at 37°C in Luria-Bertani medium (L-broth or 1.5% L-agar). *S. flexneri* serotype five strains M90TS (wild-type), Sh4 (*dsbA::kan*), Sh42 (*dsbA33G*), and Sh54 (*yihEΔ398*) were routinely grown at 37°C overnight on tryptic soy agar (TSA) containing 0.01% Congo red. Red colonies were inoculated into tryptic soy broth (TSB) and grown to an appropriate turbidity at 37°C with shaking (200 rpm) for subsequent experiments. Antibiotics, when necessary, were added to final concentrations: streptomycin 100 μg/ml, and ampicillin 200 μg/ml.

Construction of Sh54 (yihEΔ398). The *yihE-dsbA* region of 2722 bp was amplified from M90TS genomic DNA by the polymerase chain reaction (PCR) using primers OA1 (5′-cgtgtctgtctcaagagtaa-3′) and OB1(5′-gccttcaatatccaggttag-3′), and cloned into the plasmid vector pGEM-T (Promega). An inverse PCR was performed on the resultant plasmid using the primers OA3 (5′-gctctagaccggaacttcatca-3′) and OA4 (5′-gctctagatctggatgatgcacgt-3′) to generate a product of 5734 bp with *Xba*I restriction sites at each end (underlined bases in the primers). After *Xba*I digestion, the DNA was circularized using T4 DNA ligase and transformed into XL-1 Blue, producing a clone with a 398-bp internal deletion and stop codon in the *Shigella yihE* gene, confirmed by DNA sequencing. The insert was then subcloned as a *Sac*I–*Sph*I fragment into the suicide vector pCVD422, and the resultant clone was used for allelic exchange to replace the wild-type *yihE* gene in the chromosome of M90TS. This gave rise to a *yihEΔ398* mutant designated Sh54, confirmed by PCR and DNA sequencing. Sh54 can be anticipated to produce a truncated YihE composed of the first 84 amino acid residues, a quarter of the mature protein sequence (328 amino acids).

Total RNA isolation, radioactive labeling, and array hybridization. All *Shigella* strains were cultured in TSB to an optical density (OD_{600nm}) of approximately 0.8, and total RNA isolated using the RNeasy mini kit (Quigen). Preparations were treated with DNase I (Life Technologies) to eliminate genomic DNA contamination. Concentrations of the nucleic acids were determined by measuring light absorption of 260 nm (A_{260nm}), and 1 μg of total RNA from each preparation was reverse-transcribed using Superscript II reverse-transcriptase to generate "first-strand" cDNAs (Life Technologies). The cDNAs were labeled with [α-^{33}P]-dCTP (Amersham Biotech) using an *E. coli* gene-specific primer mix (Sigma-Genosys), and purified for use as probes using MicroSpin columns (Pharmacia Biotech).

Panorama membranes of the *E. coli* DNA high-density arrays were purchased from Sigma-GenoSys. The array membranes were prehybridized at 42°C for 2 h in 15 ml Denhardt's solution (14), and then hybridized with the above probes at 42°C for 60 h in 10-ml Denhardt's solution. After stringent washing at 65°C for 1 h in a 2 liter solution containing 0.5% SDS and 0.1 × SSC, the membranes were exposed to a phosphorimage screen for 4 days. Electronic images

were obtained by scanning the screen at 50 microcon using a Storm-860 phosphorimager (Molecular Dynamics).

Array data analysis. The TIFF image files generated with the PhosphorImager were processed using the Array Vision software at Sigma-Genosys (TX). The pixel density (intensity) of duplicated spots of each of the 4290 genes was determined, and the average pixel density of the duplicated spots of each gene was normalized with regard to the specific activity of the probes used, and expressed as a percentage of total genomic pixels. The percentage expressions of all the 4290 genes hybridized with probes of Sh4, Sh42, and Sh54 were then respectively plotted against the percentage expressions of all the 4290 genes hybridized with probes of M90TS. As the majority of genes were transcribed at the same level in all strains, the means of the plots (Sh4 vs M90TS, Sh42 vs M90TS, and Sh54 vs M90TS) were all set to zero. Thus, positive values indicate that the transcription of genes in Sh4, Sh42, and Sh54 was higher than in M90TS, while negative values indicate lower transcription in the mutants than in M90TS. Genes with either positive or negative values equal or greater than 2X standard deviation from the mean zero were considered significantly changed in their transcription under the growth conditions used.

RESULTS AND DISCUSSION

Phenotypic Characteristics of Sh54

Figure 1 illustrates the *yihE-dsbA* locus in M90TS, Sh4, Sh42, and Sh54. Sh54 formed small colonies on TSA and had a prolonged doubling time similar to Sh4 (45 min). Like the *dsbA* mutants, Sh54 did not cause keratoconjunctivitis in guinea pigs. However, Sh54 exerted unusual cytotoxicity to cultured epithelial cells (data not shown), which is a characteristic of neither the *dsbA* mutants nor the virulent *S. flexneri* (15), suggesting that *yihE* functions differently than *dsbA*.

Quality of the Hybridization and the Whole Genomic Perspective

A correlation analysis was applied to the raw pixel volumes of the duplicated spots of all 4290 genes hybridized with probes of M90TS and the three mutants. In all cases, the coefficients of correlation were found to be 0.99, indicating that the array membranes were of good quality and that the hybridization was reproducible.

The correlation analysis was then applied between arrays, i.e., to correlate the percentage expressions of all the 4920 genes hybridized with probes of M90TS to the percentage expressions of all the 4290 genes hybridized with probes of Sh4, Sh42, and Sh54, respectively. The correlation between M90TS and Sh42 was found to be highest (coefficients = 0.95), while the correlation coefficients with Sh4, and Sh54 were poorer (coefficients = 0.85 and 0.81 respectively). In other words, under the growth conditions used, Sh42 had fewer, and Sh4 and Sh54 more, genes that were expressed at different levels compared to the levels in M90TS. When the percentage expression of each of the 4290 genes from Sh4, Sh42, and Sh54 was plotted respectively against that of M90TS, a normal distribution with the mean at zero became apparent in all the cases (Fig. 2). Only a few genes are apparently affected by mutation in *dsbA* or *yihE*, the majority of the genes being unaffected in their expression. The *dsbA33G* mutation had the smallest impact, as indicated by the smallest standard deviation of 0.29. Just 20 genes fulfil the criterion of significance (≥2 standard deviation = 0.58), with 10 up- and 10 down-regulated in transcription. The *dsbA::kan* and the *yihEΔ398* mutations apparently had greater impact, with 99 and 92 genes fulfilling the criterion of significance, respectively (≥2× standard deviation = 1.5). Sh4 had 50 up- and 49 down-regulated genes, while Sh54 had 22 up- and 70 down-regulated genes, a distribution skewed to the negative side (Fig. 2). This suggests that *yihEΔ398* and *dsbA::kan* have different effects on the genomic expression.

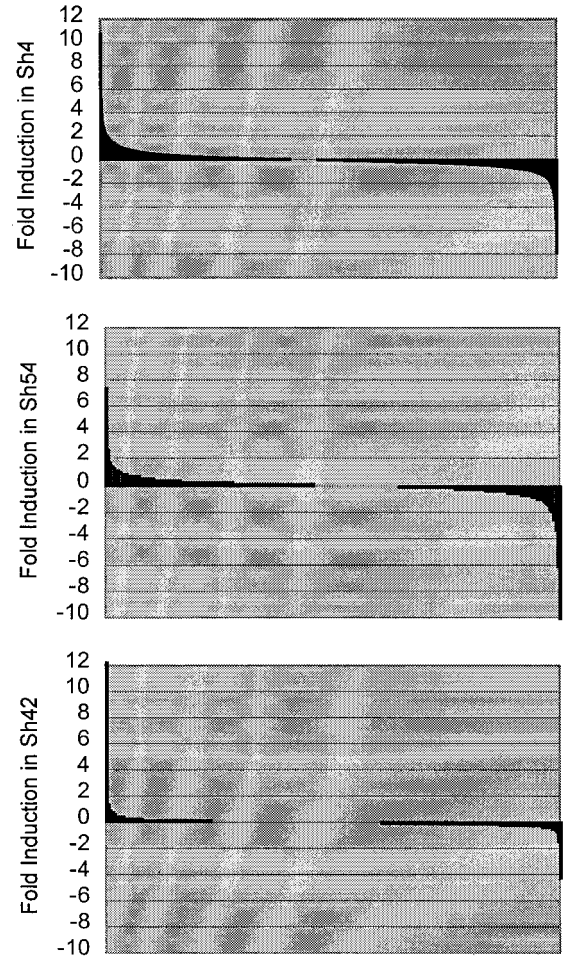

FIG. 2. Whole genomic perspective. The percentage expression of the average pixel volumes of each of the 4290 genes expressed in Sh4, Sh54, and Sh42 was plotted against that of M90TS to give rise to the Fold Induction indicated by the Y axes. Genes up- and down-regulated are expressed as positive and negative values respectively from the mean (0.00, the X axis) that indicates the fact that the majority of genes did not change their expression under the conditions used. The standard deviations from the mean zero are 0.29 for Sh42 and 0.75 for both Sh4 and Sh54.

TABLE 1

Up and Down-Regulated Genes Involved in Small-Molecule Metabolism

Pathway/gene name	Fold induction Sh4	Sh54	Sh42	Swissprot Accession No.	Gene product description
(A) Energy metabolism					
Glycolysis and TCA cycle					
gpmA		−10.1		P31217	phosphoglycerate mutase 1
gltA	−3.6	−3.4		P00891	2-oxoglutarate dehydrogenase E1 component, synthesesing succinyl-CoA
sdhA	−3.4	−3.2		P10444	succinate dehydrogenase flavoprotein subunit
sdhB	−5.5	−4.9		P07014	succinate dehydrogenase iron-sulfur protein
sucA	−8.9	−6.5		P07015	2-oxoglutarate dehydrogenase E1 component
sucB	−6.2	−4.2		P07016	dihydrolipoamide succinyltransferase component (E2)
sucC	−3.9	−3.6		P07460	succinyl-coA synthetase beta chain
sucD		−4.4		P07459	succinyl-coA synthetase alpha chain
fruB	4.1			P24217	Transport of carbohydrates, organic acids, alcohols. Pts system, fructose-specific IIA/FPR components
fruK	4.6			P23539	1-phosphofructokinase; ATP + D-fructose 1-phosphate = ADP + D-fructose 1,6-bisphosphate
Energy metabolism					
aldH	−3.7			P23833	aldehyde dehydrogenase, second step in ethanol utilization
cyoD	−3.3	−4.4		P18403	inner membrane, cytochrome o ubiquinol oxidase C subunit
lctD	−3.7	−4.1		P33232	inner membrane, L-lactate dehydrogenase, aerobic respiration and anaerobic nitrate respiration
fdnG		−6.0		P24183	anaerobic formate dehydrogenase major subunit
fdoG	−4.7	−3.3		P32176	formate dehydrogenase-O alpha subunit
fdoH	−5.5	−4.5		P32175	formate dehydrogenase-O beta subunit
fdoI	−4.3			P32174	formate dehydrogenase-O gamma subunit
narG	−3.2	−8.8		P09152	respiratory nitrate reductase 1 alpha chain
narH		−5.4		P33934	respiratory nitrate reductase 1 beta chain
narJ		−3.7		P11351	respiratory nitrate reductase 1 delta chain, promoting the correct association of the alpha and beta subunits
narK		−3.3		P10903	nitrite extrusion protein (nitrite facilitator), involved in anaerobic nitrate respiration
appB		6.6		P26458	cytochrome oxidase subunit II, stationary phase inducible
b1501	−7.6	3.9		P77561	homologue of formate dehydrogenase from Methanobacterium, converting formate + NAD to NADH + carbon dioxide
b2373	−4.6	5.3		P78093	homologue of Oxalobater oxalyl-CoA decarboxylase, converting oxaly-CoA to formyl-CoA + carbon dioxide
(B) Carbon degradation					
bglA	3.5			Q46829	6-phospho-beta-glucosidase, family of glycosyl hydrolases
galE		−4.2		P09147	UDP-glucose 4-epimerase
galK		−11.1		P06976	galactokinase
galT		−7.8		P09148	galactose-1-phosphate uridylyltransferase
galM		−5.3		P40681	aldose 1-epimerase (mutarotase), funnels beta-galactose into reactions of alpha-galactose catabolism
glcC	−3.1			P52072	glc operon transcriptional activator
b0271	−3.8	−3.3		P77713	putative beta-xylosidasehydrolysis of 1,4-beta-D-xylans
(C) Central intermediary metabolism					
gcvH	−4.6	−7.2	−2.9	P23884	lipoylprotein, glycine cleavage system H protein
gcvP	−5.7	−5.8	−2.6	P33195	glycine dehydrogenase (decarboxylating), shared with pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase
gcvT	−4.1	−6.3	−2.7	P27284	glycine degradation, aminomethyltransferase
aspA		−4.2		P04422	aspartase, converting aspartate and glutamine to asparagine
cysD			−5.4	P21156	ATP sulfurylase, catalyses the first reaction of the cysteine synthesis pathway (ATP + sulfate = APS + pyrophosphate)
talB	−5.9			P30148	transaldolase B, important for the balance of metabolites in the pentose-phosphate pathway
agaC	3.4			P42910	sugar phosphotransferase system (pts) system, n-acetylgalactosamine-specific IIC component 1
gadA		3.5		P80063	glutamate decarboxylase-alpha, catalyses L-glutamate to 4-aminobutanoate and CO2
gadB		3.1		P28302	glutamate decarboxylase-beta, catalyses L-glutamate to 4-aminobutanoate and CO2
(F) Amino acid biogenesis					
ilvC		−4.2		P05793	Isoleucine and valine synthesis: ketol-acid reductoisomerase
PheL		6		P03057	Leader peptide of chrorismate mutase-P-prephenate dehydratase
(D) Nucleotide salvage					
deoA	−3.3	−4.8		P07650	thymidine phosphorylase, reversible pjosphorolysis of pyrimidine nucleosides
deoB		−3.0		P07651	phosphopentomutase, transfer of a phosphate group to ribose and deoxyribose, respectively

TABLE 2

Other Genes Up and Down-Regulated

Pathway/gene name	Fold induction Sh4	Sh54	Sh42	Swissprot Accession No.	Gene product description
(A) Cell process and structure					
asr	11.7		14.7	P36560	acid shock protein
b2375	−5.1	8.4	−4	P76520	homologue to a yeast cell wall protein Pir3 involved in heat shock tolerance
csgE		3.3		P52105	assembly/transport component in fibronectin- and Congo red-binding curli polymers
csgF	5.0			P52104	surface structures, assembly/transport component in curli production
dinJ			7.4	Q47150	DNA repair, damage-inducible protein J
flgL	−3.1			P29744	flagellar hook-associated protein 3
fliR	−4.6			P33135	membrane protein, flagellar biosynthetic protein FliR
hdeA		3.1		P26604	acid resistance protein, abundant periplasmic protein at stationary phase and normally represed by H-NS
hdeB		3.4		P26605	acid resistance periplasmic protein of 12.5 kD
htpX	3.8	4.6		P23894	intergral membrane protein, adaptations, atypical conditions, heat shock protein HtpX
hupB	−4.2			P02341	histone like family, stabilizing DNA and prevent its denaturation under extreme environmental conditions
mopA	−3.7			P06139	GroEL protein, Chaperones
mopB	−3.4			P05380	GroES protein, Chaperones
ompC	−5.5			P06996	outer membrane protein, forms passive diffusion pores allowing small hydrophilic molecules across the outer membrane
ompF	6.1			P02931	outer membrane protein, small hydrophilic molecules diffusion and phage T2 receptor
sodA	−4.6	−3.2		P00448	Manganese superoxide dismutase, Detoxification
mviN	3.2			P75932	integral membrane protein, adaptations, atypical conditions, o511; 94 pct identical to MVIN_SALTY SW: P37169
pmrD			3.0	P37590	polymyxin B resistance and regulatory function
slp	−8.9	3.7		P37194	outer membrane protein inducible at carbon starvation and stationary phase
b0363	3.9			Q47536	membrane protein yaiP: polysaccharide metabolism
b0938	3.8			P75855	fimbrial-like protein ycbQ precursor
(B) Regulators					
acrR	3.5			P34000	acrAB operon repressor: efflux for acriflavine & carbon starvation
cspA	3.1			P15277	major cold shock protein cspA, involving in gyrA and hns activation
iclR	−3.1			P76268	acetate operon repressor: (malate + CoA = acetyl-CoA + H2O + glyoxylate). Outer membrane constituents
phoP			2.9	P23836	global transcriptional regulatory protein
nlp	3.7			P18837	positive regulation of the metabolism of sugars, belongs to the *ner* family of transcriptional regulators
rstA	3.1			P52108	rstA/B two-component signal transduction system, alternate name urpT; GTG start
yhiX	−3.1			P37639	transcriptional regulator required in glutamate-dependent acid stress response
b1526	3.6			P77309	putative transcribtional regulator belongs to the LysR family
b2847	5.0			Q46942	24% identical (2 gaps) to 127 residues of an approx. 296 aa protein of the *Vibrio cholerae* ToxR
(C) DNA replication recombination, modification					
fis	3.2			P11028	activates ribosomal RNA transcription, directly binding to the upstream of the RRNA promoters
b0247	3.6			Q47685	radC family, DNA repair function
b0354		3.9		P51024	This 218 aa ORF is 31% identical (2 gaps) to 82 residues of an approx. 1392 aa protein MST2_DROHY SW: Q08696
(D) Phage, transposon or plasmid					
insB_6	−4.2			P03830	required for transposition of IS1
ogrK		4.0		P27057	positive regulator of phage P2 late gene transcription
pspB	3.2			P23854	phage shock protein B, inducible by heat, ethanol and osmotic shock critical for survival at nutrient starvation
sieB	4.2			P38392	superinfection exclusion protein, 97% identical to 84 aa of SIEB_ECOLI SW: P38392 (114 aa)
(E) Transport					
cycP			−2.6	P16676	ABC transporter, thiosulfate-binding protein
cycW	6.1			P16702	sulfate transport system permease protein CycW
ftn	−3.8	−4.8		P23887	transport and ferritin-like protein
lacY	−3.5	−3.1		P02920	lactose permease: lactose-proton symport
manX		−7.3		P08186	phosphotransferase enzyme II, AB component
manY	−4	−8.8		P08187	phosphotransferase enzyme II, C component
manZ	−4.6	−8.2		P08188	phosphotransferase enzyme II, D component
modA		−4.1		P37328	molybdate-binding periplasmic protein
modB	−3.8			P09834	molybdenum transport system permease protein ModB
modE		−3.2		P46930	identical to modR, a supressor of the modABC operon
nupC		−3.5		P33031	nucleoside permease NupC
ybbA	−3.9	−3.9		P31219	hypothetical, drug/analog sensitivity
bfr	3.1			P11056	transport and binding protein, bacteroferritin
celC		3.3		P17335	transport: carbohydrates, organic acids, alcohols, phosphoenolpyruvate dependent phosphotransferase enzyme III-cellobiose
prlA		3.6		P03844	translocase SecY subunit
corA	3.2			P27841	magnesium and cobalt transport
fecC	5.1			P15030	iron(III) dicitrate transport system permease FecC
sscC	3.4			P075851	also known as b0934, ABC type transporter for sulfur from aliphatic sulfonates

TABLE 3

Putative or Hypothetical Genes Up or Down-Regulated

Gene name	Fold induction Sh4	Sh54	Sh42	Swissprot Accession No.	Description
(A) Putative enzymes					
b1501 (f759)	−4.6			P77561	putative oxidoreductase, major subunit
b2373 (f564)	−5.2			P78093	probable oxylyl-CoA decarboxylase
yidJ	−4.0			P31447	putative sulafatase
b1168 (o521)	3.4			P75995	putative proteases, contains 1 DUF2 domain
b1501 (f759)		4.3		H64903	putative oxidoreductase, major subunit
b2373 (f564)		5.3		P78093	probable oxalyl-CoA decarboxylase, similarity to OCX_OXAFO SW: P40149
b2878 (o1032)	3.9	3.2		F65071	putative oxidoreductase, Fe-S subunit
yjhG	4.0			P39358	hypothetical 70.1 kD protein in fecl-fimB intergenic region
yjhP			2.7	P39367	hypothetical 27.4 kD protein in fecl-fimB intergenic region
(B) Hypothetical					
b0663 (o111)	−3.4			E64801	Hypothetical, unclassified, unknown
b0667 (o45)	−3.4			F64801	Hypothetical, unclassified, unknown
b0725 (o86)	−7.6	−3.2		P75752	Hypothetical, unclassified, unknown
b1240 (f76)	−3.4			P76024	Hypothetical, unclassified, unknown
b1346 (f79)		−4.2		P76057	Hypothetical, unclassified, unknown
b1375 (f88)		−5.1		P76073	Hypothetical, unclassified, unknown, almost identical to *E. coli* YdfK
b1555 (f103)		−4.8		P76160	Hypothetical, unclassified, unknown
b1567 (f49)		−8.1		P76164	Hypothetical, unclassified, unknown
b1936 (o92)		−3.1		P76323	Hypothetical, unclassified, unknown
b2191 (o40)		−4.7		P76451	Hypothetical, unclassified, unknown
b2372 (f314)	−5.1			P76519	YfdV protein, belongs to an uncharacterised member of the AEC family of auxin efflux transporters
b2634 (o233)		−3.6		P76500	Hypothetical, unclassified, unknown
b2637 (o155)	−3.3			P52135	Hypothetical, unclassified, unknown, almost identical to *E. coli* YdfK
b2643 (o152)		−3.8		P52139	Hypothetical, unclassified, unknown, similarity to *E. coli* YafX, and plasmids antirestriction protein KlcA/KilC
b2654 (o110)		−7.2		P76616	Hypothetical, unclassified, unknown
b2363 (o101)		−4.0		P76516	Hypothetical, unclassified, unknown
b2666 (f52)		−3.7		P77240	Hypothetical, unclassified, putative integral membrane protein
b2756 (f199)		−3.6		Q46897	Hypothetical, unclassified, unknown
b2884 (o189)		−4.2		Q46817	Hypothetical, unclassified, unknown
b3002 (f164)		−3.3		P52082	Hypothetical, unclassified, unknown, putative integral membrane protein
b3837		−3.6		O32530	Hypothetical, unclassified, unknown
hdeA			−2.7	P26604	putative periplasmic protein
ydaC		−6.1		G64736	hypothetical protein in recT 3′ region
ydbA_2		−3.5		C48399	putative ABC transporter
yebJ	−3.1			G64994	hypothetical 4.2 kD protein in prc 5′ region
yfeC		−3.1		A65014	probably an iron (chelated) ABC transporter, permease protein; shares similarity with YfeC of *H. influenzae* Rd]
yggL		−4.4		P38521	hypothetical protein in mutY 5′ region
ygjV	−3.1			P42603	hypothetical 20.5 kD protein in ebgC-exuT intergenic region
yhaB		−4.0		P11865	hypothetical 20.6 kD protein in tdcR-mpB intergenic region
yhbL		−3.9		P26428	cross-reacting with antiserum to Sigma-70 or Sigma-30
yhbT		−3.8		P45474	hypothetical 19.7 kD protein in sohA-mtr intergenic region
yhcB		−3.2		P39436	hypothetical 15.2 kD protein in rplM-hhoA intergenic region
yhhA		−4.8		P23850	o146; 100% identical to YHHA_ECOLI SW: P23850; alternate gene names o146a, orfQ
yhhG		−4.5		P28910	hypothetical 15.1 kD protein in nike-rhsb intergenic region
yhiO		−3.4		AAC76519	hypothetical 13.0 kD protein in pit-uspA intergenic region
ykgE		−6.3	−2.7	P77252	o239; 29% identical (12 gaps) to 179 residues of the 396 aa protein GLPC_ECOLI SW: P13034
ykgF			−2.9	P77536	o475; 24% identical (9 gaps) to 163 residues of approx. 432 aa protein GLPC_HAEIN SW: P43801
ykgG				P77433	26% identical (2 gaps) to 97 residues of approx. 168 aa protein FMA2_BACNO SW: P17824
yqgD		−3.5		P46879	hypothetical 9.5 kD protein in speA-metK intergenic region
b0808 (f786)	3.2			P75783	putative transport protein
b1387 (f681)	5.1			P77455	phenylacetic acid degradation protein PaaZ

TABLE 3—*Continued*

Gene name	Fold induction Sh4	Sh54	Sh42	Swissprot Accession No.	Description
b1389 (o95)	3.1			P76078	may be part of a multicomponent oxygenase involved in phenylacetyl-CoA hydroxylation
b1641 (o155)			2.9	p55741	outer membrane lipoprotein SlyB
b1825 (f95)			2.5	P76266	hypothetical 21.0 kD protein in panB-htrE intergenic region
b1826 (f47)			3.1	P76267	f47; This 47 aa ORF is 34% identical (3 gaps) to 43 residues of an approx. 176 aa protein MRED_BACSU SW: Q01467
b1978 (o2383)	4.4			P76347	strong similarity to numerous bacterial attaching and effacing proteins and invasins
b2085 (f125)		3.3		P76406	f125
b2191 (o40)			2.4	P76451	o40; This 40 aa ORF is 38% identical (1 gap) to 31 residues of an approx. 384 aa protein ILEU_HUMAN SW: P30740
b2374 (f416)		3.6		P77407	f416
b2629 (f87)	3.3			P52128	f87; This 87 aa ORF is 31% identical (6 gaps) to 58 residues of an approx. 128 aa protein YHIK_ECOLI SW: P37628
b2852		4.4		P76639	Hypothetical, homologue of Salmonella invasion lagA
b2971 (f136)	3.3			Q46835	putative envelope protein attached to the membrane by a lipid anchor
b3238 (o104)			6.5	P46477	o104
b3254 (f33)	3.5			Q47712	f33; This 33 aa ORF is 57% identical (1 gap) to 21 residues of an approx. 304 aa protein NC5R_HUMAN SW: P00387
b3263 (o59)	3.1			P45764	o59
eutl		3.7		A65021	Hypothetical ethanolamine utilization protein Eutl, degradation of small molecules: Amines
yadC	4.2			P31058	putative fimbrial-like protein
yadL	6.9			P37017	Drug/analog sensitivity, hypothetical fimbrial-like protein in panB-htrE intergenic region
yafQ	3.3			B64747	hypothetical protein in gmhA-fhiA intergenic region
ybbD	3.7			P33669	o86; 100% identical to 67 residues of YBBD_ECOLI SW: P33669 (78 aa) but differs at C-term
yhbM		4.4	3.3	AAC76519	putative control proteins 33.6 kD protein in dead-pnp intergenic region
yebE	4.0	4.1		P33218	hypothetical 23.7 kD protein in purT 5′ region
yebK	3.4			E64947	hypothetical 32.0 kD protein in pykA-zwf intergenic region
yehI	3.6			P33346	hypothetical 138.1 kD protein in molR-bglX intergenic region
yfhJ			3.0	P37096	hypothetical 7.7 kD protein in fdx 3′ region
yjbL		4.7		P32693	hypothetical 9.7 kD protein in dinf-qor intergenic region (o8)
yjdA			6.4	P16694	Outer membrane constituents, hypothetical 84.2 kD protein in phnA-proP intergenic region
yjiY	3.7			P39396	Not classified, hypothetical 77.9 kD protein in mrr-tsr intergenic region

Significantly Changed Expression of the Genes Involved in Small-Molecule Metabolism

Table 1 lists genes involved in small-molecule metabolism with significantly changed expression. Most strikingly, many genes involved in energy metabolism were down-regulated in Sh4 and Sh54 but not in Sh42, consistent with their more pronounced slowing of growth. These included genes involved in the tricarboxylic acid (TCA) cycle: *sucAB* encoding α–ketoglutarate dehydrogenase, *sucCD* encoding succinyl CoA synthetase, *sdhCDAB* encoding succinate dehydrogenase, and *gltA* encoding citrate synthase. As the TCA cycle is the most efficient route to ATP synthesis, down-regulation of these genes will undoubtedly result in limited energy supply. In Sh54, the transcription of *gmpA*, encoding phosphoglycerate mutase, was additionally decreased, limiting the supply of 2-phosphoglycerate (the substrate for pyruvate kinase and pyruvate dehydrogenase), reducing the TCA cycle activity even further. Sh4 had increased transcription of *fruB* and *fruK*, enhancing uptake and utilization of exogenous fructose via glycolysis, which may compensate for a reduction in TCA cycle efficiency.

Transcription of *cyoD* and *lctD*, involved in aerobic respiration, was reduced in both Sh4 and Sh54. Several genes involved in anaerobic respiration were also down-regulated in Sh4 and Sh54. These include *fdnG*, *fdoG*, *fdoH*, and *fdoI* that are involved in formate metabolism, and *narG*, *narH*, *narJ*, and *nirK* that are involved in nitrate reduction. Interestingly, transcription of *b1501* and *b2373* were reduced in Sh4 but increased in Sh54. These, respectively, encode homologues of the *Methanobacterium* formate dehydrogenase (16) and the *Oxalobacter* oxalyl-CoA decarboxylase (17) indicating that they also have physiologic roles in conditions such as the inactivation of *dsbA* and *yihE*. Interestingly also, Sh54 had increased transcription of

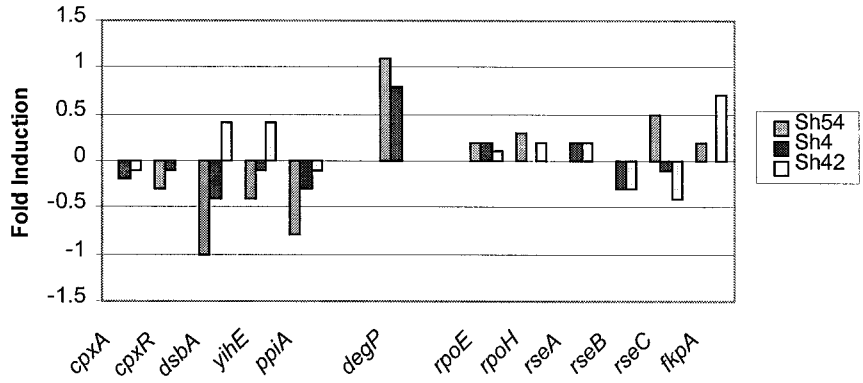

FIG. 3. Transcription of the genes belong to the *cpxRA* and σ^E regulons. The linear percentage of the average pixel volumes of each of the genes expressed in Sh4, Sh54, and Sh42 was plotted against that of M90TS to give rise to the Fold Induction indicated by the Y axis. The *yihE-dsbA* transcripts are represented by the value of *yihE*, and the value of *dsbA* represents the combination of *dsbA* and *yihE-dsbA*. As both σ^E and CpxRA control *degP* expression, the values of *degP* transcription is illustrated in the middle.

appB that encodes an alternative cytochrome oxidase (18), indicating that this gene is required under conditions such as inactivation of *yihE*.

In central intermediary metabolism, the *gcv* system for catalysis of the reversible oxidation of glycine was down-regulated in all three mutants, imposing an additional adverse effect on energy metabolism to Sh4 and Sh54. The effect in Sh42 is not clear, as its TCA and glycolysis pathways were not affected. However, since the *gcvP* product, glycine dehydrogenase (E3), is shared with pyruvate dehydrogenase, the reduction of *gcvP* can cause a negative effect on energy metabolism in general.

Other Up- and Down-Regulated Genes

Genes with known functions are listed in Table 2, and those with hypothetical or with unclassified functions are listed in Table 3. Strikingly, each of the three strains expressed a small unique set of genes involved in various stress responses, suggesting that each of the mutations they carry triggers a unique stress response (Table 2). Some of these changes are counter-intuitive. For example, in Sh4 the expression of *mopA* and *mopB* (encoding GroEL and GroES, respectively) was reduced, but the expression of *htpX* (encoding the atypical heat shock protein HtpX) was increased. This is puzzling because *mopA*, *mopB*, and *htpX* are all positively regulated by σ^{32} (19). The increase of *pmrD* in Sh42 is consistent with the increase of *phoP* (20). PmrD confers polymyxin B resistance and also acts as a secondary transcription regulator under PhoPQ. Since the genes downstream under their regulation, such as *pmrCAB*, were not included in the DNA arrays, the effect of the increase of *pmrD* is not yet clear. It is noted that genes up- or down-regulated in Sh42 are not in common with those in Sh54 in this category, suggesting that the stress imposed by defective disulfide bond formation is different than that by distorted YihE expression. On the other hand, the transcription of the heat inducible gene *htpX* (21) is increased in both Sh4 and Sh54—possibly caused by decreased *yihE* expression. In addition, *sodA* encoding the manganese-containing superoxide dismutase, crucial for protection from superoxide radical damage, was down-regulated in Sh4 and Sh54, which may account for the slow growth of these two mutants (22). *sodA* is negatively regulated by Arc (aerobic respiration control), FNR (anaerobic respiration control), IHF (integration host factor), and Fur (Fe uptake) while positively regulated by the products of *soxRS* (superoxide response) and certain *soxQ* alleles. Since only *arcR* was reduced in Sh4, and the levels of expression of the rest of these regulators did not changed significantly, the underlying regulation mechanism can not be revealed by the array data.

The Expression of the cpxRA and σ^E Regulons

Transcription of *yihE-dsbA* is controlled by the two-component signal transduction system, CpxRA, which together with σ^E governs the extracytoplasmic stress response (12). Expression of genes belonging to these two regulons was, therefore, analyzed to ascertain the effects of *yihE* and *dsbA* on gene expression (Fig. 3). With the exceptions of the increase of *dsbA* and *yihE* in Sh42, transcription of other genes appeared to be consistent with the expression patterns of the regulators σ^E, *cpxR* and *cpxA*. As expected, transcription of *dsbA* and *yihE* was deceased in Sh4 and Sh54. Additionally, transcription of *ppiA*, encoding the periplasmic peptidyl-prolyl *cis*/*trans* isomerase (PPI), also decreased in Sh4 and Sh54, consistent with the negative values of *cpxR and cpxA*. The increased transcription of *fkpA* (another periplasmic PPI) and *degP* (a major periplasmic protease) in the mutants is as expected for the positive value of σ^E. Transcription of *degP* is known to be controlled by CpxRA and σ^E (12). In the cases of

Sh4 and Sh54, however, the increased transcription can only be explained by the increased σ^E. Furthermore, the increase in the level of *degP* transcription in Sh54 was nearly statistically significant (1.1 vs 1.5), suggesting that σ^E-mediated extracytoplasmic stress response may well contribute to the overall stress-state in Sh54. Moreover, both Sh4 and Sh54 had increased transcription of *htpX* (Table 2) that is dependent on σ^{32}, indicating that the small increase of the transcription of σ^{32} must be biologically significant.

In conclusion, both *yihEΔ398* and *dsbA::kan* cause a major change in genomic expression while *dsbA33G* makes a minimal impact. This suggests that reduced expression of *yihE* is principally responsible for the global change in Sh4 and Sh54. However, these results do not represent a definitive analysis of gene regulation as a result of inactivation of *yihE* or *dsbA* as they reflect a single analysis of a single set of growth conditions at a single growth phase. Nevertheless, the quality of the array hybridization and reproducibility do offer the confidence to predict that *yihE* is an important gene in the *cpxRA* regulon. The concept that inactivation of *dsbA* triggers extracytoplasmic stress response (23) or sends a periplasmic-stress signal (24), now needs qualification, since mutants similar to Sh4 were used in these studies. The results, therefore, can also be complicated by the altered *yihE* expression. The clearly different patterns of genomic expression of the two-*dsbA* mutants have offered a basis for explanations as to why they possess different properties *in vitro* and *in vivo*. At least it explains well why Sh4 grows more slowly than Sh42: so many genes involved in energy metabolism are down-regulated in the former but not in the latter. Given its unique localization, *yihE* is probably involved in the extracytoplasmic stress response, indicated by the up-regulation of *degP* in Sh4 and Sh54. With the progress of analytic studies, we hope that the function of *yihE* will be detailed soon.

ACKNOWLEDGMENTS

We thank Antoinette Hartman for performing the Sereny test, and Ganine Bosse and Philip Butcher for critical reading the manuscript. This work is financed by a Career Development Award to Jun Yu from The Wellcome Trust, UK (Grant No. 047657/Z/96/Z). M-S Li is supported by a grant from Wellcome Trust awarded to J.S. Kroll.

REFERENCES

1. Sansonetti, P. J., Kopecko, D. J., and Formal, S. B. (1982) Involvement of a large plasmid in the invasive ability of *Shigella flexneri*. *Infect. Immun.* **35,** 852–860.
2. Rajakumar, K., Sasakawa, C., and Adler, B. (1997) Use of a novel approach, termed island probing, identifies the *Shigella* she pathogenicity island which encodes a homologue of the immunoglobulin A protease-like family of proteins. *Infect. Immun.* **65,** 4606–4604.
3. Vokes, S. A., Reeves, S. A., Torres, A. G., and Payne, S. M. (1999) The aerobatin iron transport system genes in *Shigella flexneri* are present within a pathogenicity island. *Mol. Microbiol.* **33,** 63–67.
4. Moss, J. E., Cardozo, T. J., Zychkinsky, A., and Groisman, E. A. (1999) The *selC*-associated SHI-2 pathogenicity island of *Shigella flexneri*. *Mol. Microbiol.* **33,** 74–83.
5. Way, S. S., Sallustio, S., Magliozzo, R. S., and Goldberg, M. B. (1999) Impact of either elevated or decreased levels of cytochrome bd expression on *Shigella flexneri* virulence. *J. Bacteriol.* **181,** 1229–1237.
6. Reeves, S. A., Torres, A. G., and Payne, S. M. (2000) TonB is required for intracellular growth and virulence of *Shigella dysenteriae*. *Infect. Immun.* **68,** 6329–6336.
7. Yu, J. (1998) Inactivation of DsbA, but not DsbC and DsbD, affects the intracellular survival and virulence of *Shigella flexneri*. *Infect. Immun.* **66,** 3909–3917.
8. Yu, J., Edwards-Jones, B., Neyrolles, O., and Kroll, J. S. (2000) Key role for DsbA in cell-to-cell spread of *Shigella flexneri*, permitting secretion of Ipa proteins into interepithelial protrusions. *Infect. Immun.* **68,** 6449–6456.
9. Yu, J., Emmanuelle, E. O., Stephens, A., Kroll, J. S., and Venkatesan, M. M. (2001) Inactivation of DsbA alters the behaviour of *Shigella flexneri* toward murine and human derived macrophage-like cells. *FEMS Microbiol. Lett.* (in press).
10. Watarai, M., Tobe., T., Yoshikawa, M., and Sasakawa, C. (1995) Disulfide oxidoreductase activity of *Shigella flexneri* is required for release of Ipa proteins and invasion of epithelial cells. *Proc. Natl. Acad. Sci. USA* **92,** 4927–4931.
11. Belin, P., and Boquet, P. L. (1994) The *Escherichia coli dsbA* gene is partly transcribed from the promoter of a weakly expressed upstream gene. *Microbiology* **140,** 3337–3348.
12. Raivio, T. L., and Silhavy, T. J. (1999) The σ^E and Cpx regulatory pathways: Overlapping but distinct envelope stress responses. *Curr. Opin. Microbiol.* **2,** 159–165.
13. Donnenberg, M. S., and Kaper, J. B. (1991) Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. *Infect. Immun.* **59,** 4310–4317.
14. Sambrook, J., Fritsch, E. F., and Manistis, T. (1989) Molecular Cloning, A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
15. Mantis, N., Prevost, M. C., and Sansonetti, P. (1996) Analysis of epithelial cell stress response during infection by *Shigella flexneri*. *Infect Immun.* **64,** 2474–2482.
16. White, W. B., and Ferry, J. G. (1992) Identification of formate dehydrogenase-specific mRNA species and nucleotide sequence of the *fdhC* gene of *Methanobacterium formicicum*. *J. Bacteriol.* **174,** 4997–5004.
17. Lung, H. Y., Baetz, A. L., and Peck, A. B. (1994) Molecular cloning, DNA sequence, and gene expression of the oxalyl-coenzyme a decarboxylase gene, *oxc*, from the bacterium *Oxalobacter formigenes*. *J. Bacteriol.* **176,** 2468–2472.
18. Dassa, J., Fsihi, H., Marck, C., Dion, M., Kieffer-Bontemps, M., and Boquet, P. L. (1991) A new oxygen-regulated operon in comprises the genes for a putative third cytochrome oxidase and for pH 2.5 acid phosphatase. *Mol. Gen. Genet.* **229,** 341–352.
19. Gross, C. A. (1996) Function and regulation of the heat shock proteins. *In Escherichia coli* and Salmonella (Ferderrick, C., and Neidhardt, T., Eds.), pp. 1382–1399, ASM Press, Washington DC.
20. Groisman, E. A. (2001) The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol.* **183,** 1835–1842.

21. Kornitzer, D., Teff, D., Altuvia, S., and Oppenheim, A. B. (1991) Isolation, characterization, and sequence of an *Escherichia coli* heat shock gene, *htpX*. *J. Bacteriol.* **173,** 2944–2953.

22. Lynch, A. S., and Lin, E. C. C. (1996) Responses to molecular oxygen. *In Escherichia coli* and Salmonella (Ferderrick, C., and Neidhardt, T., Eds.), pp. 1526–1538, ASM Press, Washington DC.

23. Missiakas, D., and Raina, S. (1997) Protein folding in the bacterial periplasm. *J. Bacteriol.* **179,** 2465–2471.

24. Vincent-Sealy, L. V., Thomas, J. D., Commander, P., and Salmond, G. P. C. (1999) Erwinia carotovora DsbA mutants: Evidence for a periplasmic-stress signal transduction system affecting transcription of genes encoding secreted proteins. *Microbiology* **145,** 1945–1958.