Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in Escherichia coli

Nicole Sommerfeldt, Alexandra Possling, Gisela Becker, Christina Pesavento, Natalia Tschowri and Regine Hengge

Institut für Biologie – Mikrobiologie, Freie Universität Berlin, 14195 Berlin, Germany

INTRODUCTION

Bacteria can switch between a single-cell planktonic and motile lifestyle and a sessile multicellular state, i.e. a biofilm existence (Beloin et al., 2008). The ubiquitous bacterial second messenger bis-(3′-5′)-cyclic-diguanosine monophosphate (c-di-GMP) is a key regulatory factor that interferes with motility and stimulates adhesion and biofilm matrix production (for recent comprehensive reviews, see Jenal & Malone, 2006; Ro¨mling et al., 2005; Paul et al., 2006a; Tamayo et al., 2007; Wolfe & Visick, 2008). Cellular c-di-GMP levels are controlled by antagonistically acting diguanylate cyclases (DGCs) and phosphodiesterases (PDEs). DGC activity in these enzymes is provided by GGDEF domains (Hickman et al., 2005; Paul et al., 2004; Ryjenkov et al., 2005), whereas PDE activity resides in EAL or HD-GYP domains (Chang et al., 2001; Christen et al., 2005; Ryan et al., 2006b; Schmidt et al., 2005; Tamayo et al., 2005). Active DGCs are characterized by an intact GGDEF motif that represents the active centre of the enzyme (called the A-site) and usually also carry an I-site, i.e. a secondary and inhibitory binding site for c-di-GMP (Christen et al., 2005; Malone et al., 2007). Also, in EAL domains with PDE activity, a number of functionally important amino acids are highly conserved (Rao et al., 2008; Schmidt et al., 2005). Notably, many species also contain GGDEF/EAL domain proteins, in which these signatures are degenerate, and which therefore cannot act as DGCs or PDEs, but nevertheless generate phenotypes (Rao et al., 2008; Suzuki et al., 2006). GGDEF and EAL domains can occur alone or in combination in single polypeptides. The majority of these proteins also carry N-terminal sensory-input domains, which can control the activities of the GGDEF and/or EAL domains. C-di-GMP is bound by specific effector components, which can be proteins belonging to different families (Amikam & Galperin, 2006; Benach et al., 2007; Christen et al., 2007; Hickman & Harwood, 2008; Lee et al., 2007; Merighi et al., 2007; Ryjenkov et al., 2006) or RNA molecules acting as riboswitches (Sudarsan et al., 2008). These effectors control various target processes, including transcription initiation,

Abbreviations: c-di-GMP, bis-(3′-5′)-cyclic-diguanosine monophosphate; DGC, diguanylate cyclase; PDE, phosphodiesterase; RNAP, RNA polymerase.

Two supplementary tables, listing designations of homologous GGDEF/EAL genes in Escherichia coli and Salmonella enterica (serovar Typhimurium), and oligonucleotide primers used, are available with the online version of this paper.
enzyme activity and flagellar motor function (see reviews cited above).

While this overall mechanism of c-di-GMP control and action seems straightforward, current knowledge is rather fragmentary when it comes to the precise molecular details. Moreover, there is a striking multiplicity of GGDEF/EAL proteins in single species. In particular, proteobacteria often have several dozen such proteins. This immediately raises questions: can different GGDEF/EAL proteins act independently from one another and thereby affect different targets, or do all these proteins control a common cellular pool of diffusible c-di-GMP and therefore a common set of targets? How is specificity of signalling achieved, i.e. potential crosstalk avoided, with so many proteins that make and break c-di-GMP? Is there differential expression of genes encoding GGDEF/EAL domain proteins? Also, in order to identify the output functions of distinct GGDEF/EAL domain proteins, a knowledge of their conditions of expression is essential.

Here, we present the first systematic study, to our knowledge, of the expression of all genes encoding GGDEF/EAL domain proteins in a bacterial species. Depending on the strain, Escherichia coli K-12 has 28 or 29 such genes, i.e. a number that can still be handled in toto (in terms of making reporter fusions, mutations and functional assays for each of them in parallel). Moreover, E. coli exhibits a growth-phase-dependent lifestyle switch: during the post-exponential phase of the growth cycle (in complex medium), cells express flagella and become highly motile. Yet later, during entry into stationary phase, motility is downregulated (Adler & Templeton, 1967; Amsler et al., 1993; Pesavento et al., 2008) and the expression of adhesive curli fimbriae is induced, which results in autoaggregation and adhesion to surfaces, an important early step in biofilm formation (Olsson et al., 1989; Pesavento et al., 2008; Weber et al., 2006). Both in E. coli and its close relative Salmonella, curli expression is known to be controlled by the general stress response and stationary phase sigma factor σ^σ^ (RpoS) and by c-di-GMP (Kader et al., 2006; Pesavento et al., 2008; Romling et al., 1998; Simm et al., 2007; Weber et al., 2006). Here, we show highly differential expression of all genes encoding GGDEF/EAL domain proteins in E. coli as a function of parameters relevant for switching from motility to curli-mediated adhesion, i.e. growth phase, temperature and growth on liquid or solid medium. We also demonstrate that the majority of expressed GGDEF/EAL domain-encoding genes in E. coli are under the control of σ_P, and provide evidence that six GGDEF/EAL genes differentially affect curli expression, whereas several others do not have such an influence, despite clear expression.

METHODS

Bacterial strains and growth conditions. All strains used in this study are derivatives of the E. coli K-12 strain W3110 (Hayashi et al., 2006) and also carry a ΔΔΔ deletion. Mutations were transferred by P1 transduction (Miller, 1972). The following deletion–insertion muta-

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Wed, 24 Jul 2019 08:00:33

ions have been described previously: rpoS539::Tn10 (Lange & Hengge-Aronis, 1991); yaiC::kan, ydaM::cat, yeeC::kan, yedQ::cat, yedV::Tns, mllA::kan, esgB::cat (Weber et al., 2006); yegE::kan, yeaf::kan, yehH::cat (Pesavento et al., 2008). All other mutations are deletion–insertion mutations generated by one-step inactivation according to Datsenko & Wanner (2000), using the primers listed in Supplementary Table S1 and pKD3 (cat cassette), pKD4 or pKD13 (kan cassettes) for the PCRs. In cases in which insertions could have a polar effect on downstream genes (for operon structures of certain genes encoding GGDEF/EAL proteins, see below) or insertions in multiple genes had to be combined, non-polar in-frame deletion mutations were also obtained by flipping out the insertion cassettes as described by Datsenko & Wanner (2000).

Cells were grown at 28 or 37 °C under aeration (in liquid medium if not otherwise indicated). The medium used was Luria–Bertani (LB) broth (Miller, 1972). Antibiotics were added as recommended by Miller (1972). Growth was monitored by measuring OD_{578}.

Construction of lacZ reporter fusions and their transfer into the chromosome. Single-copy lacZ reporter fusions to the following E. coli GGDEF/EAL genes have been described previously: yaiC, ydaM, yedV, yeaC, yedQ, yciR and yedV (Weber et al., 2006), and yegE and yehH (Pesavento et al., 2008). The primers used to construct all other lacZ fusions are listed in Supplementary Table S1. The appropriate PCR fragments (depending on the chromosomal context of the specific genes) were cloned into the lacZ fusion vector pLL28, as previously described (Weber et al., 2006). All fusions were translational fusions that included approximately 10 codons of the corresponding ORF (also making sure that no complete signal sequences were present in the fusion proteins) and contained at least 300 bp upstream of the coding region (standard ‘short fusions’). In cases (yeaf–yeal, yfiR, yliE–yliF, yedQ) in which a gene could be part of an operon and would then not be the promoter-proximal gene, additional lacZ fusions were generated, which included the entire promoter-proximal gene(s) as well as at least 300 bp of the non-coding upstream region (‘long fusions’). For yeaf–yeal expression, the expression of the respective ‘short’ and ‘long’ fusions was similar, indicating that these genes are monocistronic; the ‘long’ yfiR::lacZ (including the upstream gene yfiR) showed weak activity, whereas the ‘short’ fusion seemed inactive, i.e. yfiR seemed to be the second gene in an operon; the lacZ fusions in yegE (also containing yegC and yegH, which are located upstream of yegE in a putative operon; upstream of yehH is a putative transcriptional terminator) as well as those in yehF and yedQ did not exhibit significant expression, i.e. it remains to be shown whether these genes are part of operons expressed under other conditions.

All reporter fusions were transferred to the att(b) location of the chromosome via phage λRS45 or λRS74 (Simons et al., 1987). Single lysogeny was tested by a PCR approach (Powell et al., 1994).

SDS-PAGE and immunoblot analysis. Sample preparation for SDS-PAGE and immunoblot analysis were performed as described previously (Lange & Hengge-Aronis, 1994). Cellular protein, 5 or 10 µg per lane, was applied. Polyclonal sera against σ^σ^ and CsgD (Weber et al., 2006), a goat anti-rabbit IgG alkaline phosphatase conjugate (Sigma) and a chromogenic substrate (BCIP/NBT, Boehringer Mannheim) were used.

Northern blot analysis. For RNA preparation and Northern blot analysis of csgD mRNA, cells were grown in LB medium at 28 °C and harvested at OD_{578} 4.0. The procedure and materials used were exactly as described previously (Weber et al., 2006).

β-Galactosidase assay. β-Galactosidase activity was assayed by use of ONPG as a substrate and is reported as µmol o-nitrophenol min^{-1}(mg cellular protein)^{-1} (Miller, 1972). Experiments examining the expression of lacZ fusions during the entire growth cycle were done at least twice, and a representative experiment is shown for determin-
ing β-galactosidase activities in cells grown on solid medium, cells were grown overnight in patches on LB agar plates and cells growing in the central parts of the patches were resuspended in Z-buffer (Miller, 1972). OD₅₇₈ was determined and measurements were performed as with cells grown in liquid medium.

**Bacterial motility assay.** Motility was tested on swim plates containing 0.5% bacto-tryptone, 0.5% NaCl and 0.3% agar. A 3 μl volume of overnight culture (adjusted to OD₅₇₈ 4.0 in its own supernatant) was inoculated into the swim plates and cells were allowed to grow and swim for 4-6 h at the temperature indicated.

**RESULTS**

**GGDEF/EAL genes in *E. coli* and the construction of lacZ reporter fusions**

*E. coli* K-12 has 29 genes, which encode 12 proteins with GGDEF domains, 10 proteins with EAL domains and seven proteins that feature both domains. Some *E. coli* K-12 strains lack one of the EAL-only proteins [YahA; this also turned out to be the case for the W3110 strain used in our study. For a list of the 28 GGDEF/EAL genes present, see Table 1. Note that the order of GGDEF/EAL genes in all composite figures in this study is always the same as in Table 1, i.e. GGDEF-only, GGDEF+EAL and EAL-only genes (and within these subsets, in alphabetical order). As the physiological functions of most of these genes is still unknown, they carry y-designations (e.g. yaiC, ydaM, etc.; the only exception is rtn). We continue to use these y-designations as we think that currently at least there is no unifying basis for renaming these genes: (i) DGC and/or PDE activities of purified proteins in vitro have been determined for only a small minority of these proteins (Pesavento et al., 2008; Ryjenkov et al., 2005; Schmidt et al., 2005; Weber et al., 2006); (ii) whether the 'composite' GGDEF+EAL proteins have DGC and/or PDE activities (or none at all) is not unequivocally predictable; and (iii) in some proteins (YeaI, YhdA, YcgF, YdiV) the GGDEF

<table>
<thead>
<tr>
<th>Gene</th>
<th>b-Number</th>
<th>N-terminal domains</th>
<th>GGDEF domain:</th>
<th>EAL domain:</th>
<th>catalysis in LB</th>
<th>highest expression at temperature growth phase</th>
<th>regulated by α⁣²¹①</th>
</tr>
</thead>
<tbody>
<tr>
<td>yaaC</td>
<td>b2658</td>
<td>PAS</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>pos</td>
</tr>
<tr>
<td>ycdt</td>
<td>b1625</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>ydat</td>
<td>b1341</td>
<td>?-PAS</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>pos</td>
</tr>
<tr>
<td>yddv</td>
<td>b1490</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>ydeh</td>
<td>b1535</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>neg</td>
</tr>
<tr>
<td>yeaI</td>
<td>b1785</td>
<td>?</td>
<td>+++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>yecj</td>
<td>b1786</td>
<td>?</td>
<td>+++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>yecp</td>
<td>b1794</td>
<td>GFP</td>
<td>+++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>neg</td>
</tr>
<tr>
<td>yedQ</td>
<td>b1956</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>pos</td>
</tr>
<tr>
<td>yflN</td>
<td>b2664</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>yflP</td>
<td>b2634</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
<tr>
<td>yneP</td>
<td>b1522</td>
<td>?</td>
<td>++</td>
<td>++</td>
<td>28°C</td>
<td>SF</td>
<td>-</td>
</tr>
</tbody>
</table>

1I-site, inhibitory binding site for c-di-GMP; A-site, GTP-binding and enzyme active site (consensus is GGDEF or GGEEF).

According to Rao et al. (2008), the EAL motif is involved in coordinating Mg²⁺, four amino acids (Q, R, D, D) dispersed along the primary sequence are required for c-di-GMP binding (the R is the one that follows the EAL motif with one non-conserved amino acid in between), and two amino acids (T, E) are directly involved in catalysis. Note that also the proteins in which most or even all of these functionally relevant amino acids are not present, are clearly identified as EAL domains by the BLAST algorithm.

For expression data, see Fig. 1; ++++, +++, ++, +, maximal specific β-galactosidase activity measured was >0.1, >0.01 or >0.002 μmol min⁻¹ (mg total cellular protein)⁻¹, respectively.

LP, exponential growth phase; SP, stationary phase; if a gene exhibits predominant (but not exclusive) expression at one temperature or in one growth phase, this is highlighted in bold type.

Protein contains transmembrane domains.
and/or EAL motifs (which essentially contribute to the enzyme activity) are degenerate (Table 1), and therefore these proteins are highly unlikely to synthesize or degrade c-di-GMP.

In order to study the regulation of expression of all genes encoding GGDEF/EAL domain proteins in *E. coli* K-12, we generated single-copy lacZ reporter fusions to all of them. These reporter fusions mainly reflect transcription of these genes, but also contain the translational start site and approximately 10 codons of the respective genes fused to lacZ, and therefore also take into account putative differences in translation initiation efficiencies of the respective proteins. As all reporter fusions were constructed in the same way, the β-galactosidase activities measured allowed not only monitoring of the dynamics of expression of specific single genes, but also a rough comparison of expression levels among such genes. The length of upstream DNA present and fusion construction in cases in which GGDEF/EAL genes could be part of an operon (yeaI-yeaJ, yfnN, yliE-yliF, yneF) are described in detail in Methods. In the single case in which an operon structure had been clearly demonstrated earlier, i.e. for yddV-yddU (Méndez-Ortiz et al., 2006), we studied a yddV::lacZ fusion alone. Overall, at least 22 of the 28 GGDEF/EAL genes in *E. coli* K-12 seem to be expressed as single-gene transcriptional units (for details, see Methods).

**Expression patterns of all *E. coli* genes encoding GGDEF/EAL domain proteins as a function of growth phase and temperature**

For both *E. coli* and *Salmonella*, certain GGDEF/EAL proteins have been implicated in the control of motility and the expression of adhesive fimbriae. As motility genes are transiently expressed during the post-exponential phase of the growth cycle and curli expression occurs during entry into stationary phase, we monitored the expression of the lacZ reporter fusions to all GGDEF/EAL domain-encoding genes as a function of growth phase (Fig. 1). Moreover, as curli expression specifically occurs at temperatures below 30 °C (Boedeker et al., 1989), we compared the expression patterns of all fusions at 37 and 28 °C (shown in Fig. 1a and b, respectively).

First of all, the majority of the *E. coli* GGDEF/EAL domain-encoding genes were indeed expressed (21/28). Fusions to only seven genes, i.e. yeaI, yliF, yneF, yfgF, ycgG, ydiV and yliE, exhibited baseline β-galactosidase levels [0.001–0.002 μmol min⁻¹ (mg total cellular protein)⁻¹], see summarized data in Table 1]. The other genes showed a wide range of expression levels with more than two orders of magnitude difference. At the lower end, there was yhDA, which nevertheless produced a clear phenotype when knocked out (see below). Among the most strongly expressed genes were yeaI and yddV (both encoding GGDEF-only proteins), as well as yhjH (encoding an EAL-only protein). Even these genes, however, belong to the rather moderately expressed genes in *E. coli*, as judged by comparison with the highly expressed genes seen by microarray analysis as well as by the assay of similarly constructed lacZ fusions (our unpublished data).

Expression patterns of GGDEF/EAL domain-encoding genes as a function of growth phase were very different, with two major types becoming apparent (Fig. 1a, b, and summary in Table 1): (i) genes predominantly expressed during exponential or post-exponential phase (showing at least constant or increasing β-galactosidase activities in the growing cultures); and (ii) genes that exhibited induction during entry into stationary phase (showing increasing β-galactosidase activities, while culture growth slowed down and finally stopped). Representatives of the first group included yeaI, yeaP, yfeA and yhjH. Genes induced during transition into stationary phase included ydaM, yddV, yciR, yegE, yicC, ylaB and yoaD. Especially at 28 °C (Fig. 1b), at which cells grow more slowly, differences in the timing of induction during entry into stationary phase also became apparent: some genes were activated relatively early during entry into stationary phase, e.g. ydaM, yddV and yegE, whereas yaiC and yoaD were induced significantly later (with yaiC also showing expression at 28 °C alone).

At least eight of the 21 expressed GGDEF/EAL genes exhibited differential temperature regulation (compare patterns in Fig. 1a and b, summary in Table 1). While seven of these genes were increasingly or even exclusively expressed at 28 °C (yaiC, ydaM, yddV/yddU, yhjK, yfgF and yoaD), a single gene (yeaI) was slightly preferentially expressed at 37 °C. Thus, the overall tendency for temperature regulation of GGDEF/EAL domain-encoding genes is towards increased expression at temperatures below 37 °C.

**The majority of expressed genes encoding GGDEF/EAL domain proteins are under the control of the general stress sigma factor σ^S^**

Stationary phase induction of gene expression suggested regulation by the general stress and stationary phase sigma factor σ^S^, Indeed, σ^S^ control has been previously observed for a subset of *E. coli* GGDEF/EAL domain-encoding genes (although in a different strain background; Weber et al., 2006). Therefore, the systematic analysis of expression patterns of all our reporter fusions at both temperatures was also performed in a σ^S^-deficient rpoS mutant background (Fig. 1; data represented by circles in all panels). In practically all cases in which a GGDEF/EAL domain-encoding gene was clearly activated during entry into stationary phase, the increase in expression was σ^S^-dependent (yaiC, ydaM, yddV/yddU, yedQ, yciR, yegE, yicC, ylaB and yoaD). Thus, 10 out of the 21 expressed GGDEF/EAL domain-encoding genes were under the significant positive control of σ^S^.

Interestingly, several genes predominantly expressed during the exponential or post-exponential phase appeared to be negatively σ^S^-regulated. This was most apparent for
GGDEF/EAL genes in E. coli

Specific β-galactosidase activity (μmol min⁻¹ mg⁻¹)

OD₅₇₈

Time (h)
ydeH (encoding a GGDEF-only protein), which in the absence of $\sigma^S$ revealed a potential for stationary phase induction or inverse growth-rate regulation that remained cryptic in the wild-type background (at 37 °C; circles in Fig. 1 represent the rpoS mutant data). At 37 °C, yfeA, rtn and yhjH also exhibited higher expression in the rpoS mutant (all three encode proteins with EAL domains). This negative control by $\sigma^S$ became more pronounced when cells were grown on solid rather than in liquid medium (Fig. 2; in addition to the aforementioned genes, ycgF also

![Graph](image-url)

**Fig. 1.** Expression of GGDEF/EAL genes of *E. coli* during the growth cycle in liquid LB medium. Derivatives of strain W3110 carrying single-copy lacZ fusions in the GGDEF/EAL genes indicated in the respective panels were grown in LB at 37 °C (a) or 28 °C (b). For all fusions, otherwise isogenic rpoS+ (squares) and rpoS::Tn10 (circles) strains were tested. OD_{578} (open symbols) and specific $\beta$-galactosidase activities (filled symbols) were determined along the growth curve. ON, overnight culture (24 h after starting the experiment).

**Fig. 2.** Expression of GGDEF/EAL genes of *E. coli* after overnight growth on LB plates. The same lacZ fusion strains as used in Fig. 1 were grown overnight in patches on LB plates incubated at 28 °C (a) or 37 °C (b). Cells from the central part of the patches were scratched from the agar plates and resuspended, and OD_{578} and specific $\beta$-galactosidase activities (represented by the bars) were determined. As in Fig. 1, activities of all fusions were determined in otherwise isogenic rpoS+ (black bars) and rpoS::Tn10 (white bars) backgrounds. The entire experiment was done three times independently and mean activity ± SD is shown.
showed enhanced expression in the rpoS mutant under these conditions).

We conclude that σ5 directs a major reorganization of the expression of genes encoding GGDEF/EAL proteins during entry into stationary phase. While downregulated genes mainly encode EAL proteins, more strongly expressed genes encode both DGCs and PDEs (both proven and putative).

### Differential expression of GGDEF/EAL domain-encoding genes on liquid and solid medium

A bacterial colony on an agar surface represents a biofilm at a solid wet medium–air interface. Curli fimbriae and cellulose as a biofilm matrix component contribute to the highly structured morphology of such colony biofilms and the formation of these components is stimulated by high cellular c-di-GMP levels (Römling, 2005). We therefore tested whether any of the E. coli genes encoding GGDEF/EAL proteins exhibited altered expression during growth on agar plates in comparison with liquid medium (Fig. 2). Four genes were more strongly expressed on solid medium (compare data in Fig. 2 and Fig. 1): (i) yaiC (sixfold higher expression on plates incubated at 28 °C; again no expression at 37 °C); (ii) ydaM (twofold higher expression on plates at 28 °C, but tenfold higher expression on plates at 37 °C; also, expression was now similar at 28 and 37 °C, i.e. temperature regulation of ydaM could be seen only in liquid medium); (iii) yciR (twofold higher expression on plates at 28 °C, no difference at 37 °C); and (iv) yoaD (fivefold higher expression on plates at 28 °C, twofold only at 37 °C). Only a single gene, yhjH, was downregulated during growth on solid medium (fourfold at both 28 and 37 °C). This differential regulation becomes particularly significant when seen against the background of all the other GGDEF/EAL genes, which showed remarkably similar expression levels in plate-grown colonies (Fig. 2) and during entry into stationary phase in liquid medium (Fig. 1).

In addition, growth on solid medium was characterized by a clear dominance of GGDEF proteins over EAL proteins (especially at 28 °C, the GGDEF-only proteins YaiC, YdaM, YddV and YeaJ were the major proteins expressed from the entire GGDEF/EAL family; see Fig. 2a). This suggests increased total DCG activity and therefore c-di-GMP accumulation when cells grow on a solid agar surface.

### Analysis of knockout mutations in all GGDEF/EAL domain-encoding genes: six genes are differentially involved in the control of the biofilm regulator CsgD and curli fimbriae expression

Our knowledge of the expression patterns of all GGDEF/EAL genes was now able to guide the search for their specific physiological functions. For phenotypic analyses, we isolated knockout mutations in all 28 genes encoding GGDEF/EAL proteins (for details, see Methods). Positive c-di-GMP control has been shown for the expression of adhesive curli fimbriae in E. coli and Salmonella (reviewed by Römling et al., 2005). Curli expression occurs during entry into stationary phase (at temperatures below 30 °C). It is strictly dependent on the general stress sigma factor σ5, which activates the expression of MlrA, a transcriptional activator that cooperates with σ5-containing RNA polymerase (RNAP) to activate the expression of CsgD, which acts as an essential activator that together with vegetative RNAP turns on the expression of the curli structural operon csgBAC (Brown et al., 2001; Gerstel et al., 2003; Römling et al., 2000; Weber et al., 2006). In E. coli, YdaM and YegE (a proven and a hypothetical DGC, respectively) also positively control csgD transcription, whereas the PDEs YciR and YhjH play the opposite role (Pesavento et al., 2008; Weber et al., 2006). Our finding that a majority of GGDEF/EAL domain-encoding genes in E. coli are expressed between late-exponential phase and entry into stationary phase raises the question of whether any other genes besides ydaM, yegE, yciR and yhjH also modulate curli formation.

To address this question, we first determined expression of the curli structural operon (i.e. the output of the entire cascade) with a single-copy csgB::lacZ reporter fusion. Besides the four mutations previously observed to affect csgB::lacZ expression, mutations in yeaP and yhdA also reduced csgB::lacZ expression (Fig. 3a). The remaining 22 GGDEF/EAL knockout mutations did not affect csgB::lacZ expression under the conditions tested (data not shown). Consistent with impaired curli formation, the mutants with strongly reduced csgB::lacZ expression (ydaM, yegE and yhda) also showed reduced biofilm formation, visible as the attachment of cells in a ring-like formation when cultures were grown in rolling glass tubes (data not shown). Just like the mutations in ydaM and yegE, the yhda mutation also strongly diminished the expression of a second target gene of CsgD, yaiC (encoding a GGDEF protein required for cellulose synthesis). The yeaP mutation, however, reduced curli expression only partially (Fig. 3a) and did not affect yaiC expression (Fig. 3b).

In order to see where in the curli control cascade these mutations have an effect, we followed the accumulation of the master regulator σ5 and of the essential curli activator CsgD during entry into stationary phase by immuno blot analysis. In the wild-type background, the accumulation of σ5 is essentially complete at OD578 2, whereas induction of CsgD starts only at OD578 3 and is complete at OD578 4 and above (Fig. 4). While none of the GGDEF/EAL knockout mutations significantly affected σ5 levels, five mutations clearly altered CsgD levels and/or the kinetics of CsgD induction during entry into stationary phase (Fig. 4): besides the mutations in ydaM, yciR, yegE and yhjH (which affected CsgD levels as expected, see above), the mutation in yhdA also strongly reduced the induction of CsgD during entry into stationary phase. The yeaP mutation, however, did not significantly change CsgD accumulation (Fig. 4).
Because of its strong effect, the yhdA mutation was further analysed (Fig. 5). This mutation not only strongly reduced CsgD protein levels (Fig. 4) but also csgD mRNA levels (Fig. 5b). Besides \( \sigma^S \), at least two factors are essential for csgD transcription, the MerR-like activator MlrA and the DGC YdaM. Using lacZ fusions to the mlrA and ydaM genes, we therefore tested whether the yhdA mutation alters MlrA and/or YdaM expression. No effect was found for the mlrA::::lacZ fusion (data not shown), but ydaM::::lacZ expression was reduced in the yhdA mutant background (Fig. 5a). In the yhdA mutant, expression of the flagellar master regulator FlhDC (assayed with a translational flhDC::::lacZ fusion that also reflects post-transcriptional control) was also shut off earlier during entry into stationary phase (Fig. 5c), and the yhdA mutant showed reduced motility on swim plates (Fig. 5d). These data are in line with an earlier report that shows an indirect positive role of YhdA in FlhDC expression (Suzuki et al., 2006). The yhdA mutation was actually the only one of all our GGDEF/EAL gene knockouts that had an effect on the expression of flhDC (data not shown). We conclude that YhdA not only positively modulates motility gene expression but also is required for curli expression. In the curli control cascade, YhdA stimulates the expression of the GGDEF protein YdaM, but the finding that its effect on CsgD (both on csgD mRNA and CsgD protein; Figs 4 and 5) is stronger than its effect on YdaM expression suggests that it also plays a second, more direct role in csgD expression.

**DISCUSSION**

Expression patterns of all GGDEF/EAL genes in *E. coli* and implications for c-di-GMP-dependent regulatory scenarios

To our knowledge, our study is the first to systematically analyse the expression of all genes encoding GGDEF/EAL proteins in a bacterial species, i.e. the expression of genes whose products in general control the synthesis and degradation of the signalling molecule c-di-GMP. The majority of these genes in *E. coli* K-12 were expressed when cells were grown in liquid complex medium (21 of 28 GGDEF/EAL genes present in strain W3110), although their actual levels of expression and regulatory patterns varied greatly. Taking all data together, certain clear tendencies become apparent (Table 1): (i) there are different subsets of dominantly expressed GGDEF/EAL domain-encoding genes during different phases of the growth cycle; (ii) a majority of the expressed genes (15 of 21) are under the control of the general stress and stationary phase sigma factor \( \sigma^S \); (iii) a subset of genes exhibits stronger or even exclusive expression at reduced temperature (28 °C); and (iv) a distinct small group of genes shows higher expression during growth on solid medium in comparison with liquid medium.

Expression of GGDEF/EAL domain-encoding genes differs by more than two orders of magnitude, with just two genes dominantly expressed in growing cells: yeal, which encodes a GGDEF-only protein (with intact A- and I-site signatures, and therefore most likely DGC activity), and yhjH, which is under class 3 flagellar control and encodes

---

**Fig. 3.** Knockout mutations in several *E. coli* GGDEF/EAL genes change the expression of curli genes and the second CsgD target gene, yaiC. A W3110 derivative carrying the single-copy reporter fusions csgB::::lacZ (a) or yaiC::::lacZ (b), as well as the derivatives carrying knockout mutations in ydaM, yciR, yegE, yhjH, yhdA and yeaP, were grown in LB at 28 °C. OD578 (open symbols) and specific \( \beta \)-galactosidase activities (filled symbols) were determined along the growth curve. For symbols for the respective mutants, see the top of each panel. wt, Wild-type.
an EAL-only protein with PDE activity (Frye et al., 2006; Pesavento et al., 2008). These two proteins have been shown to antagonistically control motility in post-exponentially growing cells (at 37 °C; at 28 °C, two other putative DGCs, YegE and YedQ, substitute for YeaJ; Pesavento et al., 2008). When cells enter into stationary phase, additional GGDEF/EAL proteins are induced, i.e. the c-di-GMP control potential represented by these proteins becomes more diverse.

Consistent with these growth phase-related expression patterns, we observed regulation by $\sigma^S$ to be much more widespread among the 28 GGDEF/EAL domain-encoding genes in *E. coli* than was previously apparent when only a subset of these genes was analysed (Weber et al., 2006). Ten out of 21 expressed genes were positively $\sigma^S$-controlled and clearly stationary phase-induced. Five of the corresponding proteins have GGDEF domains with intact A- and I-site motifs, and therefore (most likely) have DGC activity (YaiC, YdaM, YddV, YedQ and YegE). The other five positively $\sigma^S$-controlled genes encode EAL-proteins (YciR, YddU, YjcC, YlaB and YoaD) with signatures consistent with PDE activity. In addition, five genes encoding GGDEF/EAL proteins were observed to be under negative control of $\sigma^S$. This negative control is most likely due to increased formation of vegetative or other alternative RNAP holoenzymes in the absence of competition by $\sigma^S$ with the other sigma factors (Gruber & Gross, 2003; Weber et al., 2005). Negative control by $\sigma^S$ was most clear for ydeH (encoding a GGDEF domain; Fig. 1a), but became more apparent also for four other genes when cells were grown on solid medium (Fig. 2). The products of the latter four genes are all EAL proteins (YfeA, Rtn, YcgF and YhjH; with YfeA also carrying an additional degenerate GGDEF domain). Interestingly, ydeH was inversely growth rate-regulated in the absence of $\sigma^S$, suggesting that it shows specific activation and a distinct function under some stress conditions that reduce growth but do not induce $\sigma^S$.

In conclusion, these data reveal an interesting overall pattern: during transition into stationary phase, *E. coli* uses $\sigma^S$ control to deploy a diverse battery of GGDEF/EAL proteins, i.e. proven or putative DGCs and PDEs, which are not expressed (or are expressed at lower levels) in rapidly growing cells. This means that an intricate control of c-di-GMP production and degradation, and therefore also of various targets that are controlled by c-di-GMP, becomes more important when cells enter into stationary phase. However, this does not automatically imply that stationary phase cells will always have increased c-di-GMP levels (although a tendency towards higher c-di-GMP levels is suggested by the downregulation of several genes for EAL proteins by $\sigma^S$). Whether c-di-GMP actually accumulates in stationary phase may very much depend on the actual conditions that are perceived by the sensory input domains of these enzymes. From what is known about stationary phase induction and the role of $\sigma^S$, c-di-GMP and temperature in curli regulation in *E. coli* and *Salmonella* (see above), it is clear that c-di-GMP levels must increase during entry into stationary phase at 28 °C. Nevertheless, it is intriguing that in parallel, $\sigma^S$-dependent EAL proteins (YciR, YddU, YjcC, YlaB and YoaD) accumulate, which represents a clear potential to rapidly break down c-di-GMP when cells are suddenly challenged with conditions under which massive curli production or other c-di-GMP-stimulated functions could become counterproductive.
Another general pattern in the control of GGDEF/EAL domain-encoding genes is temperature regulation (Fig. 1, data summarized in Table 1). Of the 21 expressed genes, seven were increasingly or even exclusively expressed at reduced temperature (28 °C; yaiC, ydam, ydv-ydU, yhjK, ycgF and yoaD). Only a single gene (yeaJ) showed somewhat higher expression at 37 °C, consistent with YeaJ being responsible for downregulation of motility at 37 °C, but not at 28 °C (Pesavento et al., 2008). The tendency for increased expression below 37 °C of many GGDEF/EAL proteins with diverse sensory input domains suggests that a multiple signal-integrating control of cellular c-di-GMP levels may be most important when *E. coli* finds itself in a highly variable environment outside the mammalian host.

Both σ5 control and temperature effects in the control of most GGDEF/EAL domain-encoding genes were even more pronounced when cells were grown on solid medium (in particular, negative regulation by σ5; compare data in Figs 1 and 2). In addition, a subset of four strongly σ5-dependent genes, ydam, yciR, yaiC and yoaD, showed increased expression on solid medium in comparison with liquid medium. Precisely these four genes play important roles in the synthesis of curli fimbriae and cellulose (see below), i.e. they are determinants for the complex 'rdar' morphology of colonies on plates (Römling, 2005). Consistently, YhjH, the PDE that maintains the low c-di-GMP levels required for motility in liquid medium (Girgis et al., 2007; Ko & Park, 2000; Pesavento et al., 2008; Ryjenkov et al., 2006; Simm et al., 2004), is downregulated on solid medium (Figs 1 and 2). Moreover, genes encoding GGDEF-only proteins (yaiC, ydaM, ydv and yeaJ) clearly become dominantly expressed in cells growing on solid medium (Fig. 2), suggesting c-di-GMP accumulation under these conditions.

**Functions of a specific subset of GGDEF/EAL domain-encoding genes in the control of aggregative curli fimbriae**

A major target of positive c-di-GMP control is the synthesis of aggregative curli fimbriae (Römling et al., 2005). In *E. coli*, the DGC YdaM and the PDE YciR represent the c-di-GMP control module that specifically and apparently exclusively regulates the transcription of csgD (Weber et al., 2006), which encodes an activator that is essential for expressing the curli structural operon csgBAC. In addition, YegE and YhjH antagonistically modulate csgD transcription, and currently it is not clear how this influence is integrated with that of YdaM/YciR (Pesavento et al., 2008). CsgD also activates yaiC (aggD in *Salmonella*), which encodes the DGC that controls cellulose synthase activity (Römling et al., 2000; Simm et al., 2007). The relevant antagonist to YaiC may be the EAL protein YoaD, which shows a similar late timing of induction to that of YaiC (see Fig. 1b) and has been implicated in downregulating cellulose production (Brombacher et al., 2006).

**Fig. 5.** The yhdA mutation affects ydam expression, csgD mRNA levels, the expression of the flagellar master regulator FlhDC and motility. W3110 derivatives carrying the single-copy fusions ydam::lacZ (a, squares) or flhDC::lacZ (c, squares), as well as the respective yhdA knockout derivatives (circles in a and c) were grown in LB at 28 °C, and OD578 (open symbols) and specific β-galactosidase activities (filled symbols) were determined along the growth curve. For the determination of csgD mRNA levels by Northern analysis (b), W3110 and its yhdA derivative were grown in LB at 28 °C to OD578 4.0. In (d), motility of W3110 and its yhdA and yhjH derivatives was assayed on swim plates. wt, Wild-type.
Here, we observed that two more genes encoding GGDEF/EAL proteins, *yhdA* and *yeaP*, are also involved in the regulation of curli expression in *E. coli* (Figs 3–5). YhdA plays a positive role in YdaM and CsgD expression during entry into stationary phase, which may reflect two separate effects that add up along the curli control cascade, as expression of CsgD seemed more strongly affected than YdaM expression. YhdA is unlikely to act via c-di-GMP, as both its GGDEF and its EAL domains are degenerate at amino acid positions crucial for DGC and PDE activities. In an earlier study (Suzuki et al., 2006), YhdA (there termed CsrD) has been shown to stimulate the degradation of two small RNAs, CsrB and CsrC, which by sequestering the RNA-binding protein CsrA can downregulate the expression of the flagellar master regulator FlhDC at the post-transcriptional level. Consistent with these results, we found that YhdA is involved in the precise timing of the shutting-off of FlhDC expression during transition into stationary phase (Fig. 5c), which is required for curli induction (Pesavento et al., 2008). It seems unlikely that YhdA also acts via the Csr system in curli control, as *ydaM* and *csgD* were not among the genes affected by CsrA overproduction detected in a very recent microarray study (Jonas et al., 2008). Also, CsrA strongly interfered with the expression of two GGDEF domain-encoding genes (*ycdT* and *ydeH*) (Jonas et al., 2008) that did not affect curli control (this study). Thus, curli regulation may involve some still unknown small regulatory RNA(s) targeted by YhdA.

In contrast to the mutations in *ydaM*, *yegE* and *yhdA*, knocking out *yeaP* (which encodes a known DGC; Ryjenkov et al., 2005) only partially reduced curli expression (Fig. 3a), and did not alter the expression of another CsgD target gene, *yaiC* (Fig. 3b). As the *yeaP* mutation also did not significantly affect CsgD expression (Fig. 4), YeaP seems to have a specific effect on curli fimbriae expression only. This effect may be linked with a specific requirement for CsgD-stimulated transcription at the very strong csgB promoter by vegetative RNAP (Weber et al., 2006); the weakly active *yaiC* promoter is transcribed by *σ^-*-containing RNAP; C. Pesavento and R. Hengge, unpublished data). Alternatively, YeaP may play a role in a post-transcriptional control of the *csgB* operon (*csgB: lacZ* used here is a translational reporter fusion). While these mechanistic details have to be unravelled by future work, it is interesting to note that the DGC YeaP is already expressed in exponential phase (Fig. 1), but may have to be activated by unknown signals perceived by its GAF domain.

**What are the functions of GGDEF/EAL domain-encoding genes that are expressed but do not affect motility and/or the curli control pathway?**

Fourteen GGDEF/EAL genes were expressed to varying degrees in growing cells (Table 1), but only five were involved in motility control (*yegE, yedQ, yeaI, yjhH* and *yhdA*) (Girgis et al., 2007; Pesavento et al., 2008; Suzuki et al., 2006); 13 GGDEF/EAL genes continued to be expressed or were newly induced during entry into stationary phase, but only five played a role in curli control (*ydaM, yciR, yegE, yhdA* and *yeaP*; Pesavento et al., 2008; Weber et al., 2006; this study). What is the role of the other genes?

Two of these other genes (*yegF* and *ydiV*) encode EAL-only proteins with degenerate PDE signatures, i.e. these proteins most likely do not control c-di-GMP levels. The remaining genes could be activated at the genetic level, but the corresponding proteins may not be present due to post-transcriptional regulation or proteolysis. Alternatively, the proteins may be there but may be inactive in the absence of appropriate signals detected by their sensory input domains. Examples seem to be: (i) YeJ, which plays a role in motility control at 37 °C, but at 28 °C, despite being expressed, affects neither motility nor curli formation (Pesavento et al., 2008); and (ii) YeAP, which is already expressed in exponential phase, but does not affect motility, and only later stimulates curli expression (Pesavento et al., 2008, and see above). Also, the influence of a GGDEF/EAL protein with relatively low expression may remain cryptic due to the functional redundancy of several such expressed proteins.

In addition, the multitude of GGDEF/EAL proteins in many species has led to the suggestion that these proteins are not only temporally sequestered (by means of differential control of expression and activity as discussed above), but also in a functional way, i.e. different GGDEF/EAL systems present and active at the same time would operate in parallel in a locally separate manner. This involves a concept of ‘microcompartimentation’ and locally separate pools of c-di-GMP, which is supported by some evidence (Jenal & Malone, 2006; Kader et al., 2006; Kulasakara et al., 2006; Pesavento et al., 2008; Ryan et al., 2006a; Weber et al., 2006). Ultimately, the question of c-di-GMP control systems operating in parallel will have to be solved by demonstrating that the proteins involved (i) do have DGC/PDE activity, (ii) are expressed and (iii) affect clearly different targets under the same growth conditions. Our data about the expression and function of all GGDEF/EAL domain-encoding genes in *E. coli* can indicate interesting candidates for such parallel-operating systems.

**A comparison of the functions of GGDEF/EAL proteins in the closely related *E. coli* and *Salmonella***

The overall architecture of the curli control cascade is clearly the same in the closely related *E. coli* and *Salmonella enterica* (serovar Typhimurium) (Romling et al., 1998). However, c-di-GMP control in curli control seems rather different. Not only the numbers (29 or 28 versus 19) but also the identities of GGDEF/EAL proteins in the two species are surprisingly different (see Supplementary Table S2). In *Salmonella*, in which YdaM does not exist, STM3388 (a GGDEF+EAL protein not present in *E. coli*)

http://mic.sgmjournals.org
and STM2133 (a YegE homologue) positively modulates CsgD expression, whereas the YciR homologue STM1703 and STM4264 (a homologue of the EAL-only protein YjcC) downregulate CsgD levels (Kader et al., 2006; Simm et al., 2007). In contrast to the situation in E. coli, YhjH seems to play a minor role in Salmonella: yhjH mutants show only partially reduced motility (Frye et al., 2006; Rydelik et al., 2002; Ryjenkov et al., 2006; Simm et al., 2007), and YhjH has a very minor effect if any on Salmonella curli expression (Simm et al., 2007). Thus, YjcC (STM4264) may be the major PDE for maintaining a low cellular c-di-GMP pool in the post-exponential phase. In E. coli, however, YhjH plays this role: its knockout strongly reduces motility (Pesavento et al., 2008), YhjH (but not YjcC) reduces the induction of CsgD and curli during entry into stationary phase (Figs 3 and 4), and yjcC expression is much lower than that of yhjH and starts during entry into stationary phase only (Fig. 1). One may speculate that a simple variation in the promoters of yjcC (to vary its $\sigma^+$ dependence) would suffice to generate these differences.

Also, the yhdA gene is present in Salmonella. Although the effects on gene expression of a yhdA mutation have not been studied in Salmonella, the mutation reduces cellular aggregation in liquid medium and results in a slight swimming defect (Simm et al., 2007), suggesting that it plays a similar role in controlling FlhDC and curli expression to that shown here for E. coli. Interestingly, the GGDEF protein YeaP, which specifically modulates the expression of the csgBAC curli operon in E. coli (Figs 3–5), does not exist in Salmonella. There is indirect evidence, however, that in Salmonella some unidentified c-di-GMP input also occurs downstream of the regulation of CsgD expression (Simm et al., 2007). In conclusion, a comparison of these two closely related bacterial species reveals that evolution can rapidly rewire conserved c-di-GMP-controlled functions such as curli fimbriae expression to different signal inputs by horizontally acquiring and/or altering expression levels of genes encoding GGDEF/EAL domain proteins.

ACKNOWLEDGEMENTS

Financial support was provided by the Deutsche Forschungsgemeinschaft (He 1556/13-1), the Dr Hans-Messner-Stiftung and the Fonds der Chemischen Industrie.

REFERENCES


Edited by: J. Green

http://mic.sgmjournals.org