Regulation of the *Escherichia coli* csgD promoter: interplay between five transcription factors

Hiroshi Ogasawara,1 Kayoko Yamada,1 Ayako Kori,1 Kaneyoshi Yamamoto1,2 and Akira Ishihama1,2

1Department of Frontier Bioscience, Hosei University, Koganei, Tokyo 184-8584, Japan
2Research Center for Micro-Nano Technology, Hosei University, Koganei, Tokyo 184-8584, Japan

Under stressful conditions in nature, *Escherichia coli* forms biofilms for long-term survival. Curli fimbriae are an essential architecture for cell–cell contacts within biofilms. Structural components and assembly factors of curli are encoded by two operons, csgBA and csgDEFG. The csgD gene product controls transcription of both operons. Reflecting the response of csgD expression to external stresses, a number of transcription factors participate in the regulation of the csgD promoter. Analysis of the csgD mRNA obtained from *E. coli* mutants in different transcription factors indicated that CpxR and H-NS act as repressors while OmpR, RstA and IHF act as activators. An acid-stress response regulator, RstA, activates csgD only under acidic conditions. These five factors bind within a narrow region of about 200 bp upstream of the csgD promoter. After pair-wise promoter-binding assays, the increase in csgD transcription in the stationary phase was suggested to be due, at least in part, to the increase in IHF level cancelling the silencing effect of H-NS. In addition, we propose a novel regulation model of this complex csgD promoter through cooperation between the two positive factors (OmpR–IHF and RstA–IHF) and also between the two negative factors (CpxR–H-NS).

INTRODUCTION

Bacteria can switch from a single-cell planktonic growth mode to a multicellular community (biofilm) mode. During such a growth transition state, cell morphology, physiology and metabolism are markedly altered (Prigent-Combaret et al., 2001; Schembri et al., 2003; Beloin et al., 2004; Ren et al., 2004). In single-cell growth mode, cell motility using flagella is critical for adaptation to environments. The master regulator FlhCD complex plays a key role in controlling transcription of a set of genes for flagella formation (Claret & Huges, 2000). When *Escherichia coli* cells switch their life mode from single planktonic cell growth to multicellular community (biofilm) mode, the genetic system for flagella formation is turned off and the genes involved in cell–cell adhesion are activated. The biofilm matrix is a complex architecture of cell aggregates that are attached on the surface of inorganic solid materials in nature or on eukaryotic tissues in host animals. After a comprehensive analysis of a set of *E. coli* mutants, each lacking one of 3985 non-essential genes, a total of 110 genes were indicated to be involved in biofilm formation (Niba et al., 2007).

During biofilm development, curli fimbriae, the major biofilm component, play a key role in both initial adhesion to solid surfaces and subsequent cell–cell interactions (Vidal et al., 1998; Chapman et al., 2002; Prigent-Combaret et al., 2000). Curli fimbriae also mediate bacterial adhesion to host cells and invasion, and activation of both the proinflammatory response and the immune system (Bian et al., 2000; Zoga et al., 2001). Accordingly, curli participate in virulence phenotypes (Hammar et al., 1995; Vidal et al., 1998; Bian et al., 2000; Gophna et al., 2001; Cookson et al., 2002). A set of polysaccharide polymers such as colanic acid and cellulose are also involved in the formation of cell aggregates and their attachment to solid surfaces, together leading to biofilm formation (Arnaquist et al., 1994; Prigent-Combaret et al., 2001; Sutherland, 2001; Zogaj et al., 2001; Wang et al., 2004). Co-production of curli and polysaccharides is required for efficient biofilm formation (Prigent-Combaret et al., 2000; Zogaj et al., 2001). Interestingly, bacterial curli fibres share biochemical and structural characteristics with eukaryotic amyloid fibres and thus provide a model system for understanding amyloid fibre formation (Wang et al., 2007).

Genes for curli formation are organized in two adjacent divergently transcribed operons: csgBA, encoding the structural components, and csgDEFG, encoding proteins for curli assembly and transport (Loferer et al., 1997). Between two operons, there is one of the largest intergenic spacers in the *E. coli* genome without coding capacity; it is 755 bp in length with unique features such as high curvature and low stability (Pedersen et al., 2000). Expression of the curli operons is under the control of...
the CsgD protein, the first gene product of the csgDEFG operon (Hammar et al., 1995). The CsgD protein, a FixJ/LuxR-family transcription regulator, modulates the expression of not only the csg operon but also a set of genes for adaptation of cell physiology to the biofilm life style (Chirwa & Herrington, 2003; Brombacher et al., 2006), including bapA, encoding a large secreted protein (Latasa et al., 2005), and adrA, encoding the enzyme for synthesis of cyclic di-GMP, a bacterial second messenger (Simm et al., 2004) for enhancement of cellulose production (Romling et al., 2000; Zogaj et al., 2001). The positive regulation of cellulose production by CsgD is mediated through the transcriptional regulation of adrA. In turn, the AdrA protein activates cellulose synthesis at a post-transcriptional level by controlling the synthesis of c-di-GMP. CsgD also regulates youD, encoding one of the c-di-GMP phosphodiesterases (Brombacher et al., 2006), suggesting that CsgD is involved in control of the intracellular level of c-di-GMP. So far CsgD-binding sites have been identified in the CsgD-dependent csgB, adrA and pepD promoters (Brombacher et al., 2003). CsgD upregulates glyA to facilitate the synthesis of curli, which contain more glycine than normal E. coli proteins (Chirwa & Herrington, 2003). Deletion of the GlcNAc-6P deacetylase gene, nagA, results in decreased transcription from the curli-specific promoters, csgBA and csgDEFG, and a corresponding decrease in curli production (Barnhart & Chapman, 2006). Disruption of nagC, encoding the regulator of nag operon for utilization of GlcNAc, leads to a reduction in curli production, while an increase in intracellular GlcNAc-6P level leads to downregulation of curli gene expression. Together, these observations indicate that CsgD is the master regulator of a set of genes for biofilm formation.

In concert with the complex mode of regulation for biofilm formation, the expression of csgD has been proposed to be under the control of various stress-sensing regulatory proteins such as OmpR (Romling et al., 1998; Vidal et al., 1998; Prigent-Combaret et al., 1999, 2000, 2001; Gerstel & Romling, 2001; Gerstel et al., 2003), CpxR (Prigent-Combaret et al., 2001; Jubelin et al., 2005), Crl (Boudour et al., 2004; Robbe-Saule et al., 2006), CRP (Zheng et al., 2004), H-NS (Arnvist et al., 1994; Gerstel et al., 2003; Olsen et al., 1993), IHF (Gerstel et al., 2003; Gerstel & Romling, 2001), MlrA (Brown et al., 2001) and RcsB (Ferrieres & Clarke, 2003; Vianney et al., 2005). Recently, we found that the phosphorelay network from the Mg(II)-sensing PhoQP to the low-pH-sensing RstBA is directly involved in regulation of the csgD promoter (Ogasawara et al., 2007a). At present, however, the molecular mechanisms of transcription regulation of the csg operons by so many transcription factors and possible interplay between these regulatory proteins remain totally unsolved. Knowledge of how multiple factors, each monitoring a specific environmental condition, influence csgD promoter activity could be a clue toward understanding the pathway of E. coli differentiation leading to formation of biofilm. We have therefore performed a systematic study of transcription regulation of a single csgD promoter by as many as 10 reported or predicted transcription factors.

In this report, we describe the regulatory modes of five transcription factors, CpxR, OmpR, RstA, IHF and H-NS, and their interplays in binding in vitro to the csgD promoter. Based on the results described herein, we propose a novel regulation model of cooperation between two positive factors (OmpR and IHF) and RstA and IHF and also between two negative factors (CpxR and H-NS). On-going studies on the regulation of the csgD promoter by other regulators, CRP, Crl, RcsB and MlrA, will be described elsewhere.

**METHODS**

**Bacterial strains.** Escherichia coli BL21(DE3) [F-ompT hsd (r_{E. coli} m_{E. coli}) dcm gal λ(DE3)] was used for expression and purification of all the transcription factors examined in this study. E. coli K-12 wild-type BW25113 and single-gene knockout mutant strains, JW3368 (ompR), JW1702 (cpxR), JW3883 (rsta), JW1600 (ihfA), JW0895 (ihfB) and JW1225 (hns), were obtained from the Keio collection in the National Bio-Resource Project (National Institute of Genetics, Japan). Cells were cultured in LB medium, YESCA medium (Pratt & Silhavy, 1998) or low-phosphate minimal glucose medium (LPM) (Suziedeliene et al., 1999) supplemented with peptone (0.6 mg ml^{-1}) and at 28 °C or 37 °C. When necessary, 100 μg ampicillin ml^{-1} and 50 μg kanamycin ml^{-1} were added to the medium.

**Plasmid construction.** For the construction of a lacZ reporter vector, a DNA fragment containing the csgD promoter region was prepared by PCR using E. coli KP7600 genomic DNA as a template and primers csgD-EcoRI-F and csgD-BamHI-R2 (for sequences see Table 1). After digestion with EcoRI and BamHI, the PCR-amplified fragment was inserted into pRS551 (Simons et al., 1987) at the corresponding site to generate the plasmid pRScsgD. For construction of an arabinose-inducible NlpE expression plasmid, a DNA fragment (774 bp) containing the NlpE-coding region was prepared by PCR using KP7600 genomic DNA as a template and primers csgD-EcoRI-F and csgD-BamHI-R2 (for sequences see Table 1). After digestion with EcoRI and XhoI, the PCR-amplified fragment was inserted into pBAD18 at the corresponding site to generate the plasmid pBADnlpE.

**Expression and purification of transcription factors.** All the transcription factors used in this study were overexpressed in and purified from E. coli as His-tagged forms except for IHF, which was purified in native form without His-tag as described previously (Murakami et al., 1996; Azam & Ishihama, 1999; Yamamoto et al., 2005). For construction of plasmids for His-tagged transcription factors, pCpxR, pH-N, pOmpR and pRstA, DNA fragments corresponding to the coding sequences of the respective transcription factors were amplified by PCR using E. coli W3110 genomic DNA as a template, with primers BAD-nlpE-EcoRI and BAD-nlpE-XbaI-R (for sequences see Table 1). After digestion with EcoRI and XhoI, the PCR-amplified fragment was inserted into pBAD18 at the corresponding site to generate the plasmid pBADnlpE.
washed with lysis buffer (50 mM Tris/HCl, pH 8.0 at 4 °C, and 100 mM NaCl), and then stored at ~80 °C until use.

For protein purification, frozen cells were lysed by treatment with lysozyme in lysis buffer containing 100 mM PMSF followed by sonication for cell disruption. His-tagged transcription factors were affinity-purified from the supernatant fractions by use of a Ni-NTA column (Qiagen) (Murakami et al., 2005). The purity of purified transcription factors was checked by agarose gel electrophoresis on a 6% polyacrylamide gel containing 7 M urea using the SHIMADZU slab gel electrophoresis system (DSQ-500L).

**Table 1. Primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hns-F</td>
<td>GAGATTACTCATATGAGCGGACG</td>
</tr>
<tr>
<td>hns-R</td>
<td>GCAATCTACAGGCGCTTCCGTCATC</td>
</tr>
<tr>
<td>csgD-EcoRI-F</td>
<td>AGACAGGAATTCCTCCTGCCCCGTCGCT</td>
</tr>
<tr>
<td>csgD-BamHI-R2</td>
<td>CTGACAGGATCCTCTTAAATTAGAAGCACA</td>
</tr>
<tr>
<td>BAD-nlpE-EcoRIF</td>
<td>CAAGCGCTATGATGCGCCGCGAAAGTGC</td>
</tr>
<tr>
<td>BAD-nlpE-XbaI-R</td>
<td>TATATCGCTTACTAGGTTTCTGTTT</td>
</tr>
<tr>
<td>CD1F</td>
<td>AATTATTTTATATGACATT</td>
</tr>
<tr>
<td>CD1R-FITC-R</td>
<td>CAATTTTTAAAAATCATACATAC</td>
</tr>
<tr>
<td>CD3F</td>
<td>TTAATAAAAAACCTTAAGGTATAA</td>
</tr>
<tr>
<td>CD3R-FITC-R</td>
<td>TTAAATGATGACTAAATGCGT</td>
</tr>
<tr>
<td>CD6F</td>
<td>TTCTGGCCCTGCTGATTGCT</td>
</tr>
<tr>
<td>CD6R-FITC-R</td>
<td>GCACTGCTTGTTGATGAATAT</td>
</tr>
<tr>
<td>asr-F</td>
<td>ATGAAAGAGTATTACATGCCTGTTGTTG</td>
</tr>
<tr>
<td>asr-R</td>
<td>CGGTTTGGCGAACAGGTTTT</td>
</tr>
<tr>
<td>ompC-F</td>
<td>GCCGTTCATACCTATGATTAGGAAG</td>
</tr>
<tr>
<td>ompC-R</td>
<td>TTAGAACGTGAACACGACAGCAG</td>
</tr>
<tr>
<td>cpxP-F</td>
<td>TCAAGCGTGCGCTGACTGATTCA</td>
</tr>
<tr>
<td>cpxP-R</td>
<td>GGAAACGTAGTTGCTACCTACAAAT</td>
</tr>
<tr>
<td>csgD-S</td>
<td>TTATCGCCTGAGGTATGTTG</td>
</tr>
<tr>
<td>csgD-T</td>
<td>TTTCAGGCTTATATCCTGCGATAT</td>
</tr>
</tbody>
</table>

*FITC label at 5’ end.

**Northern blot analysis.** For preparation of total RNA for Northern blot analysis, overnight cultures were diluted 100-fold with 30 ml YESCA medium or LPM supplemented with peptone (0.6 mg ml⁻¹) at either pH 7.0 or pH 4.5 and cells were incubated for 6 or 12 h (until late stationary phase) at 28 °C. RNA purification (Ogasawara et al., 2007a, b) and Northern blotting (Shimada et al., 2007; Umezawa et al., 2008) were performed as described previously. The DIG-labeled DNA fragments were amplified by PCR using W3110 genomic DNA (50 ng) as template, DIG-11-dUTP (Roche) and dNTP as substrates, gene-specific forward and reverse primers, and Ex Taq DNA polymerase (Takara). In brief, 4 μg total RNA was incubated in formaldehyde-MOPS gel-loading buffer for 10 min at 65 °C for denaturation, separated by electrophoresis on formaldehyde-containing 2% agarose gel, and then transferred on to nylon membrane (Roche). Hybridization with a DIG-labeled probe was performed with the DIG easy Hyb system (Roche) at 50 °C overnight. For detection of the DIG-labeled probe, the membranes were treated with anti-DIG-AP Fab fragments and CDP-Star (Roche), and the images were scanned with LAS-4000 (Fuji Film). The product size on the membrane was estimated on the basis of migration of RNA markers (Toyobo).

**Quantitative Western blot analysis.** Intracellular concentrations of transcription factors were determined by using the quantitative Western blot system, which was employed for determination of sigma factors (Jishage et al., 1996; Maeda et al., 2000) and nucleoid proteins (Azam et al., 1999). Briefly, E. coli whole-cell lysates were prepared by sonication after lysozyme treatment, and treated with antibodies, which were prepared in rabbits against purified transcription factors, CsgD, RstA, OmpR, CpxR, IHF and H-NS. Immunoprecipitates were recovered with purified Protein-A and subjected to SDS-PAGE in parallel with various amounts of the respective purified proteins to make the standard curve.
RESULTS AND DISCUSSION

Promoter organization of the csgDEFG operon

Transcription start site P1 of the csgDEFG operon has been identified at 148 bp upstream of the csgD translation initiation codon in wild-type E. coli K-12 Y-mel cells grown in CFA agar for 48 h at 26 °C (Hammar et al., 1995). In addition, we identified another promoter, P2, 10 bp downstream from P1 (Ogasawara et al., 2007a). These csgD promoters carry the typical feature of RpoD-dependent promoters, but transcription of the csgD operon is functionally dependent on the stationary-phase-specific RpoS sigma factor as well as the major RpoD sigma factor (Hammar et al., 1995; Ogasawara et al., 2007a). RpoD-dependent promoters are often recognized by not only RpoD but also RpoS sigma factors (Kolb et al., 1995; Tanaka et al., 1995).

CsgD expression leads to an increase in RpoS level in various ways (Ren et al., 2004; Gualdi et al., 2007). For instance, CsgD induces the expression of iraP, which encodes a stabilization factor of the RpoS sigma factor, by interfering with the RssB-mediated proteolysis of RpoS by the ClpXP protease (Bougdour & Gottesman, 2007; Gualdi et al., 2007). Consequently, expression of a number of the RpoS-regulated genes is induced after CsgD expression (Lacqua et al., 2006; Gualdi et al., 2007).

Transcription factors involved in regulation in vivo of the csgD promoter

Expression of the curli operons, csgBA and csgDEFG, is induced in response to various environmental stresses such as low temperature, low osmolarity and nutrient deficiency (for instance see Olsen et al., 1993). Under laboratory culture conditions in the absence of specific external stresses, the expression of csgD is low in the exponential growth phase but increases in the stationary phase as detected by Northern blot analysis (Fig. 1a). For regulation of the csgD promoter, more than 10 transcription factors have been indicated to be involved besides two sigma factors. Among these transcription factors, direct binding to the csgD promoter has been identified for five factors, CpxR (Jubelin et al., 2005), H-NS (Gerstel et al., 2003), IHF (Gerstel et al., 2003), OmpR (Vidal et al., 1998) and RstA (Ogasawara et al., 2007a). The influence of these five transcription factors on transcription in vivo from the csgD promoter was examined by Northern blot analysis for a set of mutants, each lacking one of these factors, cultured at 28 °C in YESCA medium, in which curli expression has been identified (Pratt & Silhavy, 1998; Romling et al., 1998). In ompR, ihfA and ihfB mutants, csgD mRNA was not detected in either the exponential (6 h) or the stationary (12 h) phase (Fig. 1b, lanes 2, 5 and 6) [note that IHF is composed of an α–β heterodimer encoded by ihfA and ihfB, respectively], indicating that OmpR and IHF are the positive factors of the csgD promoter, in agreement with the previous studies (Vidal et al., 1998; Gerstel et al., 2003).

On the other hand, a marked increase in csgD mRNA level was observed for mutants lacking CpxR and H-NS, particularly in the stationary phase (Fig. 1b, lanes 3 and 7), indicating that CpxR and H-NS are the negative regulators of the csgD promoter. Activation of the Cpx signalling pathway in response to surface interactions such as adhesion requires the outer-membrane protein NlpE (Snyder et al., 1995; Otto & Silhavy, 2002). When NlpE was overexpressed, transcription of cpxP, one of the known targets of CpxR, was indeed activated, but transcription of csgD was markedly reduced (Fig. 2a, lanes 1 and 2). In the absence of CpxR, NlpE overexpression did not lead to reduction of csgD transcription (Fig. 2a, lanes 3 and 4). Taking these results together we concluded that CpxR plays a negative role in csgD transcription. Nucleoid protein H-NS is known as a general silencer of transcription in E. coli (reviewed by Ishihama, 2009). In the absence of H-NS, the level of csgD

![Fig. 1](image-url)
mRNA significantly increased (Fig. 2a, lanes 1 and 5), indicating that H-NS is a repressor of the csgD promoter. The reduction of csgD mRNA mediated by CpxR-NlpE was, however, observed in this hns knockout mutant (Fig. 2a, lanes 5 and 6).

In the rstAB mutant grown at pH 7.0, the level of csgD mRNA increased slightly (Fig. 2b, lane 5) while overproduction of RstA resulted in significant reduction of csgD mRNA (Ogasawara et al., 2007a), these results together indicating a negative role of RstA in csgD transcription. In the absence of RstA, however, translation of rpoS is enhanced (Sugiura et al., 2003) while overproduction of RstA induces degradation of RpoS in Salmonella (Cabeza et al., 2007). Thus the observed change in csgD transcription at neutral pH might not be the direct consequence of the change in repressor RstA level but a result of the change of RpoS level. Previously we noticed that RstA is involved in asr activation for the acid stress response (Ogasawara et al., 2007a). We therefore examined the possible influence of acidic conditions on csgD transcription. In the wild-type strain, the level of csgD transcription was essentially the same between pH 7.0 and pH 4.5 (Fig. 2b, lanes 1 and 2), in agreement with the published observation (Gerstel & Romling, 2001). In contrast, however, csgD mRNA was virtually undetectable in the rstA knockout mutant grown at pH 4.5 (Fig. 2b, lanes 5 and 6). This finding supports the interpretation that RstA is required for csgD transcription at low pH.

**Intracellular concentrations of transcription factors**

Growth-dependent change in the csgD promoter in wild-type E. coli, shown in Fig. 1(b), might be attributable to changes in the intracellular concentrations of transcription factors involved in the regulation of the csgD promoter. The intracellular level of the test transcription factors was therefore determined at various times of cell growth by using quantitative immunoblotting (Jishage et al., 1996; Maeda et al., 2000). In growing cells, the intracellular concentrations of two nucleoid proteins, IHF and H-NS, are similar and higher than those of the specific regulators, RstA, OmpR and CpxR (Fig. 3). Upon entry to the stationary phase, a 2.5- to 3-fold increase was observed for IHF, reaching a level of 25 000–30 000 molecules per cell (Fig. 3). In contrast, H-NS stayed constant at a level of approximately 10 000 molecules per cell (Fig. 3). The levels of these two nucleoid proteins are in good agreement with our previous determination (Ishihama, 1993; Azam et al.,...
1999). The enhancement of csgD transcription in the stationary phase might be due, at least in part, to the increase in the level of a positive factor IHF (see below for the competition between IHF and H-NS). Upon entry to the stationary phase, the levels of three specific regulators also increased slightly, but their relative ratio remained constant. The relative order, RstA>OmpR>CpxR, was also unaltered from exponential to stationary phase (Fig. 3). Thus, except for the marked increase of IHF, the change in the intracellular concentrations of other transcription factors might not be the major cause of the activation of csgD promoter in the stationary phase.

Interaction in vitro of five test transcription factors on the csgD promoter

In order to examine the mechanisms of regulation of the csgD promoter by these five factors, we purified all five factors to near homogeneity and subjected them to the gel retardation assay using various DNA fragment probes covering the csgD promoter sequence. Among the five factors examined, the three response regulators, OmpR, RstA and CpxR, are organized in two-component systems, and are converted to active forms only when phosphorylated by the respective sensor kinases, EnvZ (Aiba et al., 1989; Yamamoto et al., 2005), RstB (Minagawa et al., 2003; Yamamoto et al., 2005) and CpxA (Raivo & Silhavy, 1997; Yamamoto et al., 2005), respectively. Thus, the DNA-binding activity of OmpR, RstA and CpxR was examined in the presence of acetyl phosphate for phosphorylation in vitro (note that phosphorylated forms are shown as OmpRp, RstAp and CpxRp in the respective figures).

All five transcription factors, CpxR, OmpR, RstA, IHF and H-NS, formed binary DNA complexes in vitro in the absence of other factors (Fig. 4). The two nucleoid proteins, IHF and H-NS, formed more than two complex bands on PAGE, indicating the presence of more than two binding sites on the respective probes. The binding affinity of these five factors to the csgD promoter could be estimated based on the protein-dose-dependent disappearance pattern of probe DNA. The apparent affinity for the csgD promoter was found to be high for IHF and CpxR, followed by OmpR, H-NS and RstA (Fig. 4). Since both IHF and H-NS bind at multiple sites on the csgD promoter, the actual affinity per single binding site might be lower for these nucleoid proteins as described in our previous report (Azam & Ishihama, 1999).

Transcription-factor-binding sequences on the csgD promoter

To get insight into the mechanism by which the csgD promoter is regulated by these five transcription factors, we next determined the site(s) of recognition and binding in vitro by each purified protein. For identification of the binding sequences by all five transcription factors, DNase I footprinting mapping was performed using the respective purified proteins and the csgD promoter DNA fragment (see Fig. 5 for the gel patterns and Fig. 6 for the protected sequences). A single unique protection sequence was identified for OmpR and RstA, which was then designated hot-spot II. In addition, three factors, IHF, CpxR and H-NS, bind to the same upstream region

Fig. 3. Intracellular concentrations of the test transcription factors. (a) Overnight culture of wild-type E. coli K-12 BW25113 was transferred into fresh YESCA medium and cultured at 28 °C. At the times indicated, whole-cell lysates were prepared and subjected to quantitative Western blot analysis using antibodies against IHF, OmpR, H-NS, CpxR and RstA (upper panels), and anti-RpoA (bottom panel). (b) The intensity of immunoblot bands was measured with a LAS-4000 image analyser and IMAGE GAUGE (Fuji Film). Under the culture conditions, the intracellular concentration of RpoA was determined to be 5000 molecules per genome equivalent of DNA (Jishage et al., 1996; Maeda et al., 2000). The level of transcription factors was then determined based on the difference of immunostaining intensity between RpoA and test transcription factors. The relative level was calculated against the concentration (957 molecules per genome equivalent of DNA) of CpxR at 3 h (shown by an arrow).
between $-188$ and $-159$, designated hot-spot I. Interplay between transcription factors in binding to these two hotspots should influence the level of $csgD$ transcription.

OmpR, the response regulator of the EnvZ-OmpR two-component system, which senses the environmental osmolarity, was found to bind a 24 bp sequence ($-62$ to $-39$) (Fig. 5) including a direct repeat sequence of ($-59$)TTACATTTAgTTACATGTT($-40$) (Fig. 6), which agrees with the OmpR box (Harlocker et al., 1995; Pratt & Silhavy, 1995; Ishihama, 2010). The location of the OmpR-binding site on the $csgD$ promoter is consistent with the determinations by other laboratories (Vidal et al., 1998; Jubelin et al., 2005). After genomic SELEX screening, we found that the RstBA two-component system is under the control of the Mg(II)-sensing PhoQP two-component system (Ogasawara et al., 2007a). Among the regulation targets of RstA identified by genomic SELEX, we found the
consensus sequence of TACATNTNGTTACA as a common sequence recognized by RstA (Ishihama, 2010). This RstA-box-like sequence (−56)TACATTAGTTACA(−44) is present within the region protected by RstA (−61 to −43) (Fig. 6). CpxR, the response regulator of the CpxAR two-component system, which monitors denaturation of membrane proteins, binds to the consensus CpxR box consisting of a direct repeat sequence of GTAA(N5–7)GTAA (Pogliano et al., 1997; Yamamoto & Ishihama, 2006; Ishihama, 2010). On the csgD promoter, CpxR was found to bind over a long sequence (−190 to +15) including the CpxR-like sequences (Fig. 6). This CpxR-binding sequence mostly overlaps the reported CpxR sites (−107 to +9) (Jubelin et al., 2005).

IHF is one of the major nucleoid architectural proteins (Azam et al., 1999; Ishihama, 1999, 2009) but it also functions as a global regulator of transcription when it binds near promoters (for a review see Ishihama, 2009). IHF has been indicated to be involved in regulation of the csgD promoter (Gerstel et al., 2003, 2006). In fact, we identified two IHF-binding sites, one between −188 and −159 (IHF-1) and another between −96 and −37 (IHF-2), on the csgD promoter (Fig. 6; for the DNase footprinting pattern see Fig. 5). The promoter-proximal IHF-2 was not identified in Salmonella (Gerstel et al., 2003). The upstream IHF-1 overlaps with the H-NS- and CpxR-binding sites, altogether forming the transcription-factor-binding hot-spot I, while the downstream IHF-2 overlaps with the binding sites of all the other four factors, OmpR, RstA, CpxR and N-NS, forming hot-spot II (Fig. 6).

H-NS prefers curved DNA for binding and often functions as a general silencer of gene transcription by binding near the promoter of target genes (for reviews see Ishihama, 1999, 2009; Dorman, 2004). On the csgD promoter, H-NS protects a wide region (−201 to +28) including AT-rich sequences (Fig. 6), implying that the csgD promoter is a silencing target of H-NS.

**Interplay between transcription factors in binding to the csgD promoter**

All five transcription factors were found to bind within a narrow region of 229 bp in length between −201 and +28
of the csgD promoter (Fig. 6). The contact sites of these five factors overlap in various combinations within this 229 bp sequence. In particular, three factors bind at the upstream hot-spot I between −188 and −159 and five factors at the downstream hot-spot II between −61 and −43. Thus, interplay between these factors, each monitoring one specific environmental factor or condition, should influence csgD expression, ultimately leading to control biofilm formation. As an initial attempt to elucidate this, we analysed pair-wise interplay between the five transcription factors. In addition to the competition between positive and negative factors, we found unexpected cooperation between two negative factors, CpxR and H-NS, and between two pairs of positive factors, OmpR and IHF, and RstA and IHF.

**Competition between positive and negative factors.** The two nucleoid proteins, IHF and H-NS, exhibited opposite effects on transcription from the csgD promoter. As noted above, IHF acts as a positive regulator while H-NS is a potent repressor. IHF binds at two regions, one between −188 and −159 (IHF-1) and another between −96 and −37 (IHF-2), including the transcription-factor-binding hot-spots I and II, respectively. Since the IHF-2 region is almost twice the length of IHF-1, two molecules of IHF should bind on the IHF-2 site. On the other hand, H-NS binds at a long sequence, possibly including four sites, covering almost the entire csgD promoter from −201 to +28 (see Fig. 6). We thus examined possible competition between the two nucleoid proteins using two DNA fragment probes, CD1 (−198 to −140; including hot-spot I) and CD3 (−96 to −37; including hot-spot II), covering the upstream (IHF-1) and downstream (IHF-2) IHF-binding site, respectively (Fig. 7c). [Note that in Salmonella, only one IHF-binding site exists on the csgD promoter (Gerstel et al., 2003).]

When the amount of H-NS added was increased in the presence of a saturating amount of IHF, H-NS was able to bind by replacing IHF sites (Fig. 7a, b, lanes 8–10). On the other hand, the addition of increasing amounts of IHF in the presence of a high concentration of H-NS to saturate the CD1 and CD3 probes did not change the gel pattern (Fig. 7a, b, lanes 11–13). These results indicate that the two nucleoid proteins, IHF and H-NS, compete with each other for binding to the respective target sequences, and in the presence of high concentrations of both proteins, only the H-NS complexes are formed because of the higher affinity of H-NS than IHF. However, when the 400 bp probe CD6, including both the IHF-1 and the IHF-2 sites, was used, IHF–H-NS–DNA ternary complexes were detected (data not shown), implying that IHF was not ejected even in the presence of high concentrations of H-NS. The simultaneous binding of two molecules of IHF to both the IHF-1 and the IHF-2 sites must strengthen their DNA-binding affinity, possibly through IHF–IHF interaction, thereby leading to resistance against competition from H-NS. This finding suggests that the intracellular concentrations of IHF and H-NS influence csgD transcription.

In the exponential growth phase, the intracellular concentration of H-NS and IHF is similar, but upon entry into...
stationary phase, the IHF level increases 2.5–3-fold while the H-NS level stays almost constant (see Fig. 3). As a result, the increased amount of IHF bound on the csgD promoter should lead to cancellation of the H-NS-mediated silencing of csgD in stationary phase (Fig. 7d).

**Cooperation between the two negative factors CpxR and H-NS.** csgD mRNA markedly increased in a mutant lacking CpxR (see Fig. 1), in agreement with its negative regulator model (Jubelin *et al.*, 2005). Since both CpxR and H-NS bind to wide and overlapping regions, −190 and −15 for CpxR and −201 and +28 for H-NS, we used the long probe CD6 (−335 to +67; see Fig. 6) for the gel shift assay. Both CpxR and H-NS alone formed multiple CD6 complexes, in agreement with the presence of multiple binding sites for both factors (Fig. 8a, lanes 2–12). In the simultaneous presence of both CpxR and H-NS, the DNA complexes were supershifted (Fig. 8a, lanes 13–19), indicating that both these factors are able to bind simultaneously on the CD6 probe even though the binding sites mostly overlap. This finding implies cooperative repression of the csgD promoter by the two negative factors CpxR and H-NS.

**Cooperation between the two positive factors OmpR and IHF.** The binding site of the positive regulator OmpR on the csgD promoter (−62 to −39) within hot-spot II overlaps with the promoter-proximal IHF-2 site (−96 to −37). When CD6 was used as a probe, both OmpR and IHF bound to the same probe (Fig. 9b), as expected from the map of binding sites (Fig. 9a). However, even when the gel shift assay was performed using the shorter probe CD3, the complex band was again supershifted to high-molecular-mass complexes (Fig. 9c), indicating that both OmpR and IHF were capable of binding simultaneously to

---

**Fig. 8.** Cooperative binding of the two repressors CpxR and H-NS to the csgD promoter. (a) The gel shift assay was carried out using CpxR and H-NS with FITC-labelled CD6 probe DNA in the presence of acetyl phosphate (10 mM). Proteins used were: CpxR alone (lanes 1–5, 0, 5, 10, 15 and 20 pmol); H-NS alone (lanes 6–12, 10, 15, 20, 25, 30, 35 and 40 pmol); H-NS plus CpxR (lanes 13–17 as in lanes 6–12) plus CpxR (20 pmol). (b) Model of cooperation between CpxR and H-NS in binding to the csgD promoter. Both CpxR and H-NS are able to bind to the csgD promoter simultaneously.

**Fig. 9.** Cooperative binding of the two activators IHF and OmpR to the csgD promoter. (a) DNase I footprinting analysis of the csgD promoter by IHF and OmpRp. Fluorescently labelled DNA probe of the csgD promoter CD1 fragment (1 pmol) was incubated with increasing concentrations of purified IHF (lanes 1 and 2, 0 pmol; lane 3, 2.5 pmol; lane 4, 5.0 pmol; lanes 5 and 6, 10 pmol) in the absence (lane 1, 6) or presence (lanes 2–5) of OmpRp (80 pmol) and subjected to DNase I footprinting assays as described in Methods. The bars on the right indicate the IHF- and OmpR-binding sequences. The nucleotide numbers represent the distance from the transcription initiation site P1. (b) Gel shift assay in the presence of both IHF and OmpR. Fluorescently labelled csgD CD6 probe was incubated at 37 °C for 10 min with IHF (lanes 1 and 2, 0 pmol; lane 3 and 4, 1.0 pmol; lanes 5 and 6, 1.5 pmol; lanes 7 and 8, 2.0 pmol; lanes 9 and 10, 2.5 pmol; lanes 11 and 12, 3.0 pmol), and then 30 pmol OmpRp (lane 2, 4, 6, 8, 10, 12) was added together with 10 mM acetyl phosphate and incubated at 37 °C for 20 min. The protein/DNA mixtures were directly subjected to PAGE. (c) Gel shift assay in the presence of both IHF and OmpR. FITC-labelled CD3 probe DNA was incubated with IHF and/or OmpRp at 37 °C for 10 min. Proteins used were: IHF alone (lanes 1–4, 0, 0.625, 1.25 and 2.5 pmol); OmpRp alone (lanes 5–7, 10, 20 and 40 pmol); OmpRp (lanes 8–10 as in lanes 5–7) plus IHF (2.5 pmol). (d) Model of cooperation between OmpR and IHF in binding to the csgD promoter. Both OmpR and IHF are able to bind to the IHF-2 site simultaneously.
the short CD3 probe (Fig. 9c) even though the OmpR-binding site and the downstream IHF-binding site IHF-2 completely overlap on CD3. Since there are two IHF-binding sites within IHF-2, the CD3 probe formed multiple species of OmpR–IHF–CD3 complex, e.g. IHF–OmpR–CD3 and IHF–IHF–OmpR–CD3, should be formed in the simultaneous presence of two factors, thereby giving a broad distribution of complexes.

**Cooperation between the two positive factors RstA and IHF.** The binding site of the positive regulator RstA on the csgD promoter (−61 to −43) also overlaps with the promoter-proximal IHF-2 (−96 to −37) (see Fig. 6). When CD6 was used as a probe, both RstA and IHF bound to the probe (Fig. 10a) as expected from the map of binding sites. When the gel shift assay was performed using the shorter CD3 probe, the complex band was again supershifted to high-molecular-mass complexes (Fig. 10b), indicating that both RstA and IHF are capable of binding sites. When the gel shift assay was performed using RstA and IHF with FITC-labelled CD6 probe DNA in the presence of acetyl phosphate (10 mM). Proteins used were: RstA alone (lanes 1–4, 0.625, 1.25 and 2.5 pmol); IHF alone (lanes 5–7, 0.625, 1.25 and 2.5 pmol); RstA (lanes 8–10 as in lanes 2–4) plus IHF (2.5 pmol). (b) The gel shift assay was carried out using RstA and IHF with FITC-labelled CD3 probe DNA in the presence of acetyl phosphate (10 mM). Proteins used were: IHF alone (lanes 2–4, 0.625, 1.25 and 2.5 pmol); RstAp alone (lanes 5–7, 10, 20 and 40 pmol); RstA (lanes 8–10 as in lanes 5–7) plus IHF (2.5 pmol). (c) Model of cooperation between RstA and IHF in binding to the csgD promoter. Both RstA and IHF are able to bind to the IHF-2 site simultaneously.

**Fig. 10.** Cooperative binding of the two activators RstA and IHF to the csgD promoter. (a) The gel shift assay was carried out using RstA and IHF with FITC-labelled CD6 probe DNA in the presence of acetyl phosphate (10 mM). Proteins used were: RstA alone (lanes 1–4, 0, 5, 10 and 20 pmol); IHF alone (lanes 5–7, 0.625, 1.25 and 2.5 pmol); RstA (lanes 8–10 as in lanes 2–4) plus IHF (2.5 pmol). (b) The gel shift assay was carried out using RstA and IHF with FITC-labelled CD3 probe DNA in the presence of acetyl phosphate (10 mM). Proteins used were: IHF alone (lanes 2–4, 0.625, 1.25 and 2.5 pmol); RstAp alone (lanes 5–7, 10, 20 and 40 pmol); RstA (lanes 8–10 as in lanes 5–7) plus IHF (2.5 pmol). (c) Model of cooperation between RstA and IHF in binding to the csgD promoter. Both RstA and IHF are able to bind to the IHF-2 site simultaneously.

IHF–RstA–CD3 and IHF–IHF–RstA–CD3 complexes as in the case of OmpR–IHF interplay (see Fig. 9).

**Multi-factor promoters in prokaryotes**

Biofilm is formed under various stress conditions. Reflecting such situations in nature, the promoter for the master regulator CsgD is under the control of a number of transcription factors. Here we identified that five regulatory proteins (CpxR, OmpR, RstA, IHF and H-NS) bind within a narrow region of about 229 bp along the csgD promoter (see Fig. 6). Northern blot analysis of csgD mRNA indicated that these factors could be classified into two groups: three positive factors (OmpR, IHF and RstA) and two negative factors (CpxR and H-NS) (see Fig. 1). Of the three positive factors, RstA functions as a positive factor of csgD transcription only under acidic conditions (see Fig. 3). Interestingly, there are two hot-spots for regulator binding: three factors (IHF, CpxR and H-NS) bind at hot-spot I between −188 and −159, and all five factors bind at hot-spot II between −61 and −43 (see Fig. 6a). The simultaneous binding of multiple regulators on the csgD promoter presumably through protein–protein interaction should increase the binding affinity of each factor to the respective target sequence and in addition, should influence direct protein–protein contact with RNA polymerase αCTD or αCTD (Ishihama, 1993; Busby & Ebright, 1999) for amplification of regulatory functions. Judging from the number of transcription factors involved in regulation, the csgD promoter is the most complex promoter so far analysed in *E. coli* (Ishihama, 2010).

All the nucleoid proteins are bifunctional, playing both architectural and regulatory roles (Ishihama, 2009). Within the 200 bp sequence upstream from the csgD promoter, IHF binds at least at two sites (see Fig. 6). The functional IHF protomer is a heterodimer composed of IHFz (encoded by *ihfA*) and IHFβ (encoded by *ihfB*), each being interwound to form a compact body. IHF recognizes specific sequences for binding and facilitates DNA bending with the bend angle of about 160° as analysed by crystallography (Rice *et al.*, 1996), thereby affecting transcription. The binding of IHF to the csgD promoterproximal site IHF-2 should activate its transcription through protein–protein contacts with either αCTD or αCTD (Ishihama, 1993). IHF also binds at promoter-distal site IHF-1 (see Fig. 6). The activation of the csgD promoter by IHF associated with site IHF-1 might be attributable to a conformational change of the upstream segment of the csgD promoter, for instance through protein–protein interaction between site-1- and site-2-associated IHF molecules.

On the other hand, H-NS is a global repressor of transcription, sometimes called a silencer (Dorman, 2004; Ishihama, 2009). H-NS employs a mechanism of DNA recognition different from that of IHF. One H-NS dimer binds to each 15–20 bp (Amit *et al.*, 2003) and bridges adjacent tracts of DNA, resulting in lateral condensation of
large regions and formation of large globular structures (Dame et al., 2005). The effect of H-NS on gene expression is direct, being mediated by binding of H-NS and changing the DNA topology of the promoter region, functioning as a transcription silencer through interference with the formation of RNA polymerase–promoter complex (Navarre et al., 2006). H-NS binds to the AT-rich sequence within the 200 bp sequence upstream from the csgD promoter, and interferes with the binding of RNA polymerase (data not shown). The level of csgD silencing may be related to the number and location of H-NS binding sites among the four potential binding sites along the csgD promoter. Thus, the silencing of csgD by H-NS should be reduced in stationary phase because of the increase in IHF level (see Fig. 2).

Conclusions

More than 10 transcription factors are involved in regulation of the csgD promoter of the master regulator operon for biofilm formation. Judging from the number of transcription factors involved in regulation, the csgD promoter can be categorized as one of the most complex E. coli promoters. Each transcription factor monitors a specific factor or condition in nature for the control of biofilm formation. Based on the in vitro binding assays, we propose not only competition between the positive and negative factors but also cooperation within the positive and negative factor groups. Detailed understanding of the csgD promoter regulation system should provide a better understanding of the genome regulation for survival of E. coli in nature.

ACKNOWLEDGEMENTS

This work was supported by a Grant-in-Aid for Scientific Research on Priority Area System Cell Engineering by Multi-Scale Manipulation (17076016) to A.I. and Grants-in-Aid for Scientific Research (A) (21241047) and (B) (18310133) to A.I. and Grant-in-Aid for JSPS Fellows (218850) to H.O. from the Ministry of Education, Culture, Sports, Science and Technology of Japan. We also acknowledge the support from Project of Micro-Nano Technology Research Center of Hosei University. H.O. is a recipient of a JSPS Post-doctoral Fellowship.

REFERENCES


Gerstei, U. & Romling, U. (2001). Oxygen tension and nutrient starvation are major signals that regulate agfD promoter activity and


Edited by: V. Sperandio