JOURNAL OF BACTERIOLOGY, Aug. 2009, p. 5180–5195 Vol. 191, No. 16
0021-9193/09/$08.00+0 doi:10.1128/JB.00476-09

Sxy Induces a CRP-S Regulon in *Escherichia coli*▽†

Sunita Sinha,* Andrew D. S. Cameron,‡ and Rosemary J. Redfield

Department of Zoology, University of British Columbia, Vancouver V6T 3Z4, Canada

Received 7 April 2009/Accepted 1 June 2009

***Escherichia coli* is not considered naturally competent, yet it has homologues of the genes that most competent bacteria use for DNA uptake and processing. In *Haemophilus influenzae* and *Vibrio cholerae*, these genes are regulated by the Sxy and cyclic AMP receptor (CRP) proteins. We used microarrays to find out whether similar regulation occurs in *E. coli*. Expression of *sxy* strongly induced 63 transcriptional units, 34 of which required CRP for transcriptional activation and had promoter sites resembling the Sxy- and CRP-dependent CRP-S motif previously characterized in *H. influenzae*. As previously reported, *sxy* expression also induced the sigma-H regulon. Flagellar operons were downregulated by *sxy* expression, although motility remained unaffected. The CRP-S regulon included all of *E. coli*'s known competence gene homologues, so we investigated Sxy's effect on competence-associated phenotypes. A *sxy* knockout reduced both "natural" plasmid transformation and competitive fitness in long-term culture. In addition, expression of plasmid-borne *sxy* led to production of type IV pilin, the main subunit of the DNA uptake machinery of most bacteria. Although *H. influenzae* Sxy only weakly activated the *E. coli* Sxy regulon, induction was dramatically improved when it was coexpressed with its cognate CRP, suggesting that intimate interactions between Sxy and CRP are required for transcriptional activation at CRP-S sites.**

Natural competence is the genetically programmed ability of bacteria to actively take up DNA from their environment. Although several studies of *Escherichia coli* have identified conserved homologues of all genes required by most competent bacteria to take up DNA (12, 15, 22, 59), researchers have for decades struggled to demonstrate that *E. coli* takes up and recombines chromosomal DNA, and introduction of plasmids into *E. coli* cells relies on artificial permeabilization by calcium chloride or on electroporation. Although more "natural" variants of these techniques have been reported (7, 8, 10, 82), no competence genes have been implicated, and a recent study found that plasmid transformation of *E. coli* on agar plates does not require four putative competence genes whose homologues are essential for DNA uptake in other bacteria (70).

Finkel and coworkers have, however, demonstrated that *E. coli* can use external DNA as its sole carbon source and that this requires homologues of the *Haemophilus influenzae com* operon, strongly suggesting that the products of these homologues can be used for DNA uptake (22, 48). Moreover, although *E. coli* K-12 has all of the genes required for type IV pilus (T4P) biogenesis and function (whose products are essential for DNA uptake in most bacteria), T4P are not produced, and Sauvonnet et al. were unable to find conditions under which they are significantly expressed (59). Both of these groups of workers have hypothesized that *E. coli* fails to express T4P and develop competence because it does not fully induce the necessary genes (48, 59, 71).

In *H. influenzae*, the model organism for competence regulation in gram-negative bacteria, induction of competence genes requires two positive regulators of transcription, cyclic AMP (cAMP) receptor protein (CRP) and Sxy (56). CRP is the master regulator of the carbon-energy starvation response, but it takes on a distinct role when Sxy is present, upregulating genes necessary for DNA uptake, transport, and processing, as well as other genes whose role in competence has not been determined experimentally (56). Each promoter in the *H. influenzae* Sxy-CRP regulon contains a CRP-S site (termed "S" for Sxy-dependent), variant CRP-binding sites that match the canonical CRP consensus sequence except for two positions where nonconsensus bases hinder stable CRP-DNA interactions in the absence of Sxy (11). Sxy is thus thought to enable CRP to activate transcription at these unusual binding sites, although its toxicity when it is overexpressed for purification has hindered attempts to characterize its exact mode of action (11, 35, 45, 56).

Similar regulatory mechanisms are likely to exist in *E. coli*: in addition to its well-characterized *crp* gene, *E. coli* has a *sxy* orthologue (b0959, also known as *yccR* or *tfoX*) and has predicted CRP-S sites in 9 of the 12 transcriptional units that contain homologues of *H. influenzae* CRP-S genes (12). Two of these sites have been shown to require both Sxy and CRP for activation (12), but the other sites have not been examined. Because our previous work examined only homologues of *H. influenzae* CRP-S genes, additional *E. coli* CRP-S sites may exist either in genes not associated with competence or in genes not present in *H. influenzae*. Furthermore, although in *H. influenzae* Sxy acts only at CRP-S sites, we do not know the scope of its action in *E. coli*.

If Sxy is indeed required for transcription of *E. coli*'s competence gene homologues, then a failure to induce *sxy* expression may explain why these genes were poorly expressed under any of the culture conditions tested by Sauvonnet et al. (59) and Palchevskiy and Finkel (48). In *H. influenzae*, *sxy* expres-

* Corresponding author. Mailing address: Department of Zoology, University of British Columbia, Life Sciences Center, Vancouver, Canada V6T 3Z4. Phone: (604) 822-6323. Fax: (604) 827-4135. E-mail: sinha@zoology.ubc.ca.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: Department of Microbiology, Trinity College Dublin, Dublin 2, Ireland.

▽ Published ahead of print on 5 June 2009.

sion is limited by both transcriptional and posttranscriptional regulation (13), but nothing is known about its regulation in *E. coli*. As described below, we have not been able to identify culture conditions that induce *E. coli sxy*. Consequently, we used artificially induced *sxy* to identify genes regulated by *E. coli* Sxy using whole-genome microarrays, and we analyzed promoters of member genes to find regulatory motifs. We then investigated the effect of *sxy* induction on competence-associated phenotypes and compared the activities of the *H. influenzae* and *E. coli* Sxy proteins.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. *E. coli* K-12 strain BW25113 (*lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$) and its derivatives JW0942 (*sxy*::kan), JW3778 (*cya*::kan), JW5702 (*crp*::kan), and JW1908 (*fliC*::kan) were obtained from the Keio KO collection (4). A spontaneous nalidixic acid-resistant (Nal^r) mutant of the parent strain was isolated by growth on increasing concentrations of the antibiotic and was used in batch culture competition assays and as the DNA source for chromosomal transformation assays. Plasmids p*Ecsxy* and p*EcppdD* (Cm^r) were obtained from Genobase ASKA/GFP(-) (35); in both of these plasmids, the gene is cloned with an N-terminal six-His tag under control of the isopropyl-β-D-1-thiogalactopyranoside (IPTG)-inducible T5-*lac* promoter. All *E. coli* strains were grown in Luria-Bertani (LB) broth or agar (1.2%) or in liquid M9 minimal medium (6.78 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 0.4% glucose) at 37°C. When required, antibiotics were used at the following concentrations: kanamycin, 10 μg/ml; chloramphenicol, 20 μg/ml; nalidixic acid, 20 μg/ml; ampicillin, 100 μg/ml; and tetracycline, 10 μg/ml.

To create plasmid p*Hisxy*, the coding sequence of the *H. influenzae sxy* gene HI0601 (without the start and stop codons) was amplified from the genome of strain Rd KW20. Primer sequences are shown in Table S1 in the supplemental material. Plasmid p*Ecsxy* was digested with SfiI, and the released *E. coli sxy* insert was replaced with the *sxy* gene from *H. influenzae* by ligation. A "no-insert" control plasmid, designated pnoins, was also created by religating the SfiI-linearized vector.

Plasmids p*Eccrp* and p*Hicrp* were constructed by PCR amplifying the predicted promoter region and coding sequence of the *crp* genes from the genomes of *E. coli* strain BW25113 and *H. influenzae* strain Rd KW20, respectively, and cloning them in the vector pACYC184 (Tet^r). The primer sequences are shown in Table S1 in the supplemental material.

Growth conditions for RNA preparation. (i) Monitoring *sxy* expression under different conditions. Samples of broth- and agar-grown bacteria were obtained after overnight culture by removing 2 ml liquid culture and by scraping bacteria from the agar and resuspending them in 2 ml phosphate-buffered saline (PBS), respectively. To obtain log-phase cells, an overnight culture was diluted 1:100 in LB broth and grown at 37°C, and 2-ml samples were removed after 1 h, 3 h, and 5 h of growth. To assay for *sxy* induction by nutrient limitation, cells from 10 ml of this culture at an optical density at 600 nm of 0.2 were collected on a nitrocellulose filter, washed with and resuspended in 10 ml M9 minimal medium, and grown at 37°C. Two-milliliter samples were removed from the LB medium culture (zero time) and after 100 and 200 min of incubation in M9 minimal medium.

(ii) Global gene expression analysis. Bacteria were grown in LB broth to an optical density at 600 nm of 0.5, when expression of cloned genes was induced by addition of 1 mM IPTG. After 30 min of expression, 2 ml of the bacterial culture was removed for each strain.

(iii) Reciprocal complementation experiments. The conditions used for the reciprocal complementation experiments were the same as those used for the global gene expression analysis except that the strain was the *E. coli crp*::kan strain carrying two plasmids, one with the *E. coli* or *H. influenzae sxy* gene (p*Ecsxy* or p*Hisxy*) and one with the *E. coli* or *H. influenzae crp* gene (p*Eccrp* or p*Hicrp*).

RNA sample preparation. Each sample was mixed with 2 volumes of RNAProtect bacterial reagent (Qiagen) and incubated at room temperature for 5 min to stabilize RNA. Cells were then pelleted by centrifugation for 5 min at 3,500 × *g*, and total RNA was extracted using an RNeasy mini kit (Qiagen). DNA was removed by two successive treatments with Turbo DNase (Ambion Inc.). RNA purity was assessed by electrophoresis on 1% Tris-acetate-EDTA agarose, and concentrations were determined with a Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, United States). The absence of DNA contamination in all samples was verified by PCR.

Microarray analysis and qPCR. *E. coli* 3x6K microarray slides were obtained from the Microarray and Proteomics Facility of the Department of Biological Sciences at the University of Alberta. The arrays contained all open reading frames from three *E. coli* genomes (K-12, O157:H7 EDL933, and O157:H7 Sakai genomes). The arrays were hybridized and washed using the instructions in the Corning epoxide-coated slides manual (http://www.corning.com/lifesciences/pdf/epoxide_coated_slides_ss.pdf). Test and control RNA samples (2 μg) were labeled using Superscript II reverse transcriptase (Invitrogen) with Cy3-dCTP and Cy5-dCTP (Amersham), respectively. Samples were then mixed and purified using a MiniElute reaction cleanup kit (Qiagen). The hybridized arrays were scanned using a GenePix 4000B array scanner (Axon), and intensity fluorescence data were acquired using GenePix 6.0 (Molecular Devices). Expression data were analyzed using GeneSpring GX (Agilent). For each comparison, gene expression profiles from four independent microarray experiments were analyzed, using RNA samples from independent cultures. All data were filtered using Volcano plots with a *P* value cutoff of 0.1. The standard twofold-change cutoff was used for all arrays except the ±*sxy* arrays, where greater changes in gene expression allowed use of a more stringent fourfold-change cutoff. For confirmation of microarray data, quantitative PCR (qPCR) was performed as described by Cameron and Redfield (8, 12); one *sxy*-independent gene (*murA*) and two Sxy-induced genes (*hofM* and *ppdD*) were used as negative and positive controls, respectively. The number of 23S mRNA copies was also determined for each sample and used to correct for differences in the amount of RNA present. Primer sequences are shown in Table S1 in the supplemental material.

Promoter motif analysis. The promoter region DNA sequence (from −300 bp to the start codon of the first gene in each transcriptional unit, allowing overlap with upstream genes) was retrieved using RSATools (http://rsat.ccb.sickkids.ca/) (78). The unbiased motif-finding programs Gibbs Recursive Sampler (73) and Meme (5) were used to detect overrepresented motifs in promoter regions; Gibbs Recursive Sampler (73) was run at http://bayesweb.wadsworth.org/cgi-bin/gibbs.12.pl?data_type=DNA, and Meme was run at http://meme.nbcr.net/meme/meme.html. Parameters were set to search for optimal motifs over a width range of 16 to 50 bp and to allow for promoters with multiple or no sites matching a motif. The background DNA sequence required for calculation of motif strength was always modeled from the input promoter sequences. Only DNA sites identified by both programs as matching a significant motif were included in the analysis presented here.

Protein sample preparation and immunoblotting. Whole-cell extracts were prepared by resuspending plate-grown bacteria or pelleted broth-grown bacteria in PBS, followed by lysis for 1 h at 55°C. Samples were separated by electrophoresis in a 15% sodium dodecyl sulfate-polyacrylamide gel electrophoresis gel and electrotransferred to a polyvinylidene difluoride membrane (Hybond). The membrane was then incubated first with polyclonal antiserum raised against a fusion of *E. coli* PpdD and MalE (kindly provided by Olivera Francetic, Institut Pasteur, Paris, France, and described in reference 60) and second with an alkaline phosphatase-conjugated anti-rabbit immunoglobulin G (Sigma) used at 1:2,000 and 1:10,000 dilutions, respectively. Bound antibodies were detected with 5-bromo-4-chloro-3-indolyl phosphate tablets (Sigma).

Transmission electron microscopy. Cell surface structures were visualized with an Hitachi H7600 transmission electron microscope. Samples were prepared essentially as described by Sinha et al. (64), except that cells were fixed at 37°C for 15 min in 2.5% glutaraldehyde (prepared in 200 mM sodium cacodylate, pH 7.35). Briefly, grids were floated onto the bacterial suspension and transferred to droplets of PBS containing 1% bovine serum albumin (BSA) (PBS-BSA) and a 1:100 dilution of PpdD antiserum for 30 min, droplets of PBS-BSA for three times for 1 min, and droplets of PBS-BSA and a 1:20 dilution of gold-conjugated goat antiserum (Sigma) for 30 min. This was followed by three further PBS-BSA washes and then fixation in glutaraldehyde and counterstaining in uranyl acetate with two water washes between these procedures. For each sample, gold particles were counted in four independent areas of the grid adjusted to the same size, each with comparable numbers of cells.

Transformation assays. Transformation with plasmid pCR2.1-TOPO DNA (Invitrogen) (Amp^r) was performed as described by Baur et al. (8). Briefly, bacteria were grown in broth to mid-log phase, washed and resuspended in 1 ml of 1 mM $CaCl_2$, and incubated with 1 μg/ml of plasmid DNA for 20 min at 4°C. Five hundred microliters of LB broth was added, and the cells were allowed to recover for 45 min at 37°C before plating. The same protocol was used for assays of transformation with chromosomal DNA, with 1 to 10 μg/ml of DNA from the Nal^r strain described above. Artificial transformation with plasmid pCR2.1-TOPO DNA was performed by electroporation (2.5 kV, 200 Ω, 25 μF) or by heat shock (42°C for 90 s). Chemically competent or electrocompetent cells were prepared as described elsewhere (58). In all transformation experiments, four

colonies per strain were checked for the presence of the plasmid by the miniprep procedure.

Long-term survival and batch culture competition assays. The long-term survival in coculture of the Nal^r parent and *sxy*::kan *E. coli* strains was assessed as described by Palchevskiy and Finkel (48). Briefly, the strains were pregrown overnight in LB medium, diluted into tubes containing fresh LB medium (either singly or using equal numbers of CFU of each strain), and incubated at 37°C with gentle mixing.

Motility assay. Motility was assessed after 6 h of incubation on 0.4% LB agar, as described by Sinha et al. (65). Bacteria were pregrown in LB medium, and 50-μl portions of cells in LB media with and without 1 mM IPTG were spotted onto 2.5-cm filter disks.

RESULTS

***sxy* expression.** We first sought to find natural conditions that induce *E. coli sxy* expression. Sauvonnet et al. indirectly investigated this in their search for treatments that could induce expression of the T4P pilin gene *ppdD* (59). They tested a variety of conditions and treatments (e.g., temperature shift, anaerobic conditions, growth in minimal medium, treatment with known inducers of pathogenicity genes) but found that none of the conditions and treatments were effective. Since Sxy is the only known inducer of *ppdD* homologues in other bacteria (12), one simple interpretation is that the conditions used failed to induce *ppdD* because they failed to induce *sxy*. We supplemented this analysis by testing whether *E. coli sxy* could be induced by a starvation shock like that used to induce *H. influenzae sxy* (27, 56), using qPCR to directly measure *sxy* expression. As shown in Fig. 1, transfer of exponentially growing *E. coli* cells to minimal medium M9 did not change *sxy* expression. In *H. influenzae sxy* expression varies with the growth state in rich medium (56), so we also monitored *E. coli sxy* expression in cells growing exponentially and after overnight culture in broth and on agar. The expression levels were consistently low and did not differ significantly between conditions (Fig. 1).

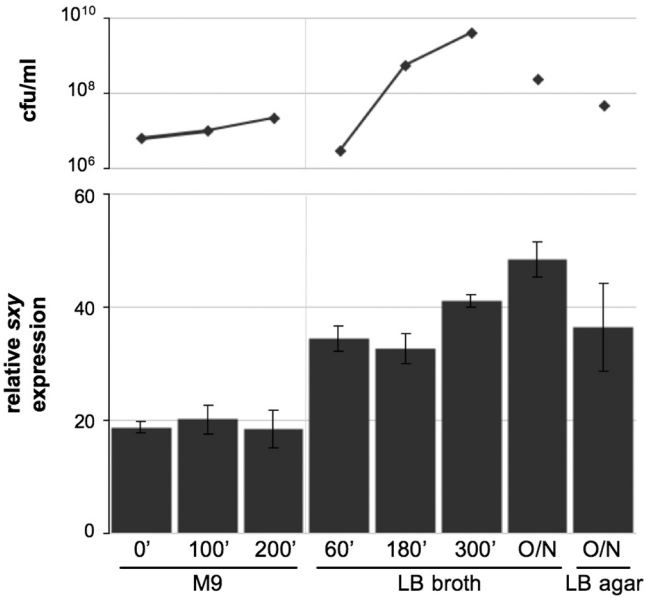

FIG. 1. *E. coli sxy* expression: CFU (upper panel) and qPCR analysis of *sxy* expression (lower panel) under different growth conditions. The levels of expression shown are for bacteria after transfer to minimal medium M9 (left side) and during growth in LB broth or on LB agar (right side). Normalized numbers of copies of *sxy* mRNA are shown (see Materials and Methods). Each data point represents the mean of two independent biological replicates, and the error bars indicate the standard deviations from the means. O/N, overnight.

Since we could not identify conditions that substantially increased the expression of the chromosomal *sxy* gene, we instead used cells carrying *sxy* on an inducible plasmid (p*Ecsxy*) to identify the genes that Sxy regulates. Addition of IPTG to these cells resulted in strong induction of *sxy* (200-fold increase in transcription after 30 min), and sodium dodecyl sulfate-polyacrylamide gel electrophoresis of protein extracts revealed an IPTG-induced band at 26 kDa, the expected size of His-tagged Sxy protein (see Fig. S1A in the supplemental material). Longer incubation times with IPTG did not increase the amount of protein produced, but the growth of the culture slowed dramatically; smaller but still substantial decreases in the growth rate were also observed when *sxy* expression was induced with lower concentrations of IPTG (see Fig. S1B in the supplemental material). Other groups have reported such toxicity when *E. coli sxy* was expressed (35, 45).

Changes in gene expression as a result of *sxy* overexpression. Sxy's only known activity in *H. influenzae* is the induction of the 13 CRP-S transcription units. To identify its activity in *E. coli*, we compared gene expression in cells carrying p*Ecsxy* and in control cells carrying the empty plasmid vector (pnoins). For this comparison we used four microarrays with RNA samples from independent cultures. Induction of *sxy* caused 81 genes to be upregulated at least fourfold and 41 genes to be repressed at least fourfold. These 122 differentially expressed genes belong to the 86 putative transcriptional units (TUs) shown in Fig. 2A and listed in Table 1. Their predicted functions and classifications according to the Clusters of Orthologous Groups of proteins (COG) database (72) are shown in Table S2 in the supplemental material.

As a first step in characterizing the regulatory signals responding to *sxy* expression, we searched the promoter regions of the 86 TUs for significantly overrepresented motifs, using the unbiased-search algorithms Gibbs Recursive Sampler and Meme. The EcoCyc database (34) was also used to examine published regulatory information for each promoter. These analyses identified three distinct classes of TUs: 34 TUs with CRP-S sites, 14 TUs of the sigma-H (σ^{32}) regulon, and 13 TUs of the FlhDC-FliA regulon. The remaining 21 TUs had no detected motif and no characterized regulator. Each class is discussed individually below. The protein-binding sites identified by motif searching are shown in Table 1, and the corresponding sequence logos are shown in Fig. 3.

(i) Genes with CRP-S sites. Unbiased searches of all upregulated promoters identified a 22-bp motif present in the promoters of 34 TUs (TUs 1 to 34 in Fig. 2; logos are shown in Fig. 3A). This motif resembles the previously characterized *H. influenzae* CRP-S motif but has some distinct features that are discussed below. Two of the CRP-S sites that we found have been characterized previously in *E. coli* (*ppdD* and *yrfD*), and both are known to require CRP and Sxy for activation (12). To test the CRP dependence of the other sites, we used additional microarrays to compare the effects of expressing *sxy* in the presence and in the absence of cAMP, CRP's essential cofactor (Fig. 2B). Expression of 30 of the 34 TUs decreased at least

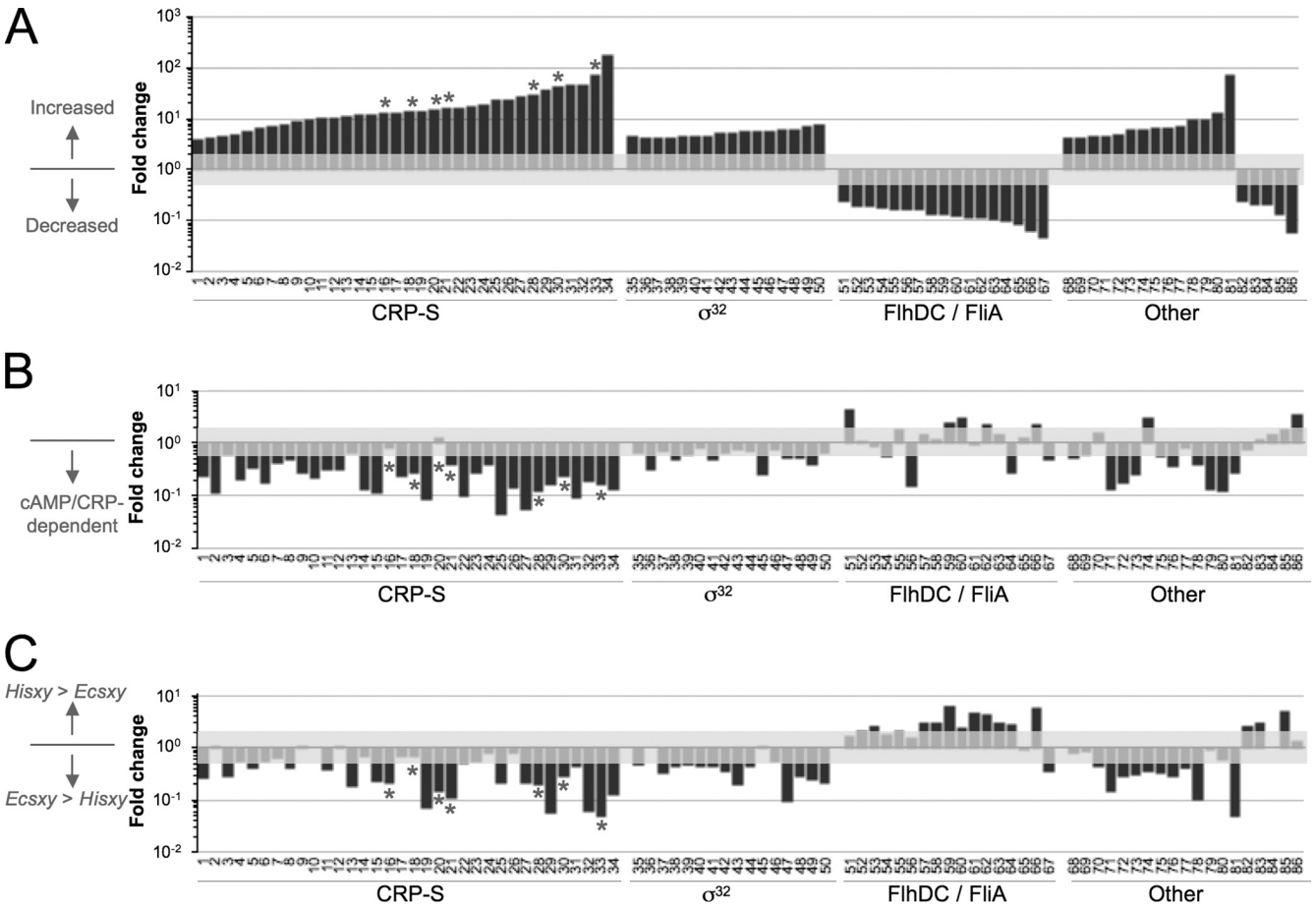

FIG. 2. Differentially expressed genes in different microarray comparisons. (A) *E. coli sxy*::kan/p*Ecsxy* over *E. coli sxy*::kan/pnoins. (B) *E. coli cya*::kan/p*Ecsxy* over *E. coli cya*::kan/p*Ecsxy* plus cAMP. (C) *E. coli sxy*::kan/p*Hisxy* over *E. coli sxy*::kan/p*Ecsxy*. Each bar represents one putative TU (see Table 1 and Table S2 in the supplemental material for a description of each TU). The expression level for each TU is the mean of the expression levels for all genes in that TU. The shaded regions indicate the ranges for a twofold change in expression. The four classes of genes identified by motif searching are indicated. Homologues of *H. influenzae* CRP-S TUs are indicated by asterisks.

twofold in the absence of cAMP (average decrease, 4.5-fold), confirming that the *E. coli* and *H. influenzae* CRP-S genes are similarly regulated. Two of the exceptions (*ycaI* and *hofMNOPQ*) were CRP dependent when they were retested using qPCR (Fig. 4A). The other two, *yfc1* and *yhgA*, encode predicted transposases and were not retested; their levels of induction by Sxy were 5- and 11-fold, respectively, and their levels of reduction in the absence of cAMP were about 1.5-fold.

The *E. coli* CRP-S regulon contains many more genes than its *H. influenzae* counterpart (64 genes versus 25 genes), and, not surprisingly, the predicted functions of the genes are much more diverse (see Table S2 in the supplemental material). Below we first consider the overlap between the *E. coli* and *H. influenzae* CRP-S regulons and then consider genes unique to the *E. coli* regulon.

Twenty of the 25 genes in the *H. influenzae* CRP-S regulon have homologues in *E. coli* (12), and microarray and qPCR analyses showed that all but one of these genes (*pppA*) belong to the *E. coli* CRP-S regulon. Fifteen of the 20 genes belong to 7 of the 34 *E. coli* CRP-S TUs (Fig. 2). qPCR showed that four of the remaining five genes (*hofQ*, *dprA*, *gntX*, and *yicR*) are also CRP-S regulated. Although the *hofQ* gene initially appeared to be highly expressed even in control cells, qPCR showed that it is Sxy induced and CRP dependent (Fig. 4A and 4B), so the array spots for this gene are likely to have been defective. Three other genes homologous to *H. influenzae* CRP-S genes (*gntX*, *yicR*, and *dprA*/*smf*) were not induced fourfold in our arrays but fit all the criteria: they have CRP-S sites (Table 1), are Sxy induced either ≥3-fold in the microarray analysis (Table 1) or 10-fold as determined by qPCR (Fig. 4B), and are at least ≥2-fold CRP dependent as determined by microarray analysis (not shown) or qPCR (Fig. 4A). The only homologue of an *H. influenzae* CRP-S gene that is not part of the *E. coli* CRP-S regulon is the prepilin peptidase gene *pppA*. The expression of this gene was not detectably induced by Sxy in arrays or qPCR experiments (Fig. 4B) and was independent of CRP (data not shown), consistent with the lack of a CRP-S site in its promoter.

The *E. coli* CRP-S regulon includes one other gene predicted to play a central role in both DNA uptake and T4P function: the *pilT* homologue *yggR* (TU 21). This gene, predicted to encode the pilus retraction ATPase, is not present in *H. influenzae* but is essential for DNA uptake in other competent bacteria (51, 60). The other 26 TUs (32 genes) in the *E.*

TABLE 1. TUs differentially regulated by *E. coli sxy* expression and their promoter motifs[a]

TU	Gene		Induction (fold)	Site or regulation
	No.	Designation		
CRP-S genes				
1	b3645	*dinD*	4.15	TTTTGCGTGCCTGCTCCAGATT
2	b3934	*cytR*	4.28	CGATGCGAGGCGGATCGAAAAA
3	b3411	*yhgA*	4.58	TTTTGCGTGCGGCTTTCCATAA
				CCCTGCGAGATGCTTTCAAAGG
4	b0799	*dinG*	4.79	TATTGCGAGCCGCTTTCCAGAA
5	b2569	*lepA*	5.74	AAATGCGAGGACGTTTCCAGAT
6	b2898	*ygfZ*	6.55	TTCTGCGAGGCGACTTCCAAAA
7	b2897	*ygfY*	8.95	TTCTGCGAGGCGACTTCCAAAA
7	b2896	*ygfX*	5.24	—[b]
8	b4059	*ssb*	7.84	TGTTGTGACCTCGGTTCCGGGA
9	b3129	*sohA*	5.14	TTTTGCGAGGGACGTTCCAAAA
9	b3130	*yhaV*	13.04	—[b]
10	b3648	*gmk*	9.89	GTCTGCGAGGACGCTTCCTGAA
				TTTTGCGTAGCTGCTTCCGCAT
11	b2733	*mutS*	10.26	TTTTGCGTACTTGCTTCATAAG
12	b1420	*mokB*	10.85	AATTGCGTGCTGTTTTCCAGAA
13	b2305	*yfcI*	11.12	ATTTGCGAGTCAGCTTCCGCTG
14	b2936	*yggG*	11.78	ACATGCGAGCATGATCCAGAGA
15	b0228	*yafM*	11.84	ATTTGCGAGCCGCTTTCCCGAT
16	b0913	*ycaI*	12.92	CTCTGCGAGGCAGCTTCCAGTT
				TTCCCCGAGGCATATTCAGAAG
17	b1289	*ycjD*	13.27	TTTCTGAATCCTCCTTCCAGAT
18	b0106	*hofC*	18.59	—[b]
18	b0107	*hofB*	11.59	—[b]
18	b0108	*ppdD*	12.55	ATTTGCGAGGCGTTACGAAGAA
19	b0018	*mokC*	14.61	AAATGCGAGGCGTCTTCAGGAT
20	b3391	*hofQ*	1.28	—[b]
20	b3392	*hofP*	5.09	—[b]
20	b3393	*hofO*	14.13	—[b]
20	b3394	*hofN*	11.25	—[b]
20	b3395	*hofM*	30.89	TAAATCGAGCCTGCTCCCAGCA
21	b2950	*yggR*	16.39	ACCTGCGAGCGCCATCGCAGAA
22	b2700	*ygaD*[c]	16.97	TTTTACGAGGAGGATTCAGAAA
				CTCTCCGAGCCGCTTTCCATAT
23	b1438	*ydcQ*	27.33	—[b]
23	b4532[d]	*yncN*	6.83	TTCTGGAATCTTCCTTCCTGAT
24	b3082	*ygjM*	23.78	—[b]
24	b3083	*ygjN*	14.71	TTCTGGAATCGCGCTTCCATAA
25	b4224	*chpS*	20.00	TTTTGCGAGACGACTCGCATTC
25	b4225	*chpB*	25.86	—[b]
26	b3890	*yiiF*	23.91	TTTTGCGAGGCGTTTTCCAGAT
27	b2798	*xni*	27.97	TACCGCGAAACCTCTCGCGGCG
28	b3765	*yifB*	30.59	TTTTGCGAGCATCATTCCACCG
29	b4338	*yjiP*	39.37	AAACGCGAGCTTGTTTCCGGAA
				TTCCCCGAGGCGCTTTCCAGGC
29	b4339	*yjiP*	34.81	—[b]
30	b2823	*ppdC*	9.93	—[b]
30	b2824	*ygdB*	31.05	—[b]
30	b2825	*ppdB*	19.50	—[b]
30	b2826	*ppdA*	109.9	TTCTTCGAGACGCCTTCCCGAA
31	b2272	*yfbM*	45.84	TTTCTCGAGGCGGGTCCAGAAA
32	b3647	*ligB*	46.38	GTCTGCGAGGACGCTTCCTGAA
				TTTTGCGTAGCTGCTTCCGCAT
33	b0442	*ybaV*	70.72	TTTTGCGAGGCGGCTTCAGGAT
34	b2244	*yfaD*	184.30	CTTTGCGAGACGTTTTCAGAGG
	b3413[e]	*gntX*	3.36	AAATGCGAGCTAAGTTCCTCGT
	b3638[e]	*yicR*	3.13	CTTTGCGAGGCGCTTTCCAGGA
	b4473[e]	*dprA*	1.15	CTTTGCGAAGCCGCTCGTCCGG
σ^{32} regulon				
35	b3931	*hslU*	4.62	σ^{32} regulated[f]
36	b3498	*prlC*	4.21	GACTCGATGATCAGGCTCTGAACGGTATGTT
37	b0439	*lon*	4.35	GCGTTGAATGTGGGGGAAACATCCCCATATA
38	b0473	*htpG*	4.38	AGCTGGAAATGGTCGACGGCACCACGTTGTC
39	b0492	*ybbN*	4.46	GAGTTGAAGCGCGCGCTTAAGCCCCCATGTC

Continued on following page

TABLE 1—*Continued*

TU	Gene		Induction (fold)	Site or regulation
	No.	Designation		
40	b1060	*bssS*	4.48	CCCTTTAAAAATTCGGTGAATACCCTTACTT
41	b1322	*ycjF*	4.64	σ^{32} regulated[f]
42	b2592	*clpB*	5.43	ACCTTGAATAATTGAGGGATGACCTCATTTA
43	b2614	*grpE*	5.47	CCCTTGAAACCCTGAAACTGATCCCCATAAT
44	b0014	*dnaK*	5.54	CCCTTGATGACGTGGTTTACGACCCCATTTA
45	b2699	*recA*	5.81	CACTTGATACTGTATGAGCATACAGTATAAT
46	b0966	*hspQ*	5.97	AACTTGAAATGTGGAATTTTACCCTTATATC
47	b3686	*ibpB*	52.38	GGCTTGAAAAGTTCATTTCCAGACCCATTTT
47	b3687	*ibpA*	7.57	—[b]
48	b3635	*mutM*	6.13	GGGTTTTTTGTTATCTGCTTGCCCCCATATT
49	b4140	*fxsA*	6.96	CACTTTAAGTGTCGGTTTTTACCCCTTAATT
50	b0631	*ybeD*	7.88	CACTTGAAAGTGTAATTTCCGTCCCCATATA
FlhDC/FliA regulon				
51	b3417	*malP*	0.25	GAGATCACATTTCCTTGCTCATCCCCGCAACTCC
52	b4355	*tsr*	0.20	TTCATAAAGTTTTTCCTTTCCAGGCCGAAAATCT
53	b3072	*aer*	0.19	FlhDC regulated[g]
54	b1881	*cheZ*	0.22	GCAATAAAGTTTCCCCCCTCCTTGCCGATAACGA
54	b1885	*tap*	0.16	—[b]
55	b1924	*fliD*	0.15	AACGTAAACTTTGCGCAATTCAGACCGATAACC
55	b1925	*fliS*	0.15	—[b]
55	b1926	*fliT*	0.23	—[b]
56	b1421	*trg*	0.17	GCGACCCATTTTGCGTTTATTCCGCCGATAACGC
57	b1887	*cheW*	0.15	GACGTAAACTTTCCCAGAATCCTGCCGATATTAT
57	b1888	*cheA*	0.16	—[b]
57	b1889	*motB*	0.19	—[b]
58	b1070	*flgN*	0.16	FlhDC regulated[h]
58	b1071	*flgM*	0.12	—[b]
58	b1072	*flgA*	0.13	—[b]
59	b1944	*fliL*	0.19	FlhDC regulated[h]
59	b1945	*fliM*	0.05	—[b]
59	b1948	*fliP*	0.15	—[b]
59	b1950	*fliR*	0.14	—[b]
60	b1566	*flxA*	0.13	CGATTAAAGATTTTTTTGTGCATGCCGATAGTGC
61	b1923	*fliC*	0.12	AACGTAAACTTTGCGCAATTCAGACCGATAACC
62	b1938	*fliF*	0.14	FliA regulated[i]
62	b1939	*fliG*	0.10	—[b]
62	b1940	*fliH*	0.11	—[b]
63	b1921	*fliZ*	0.11	FlhDC regulated[h]
63	b1922	*fliA*	0.11	—[b]
64	b3525	*yhjH*	0.10	GCGGTAAAGTTCTGCCCTTACGCGCCGATAATCT
65	b4035	*malK*	0.13	GGCAACCTCTTTCCATCCTCCTTGCCCCTACGCC
65	b4036	*lamB*	0.04	—[b]
66	b1073	*flgB*	0.04	GCGACCCATTTTGCGTTTATTCCGCCGATAACGC
66	b1074	*flgC*	0.03	—[b]
66	b1075	*flgD*	0.14	—[b]
66	b1076	*flgE*	0.04	—[b]
66	b1077	*flgF*	0.08	—[b]
66	b1078	*flgG*	0.06	—[b]
67	b4034	*malE*	0.05	GGCAACCTCTTTCCATCCTCCTTGCCCCTACGCC
Other genes				
68	b3863	*polA*	4.18	
69	b0294	*matA*	4.19	
70	b3334	*gspM*	4.62	
71	b0245	*ykfI*	4.69	
72	b4326	*yjiD*	4.96	
73	b1562	*hokD*	5.98	
74	b0235	*ykfJ*	6.19	
75	b0325	*yahK*	6.71	
76	b3554	*yiaF*	6.72	
77	b4327	*yjiE*	7.17	
78	b0032	*carA*	9.41	
79	b2821	*ptrA*	14.91	
79	b2819	*recD*	6.36	

Continued on following page

TABLE 1—*Continued*

TU	Gene		Induction (fold)	Site or regulation
	No.	Designation		
79	b2820	*recB*	7.65	
80	b1439	*ydcR*	12.84	
81	b0443	*ybaW*	73.00	
82	b1015	*putP*	0.24	
83	b2752	*cysD*	0.22	
84	b2778	*ygcG*	0.21	
85	b4109	*yjdA*	0.14	
86	b3988	*rpoC*	0.06	

[a] The TU numbers correspond to those in Fig. 2. The induction data are data for *E. coli sxy* expression. Further information about each gene is given in Table S1 in the supplemental material.
[b] The gene is cotranscribed with another gene(s) in the same TU and therefore has no CRP-S site.
[c] *ygaD* also had an identified σ^{32} site (TTTTACGAGGAGGATTCAGAAA).
[d] Replaces b1437.
[e] Not shown in Fig. 2 but CRP-S regulated (see text).
[f] See reference 45.
[g] See reference 53.
[h] See reference 68.
[i] See reference 49.

coli CRP-S regulon include genes involved in genome maintenance (7 genes) and transcription (5 genes) and genes encoding transposases (6 genes) and toxin-antitoxin pairs (8 genes) (Table 1). Induction of these genes by *sxy* expression required cAMP-CRP (Fig. 2B), consistent with the presence of CRP-S sites in their promoters (Table 1). The presence of these genes in a single tightly controlled regulon suggests that they may contribute to fitness under conditions that encourage DNA uptake.

(ii) Genes of the sigma-32 (σ^H) regulon. Motif searching identified σ^{32}-binding sites in the promoters of 14 Sxy-induced genes (Table 1 and Fig. 3B). The σ^{32} regulon responds to the accumulation of improperly folded proteins in the cytoplasm (28), so its induction is likely to be a response to the toxic effects of Sxy accumulation in the cytoplasm. The overexpression of toxic proteins is known to induce the σ^{32} response (25, 47, 50), and a screen for proteins that induced this response when they were overexpressed repeatedly identified Sxy (45). Examination of all 124 previously reported members of the σ^{32} regulon (45) showed that most genes were induced at least twofold (mean induction, 2.6-fold) (Fig. 5A). In contrast, no change was seen in the σ^E and CpxR stress regulons, which respond to misfolded proteins in the periplasm and cell envelope (Fig. 5A). If induction of the σ^{32} regulon resulted from Sxy's action at CRP-S promoters, induction of σ^{32} genes would be expected to depend on CRP. However, the absence of cAMP-CRP did not change expression of the Sxy-induced σ^{32} genes (TUs 35 to 50 in Fig. 2B) or of all members of the σ^{32} regulon (data not shown), suggesting that induction of this regulon is a response to Sxy's presence rather than to its biological activity.

(iii) Genes of the flagellar-chemotaxis regulon. Seventeen of the 22 TUs that showed decreased expression upon *sxy* expression are known members of the flagellar regulon (TUs 51 to 67 in Fig. 2), which is tightly controlled by the transcriptional activator FlhDC and the σ factor FliA (38). Consistent with this, motif searching identified FliA sites in 13 of these 17 promoters (Fig. 3C and Table 1); similar sites have been described in the remaining four promoters (49, 53, 68). Expression of *flhD* and *flhC* was unchanged in our array experiments, but most of the other genes in the FlhDC regulon were repressed by *sxy* expression (Fig. 5B). Absence of cAMP-CRP did not affect expression of these 22 TUs (Fig. 2B), suggesting that downregulation of the flagellar regulon by Sxy is indirect. Consistent with this interpretation, the σ^{32}-induced Lon protease is known to degrade the FlhDC regulator (14, 62, 63).

(iv) Other genes. Nineteen other TUs showed at least fourfold changes in expression with *sxy* induction; the expression of 14 genes increased, and the expression of 5 genes decreased (TUs 68 to 86 in Fig. 2). None of these TUs had identifiable promoter motifs, and they showed various degrees of dependence on CRP (Fig. 2B). They have no discernible relationship to each other or to other groups described above, and the significance of their regulation by Sxy is unclear.

Phenotypic correlates of changes in gene expression. We began this work with the hypothesis that Sxy is needed for transcription of *E. coli*'s competence gene homologues. In other bacteria, the products of these genes have established roles either in DNA uptake and processing (HofMNOP, YifB, DprA, YbaV, YcaI, and GntX) or in biogenesis of the T4P machinery required for DNA uptake (PpdD-HofB-HofC, HofQ, YggR, and PpdA-PpdB-YdgB-PpdC) (Table 2). Although most of their functions have not been characterized in *E. coli*, some of these genes have demonstrated roles in T4P production or are required by *E. coli* for use of DNA as a carbon source (Table 2). Because our microarray analysis showed that these genes are all in the Sxy-induced CRP-S regulon, we examined the role of Sxy in competence-related functions.

(i) *sxy* expression enables pilin production. Athough *E. coli* K-12 has homologues of all T4P biogenesis genes, it does not express pilin and does not produce T4P fibers even when pilin is artificially expressed from a plasmid (40). As our experiments showed that all T4P biogenesis genes except *pppA* are induced by Sxy and CRP, we tested whether cells expressing *sxy* can produce T4P.

Western blotting first confirmed that cells expressing *sxy* produce pilin. A strain carrying a cloned IPTG-inducible copy

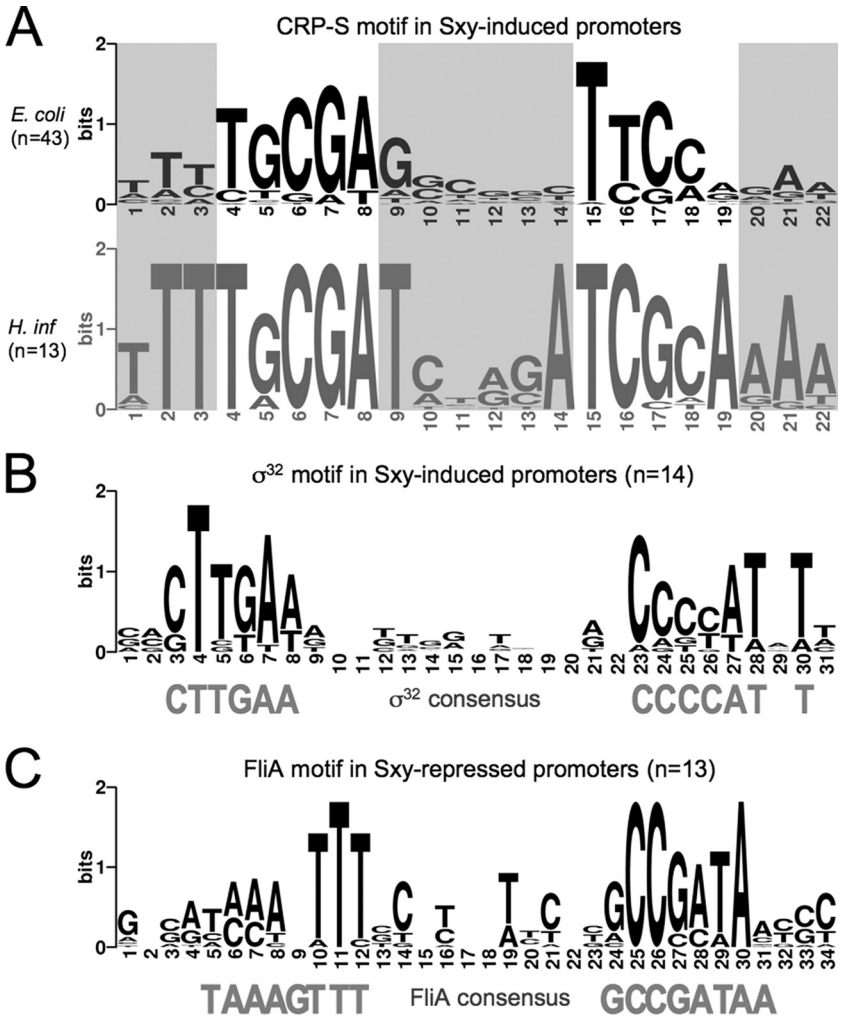

FIG. 3. Significantly overrepresented motifs identified in this work. (A) CRP-S motif found in the promoters of 34 TUs induced by *E. coli sxy* expression (43 sites). The *H. influenzae* CRP-S motif is shown for comparison. (B) σ^{32}-binding site found in the promoters of 14 TUs induced by *E. coli sxy* expression (14 sites). (C) FliA binding site found in the promoters of 13 TUs induced by *E. coli sxy* expression (13 sites). Sequence logos were generated using WebLogo (http://weblogo.berkeley.edu/logo.cgi) (16). *H. inf*, *H. influenzae*.

of the *ppdD* gene (p*EcppdD*) was used as a positive control (17.5-kDa His-tagged protein) (Fig. 6A, lane 1). A single band that was the expected size of processed pilin (15 kDa) was detected in cells expressing *E. coli sxy* (Fig. 6A, lane 3). As expected, this was dependent on the presence of active CRP (Fig. 6, lane 5). No unprocessed pilin band was seen, confirming that prepilin peptidase was active and suggesting that basal expression of the *pppA* and/or *gspD* peptidase gene is sufficient for prepilin processing (neither gene was induced by Sxy). As has been previously reported (40), no pilin was detected in the parent strain (Fig. 6A, lane 6).

To determine whether the pilin could be assembled into fibers, the surfaces of cells carrying p*Ecsxy* or the empty vector control were incubated with PpdD antiserum and examined by transmission electron microscopy. No T4P were seen, suggesting that expression of the predicted T4P biogenesis genes is not sufficient for T4P production. It may be that one or more of these genes encode nonfunctional proteins, that other unidentified components are missing or not expressed, or that artificial *sxy* induction leads to abnormally high levels of T4P proteins which the cell is unable to process efficiently. Cells expressing *sxy* had almost twice as many gold particles on their surfaces as control cells (p*Ecsxy*, 112 ± 24 particles per 10 μm^2; pnoins, 60 ± 3 particles per 10 μm^2; *ppdD*::kan, 43 ± 6 particles per 10 μm^2), suggesting that some pilin was excreted from induced cells even though fibers were not assembled.

(ii) Sxy is required for plasmid transformation. In *H. influenzae*, Sxy is essential for natural competence; a strain lacking *sxy* is not transformable, and mutations that upregulate *sxy* make competence constitutive (13, 55, 85). In *E. coli*, we could not test whether inducing plasmid-borne *sxy* allows transformation because long-term overexpression of *sxy* is toxic. We also could not test whether knocking out the chromosomal *sxy* gene abolishes transformation with linear chromosomal DNA, because wild-type cells cannot be detectably transformed (data not shown). We instead investigated whether the chromosomal copy of *sxy* contributes to transformation with plasmid DNA,

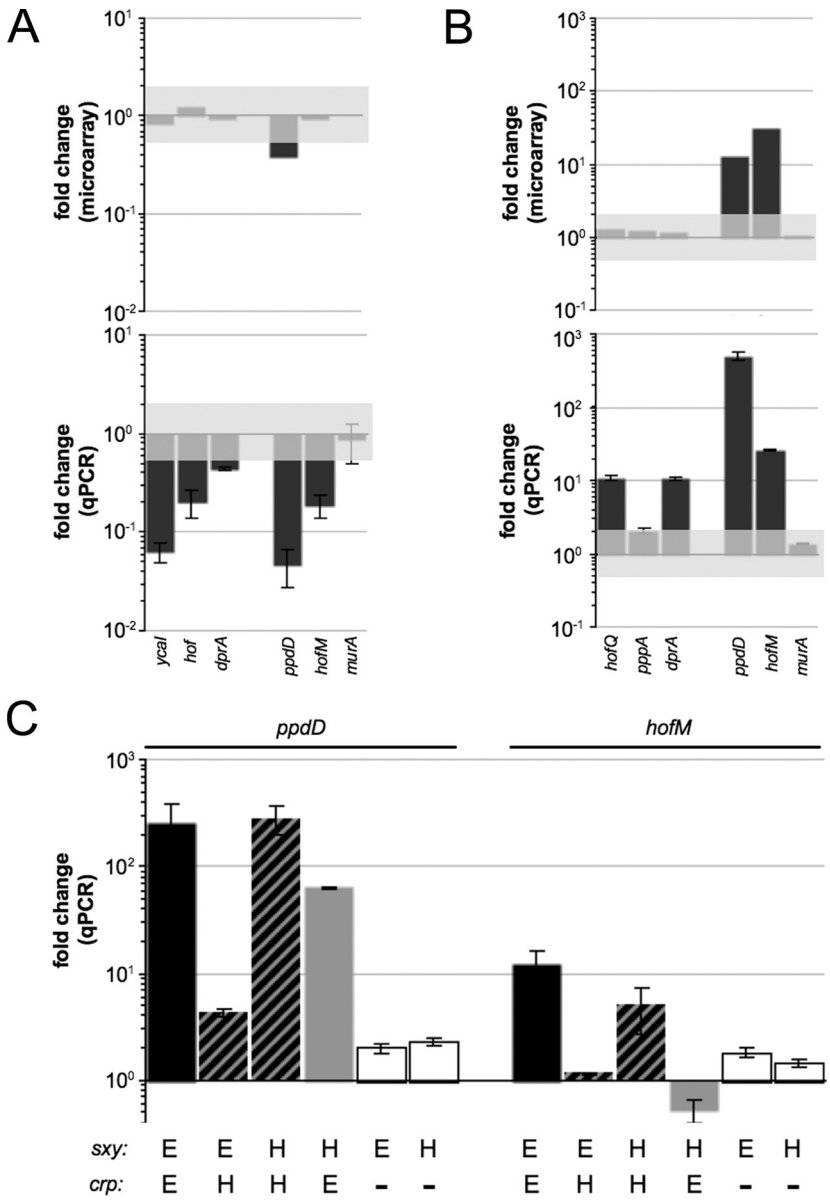

FIG. 4. Expression levels of specific genes upon *sxy* expression and their dependence on cAMP-CRP for expression. (A) Changes determined by microarray analysis (upper panel) or qPCR (lower panel) in a comparison of *E. coli cya*::kan/p*Ecsxy* and *E. coli sxy*::kan/p*Ecsxy* plus cAMP. (B) Changes determined by microarray analysis (upper panel) or qPCR (lower panel) in a comparison of *E. coli sxy*::kan/p*Ecsxy* and *E. coli sxy*::kan/pnoins. (C) Changes in expression of b0108 (*ppdD*) and b3395 (*hofM*) determined by qPCR when *E. coli* (E) or *H. influenzae* (H) *sxy* was expressed with the cognate *crp*. For qPCR data, each bar indicates the average of at least two independent biological replicates for each gene or group of genes. The error bars indicate the standard deviations from the means. The expression levels shown were adjusted for each RNA sample using 23S rRNA levels. Gene designations: b0913, *ycaI*; b3395 to b3391, *hofMNOPQ*; b2972, *pppA*; b4473, *dprA*.

using the "natural plasmid transformation" procedure of Baur et al. (8). Transformation of wild-type cells with Amp^r plasmid DNA in 1 mM $CaCl_2$ gave a transformation frequency of $9.1 \times 10^{-8} \pm 1.5 \times 10^{-8}$ (41, 37, and 29 transformant colonies from 1 ml in three independent experiments). These values are similar to the transformation frequencies (10^{-8} to 10^{-7}) reported by by Baur et al. (8) for this concentration of $CaCl_2$. In contrast, transforming equal numbers of cells lacking Sxy (*sxy*::kan) gave no transformants in any experiment (limit of detection, 2.3×10^{-9} CFU/ml). This very significant difference ($P = 0.001$) was not due to poor growth of the mutant, as the two strains had comparable doubling times and viable counts (parent, $3.9 \times 10^8 \pm 1.0 \times 10^7$ CFU/ml; *sxy*::kan strain, $4.4 \times 10^8 \pm 3.2 \times 10^7$ CFU/ml). Similar results were obtained using a plasmid carrying a different antibiotic resistance marker (data not shown). The differences in transformation were slightly smaller (10- to 15-fold) when the $CaCl_2$ concentration was increased to 10 or 50 mM. In contrast, deletion of *sxy* did not affect artificial plasmid transformation using electroporation or the classical $RbCl_2$-plus-heat shock protocol (data not shown). We were unable to detect any transformation with chromosomal DNA using the protocol of Baur et al.

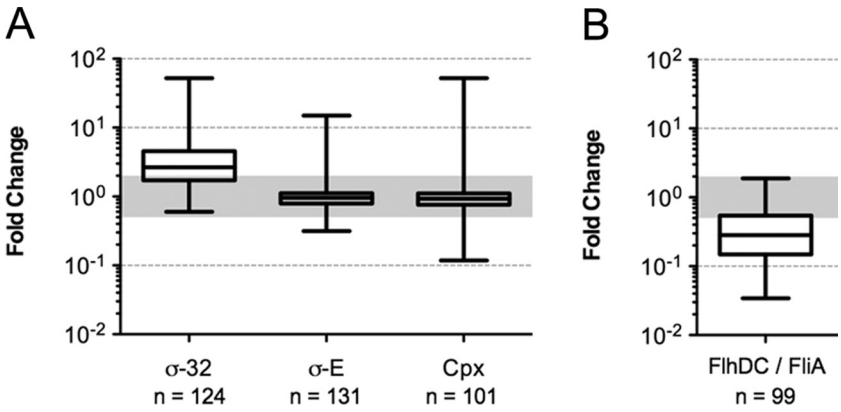

FIG. 5. Expression levels in a microarray of genes from different regulons as a result of *E. coli sxy* expression. (A) Global stress responses. Sigma-H (σ^{32}) regulon genes were obtained from references 84 and 45. Sigma-E regulon genes were obtained from references 18 and 57. Cpx regulon genes were obtained from references 19 and 20. (B) FlhDC regulon. FlhDC regulon genes were obtained from references 54 and 83. n is the number of genes in each regulon. The shaded regions indicate the ranges for a twofold change in expression. The boxes indicate the interquartile ranges; the horizontal line in each box indicates the median, and the whiskers extend to the maximum and minimum values in each data set. The data were obtained in our microarray analysis of four biological replicates.

(iii) Sxy is required for competitive fitness. The competitive fitness of *E. coli* strains decreases when homologues of *H. influenzae com* genes are inactivated, because such strains are unable to use DNA as a sole source of carbon and energy (32). Since Sxy regulates these genes in *H. influenzae*, we predicted that an *E. coli* strain lacking *sxy* would also compete poorly with the parent strain. Figure 6B shows that the survival of the *sxy* mutant was reduced when it was cocultured with the parent strain, but the mutant survived as well as the parent strain when each strain was cultivated separately. This is consistent with Sxy's role as a regulator of the *com* homologues identified by Palchevskiy and Finkel in *E. coli* (48).

(iv) Expression of *sxy* does not reduce motility. Our microarray analysis showed that *sxy* expression downregulated the flagellar and chemotaxis operons. However, when motility was assessed on soft agar, cells expressing *sxy* from p*Ecsxy* were as motile as cells carrying the no-insert control plasmid (Fig. 6C). As expected, a negative control strain lacking the flagellar subunit FliC was nonmotile. qPCR showed low but significant expression of *flhD*, *fliA*, and *fliC* upon *sxy* induction (Fig. 6C). This confirms that expression of the flagellar genes is reduced but not abolished by *sxy* expression and that residual expression is sufficient for flagellar function.

Regulation at CRP-S sites in *E. coli* compared with *H. influenzae*. Although the *H. influenzae* and *E. coli* Sxy proteins have only 24% sequence identity, the results presented above show that Sxy plays the same role in both species, acting with CRP to activate transcription at CRP-S sites. However, we do not know how Sxy does this in either system. To determine the degree of conservation of the Sxy–CRP-S systems in *H. influenzae* and *E. coli* and to shed light on Sxy's role at CRP-S sites, we evaluated the activity of both Sxy proteins in *E. coli* and compared the complete sets of experimentally validated CRP-S sites from both species.

(i) Ability of *H. influenzae sxy* to complement *E. coli sxy*. We used microarrays to compare *E. coli* cells expressing *H. influenzae sxy* with cells expressing *E. coli sxy*. If the two Sxy proteins have comparable activities, gene expression should be largely unchanged. Instead, *H. influenzae* Sxy had a weaker effect on the transcription of most TUs, giving a net negative signal for *E. coli* Sxy-induced TUs in the array (Fig. 2C). Only 21 of the 64 TUs induced by *E. coli sxy* expression were comparably induced by *H. influenzae* Sxy. A similar but inverse effect was seen for the TUs that had been downregulated by *E. coli* Sxy.

Because *H. influenzae* Sxy only weakly induced most of the T4P biogenesis homologues, we examined its effect on pilin production. The two Sxy proteins induced the pilin-encoding TU *ppdD-hofBC* equally well, and the two strains produced equal amounts of PpdD pilin protein (Fig. 6A, lane 4). However, unlike cells expressing *E. coli sxy*, cells expressing *H. influenzae sxy* had no more external pilin visible in electron micrographs than control cells (p*Hisxy*, 55 ± 7 particles per 10 μm^2). This absence of surface pilin is consistent with the lack of induction of most genes required for T4P biogenesis by *H. influenzae* Sxy.

(ii) CRP-S motifs in *E. coli* and *H. influenzae*. The inability of *H. influenzae* Sxy to induce most of the genes regulated by *E. coli* Sxy may be due to differences in CRP-S site structure. The *E. coli* CRP-S motif identified by the arrays described above strongly resembles the previously characterized *H. influenzae* CRP-S motif in overrepresentation of the CRP binding bases G5, G7, A8, T15, and C18, the AT-rich sequence at positions 1 to 3 and 20 to 22, and the noncanonical bases at positions 6 and 17 (11). However, the *E. coli* motif favors T rather than C at position 16 while retaining the strong G at position 7. This novel feature decreases the symmetry of these core sites, which is predicted to hinder binding of the CRP homodimer to DNA.

To further investigate a possible role for differences in CRP-S sites, we generated separate logos for the CRP-S sites that responded poorly to *H. influenzae* Sxy and for the CRP-S sites that responded similarly to the two Sxy proteins (see Fig. S2 in the supplemental material). The only apparent difference between these logos was that sites efficiently induced by *H. influenzae* Sxy were more AT rich at their 3′ ends and at reciprocal positions 4 and 19, features that are both predicted to facilitate CRP binding (17).

(iii) Sxy requires its cognate CRP for full activity. Although Sxy and CRP are both required for activation of CRP-S pro-

TABLE 2. Competence gene homologues in *E. coli*[a]

E. coli gene	*P. aeruginosa* homologue[b]	*H. influenzae* homologue[c]	Function in *P. aeruginosa* and/or *H. influenzae*[d]	Function in *E. coli* (confirmed or predicted)[d]	Phenotypic characterization for *E. coli*					Notes
					Sxy induced[e]	Functional in T4P production[f]	Required for use of DNA as carbon source[g]	Required for spontaneous plasmid transformation on nutrient plates with a high agar level[h]	Other phenotypes[d]	
hofM (b3395)	*pilM*	*comA*	Genes in the *pilMNOPQ* or *comABCDE* operon are required for T4P biogenesis in *P. aeruginosa* (41) and for competence in *H. influenzae* (21, 75, 76)	Predicted pilus assembly protein	Y	NT	Y	NT		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one or more genes in the operon
hofN (b3394)	*pilN*	*comB*	Genes in the *pilMNOPQ* or *comABCDE* operon are required for T4P biogenesis in *P. aeruginosa* (41) and for competence in *H. influenzae* (21, 75, 76)	Predicted pilus assembly protein	Y	NT	Y	NT		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one or more genes in the operon
hofO (b3393)	*pilO*	*comC*	Genes in the *pilMNOPQ* or *comABCDE* operon are required for T4P biogenesis in *P. aeruginosa* (41) and for competence in *H. influenzae* (21, 75, 76)	Predicted conserved membrane protein	Y	NT	Y	NT		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one or more genes in the operon
hofP (b3392)	*pilP*	*comD*	Genes in the *pilMNOPQ* or *comABCDE* operon are required for T4P biogenesis in *P. aeruginosa* (41) and for competence in *H. influenzae* (21, 75, 76)	Predicted protein	Y	NT	Y	NT		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one or more genes in the operon
hofQ (b3391)	*pilQ*	*comE*	Genes in the *pilMNOPQ* or *comABCDE* operon are required for T4P biogenesis in *P. aeruginosa* (41) and for competence in *H. influenzae* (21, 75, 76); PilQ is the outer membrane secretin in *P. aeruginosa* (40, 41) and other bacteria	Predicted outer membrane porin	Y	NT	Y	N		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one or more genes in the operon
gntX (b3413)	NA	*comF/com101*	Cotranscribed with *comABCDE* in *H. influenzae*; encodes a transformation protein (36, 37)	Gluconate periplasmic binding protein	Y	NT	Y	NT	Based on complementation experiments, GntX is thought to be involved in high-affinity gluconate transport (52)	Phenotypes in *E. coli* could be due to polar effects on *gntY*
ppdD (b0108)	*pilA*	*pilA*	Genes in the *pilABC* operon are required for pilin processing and pilus assembly in *P. aeruginosa* (46) and for T4P biogenesis and competence in *H. influenzae* (6, 21, 31, 77); the operon includes the main T4P subunit pilin PilA gene (6, 31, 77, 80)	Major type IV pilin, prelipin peptidase dependent (59, 60)	Y	Y	NT	N		Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one ore more genes in the operon
hofB (b0107)	*pilB*	*pilB*	Genes in the *pilABC* operon are required for pilin processing and pilus assembly in *P. aeruginosa* (46) and for T4P biogenesis and competence in *H. influenzae* (6, 21, 31, 77)	Predicted prelipin peptidase-dependent protein	Y	NT	NT	NT	Protein also known as HopB (host function of plasmid maintenance protein B); a *hopB* mutant shows defects in plasmid replication and maintenance (44)	Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one ore more genes in the operon
hofC (b0106)	*pilC*	*pilC*	Genes in the *pilABC* operon are required for pilin processing and pilus assembly in *P. aeruginosa* (46) and for T4P biogenesis and competence in *H. influenzae* (6, 21, 31, 77)	Predicted type IV pilus assembly protein	Y	NT	NT	NT	Protein also known as HopC (host function of plasmid maintenance protein C); a *hopC* mutant shows defects in plasmid replication and maintenance (44)	Phenotypes in *P. aeruginosa*, *H. influenzae*, and *E. coli* could be due to polar effects on one ore more genes in the operon

pppA (b2972)	*pilD*	*pilD*	Encodes the prepilin peptidase in *P. aeruginosa* (46) and *H. influenzae* (6); cotranscribed with *pilABC* in *P. aeruginosa* and *H. influenzae*	Functional prepilin peptidase (23)	N	Y	NT	NT	
ycaI (b0913)	NA	*rec-2*	Required for DNA translocation into cytoplasm in *H. influenzae* (9)	Predicted conserved inner membrane protein	Y	NT	NT	N	
ybaV (b0442)	NA	*comE1*	Homolgous to *Bacillus subtilis comEA*; encodes protein with predicted role in DNA uptake in *H. influenzae* (56)	Predicted DNA uptake protein with homology to ComEA	Y	NT	NT	NT	
ppdA (b2826)	*fimT*	*pulG/comN*	Genes in the *fimT-fimU-pilV* operon are required for T4P biogenesis in *P. aeruginosa* (1, 2), and genes in the *pulG/comN-pulJ/comO-comP-comQ* operon are required for competence in *H. influenzae* (75, 76, 79; S. Molnar and R. Redfield unpublished results)	Predicted prepilin peptidase-dependent protein A	Y	NT	NT	NT	A *ppdAB* mutant does not exhibit defects in chemotaxis or flagellar motility (49)
ppdB (b2825)	*fimU*	*pulJ/comO*	Genes in the *fimT-fimU-pilV* operon are required for T4P biogenesis in *P. aeruginosa* (1, 2), and genes in the *pulG/comN-pulJ/comO-comP-comQ* operon are required for competence in *H. influenzae* (75, 76, 79; S. Molnar and R. Redfield unpublished results)	Predicted prepilin peptidase-dependent protein B	Y	NT	NT	NT	A *ppdAB* mutant does not exhibit defects in chemotaxis or flagellar motility (49)
ydgB (b2824)	NA	*comP*	Genes in the *fimT-fimU-pilV* operon are required for T4P biogenesis in *P. aeruginosa* (1, 2), and genes in the *pulG/comN-pulJ/comO-comP-comQ* operon are required for competence in *H. influenzae* (75, 76, 79; S. Molnar and R. Redfield unpublished results)	Predicted protein type II secretion system, pilus	Y	NT	NT	NT	
ppdC (b2823)	*pilV*	*comQ*	Genes in the *fimT-fimU-pilV* operon are required for T4P biogenesis in *P. aeruginosa* (1, 2), and genes in the *pulG/comN-pulJ/comO-comP-comQ* operon are required for competence in *H. influenzae* (75, 76, 79; S. Molnar and R. Redfield unpublished results)	Predicted protein type II secretion system, pilus	Y	NT	NT	NT	
smf (b4473)	NA	*dprA*	Encodes an inner membrane protein essential for competence in *H. influenzae* that limits degradation of incoming DNA (32, 33)	DNA processing protein (66)	Y	NT	NT	N	Functional in *H. influenzae* but no apparent role in transformation, conjugation, DNA repair, or recombination in *E. coli* (66)
yifB (b3765)	NA	*comM*	Encodes an ATPase required for DNA translocation into cytoplasm (26)	Predicted ATP-dependent protease (29)	Y	NT	NT	NT	
yicR (b3638)	NA	*radC*	Involved in DNA synthesis and repair (56)	DNA repair protein (61)	Y	NT	NT	NT	
ssb (b4059)	NA	*ssb*	Encodes a single-stranded DNA-binding protein in *H. influenzae* (30)	Single-stranded DNA-binding protein; involved in replication, recombination, and DNA repair (43)	Y	NT	NT	NT	
gntY (b3414)	NA	*comG*	Downstream of *comABCDEF* in *H. influenzae* (75); no known role in competence; not competence induced in *H. influenzae*	Predicted iron-sulfur cluster scaffold protein	Y	NT	Y	NT	Protein also known as NfuA; an *nfuA* null mutant does not have a defect in gluconate metabolism (3)

Continued on following page

TABLE 2—*Continued*

E. coli gene	*P. aeruginosa* homologue[b]	*H. influenzae* homologue[c]	Function in *P. aeruginosa* and/or *H. influenzae*[d]	Function in *E. coli* (confirmed or predicted)[d]	Phenotypic characterization for *E. coli*				Notes
					Sxy induced[e]	Functional in T4P production[f]	Required for use of DNA as carbon source[g]	Required for spontaneous plasmid transformation on nutrient plates with a high agar level[h]	Other phenotypes[d]
yhiR (b3499)	NA	*comJ*	*orfJ* in (75); no known role in competence; not competence induced *H. influenzae*	Predicted conserved protein	N	NT	Y	N	
yggR (b2950)	*pilT*	NA	Encodes a pilus retraction protein in *P. aeruginosa* (81)	Predicted transporter	Y	NT	NT	NT	
yggT (b2952)	*mshC* (*V. cholenae*)	NA	Encodes a minor pilin in *Vibrio cholenae* (74)	Predicted inner membrane protein	N	NT	NT	NT	
ycgB (b1188)	*pilB*	NA	Required for pilin processing and pilus assembly in *P. aeruginosa* (46)	Predicted conserved protein; second *pilB* homologue	N	NT	NT	NT	

[a] Abbreviations: Y, yes; N, no; NT, not tested; NA, not applicable.
[b] Data from reference 59.
[c] Data from references 12 and 22.
[d] The numbers in parentheses are references.
[e] Data from this study.
[f] Data from references 23, 59, and 60.
[g] Data from references 22 and 48.
[h] Data from reference 69.

moters, we do not know whether these two proteins directly interact. If they do, failed heterospecific interactions might explain why *H. influenzae* Sxy functions poorly in *E. coli*. We therefore tested whether Sxy is a stronger inducer when its cognate CRP is present. We generated *E. coli* strains carrying both a *sxy* plasmid and a *crp* plasmid, with the *E. coli* and *H. influenzae* genes in all combinations. We then used qPCR to measure expression of *ppdD*, whose expression was induced equally well by the two Sxy proteins, and of *hofM*, whose expression was strongly induced only by *E. coli* Sxy (Fig. 4C). The control results (*E. coli* Sxy and CRP; *H. influenzae* Sxy and *E. coli* CRP) confirmed the results of the array analysis. However, the ability of *H. influenzae* Sxy to induce both *ppdD* and *hofM* was dramatically improved when the cognate CRP was present. Conversely, the ability of *E. coli* Sxy to induce these genes was dramatically reduced in the presence of *H. influenzae* CRP. This result provides the strongest evidence to date that transcriptional activation at CRP-S sites requires direct interaction between Sxy and CRP.

DISCUSSION

Our demonstration that *E. coli* has a functional CRP-S regulon like that of *H. influenzae* strengthens the hypothesis that CRP-S regulons are shared by all *Pasteurellaceae* and *Enterobacteriaceae*, as proposed by Cameron and Redfield (12). Furthermore, the demonstration of partial cross-complementation shows that Sxy and CRP have similar functions in the two species. The differences in CRP-S motifs and the absence of full cross-complementation provide additional insights into the action of Sxy and CRP at CRP-S sites and suggest that both CRP and Sxy have evolved species-specific features. The CRP-N sites (canonical CRP binding sites) of *H. influenzae* and *E. coli* are very similar, and their CRPs are identical at the residues known to interact with DNA (11). In contrast, their CRP-S motifs differ at two core positions and in flanking sequences (Fig. 3). These differences make the two halves of the *E. coli* CRP-S site core nonpalindromic. This asymmetry may affect CRP's ability to bind as a dimer or, since intimate contacts between CRP and Sxy appear to be needed for transcriptional activation at CRP-S sites, may create a need for Sxy to asymmetrically modulate CRP-DNA binding to one half-site. AT-rich regions upstream of *H. influenzae* CRP-S sites have been proposed as binding sites for Sxy, enabling it to assist in RNA polymerase recruitment, but such regions were absent from the *E. coli* motif (data not shown), so their role must be dispensable.

All *H. influenzae* CRP-S genes required for competence are also CRP-S regulated in *E. coli*. Although the function of these genes in competence remains to be demonstrated, this strongly suggests that *E. coli* can produce T4P and naturally transform and that both processes require Sxy. The toxicity resulting from *sxy* expression in this study may have prevented us from detecting these phenotypes. Expression of plasmid-borne *sxy* at a lower level was also unsuccessful (data not shown), strongly reinforcing the need to identify conditions that naturally induce *sxy* expression. Conserved homologues of *sxy* and of all competence gene homologues are found in most sequenced strains of *E. coli*. Given both the known sporadic distribution of natural competence among strains of competent bacteria

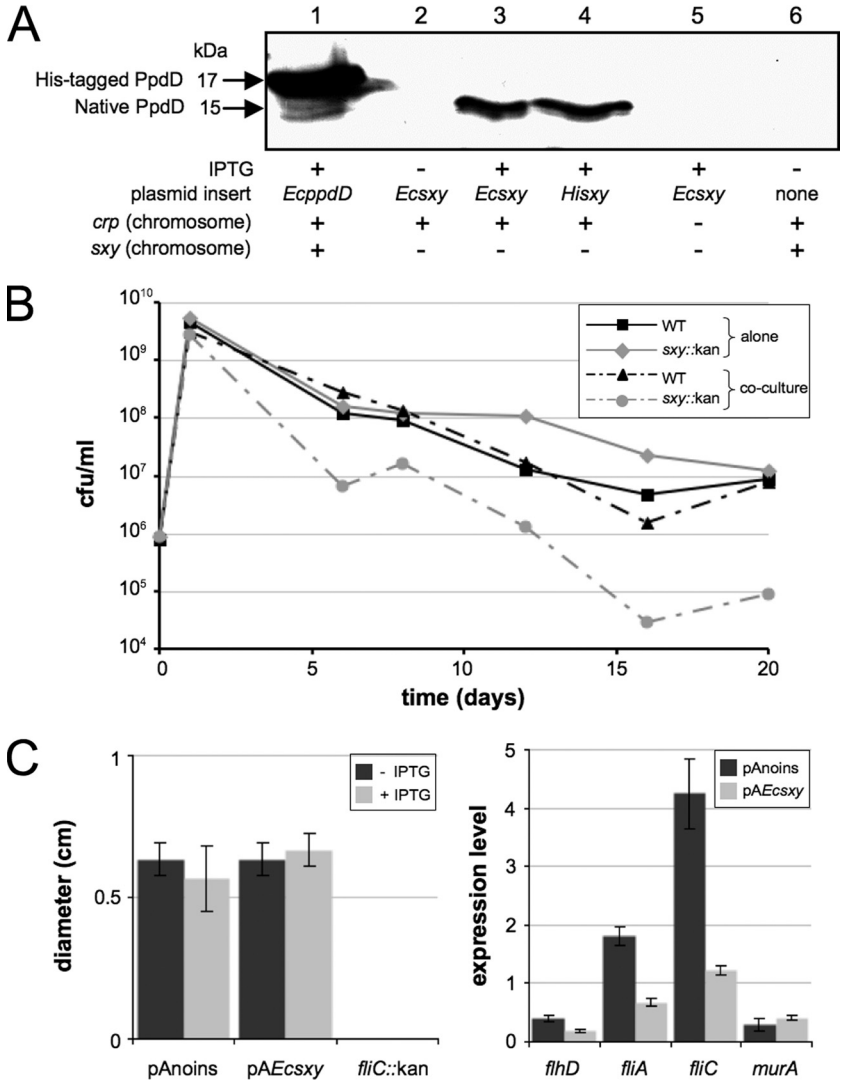

FIG. 6. Sxy-dependent phenotypes. (A) Expression of the T4P subunit PpdD in various strains. Western blotting with PpdD antiserum was performed using whole-cell extracts from broth-grown bacteria. The positions of native processed (15-kDa) and His-tagged (17-kDa) PpdD proteins are indicated by arrows. Lane 1, *E. coli* BW25113/p*EcppdD* plus IPTG; lane 2, *E. coli sxy::*kan/p*Ecsxy*; lane 3, *E. coli sxy::*kan/p*Ecsxy* plus IPTG; lane 4, *E. coli sxy::*kan/p*Hisxy* plus IPTG; lane 5, *E. coli crp::*kan/p*Ecsxy* plus IPTG; lane 6, *E. coli* BW25113. (B) Long-term survival of the *E. coli* BW25113 (parent) (black lines) and *sxy*::kan (gray lines) strains in LB broth, alone (solid lines) or in coculture (dashed lines). A decrease in survival was consistently observed in three independent experiments. The results of a representative experiment are shown. WT, wild type. (C) Effect of *sxy* expression on bacterial motility (left panel) and on the expression of flagellar genes (right panel). For motility, the diameter of each zone of motility was measured after 6 h, and the diameter of the filter disk was subtracted. Each bar indicates the average of three independent replicates, and the error bars indicate standard deviations. For gene expression, expression of *flhD* (b1892), *fliA* (b1922), and *fliC* (b1923) was measured by qPCR in *E. coli sxy*::kan/pnoins (gray bars) and in *E. coli sxy*::kan/p*Ecsxy* (black bars). Each bar indicates the mean of at least two independent biological replicates for each gene. The error bars indicate the standard deviations from the means. The expression levels are normalized to 1:1,000 23S rRNA levels.

(42, 67) and Sxy's demonstrated ability to induce expression of an apparently complete *E. coli* competence regulon, strains other than K-12 will be examined for natural competence. Because our choice of strain may also explain our inability to naturally induce *sxy*, these strains will also be tested for *sxy* expression and induction.

The discovery that some phenotypes depend on the chromosomal *sxy* gene indicates that basal expression of *sxy* and of the genes that it regulates has significant effects. Sxy may contribute to long-term survival only because Sxy induces genes allowing use of DNA as a nutrient (all but one of the genes implicated by Palchevskiy and Finkel are CRP-S regulated), but its contribution could also reflect contributions of other CRP-S-regulated genes. Although Sun et al. (69) found no contribution of individual competence gene homologues to plasmid DNA uptake, we found that Sxy contributes to plasmid transformation, most likely through its role as a regulator of these genes (46). The protocols used in the two assays were very different, which may explain the contrasting results. In light of Sxy's induction of the predicted DNA uptake machin-

ery and its use in plasmid transformation, our failure to detect transformation with chromosomal DNA may be due to cytoplasmic DNA degradation or other blocks to recombination.

Both *H. influenzae* and *E. coli* CRP-S regulons are clearly not strictly "competence regulons," as they contain genes for functions other than T4P biogenesis and DNA uptake (*E. coli* more so than *H. influenzae*). In *H. influenzae* competence is induced by depletion of nucleotide pools, consistent with the hypothesis that cells take up DNA mainly to obtain nucleotides (39, 56). The presence in its CRP-S regulon of *ssb* and *radC* led Redfield et al. (56) to suggest that these and other noncompetence genes contribute to a second arm of the nucleotide starvation response by protecting and resolving stalled replication forks (8, 38). Both of these genes are also CRP-S regulated in *E. coli.* From this perspective, the CRP-S sites of genes encoding toxin-antitoxin proteins are particularly interesting, as they have been proposed to "regulate the synthesis of macromolecules (i.e., proteins and DNA) at rates compatible with the external supply of nutrients" (24). Further investigation of the functions of the induced genes should shed light on this hypothesis.

ACKNOWLEDGMENTS

Transmission electron microscopy was performed at the University of British Columbia's Bioimaging Facility with the help of Bradford Ross.

Funding for this work was provided to R.J.R. through the Canadian Institute of Health Research.

REFERENCES

1. **Alm, R. A., and J. S. Mattick.** 1995. Identification of a gene, pilV, required for type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*, whose product possesses a pre-pilin-like leader sequence. Mol. Microbiol. **16:**485–496.
2. **Alm, R. A., and J. S. Mattick.** 1996. Identification of two genes with prepilin-like leader sequences involved in type 4 fimbrial biogenesis in *Pseudomonas aeruginosa*. J. Bacteriol. **178:**3809–3817.
3. **Angelini, S., C. Gerez, S. Ollagnier-de Choudens, Y. Sanakis, M. Fontecave, F. Barras, and B. Py.** 2008. NfuA, a new factor required for maturing Fe/S proteins in *Escherichia coli* under oxidative stress and iron starvation conditions. J. Biol. Chem. **283:**14084–14091.
4. **Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori.** 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. **2:**2006.0008.
5. **Bailey, T. L., N. Williams, C. Misleh, and W. W. Li.** 2006. MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res. **34:**W369–W373.
6. **Bakaletz, L. O., B. D. Baker, J. A. Jurcisek, A. Harrison, L. A. Novotny, J. E. Bookwalter, R. Mungur, and R. S. Munson, Jr.** 2005. Demonstration of type IV pilus expression and a twitching phenotype by *Haemophilus influenzae*. Infect. Immun. **73:**1635–1643.
7. **Bauer, F., C. Hertel, and W. P. Hammes.** 1999. Transformation of *Escherichia coli* in foodstuffs. Syst. Appl. Microbiol. **22:**161–168.
8. **Baur, B., K. Hanselmann, W. Schlimme, and B. Jenni.** 1996. Genetic transformation in freshwater: *Escherichia coli* is able to develop natural competence. Appl. Environ. Microbiol. **62:**3673–3678.
9. **Beattie, K. L., and J. K. Setlow.** 1971. Transformation-defective strains of *Haemophilus influenzae*. Nat. New Biol. **231:**177–179.
10. **Bergmans, H. E., I. M. van Die, and W. P. Hoekstra.** 1981. Transformation in *Escherichia coli*: stages in the process. J. Bacteriol. **146:**564–570.
11. **Cameron, A. D., and R. J. Redfield.** 2008. CRP binding and transcription activation at CRP-S sites. J. Mol. Biol. **383:**313–323.
12. **Cameron, A. D., and R. J. Redfield.** 2006. Non-canonical CRP sites control competence regulons in *Escherichia coli* and many other gamma-proteobacteria. Nucleic Acids Res. **34:**6001–6014.
13. **Cameron, A. D., M. Volar, L. A. Bannister, and R. J. Redfield.** 2008. RNA secondary structure regulates the translation of *sxy* and competence development in *Haemophilus influenzae*. Nucleic Acids Res. **36:**10–20.
14. **Claret, L., and C. Hughes.** 2000. Functions of the subunits in the FlhD(2)C(2) transcriptional master regulator of bacterial flagellum biogenesis and swarming. J. Mol. Biol. **303:**467–478.
15. **Claverys, J. P., and B. Martin.** 2003. Bacterial "competence" genes: signatures of active transformation, or only remnants? Trends Microbiol. **11:**161–165.
16. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res. **14:**1188–1190.
17. **Dalma-Weiszhausz, D. D., M. R. Gartenberg, and D. M. Crothers.** 1991. Sequence-dependent contribution of distal binding domains to CAP protein-DNA binding affinity. Nucleic Acids Res. **19:**611–616.
18. **Dartigalongue, C., D. Missiakas, and S. Raina.** 2001. Characterization of the *Escherichia coli* sigma E regulon. J. Biol. Chem. **276:**20866–20875.
19. **De Wulf, P., A. M. McGuire, X. Liu, and E. C. Lin.** 2002. Genome-wide profiling of promoter recognition by the two-component response regulator CpxR-P in *Escherichia coli*. J. Biol. Chem. **277:**26652–26661.
20. **Dorel, C., P. Lejeune, and A. Rodrigue.** 2006. The Cpx system of *Escherichia coli*, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? Res. Microbiol. **157:**306–314.
21. **Dougherty, B. A., and H. O. Smith.** 1999. Identification of *Haemophilus influenzae* Rd transformation genes using cassette mutagenesis. Microbiology **145:**401–409.
22. **Finkel, S. E., and R. Kolter.** 2001. DNA as a nutrient: novel role for bacterial competence gene homologs. J. Bacteriol. **183:**6288–6293.
23. **Francetic, O., S. Lory, and A. P. Pugsley.** 1998. A second prepilin peptidase gene in *Escherichia coli* K-12. Mol. Microbiol. **27:**763–775.
24. **Gerdes, K.** 2000. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. J. Bacteriol. **182:**561–572.
25. **Goff, S. A., and A. L. Goldberg.** 1985. Production of abnormal proteins in *E. coli* stimulates transcription of lon and other heat shock genes. Cell **41:**587–595.
26. **Gwinn, M. L., R. Ramanathan, H. O. Smith, and J. F. Tomb.** 1998. A new transformation-deficient mutant of *Haemophilus influenzae* Rd with normal DNA uptake. J. Bacteriol. **180:**746–748.
27. **Herriott, R. M., E. M. Meyer, and M. Vogt.** 1970. Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. J. Bacteriol. **101:**517–524.
28. **Hughes, K. T., and K. Mathee.** 1998. The anti-sigma factors. Annu. Rev. Microbiol. **52:**231–286.
29. **Iyer, L. M., D. D. Leipe, E. V. Koonin, and L. Aravind.** 2004. Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. **146:**11–31.
30. **Jarosik, G. P., and E. J. Hansen.** 1994. Cloning and sequencing of the *Haemophilus influenzae ssb* gene encoding single-strand DNA-binding protein. Gene **146:**101–103.
31. **Jurcisek, J. A., and L. O. Bakaletz.** 2007. Biofilms formed by nontypeable *Haemophilus influenzae* in vivo contain both double-stranded DNA and type IV pilin protein. J. Bacteriol. **189:**3868–3875.
32. **Karudapuram, S., and G. J. Barcak.** 1997. The *Haemophilus influenzae dprABC* genes constitute a competence-inducible operon that requires the product of the *tfoX* (*sxy*) gene for transcriptional activation. J. Bacteriol. **179:**4815–4820.
33. **Karudapuram, S., X. Zhao, and G. J. Barcak.** 1995. DNA sequence and characterization of *Haemophilus influenzae dprA*$^+$, a gene required for chromosomal but not plasmid DNA transformation. J. Bacteriol. **177:**3235–3240.
34. **Keseler, I. M., J. Collado-Vides, S. Gama-Castro, J. Ingraham, S. Paley, I. T. Paulsen, M. Peralta-Gil, and P. D. Karp.** 2005. EcoCyc: a comprehensive database resource for *Escherichia coli*. Nucleic Acids Res. **33:**D334–D337.
35. **Kitagawa, M., T. Ara, M. Arifuzzaman, T. Ioka-Nakamichi, E. Inamoto, H. Toyonaga, and H. Mori.** 2005. Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. DNA Res. **12:**291–299.
36. **Larson, T. G., and S. H. Goodgal.** 1992. Donor DNA processing is blocked by a mutation in the *com101A* locus of *Haemophilus influenzae*. J. Bacteriol. **174:**3392–3394.
37. **Larson, T. G., and S. H. Goodgal.** 1991. Sequence and transcriptional regulation of *com101A*, a locus required for genetic transformation in *Haemophilus influenzae*. J. Bacteriol. **173:**4683–4691.
38. **Liu, X., and P. Matsumura.** 1994. The FlhD/FlhC complex, a transcriptional activator of the *Escherichia coli* flagellar class II operons. J. Bacteriol. **176:**7345–7351.
39. **MacFadyen, L. P., D. Chen, H. C. Vo, D. Liao, R. Sinotte, and R. J. Redfield.** 2001. Competence development by *Haemophilus influenzae* is regulated by the availability of nucleic acid precursors. Mol. Microbiol. **40:**700–707.
40. **Martin, P. R., M. Hobbs, P. D. Free, Y. Jeske, and J. S. Mattick.** 1993. Characterization of *pilQ*, a new gene required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. Mol. Microbiol. **9:**857–868.
41. **Martin, P. R., A. A. Watson, T. F. McCaul, and J. S. Mattick.** 1995. Characterization of a five-gene cluster required for the biogenesis of type 4 fimbriae in *Pseudomonas aeruginosa*. Mol. Microbiol. **16:**497–508.
42. **Maughan, H., and R. J. Redfield.** 2009. Extensive variation in natural competence in *Haemophilus influenzae.* Evolution **63:**1852–1866.

43. **Meyer, R. R., and P. S. Laine.** 1990. The single-stranded DNA-binding protein of *Escherichia coli*. Microbiol. Rev. **54:**342–380.
44. **Niki, H., C. Ichinose, T. Ogura, H. Mori, M. Morita, M. Hasegawa, N. Kusukawa, and S. Hiraga.** 1988. Chromosomal genes essential for stable maintenance of the mini-F plasmid in *Escherichia coli*. J. Bacteriol. **170:** 5272–5278.
45. **Nonaka, G., M. Blankschien, C. Herman, C. A. Gross, and V. A. Rhodius.** 2006. Regulon and promoter analysis of the *E. coli* heat-shock factor, sigma32, reveals a multifaceted cellular response to heat stress. Genes Dev. **20:**1776–1789.
46. **Nunn, D. N., and S. Lory.** 1991. Product of the *Pseudomonas aeruginosa* gene *pilD* is a prepilin leader peptidase. Proc. Natl. Acad. Sci. USA **88:**3281–3285.
47. **Oh, M. K., and J. C. Liao.** 2000. DNA microarray detection of metabolic responses to protein overproduction in *Escherichia coli*. Metab. Eng. **2:**201–209.
48. **Palchevskiy, V., and S. E. Finkel.** 2006. *Escherichia coli* competence gene homologs are essential for competitive fitness and the use of DNA as a nutrient. J. Bacteriol. **188:**3902–3910.
49. **Park, K., S. Choi, M. Ko, and C. Park.** 2001. Novel sigmaF-dependent genes of *Escherichia coli* found using a specified promoter consensus. FEMS Microbiol. Lett. **202:**243–250.
50. **Parsell, D. A., and R. T. Sauer.** 1989. Induction of a heat shock-like response by unfolded protein in *Escherichia coli*: dependence on protein level not protein degradation. Genes Dev. **3:**1226–1232.
51. **Pelicic, V.** 2008. Type IV pili: e pluribus unum? Mol. Microbiol. **68:**827–837.
52. **Porco, A., G. Alonso, and T. Isturiz.** 1998. The gluconate high affinity transport of GntI in *Escherichia coli* involves a multicomponent complex system. J. Basic Microbiol. **38:**395–404.
53. **Pruss, B. M., J. W. Campbell, T. K. Van Dyk, C. Zhu, Y. Kogan, and P. Matsumura.** 2003. FlhD/FlhC is a regulator of anaerobic respiration and the Entner-Doudoroff pathway through induction of the methyl-accepting chemotaxis protein Aer. J. Bacteriol. **185:**534–543.
54. **Pruss, B. M., X. Liu, W. Hendrickson, and P. Matsumura.** 2001. FlhD/FlhC-regulated promoters analyzed by gene array and *lacZ* gene fusions. FEMS Microbiol. Lett. **197:**91–97.
55. **Redfield, R. J.** 1991. *sxy-1*, a *Haemophilus influenzae* mutation causing greatly enhanced spontaneous competence. J. Bacteriol. **173:**5612–5618.
56. **Redfield, R. J., A. D. Cameron, Q. Qian, J. Hinds, T. R. Ali, J. S. Kroll, and P. R. Langford.** 2005. A novel CRP-dependent regulon controls expression of competence genes in *Haemophilus influenzae*. J. Mol. Biol. **347:**735–747.
57. **Rhodius, V. A., W. C. Suh, G. Nonaka, J. West, and C. A. Gross.** 2006. Conserved and variable functions of the sigmaE stress response in related genomes. PLoS Biol. **4:**e2.
58. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
59. **Sauvonnet, N., P. Gounon, and A. P. Pugsley.** 2000. PpdD type IV pilin of *Escherichia coli* K-12 can be assembled into pili in *Pseudomonas aeruginosa*. J. Bacteriol. **182:**848–854.
60. **Sauvonnet, N., G. Vignon, A. P. Pugsley, and P. Gounon.** 2000. Pilus formation and protein secretion by the same machinery in *Escherichia coli*. EMBO J. **19:**2221–2228.
61. **Saveson, C. J., and S. T. Lovett.** 1999. Tandem repeat recombination induced by replication fork defects in *Escherichia coli* requires a novel factor, RadC. Genetics. **152:**5–13.
62. **Shi, W., C. Li, C. J. Louise, and J. Adler.** 1993. Mechanism of adverse conditions causing lack of flagella in *Escherichia coli*. J. Bacteriol. **175:**2236–2240.
63. **Shin, S., and C. Park.** 1995. Modulation of flagellar expression in *Escherichia coli* by acetyl phosphate and the osmoregulator OmpR. J. Bacteriol. **177:** 4696–4702.
64. **Sinha, S., O. H. Ambur, P. R. Langford, T. Tonjum, and J. S. Kroll.** 2008. Reduced DNA binding and uptake in the absence of DsbA1 and DsbA2 of *Neisseria meningitidis* due to inefficient folding of the outer-membrane secretin PilQ. Microbiology **154:**217–225.
65. **Sinha, S., P. R. Langford, and J. S. Kroll.** 2004. Functional diversity of three different DsbA proteins from *Neisseria meningitidis*. Microbiology **150:**2993–3000.
66. **Smeets, L. C., S. C. Becker, G. J. Barcak, C. M. Vandenbroucke-Grauls, W. Bitter, and N. Goosen.** 2006. Functional characterization of the competence protein DprA/Smf in *Escherichia coli*. FEMS Microbiol. Lett. **263:**223–228.
67. **Solomon, J. M., and A. D. Grossman.** 1996. Who's competent and when: regulation of natural genetic competence in bacteria. Trends Genet. **12:**150–155.
68. **Stafford, G. P., T. Ogi, and C. Hughes.** 2005. Binding and transcriptional activation of non-flagellar genes by the *Escherichia coli* flagellar master regulator FlhD2C2. Microbiology **151:**1779–1788.
69. **Sun, D., X. Zhang, L. Wang, M. Prudhomme, Z. Xie, B. Martin, and J. P. Claverys.** 2009. Transforming DNA uptake gene orthologs do not mediate spontaneous plasmid transformation in *Escherichia coli*. J. Bacteriol. **191:** 713–719.
70. **Sun, D., Y. Zhang, Y. Mei, H. Jiang, Z. Xie, H. Liu, X. Chen, and P. Shen.** 2006. *Escherichia coli* is naturally transformable in a novel transformation system. FEMS Microbiol. Lett. **265:**249–255.
71. **Taoka, M., Y. Yamauchi, T. Shinkawa, H. Kaji, W. Motohashi, H. Nakayama, N. Takahashi, and T. Isobe.** 2004. Only a small subset of the horizontally transferred chromosomal genes in *Escherichia coli* are translated into proteins. Mol. Cell. Proteomics **3:**780–787.
72. **Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin.** 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res. **28:**33–36.
73. **Thompson, W., E. C. Rouchka, and C. E. Lawrence.** 2003. Gibbs recursive sampler: finding transcription factor binding sites. Nucleic Acids Res. **31:** 3580–3585.
74. **Toma, C., H. Kuroki, N. Nakasone, M. Ehara, and M. Iwanaga.** 2002. Minor pilin subunits are conserved in *Vibrio cholerae* type IV pili. FEMS Immunol. Med. Microbiol. **33:**35–40.
75. **Tomb, J. F., G. J. Barcak, M. S. Chandler, R. J. Redfield, and H. O. Smith.** 1989. Transposon mutagenesis, characterization, and cloning of transformation genes of *Haemophilus influenzae* Rd. J. Bacteriol. **171:**3796–3802.
76. **Tomb, J. F., H. el-Hajj, and H. O. Smith.** 1991. Nucleotide sequence of a cluster of genes involved in the transformation of *Haemophilus influenzae* Rd. Gene **104:**1–10.
77. **Tracy, E., F. Ye, B. D. Baker, and R. S. Munson, Jr.** 2008. Construction of non-polar mutants in *Haemophilus influenzae* using FLP recombinase technology. BMC Mol. Biol. **9:**101.
78. **van Helden, J.** 2003. Regulatory sequence analysis tools. Nucleic Acids Res. **31:**3593–3596.
79. **VanWagoner, T. M., P. W. Whitby, D. J. Morton, T. W. Seale, and T. L. Stull.** 2004. Characterization of three new competence-regulated operons in *Haemophilus influenzae*. J. Bacteriol. **186:**6409–6421.
80. **Watts, T. H., C. M. Kay, and W. Paranchych.** 1983. Spectral properties of three quaternary arrangements of *Pseudomonas* pilin. Biochemistry **22:**3640–3646.
81. **Whitchurch, C. B., M. Hobbs, S. P. Livingston, V. Krishnapillai, and J. S. Mattick.** 1991. Characterisation of a *Pseudomonas aeruginosa* twitching motility gene and evidence for a specialised protein export system widespread in eubacteria. Gene **101:**33–44.
82. **Woegerbauer, M., B. Jenni, F. Thalhammer, W. Graninger, and H. Burgmann.** 2002. Natural genetic transformation of clinical isolates of *Escherichia coli* in urine and water. Appl. Environ. Microbiol. **68:**440–443.
83. **Zhao, K., M. Liu, and R. R. Burgess.** 2007. Adaptation in bacterial flagellar and motility systems: from regulon members to 'foraging'-like behavior in *E. coli*. Nucleic Acids Res. **35:**4441–4452.
84. **Zhao, K., M. Liu, and R. R. Burgess.** 2005. The global transcriptional response of *Escherichia coli* to induced sigma 32 protein involves sigma 32 regulon activation followed by inactivation and degradation of sigma 32 in vivo. J. Biol. Chem. **280:**17758–17768.
85. **Zulty, J. J., and G. J. Barcak.** 1995. Identification of a DNA transformation gene required for *com101A*$^+$ expression and supertransformer phenotype in *Haemophilus influenzae*. Proc. Natl. Acad. Sci. USA **92:**3616–3620.