Cellulose production is coupled to sensing of the pyrimidine biosynthetic pathway via c-di-GMP production by the DgcQ protein of Escherichia coli.

Journal:	*Environmental Microbiology and Environmental Microbiology Reports*
Manuscript ID	EMI-2017-1046.R1
Journal:	Environmental Microbiology
Manuscript Type:	EMI - Research article
Date Submitted by the Author:	18-Aug-2017
Complete List of Authors:	Rossi, Elio Motta, Sara; CNR, ITB Aliverti, Alessandro; University of Milan, Biosciences Cossu, Federica; University of Milan, Biosciences Gourlay, Louise; University of Milan, Biosciences Mauri, Pierluigi; CNR, ITB Landini, Paolo; University of Milan, Biosciences
Keywords:	pyrimidine de novo biosynthesis, salvage pathway, c-di-GMP, diguanylate cyclase, cellulose, signal transduction, environmental signal/stress responses

1 **Cellulose production is coupled to sensing of the pyrimidine biosynthetic pathway via c-**
2 **di-GMP production by the DgcQ protein of *Escherichia coli*.**
3
4 Elio Rossi[1,$], Sara Motta[2], Alessandro Aliverti[1], Federica Cossu[1], Louise Gourlay[1], Pierluigi
5 Mauri[2] and Paolo Landini[1]*
6
7 [1] Department of Biosciences
8 Università degli Studi di Milano
9 Milan, Italy
10 [2] Institute of Biomedical Technologies
11 National Research Council,
12 Segrate, Milan, Italy
13 *) corresponding author:
14 Tel. +39-02-50315028
15 paolo.landini@unimi.it
16
17 [$]) present address:
18 Elio Rossi: Department of Clinical Microbiology, Rigshospitalet, Copenhagen, Denmark
19
20 Running title: Cellulose production control by pyrimidines
21 Keywords: pyrimidine *de novo* biosynthesis, salvage pathway, c-di-GMP, diguanylate
22 cyclase, cellulose, signal transduction, environmental signal/stress response
23

Originality-Significance Statement

How bacteria sense and react to specific signals has deep implications for their interaction with their environment. In *Escherichia coli*, cellulose production (a stress response mechanism in this bacterium) is coupled to the pyrimidine salvage pathway, in turn activated by the availability of exogenous pyrimidines (*e.g.*, uracil), via a signal transduction pathway involving the second messenger c-di-GMP. The main sensor is the diguanylate cyclase DgcQ, which is inhibited by N-carbmoyl aspartate, while being activated by UTP. This provides an elegant mechanism to monitor which pathway the cell is using to make UTP, and to respond accordingly. Since uracil positively affects the innate immune response, we speculate that the connection between uracil sensing and cellulose production might have developed as a preemptive defense mechanism against the host immune systems.

Abstract

Production of cellulose, a stress response-mediated process in enterobacteria, is modulated in *Escherichia coli* by the activity of the two pyrimidine nucleotide biosynthetic pathways, namely, the *de novo* biosynthetic pathway, and the salvage pathway, which relies on the environmental availability of pyrimidine nitrogenous bases. We had previously reported that prevalence of the salvage over the *de novo* pathway triggers cellulose production via synthesis of the second messenger c-di-GMP by the DgcQ (YedQ) diguanylate cyclase. In this work, we show that DgcQ enzymatic activity is enhanced by UTP, whilst being inhibited by N-carbamoyl-aspartate, an intermediate of the *de novo* pathway. Thus, direct allosteric control by these ligands allows full DgcQ activity exclusively in cells actively synthesizing pyrimidine nucleotides via the salvage pathway. Inhibition of DgcQ activity by N-carbamoyl-aspartate appears to be favored by protein-protein interaction between DgcQ and PyrB, a subunit of aspartate transcarbamylase, which synthesizes N-carbamoyl-aspartate. Our results suggest that availability of pyrimidine bases might be sensed, somehow paradoxically, as an environmental stress by *E. coli*. We hypothesize that this link might have evolved since stress events, leading to extensive DNA/RNA degradation or lysis of neighbouring cells, can result in increased pyrimidine concentrations and activation of the salvage pathway.

Introduction

In the majority of bacteria, even the most basic cellular processes are directed by environmental conditions: for instance, the main regulatory mechanism of DNA replication is mediated by sensing of intracellular ATP concentrations (Skarstad and Katayama, 2013), which is in turn a function of the energy sources available in the growth milieu. Similarly, organic compounds, such as sugars or amino acids, can act as signal molecules to regulate the expression of the genes involved in their own metabolism, in systems often known as paradigmatic examples of gene regulation, such as the lactose operon (Miller, 1980). Interestingly, however, in addition to regulating their own metabolic genes, organic molecules often control the expression of genes involved in seemingly unrelated functions: for instance, glucose availability in the g/L concentration range regulates production of virulence factors in several pathogenic bacteria (Fleming and Camilli, 2016; Eisenreich and Heuner, 2016; Wu *et al.*, 2016; Rossi *et al.*, 2016), due to the fact that glucose at such concentrations can almost exclusively be found in environmental niches associated with a host. Likewise, several other environmental conditions, for example temperature or iron availability, not only control the regulation of genes required to deal with the specific stress, but also impact genes involved in adaptation to the host and/or of virulence factors through various forms of transcriptional and post-transcriptional regulation (Falconi *et al.*, 1998; Vasil and Ochsner, 1999; Llamas *et al.*, 2014; Righetti *et al.*, 2016; Gu *et al.*, 2016).

In motile bacteria such as *Escherichia coli*, one of the most important adaptation responses to environmental signals is the switch between planktonic mode, characterized by flagellar motility, and sessile (biofilm) mode. Flagellar expression and activity negatively correlate with production of extracellular polysaccharides (EPS), thus promoting a switch

from motile/planktonic cells to sessile cell aggregates, in a process directed by the bacterial second messenger c-di-GMP (Simm *et al.*, 2004; Hickman and Harwood, 2008; Boehm *et al.*, 2010; Irie *et al.*, 2012; Purcell and Tamayo, 2016). Intracellular c-di-GMP levels are determined by the activity of diguanylate cyclases (DGCs), which synthesize the signal molecule, and c-di-GMP phosphodiesterases (PDEs) that degrade it. Genes encoding proteins involved in c-di-GMP turnover can be present in high numbers, especially in the genomes of Gram negative bacteria (Römling *et al.*, 2005); although a precise function has been assigned to only a minority of c-di-GMP turnover genes, many of them appear to be involved in the switch between planktonic and sessile lifestyle, with a variety of different mechanisms and in response to various environmental signals (Römling *et al.*, 2013).

In a previous report (Garavaglia *et al.*, 2012), we provided genetic evidence that production of curli and cellulose, two main determinants for cell aggregation in *E. coli*, is linked to the pyrimidine biosynthetic pathway. Cellulose production is activated in the presence of exogenous uracil and inhibited by N-carbamoyl-aspartate, an intermediate of the *de novo* pyrimidine biosynthesis pathway, in a manner dependent on the DgcQ (formerly YedQ) diguanylate cyclase. In this work, we show that the DgcQ protein can bind both UTP and N-carbamoyl-aspartate, which affect its DGC activity in opposite fashions. Direct sensing of these two molecules by DgcQ provides an elegant mechanism for coupling cellulose production to the availability of exogenous pyrimidines and to the activity of the salvage pyrimidine biosynthetic pathway, likely perceived as a stress condition by the bacterial cell. We discuss the possible evolutionary significance of this regulatory mechanism.

Results

Cellulose production is affected by exogenous uracil and by mutations in the *de novo* pyrimidine biosynthetic pathway in a *dgcQ*-dependent manner. Cellulose is produced by many enterobacteria, including *E. coli*, mostly as a defense mechanism against environmental stresses, such as nutrient starvation or desiccation (White *et al.*, 2006; Gualdi *et al.*, 2008). Consistently, production of cellulose, and of other extracellular factors such as curli fibers, typically takes place at suboptimal growth temperatures (30°C or lower, (Zogaj *et al.*, 2001)). However, this can vary in different *E. coli* isolates; for example, fecal isolates constitutively produce cellulose at both 30°C and 37°C, while UTI isolates preferentially synthetize cellulose at 37°C (Bokranz *et al.*, 2005). The *E. coli* laboratory strain MG1655 produces very small amounts of cellulose, and exclusively at 30°C or lower; such low level cellulose production is due to the presence of a stop codon in the *bcsQ* gene in some K-12 strains, which results in reduced expression of the genes encoding cellulose synthase (Serra *et al.*, 2013). However, overexpression of the AdrA diguanylate cyclase, a positive regulator of cellulose synthase activity (Zogaj *et al.*, 2001) or of CsgD, an activator of *adrA* transcription, can overcome the effects of the *bcsQ* mutations, enhancing cellulose production (Gualdi *et al.*, 2008). In a previous work, we showed that, when grown in the presence of an excess of exogenous uracil (0.25 mM), MG1655 mutants impaired in either the *carB* or the *pyrB* genes, which encode proteins involved in the first two steps of the *de novo* pyrimidine biosynthetic pathway (summarized in Figure 1A), display a white mucous phenotype on media supplemented with Congo red and are fluorescent in the presence of Calcofluor (Garavaglia *et al.*, 2012), two dyes able to binding various extracellular structures. As shown in Figure 1, phenotypes on Congo red (CR)- and Calcofluor (CF)-supplemented media correlate with cellulose production (Figure 1D) and are totally abolished by deletion of the *bcsA* gene,

encoding the major subunit of cellulose synthase (Figure 1B, 1C), thus indicating that they are dependent on cellulose production. Albeit at a much lower extent, exogenous uracil promotes cellulose synthesis also in the MG1655 wild type, as well as in mutants in the downstream steps of *de novo* pathway, such as *pyrC* ((Garavaglia *et al.*, 2012); Figure 1). Interestingly, addition of 0.25 mM uracil, but not adenine, to LB1/4 medium supplemented with Calcofluor strongly enhanced fluorescence in LF82, an adhering-invasive *E. coli* (AIEC) strain, suggesting that induction of cellulose production by exogenous uracil is indeed a *bona fide* regulatory mechanism in *E. coli* strains other than MG1655 (Figure 1C).

These observations suggest that uracil-dependent stimulation of cellulose production is strongly enhanced in strains unable to synthesize N-carbamoyl-aspartate, regardless whether they are unable to synthesize carbamoyl-phosphate (*carB* mutant) or if they accumulate this intermediate (*pyrB* mutant). In *E. coli*, the cellulose biosynthetic machinery responds to c-di-GMP, and more specifically to the activity of two distinct DGCs, DgcC (formerly known as YaiC or AdrA; (Zogaj *et al.*, 2001; Monteiro *et al.*, 2009)) and DgcQ (YedQ), which appears to play a particularly important role in cellulose regulation in the commensal *E. coli* strain 1094 (Da Re and Ghigo, 2006). DgcQ, but not DgcC, is required for induction of cellulose by exogenous uracil ((Garavaglia *et al.*, 2012), Figure 1), suggesting a specific link between DgcQ activity and pyrimidine biosynthesis.

Identification of DgcQ interactors by co-purification experiments. Expression levels of the *dgcQ* gene were not affected by addition of 0.25 mM uracil to the growth medium (Figure S1A) indicating that exogenous uracil might affect DgcQ activity, rather than regulate its expression. Thus, we sought for possible uracil-sensing regulatory proteins that might interact with DgcQ, using pull-down experiments. To this aim, we expressed the DgcQ cytoplasmatic domain, corresponding to the amino acid residues 381-564, located immediately at the C-terminus of the second transmembrane domain ($DgcQ_{cyt}$, Figure 2A). The $DgcQ_{cyt}$ domain,

linked to a histidine tag, was expressed and purified using a cobalt affinity column. The overexpressed $DgcQ_{cyt}$ domain was found in the soluble fraction of the cell extracts only upon expression at 18 °C (see Materials and Methods), remained soluble throughout purification and could be fully recovered by imidazole elution. However, after purification, it precipitated in aqueous solutions after ca. 48 hours regardless of the storage conditions (data not shown). Despite lack of stability after imidazole elution, $DgcQ_{cyt}$ remained soluble when bound to the column matrix as it could be eluted in a soluble form with 500mM imidazole even after prolonged incubation (data not shown). The column-bound $DgcQ_{cyt}$ protein was used as a "bait" to capture DgcQ interactors from cell extracts of *E. coli* MG1655 grown in LB1/4 medium either with or without 0.25 mM uracil supplementation. Proteins binding to $DgcQ_{cyt}$ were recovered by elution of the histidine-tagged $DgcQ_{cyt}$ protein from the column with 250 mM imidazole, and identified by mass spectrometry, using Multidimensional Protein Identification Technology (MudPIT; (Mauri, 2005)).

A full overview of MudPIT identification analysis is provided in the Supplementary Information File S1. Interestingly, the pattern of cellular proteins interacting with $DgcQ_{cyt}$ changed significantly when we used cell extracts from *E. coli* grown either with or without 0.25 mM uracil. As shown in Table 1, several proteins were found to co-purify with $DgcQ_{cyt}$ in a manner strictly dependent on the growth conditions. In extracts of *E. coli* cells grown in the absence of supplemented uracil, the main $DgcQ_{cyt}$ interactors appeared to be the Dcm DNA methyltransferase, the putative phospholipid transporter MlaC, and the PyrB and PyrI proteins (Table 1). The latter two proteins constitute the two subunits of aspartate carbamoyltransferase, *i.e.*, the enzyme catalyzing the synthesis of N-carbamoyl-aspartate from aspartate and carbamoyl-phosphate (Figure 1A). This observation would suggest that the apparent inhibitory activity of N-carbamoyl-aspartate on DgcQ-dependent cellulose production (Figure 1B; (Garavaglia *et al.*, 2012)) might be mediated by direct protein-protein

interaction between $DgcQ_{cyt}$ and aspartate carbamoyltransferase. In contrast, neither PyrB nor PyrI were detected among the $DgcQ_{cyt}$-interacting proteins found in extracts of cells grown in the presence of 0.25 mM uracil (Table 1): this could be expected, as UTP synthesized from exogenous uracil via salvage pathway in turn negatively controls protein activity and expression of several genes in the *de novo* pathway, including the *pyrBI* operon, through complex feedback mechanisms (Turnbough and Switzer, 2008). We verified that uracil supplementation in the medium would indeed result in downregulation of *pyrB* expression in the strain under the conditions we used. We observed that *pyrB* expression was strongly dependent on growth phase in our experimental conditions, being higher in mid-exponential phase (Figure S1B). When LB1/4 medium was supplemented with 0.25 mM uracil, we detected a 14-fold reduction in *pyrB* transcript levels in mid-exponential phase (Figure S1B), suggesting that lack of $DgcQ_{cyt}$-PyrBI interaction in cell extracts from MG1655 grown in uracil-supplemented medium (Table 1) is due to low intracellular PyrBI concentrations resulting from downregulation of *pyrBI* expression.

In extracts from cells grown under this condition, the main $DgcQ_{cyt}$ interactors were proteins involved in post-transcriptional tRNA modifications, namely, TruD (a pseudouridine synthase) and TrmA (a 54-uridine methyltransferase), and the IF-1 initiation factor of protein synthesis (Table 1). These results might suggest a possible role of DgcQ in important cell processes other than cellulose production, such as tRNA processing and protein synthesis. However, although the possible interaction between DgcQ with these proteins is intriguing, we decided to concentrate our attention on the validation of DgcQ interaction with the PyrB•PyrI aspartate carbamoyltransferase, whose biological significance, in the light of our genetic data ((Garavaglia *et al.*, 2012); Figure 1), appears more straightforward.

Confirmation of DgcQ and PyrB•PyrI interaction by bacterial two hybrid system (BACTH). In order to validate the results of the co-purification experiments, we verified

protein-protein interaction between $DgcQ_{cyt}$ and either PyrB or PyrI, using the bacterial two hybrid system (BACTH). The BACTH system relies on the expression of target proteins fused with *Bordetella pertussis* adenylate cyclase subunits T18 and T25 in the *E. coli* BTH101 strain. Upon positive interaction between two fusion proteins activity of adenylate cyclase is reconstituted leading to transcriptional activation of the LacZ reporter (Karimova *et al.*, 1998). As shown in Figure 2B and 2C, $DgcQ_{cyt}$ showed strong protein-protein interaction with PyrB, *i.e.*, the catalytic subunit of aspartate carbamoyltransferase, but not with PyrI, whose function is to regulate PyrB activity. Likewise, no interaction was detectable between $DgcQ_{cyt}$ and the PyrC protein, which catalyzes dihydroorotate formation using N-carbamoyl-aspartate as a substrate. Although the BACTH assay can only provide a semiquantitative measurement of protein-protein interaction, co-expression of $DgcQ_{cyt}$ and PyrB resulted in CyaA activity levels in the same order of magnitude as PyrB-PyrI, which were used as a positive control in our experiment (Figure 2B and 2C), suggesting strong interaction between the two proteins.

As $DgcQ_{cyt}$ interacts with PyrB, the catalytic subunit of aspartate carbamoyltransferase, but not with the regulatory subunit PyrI, we asked ourselves whether PyrB-mediated N-carbamoyl-aspartate synthesis would somehow be necessary for this interaction: thus, we performed BACTH assays in both *ΔcarB* and *ΔpyrB* isogenic mutants of the *E. coli* BTH101 strain, unable to synthesize carbamoyl-phosphate, in turn a substrate for the aspartate carbamoyltransferase PyrB, and N-carbamoyl-aspartate itself, respectively (Figure 1A). $DgcQ_{cyt}$-PyrB interaction appears unaffected in the *ΔpyrB* mutant strain, as determined by CyaA activity levels (Figure 2B and 2C); however, the T25-PyrB fusion is fully functional and complements the loss of the *pyrB* chromosomal gene (data not shown), thus allowing N-carbamoyl-aspartate synthesis and restoring a wild type-like phenotype. In contrast, in the *ΔcarB* genetic background, co-expression of $DgcQ_{cyt}$ and PyrB failed to induce any detectable

CyaA activity, strongly suggesting that lack of enzymatic activity by PyrB, due to the absence of its substrate carbamoyl-phosphate, impairs its interaction with $DgcQ_{cyt}$ (Figure 2B and 2C).

Finally, we used BACTH assays to investigate whether, in addition to downregulation of the *de novo* biosynthetic genes by exogenous uracil, it might be possible that an increased UTP availability via the salvage pathway might directly affect DgcQ-PyrB interaction. To assess this possibility, we performed BACTH assays, using the T18-$DgcQ_{cyt}$ and the T25-PyrB plasmids, in the presence of 0.25mM exogenous uracil: no significant changes in $DgcQ_{cyt}$-PyrB interaction were detected, suggesting that, despite downregulation of the *de novo* pathway by exogenous uracil, PyrB is still enzymatically active in these conditions and can fully interact with $DgcQ_{cyt}$.

Interaction of $DgcQ_{cyt}$ with small ligands *in vitro*. The observations that cellulose production is stimulated in mutants unable to produce N-carbamoyl-aspartate, and that $DgcQ_{cyt}$ interacts directly with PyrB in a manner dependent on its carbamoyltransferase activity (Figure 2B and 2C) might suggest that N-carbamoyl-aspartate might itself be a ligand for DgcQ and as such promote $DgcQ_{cyt}$-PyrB interaction. To assess protein-ligand interactions, we performed a thermal denaturation temperature shift assay (Figure 3), which relies on the stabilization of a protein (or domain) structure by ligand binding, in turn resulting in higher denaturation temperatures (Cossu *et al.*, 2010). In addition to N-carbamoyl-aspartate, we tested the possibility that uracil, which, when provided to growth media, also seems to affect DgcQ activity *in vivo* (Figure 1B) and its ability to interact with other proteins (Table 1), might also be a ligand for this protein. GTP, as the known substrate of DGC proteins, was used as a positive control in thermal denaturation temperature shift assays, and indeed it resulted in an increase in $DgcQ_{cyt}$ denaturation temperature ranging between 1.4 and 3.6 °C. Interestingly, the highest temperature shift was observed at the lowest GTP concentration tested (0.25 mM). This phenomenon might depend on different

conformations assumed by $DgcQ_{cyt}$ upon binding on either one or more GTP molecules and on initial conversion of 2 GTP molecules to c-di-GMP, with consequent allosteric product inhibition, as already observed for other DGC enzymes (Schirmer, 2016). N-carbamoyl-aspartate also induced a significant change (up to 3.9°C) in $DgcQ_{cyt}$ denaturation temperature, although at higher concentrations than GTP (2.5-5 mM). Uracil, in contrast, failed to show any detectable change in the denaturation kinetics at any concentration used (Figure 3). However, we reasoned that, once taken up by the bacterial cell, uracil is converted to UTP via the pyrimidine salvage pathway (Figure 1). When tested in thermal denaturation temperature shift assays, UTP, at a concentration as low as 0.25 mM, induced a slight shift in $DgcQ_{cyt}$ denaturation temperature (1.1-1.3 °C, Figure 3), that is, to a lower extent than GTP or N-carbamoyl-aspartate, and below the 3°C threshold considered significant for these assays (Cossu *et al.*, 2010). However, neither ATP (Figure 3) nor CTP (data not shown) induced any detectable change in $DgcQ_{cyt}$ denaturation temperature, suggesting that UTP-induced temperature shift, albeit small, might indeed be due to a conformational change in $DgcQ_{cyt}$ upon UTP binding. Thermal stabilization of the DgcQ cytoplasmic domain indicates that DgcQ might directly bind N-carbamoyl-aspartate and, possibly, UTP suggesting that its activity might be also influenced by either ligand.

UTP and N-carbamoyl-aspartate affect $DgcQ_{cyt}$ activity in opposite ways. As thermal denaturation temperature shift assays suggest that both UTP and N-carbamoyl-aspartate can bind the $DgcQ_{cyt}$ protein, we assessed their effect on $DgcQ_{cyt}$ catalytic activity in a diguanylate cyclase assay. c-di-GMP formation was monitored by HPLC, as described in Materials and Methods. Despite its poor stability, $DgcQ_{cyt}$ retained its DGC activity, showing an enzymatic activity level of 41.5 nmol c-di-GMP min^{-1} mg^{-1}, *i.e.*, similar to what previously reported for other DGCs, such as WspR and PleD (De *et al.*, 2009; Antoniani *et al.*, 2013). However, reduced stability within 48 hours precluded replicating experiments on the same

$DgcQ_{cyt}$ purification batch. Attempts to express different $DgcQ_{cyt}$ constructs in order to improve long term stability were unsuccessful, as deletion of even few amino acids after residue 381, although greatly improving protein stability, resulted in complete loss of its enzymatic activity (data not shown). Despite these difficulties, we were able to assess the effects of UTP and N-carbamoyl-aspartate on $DgcQ_{cyt}$ enzymatic activity on at least three independent fresh purifications. In the presence of 2.5 mM N-carbamoyl-aspartate, activity was reduced by more than 10-fold, consistent with direct binding of N-carbamoyl-aspartate to $DgcQ_{cyt}$ suggested by thermal denaturation experiments (Figure 3), and with its inhibitory effects on DgcQ activity *in vivo* (Figure 1). In contrast, UTP, already at 0.1 mM, resulted in a 2.5-fold increase in $DgcQ_{cyt}$ activity; higher UTP concentrations (up to 2.5 mM) did not lead to further stimulation (data not shown). Consistent with thermal denaturation assays shown in Figure 3, ATP at concentrations up to 2.5 mM did not significantly affect $DgcQ_{cyt}$ activity (data not shown).

Since UTP and N-carbamoyl-aspartate impact $DgcQ_{cyt}$ activity in opposite ways, we tested the possible dominance of either ligand: however, the results showed very high variability, suggesting that the contemporary presence of UTP and N-carbamoyl-aspartate might further destabilize the $DgcQ_{cyt}$ protein in our assay conditions.

Discussion

In this report, we have shown that allosteric control of DgcQ activity by both UTP (positive effector) and N-carbamoyl-aspartate (negative effector) allows coupling of cellulose production to the pyrimidine salvage pathway in *E. coli*. When pyrimidine nucleotides are synthesized via the *de novo* pathway, of which N-carbamoyl-aspartate is an intermediate, c-

di-GMP synthesis by DgcQ and, consequently, cellulose production, are inhibited. On the contrary, when the salvage pathway is active, leading to UTP synthesis from uracil and consequent repression of the *de novo* pathway, DgcQ activity and cellulose production are fully activated. Finally, in the absence of active pyrimidine synthesis by either pathway, DgcQ is active at reduced efficiency, and it is likely completely turned off if, in addition to UTP, also intracellular GTP (the substrate for DgcQ) concentrations drop below a given threshold. Thus, dependence of full DgcQ activity on UTP would allow modulation of cellulose production as soon as intracellular concentrations of this nucleotide decrease. It can also be expected that low UTP concentrations will result in a decreased intracellular pool for UDP-glucose, the precursor of cellulose synthesis, thus directly affecting the efficiency of the bacterial cellulose synthase complex. Arresting cellulose synthesis would save energy, and might possibly represent an initial step in promoting biofilm dispersal and release of planktonic motile cells in response to starvation. Indeed, cellulose turnover and biofilm dispersal can be increased by starvation in a manner dependent on intracellular c-di-GMP concentration (Gjermansen *et al.*, 2010).

In vitro DGC activity assays indicate that N-carbamoyl-aspartate can inhibit the DgcQ protein at a millimolar concentration range. Although a precise estimation of intracellular N-carbamoyl-aspartate concentrations is not available, it is likely that normal intracellular concentrations for this molecule are lower, even in *E. coli* cells actively synthesizing pyrimidines. However, direct protein-protein interaction between DgcQ and enzymatically active PyrB (Table 1, Figure 2) would locally increase N-carbamoyl-aspartate concentrations available to the DgcQ protein. This notion would be consistent with the mode of action of other DGC proteins, whose activity is often linked to specific protein-protein interaction (Lindenberg *et al.*, 2013; Rybtke *et al.*, 2015). Future experiments will also address how c-di-GMP synthesis by DgcQ might in turn affect its interaction with PyrB and, possibly, with

other proteins found in co-purification experiments. From our results, it could be envisaged that, as long as the *de novo* pyrimidine biosynthetic pathway is active, DgcQ activity, and thus cellulose production, cannot be fully active. However, concentrations of exogenous uracil in the order of 0.25 mM can repress expression of the *pyrBI* operon (Figure S1B), likely preventing PyrB production and its interaction with DgcQ (Figure S1B, Table 1, File S1). Consistent with feedback control of *de novo* pyrimidine synthesis, growth in media supplemented with 0.25 mM uracil results in a 2-fold increase in cellulose production in *E. coli* MG1655 (Figure 1D), and this effect is further amplified in mutants unable to synthesize N-carbamoyl-aspartate (Garavaglia *et al.*, 2012). Our observations suggest that DgcQ allosteric control by both UTP and N-carbamoyl-aspartate allows efficient induction of cellulose production in response to presence of exogenous uracil in the environment, well below millimolar concentrations. The regulatory network connecting pyrimidine biosynthesis, DgcQ activity, and cellulose production is shown in Figure 5.

Cellulose protects *E. coli* from harsh environmental conditions such as desiccation (White *et al.*, 2006; Gualdi *et al.*, 2008) and its production is part of the general stress response controlled by the *rpoS* gene (Weber *et al.*, 2006). Thus, it can appear counterintuitive that it is also induced in response to pyrimidine availability; however, it can be argued that an increase in the availability of exogenous nucleotides and nitrogenous bases might also follow stressful events such as extensive lysis of neighbouring cells, for instance due to phage infections. Likewise, it is possible that pyrimidine and purine salvage pathways might be activated by extensive DNA/RNA degradation, again following environmental insults. We have not investigated whether intermediates of the *de novo* purine synthesis can also trigger cellulose production either in a DgcQ-dependent or independent way. However, it is remarkable that, in *Pseudomonas aeruginosa*, mutations in the pyrimidine, but not in the purine biosynthetic pathway, affect biofilm formation (Ueda *et al.*, 2009). Likewise, addition

of exogenous uracil, but not adenine, results in fluorescence of Calcofluor-supplemented plates in adherent-invasive *Escherichia coli* (Figure 1C), suggesting a specific effect of pyrimidines on cellulose production. Interestingly, uracil can be found at sub-millimolar concentrations in the gastro-intestinal tract of both invertebrates and mammals (Vogel-Scheel *et al.*, 2010; Lee *et al.*, 2013), mostly as a product of the gut microbiota. Bacterial-produced uracil is a strong inducer of mucosal immunity and of generation of reactive oxygen species in *Drosophila* (Lee *et al.*, 2013). As cellulose also plays an important role as a defense mechanism against oxidative stress in Enterobacteria (White *et al.*, 2006), it is conceivable that the link between exogenous uracil sensing and cellulose production via the DgcQ protein might have evolved as a preemptive defense mechanism against the host immune system. Similarly, a recent work reported that, in *Salmonella* Typhimurium, cellulose production is triggered by exogenous arginine via the DgcQ-homologue STM1987 (Mills *et al.*, 2015): thus, it might be tempting to speculate that relaying cellulose production to the availability of exogenous metabolites is a widespread mechanism in Enterobacteria.

These results, and also our observations that the pyrimidine-sensing DgcQ protein might also interact with proteins not involved in cellulose production, such as for instance tRNA modification enzymes (Table 1), expand previous observations showing that pyrimidine biosynthesis is hardwired to important processes such as biofilm formation, virulence factor production, antimicrobial resistance, and even evolutionary strategies (Haugo and Watnick, 2002; Beaumont *et al.*, 2009; Ueda *et al.*, 2009; Guo *et al.*, 2016), thus suggesting that pyrimidine biosynthesis flow can represent a pivotal sensing mechanism in bacteria.

Materials and Methods

Strains and growth media, phenotypic assays and cellulose determination. Strains used in this work are listed in Table S1. For strain construction and routine manipulation, bacteria were grown in LB medium (10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl) or LB-agar medium (LB medium with 15 g/L agar). When necessary, antibiotics were added to the growth medium at the following concentrations: ampicillin, 100 μg/ml; chloramphenicol, 50 μg/ml; kanamycin, 50 μg/ml.

For all other experiments, bacteria were grown either in 1:4 diluted LB medium (LB1/4) or on LB1/4-agar medium (LB1/4 medium with 15 g/L agar). LB1/4 medium was used since it favors biofilm formation and adhesion factor production compared to the full strength LB (Perrin *et al.*, 2009), while retaining its complexity and supporting growth of the *ΔcarB* and *ΔpyrB* mutants, auxotroph for pyrimidines (data not shown). If not otherwise stated experiments were performed at 30°C. Uracil was dissolved in 50% dimethyl sulfoxide (DMSO) to a 10 mM concentration and supplemented at a 0.25 mM final concentration. DMSO at a 1.25% final concentration was always added to control cultures. Congo red- (CR) and calcofluor (CF)-binding phenotypic assays were essentially carried out as previously described (Garavaglia *et al.*, 2012), spotting selected overnight cultures on LB1/4-agar media containing uracil, supplemented with either 0.004% Congo red and 0.002% Coomassie blue (CR-medium) or 0.005% Calcofluor (CF-medium) after autoclaving. Bacteria were grown for 24 h at 30 °C; phenotypes were better detected after a further 24–48 h incubation at 4 °C. Cellulose determination was performed as previously described (Gualdi *et al.*, 2008)1.

Molecular biology techniques and gene expression analysis. *Escherichia coli* MG1655 and BTH101 mutant derivatives were constructed through homologous recombination using the λ-red technique, (Datsenko and Wanner, 2000); PCR products for mutant construction were obtained using primer pairs listed in Table S2. Plasmids used in this work, listed in Table S1,

were constructed by standard methods, inserting PCR fragments obtained using primers pairs reported in Table S2 into vectors using restriction enzymes. The correctness of all newly constructed plasmids was verified by sequencing.

Gene expression levels were measured through quantitative real-time qRT-PCR as described previously (Garavaglia *et al.*, 2012), using 16S RNA as reference gene. RNA was extracted from cultures grown either in LB1/4 or in LB1/4(ura) at 30°C in full aeration (constant shaking at 100 r.p.m) either after overnight growth or during late exponential phase (OD_{600} = 0.8). The complete list of primers used for amplification is reported in Table S2.

DgcQ cytoplasmic domain ($DgcQ_{cyt}$) expression and purification. *E. coli* BL21(DE3) strain harboring the pQE80L-*dgcQ*$_{cyt}$ was grown in LB medium at 37 °C under shaking condition (120 rpm), until cultures reached cellular density corresponding to OD_{600} = 0.6. The temperature was lowered to 18 °C, shaking increased to 150 rpm, and 6xHis-tagged protein expression was induced adding 0.5 mM isopropyl β-D-thiogalactopyranoside (IPTG). After 16 h of growth, cells were harvested by centrifugation at 7,000 rpm for 15 min at 4 °C, resuspended in 1/50 vol. of cold Buffer A (50 mM Tris-Cl, pH 8.0, 300 mM NaCl, 20 mM imidazole), supplemented with 100 µg/ml lysozyme, 1 mM 4-(2-Aminoethyl) benzene-sulfonyl fluoride hydrochloride (AEBSF) and 100 µl of Protease Inhibitor (P8849, Sigma-Aldrich) and incubated for 30 min on ice. After cell lysis by sonication (10 cycles of 10 s bursts with 30 s pauses between them), cell debris and non-soluble proteins were removed by centrifugation at 18,000 rpm at 4 °C for 1 h. The supernatant, containing the soluble protein fraction, was loaded on a gravity chromatography column containing 1 mL of TALON® metal affinity resin (Clontech), equilibrated with 5 volumes of cold Buffer A. To avoid non-specific interactions, the resin was washed with 10 volumes of cold Buffer A supplemented with 1 mM AEBSF, followed by 10 volumes of 90% Buffer A + 10% Buffer B (50 mM Tris-Cl, pH 8.0, 300 mM NaCl, 500 mM imidazole). Proteins were eluted on a single step using 5

volumes of 50% Buffer A + 50% Buffer B. Imidazole was immediately removed by multiple PD-100 Desalting columns (GE Healthcare Life Sciences), followed by wash and elution in Storage Buffer (50 mM Tris-Cl, pH 8.0, 300 mM NaCl, 1 mM DTT, 5% glycerol). If required, AEBSF was added to a final concentration of 1 mM. Purified $DgcQ_{cyt}$ protein in Storage Buffer was used within 24 h for biochemical assays, as the proteins precipitated over longer times.

$DgcQ_{cyt}$ co-purification (pull down) experiments. $DgcQ_{cyt}$ was expressed and loaded by gravity on a chromatography column containing metal affinity resin as described in the previous section. After the first wash with 10 volumes of cold Buffer A, the $DgcQ_{cyt}$-carrying column was loaded with 5 mL of cleared cellular lysate deriving from cultures of MG1655 strain grown for 16 h either in presence or absence of 0.25 mM uracil. The column was then washed with 10 volumes of cold Buffer A, followed by 5 vol. of 90% Buffer A + 10% Buffer B. Finally, $DgcQ_{cyt}$ and co-eluting proteins were eluted using 5 volumes of 50% Buffer A + 50% Buffer B. An aliquot of the recovered protein's complexes were treated with RapiGest SF (Waters), digested with trypsin and identified through MudPIT (Multidimensional Protein Identification Technology) analysis, as previously described (Longo *et al.*, 2016). In order to identify specific proteins co-eluting with 6xHis-tagged $DgcQ_{cyt}$ in presence or absence of 0.25 mM uracil, outputs were analyzed with the in-house software MAProMa (multidimensional algorithm protein map) (Mauri and Dehò, 2008). Different protein amounts were estimated by means of the DAve (differential average) algorithm of MAProMa (Mauri, 2005). A DAve value higher than 10 (or lower than -10) indicates that a given protein is present in different amounts in the samples analyzed; a DAve value of either 200 or -200 indicates that a protein is exclusively present in one sample.

Bacterial two-hybrid (BACTH) assays. For the Bacterial Adenylate Cyclase Two-Hybrid System (BACTH), proteins of interest were fused to T25 and T18 fragments of *B. pertussis*

adenylate cyclase in vectors pKT25 and pUT18, respectively, and the BACTH assay was performed essentially as previously described (Longo *et al.*, 2016). Briefly, pKT25 and pUT18 plasmids derivatives (see Table S2) were used to co-transform in various combinations the *E. coli* BTH101 or its isogenic *ΔcarB* and *ΔpyrB* mutants. Resulting transformants were cultured in LB medium and spotted on either LB-agar medium containing 100 µg/ml ampicillin, 50 µg/ml kanamycin, 40 µg/ml 5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside (X-gal) and 0.5 mM IPTG. Interactions between different hybrid proteins were visualized after 48 h of incubation at 30 °C. For a quantitative measurement of protein-protein interactions, ß-galactosidase activity was measured on 100 µl of the same LB cultures grown for 16 h at 30°C by standard Miller Assay (Miller, 1972).

Biochemical assays. Thermal denaturation temperature shift assays were essentially performed as previously described (Cossu *et al.*, 2010). $DgcQ_{cyt}$ (20 µM), was mixed with 3.5 µl of Sypro orange (Sigma) diluted 60-fold, 50 mM Tris-Cl, pH 7.5, and either ATP, or UTP, or uracil, or *N*-carbamoyl-aspartate, at different concentrations. Ligands were all prepared fresh in 50 mM Tris-Cl, pH 7.5 at a stock concentration of 25 mM. Tris-Cl, pH 7.5 was used in place of ligands in negative controls. The assay was carried out using a MiniOpticon Real Time PCR Detection System (Bio-Rad), designed originally for PCR; sample plates were heated from 4 °C to 99 °C at a heating rate of 0.2 °C/5 s. The fluorescence emission intensity was measured over the 540–700 nm wavelength range by excitation at 470–505 nm. For analysis and visualization purposes, output melting curves were manually inspected, then normalized, smoothed and fitted with a 4-parameters logistic function using a custom python script in order to calculate the melting temperature (T_m).

Diguanylate cyclase (DGC) activity of $DgcQ_{cyt}$ was measured following the *in vitro* production of the reaction product c-di-GMP at 30 °C using reverse-phase HPLC. Briefly, freshly purified $DgcQ_{cyt}$ (20 µM) in Storage Buffer was added to a reaction solution

containing 10 mM $MnCl_2$, 200 mM NaCl, 50 mM Tris, pH 7.5. After an incubation time of 5 min at 30 °C, 0.5 mM GTP was added to start the reaction. Aliquots were taken at different time points between 0 and 3 h, showing that c-di-GMP synthesis was linear during at least the first 2 h of reaction; thus, a 2 h-incubation time was chosen for the DGC assays in the presence of various ligands. The reaction mixture was separated on reverse-phase 12.5 cm Supelcosil LC-18-DB, 3 μm particle size, column using a methanol-phosphate gradient (Buffer A: 100 mM potassium phosphate buffer, pH 6.0; Buffer B: Buffer A containing 20% methanol) as previously described (Antoniani *et al.*, 2010). Reaction products were identified by comparison to standard nucleotides GTP (Fermentas) and synthetic c-di-GMP (Biolog, Bremen, Germany). To test the effects of different ligands on $DgcQ_{cyt}$ activity, ATP, UTP, uracil, *N*-carbamoyl-aspartate, and combinations of them, were added, at different concentrations to *in vitro* reaction mixtures and then incubated for 5 min at 30 °C, after which the reaction was started by addition of the substrate GTP, as before. All ligands were freshly prepared in 50 mM Tris-Cl, pH 7.5, at stock concentration of 25 mM and used within 1 h.

Acknowledgments

We would like to thank Chiara Speroni for performing *dgcQ* and *pyrB* expression experiments. Elio Rossi was supported by a post-doctoral Fellowship from Fondazione Fratelli Confalonieri.

References

Antoniani, D., Bocci, P., Maciag, A., Raffaelli, N., and Landini, P. (2010) Monitoring of

diguanylate cyclase activity and of cyclic-di-GMP biosynthesis by whole-cell assays suitable for high-throughput screening of biofilm inhibitors. *Applied microbiology and biotechnology* **85**: 1095–1104.

Antoniani, D., Rossi, E., Rinaldo, S., Bocci, P., Lolicato, M., Paiardini, A., et al. (2013) The immunosuppressive drug azathioprine inhibits biosynthesis of the bacterial signal molecule cyclic-di-GMP by interfering with intracellular nucleotide pool availability. *Applied microbiology and biotechnology* **97**: 7325–7336.

Beaumont, H.J.E., Gallie, J., Kost, C., Ferguson, G.C., and Rainey, P.B. (2009) Experimental evolution of bet hedging. *Nature* **462**: 90–93.

Boehm, A., Kaiser, M., Li, H., Spangler, C., Kasper, C.A., Ackermann, M., et al. (2010) Second messenger-mediated adjustment of bacterial swimming velocity. *Cell* **141**: 107–116.

Bokranz, W., Wang, X., Tschäpe, H., and Römling, U. (2005) Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *Journal of Medical Microbiology* **54**: 1171–1182.

Cossu, F., Malvezzi, F., Canevari, G., Mastrangelo, E., Lecis, D., Delia, D., et al. (2010) Recognition of Smac-mimetic compounds by the BIR domain of cIAP1. *Protein Sci.* **19**: 2418–2429.

Da Re, S. and Ghigo, J.-M. (2006) A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. *Journal of Bacteriology* **188**: 3073–3087.

Datsenko, K.A. and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA* **97**: 6640–6645.

De, N., Navarro, M.V.A.S., Raghavan, R.V., and Sondermann, H. (2009) Determinants for the activation and autoinhibition of the diguanylate cyclase response regulator WspR. *J. Mol. Biol.* **393**: 619–633.

Eisenreich, W. and Heuner, K. (2016) The life stage-specific pathometabolism of *Legionella pneumophila*. *FEBS Lett.*

Falconi, M., Colonna, B., Prosseda, G., Micheli, G., and Gualerzi, C.O. (1998) Thermoregulation of *Shigella* and *Escherichia coli* EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. *EMBO J.* **17**: 7033–7043.

Fleming, E. and Camilli, A. (2016) ManLMN is a glucose transporter and central metabolic regulator in *Streptococcus pneumoniae*. *Mol. Microbiol.*

Garavaglia, M., Rossi, E., and Landini, P. (2012) The pyrimidine nucleotide biosynthetic pathway modulates production of biofilm determinants in *Escherichia coli*. *PLoS ONE* **7**: e31252.

Gjermansen, M., Nilsson, M., Yang, L., and Tolker-Nielsen, T. (2010) Characterization of starvation-induced dispersion in *Pseudomonas putida* biofilms: genetic elements and molecular mechanisms. *Mol. Microbiol.* **75**: 815–826.

Gu, D., Guo, M., Yang, M., Zhang, Y., Zhou, X., and Wang, Q. (2016) A σE-mediated temperature gauge controls a switch from LuxR-mediated virulence gene expression to thermal stress adaptation in *Vibrio alginolyticus*. *PLoS Pathog.* **12**: e1005645.

Gualdi, L., Tagliabue, L., Bertagnoli, S., Ieranò, T., De Castro, C., and Landini, P. (2008) Cellulose modulates biofilm formation by counteracting curli-mediated colonization of solid surfaces in *Escherichia coli*. *Microbiology* **154**: 2017–2024.

Guo, Q., Wei, Y., Xia, B., Jin, Y., Liu, C., Pan, X., et al. (2016) Identification of a small molecule that simultaneously suppresses virulence and antibiotic resistance of *Pseudomonas aeruginosa*. *Scientific reports* **6**: 19141.

Haugo, A.J. and Watnick, P.I. (2002) *Vibrio cholerae* CytR is a repressor of biofilm development. *Mol. Microbiol.* **45**: 471–483.

Hickman, J.W. and Harwood, C.S. (2008) Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. *Mol. Microbiol.* **69**: 376–389.

Irie, Y., Borlee, B.R., O'Connor, J.R., Hill, P.J., Harwood, C.S., Wozniak, D.J., and Parsek, M.R. (2012) Self-produced exopolysaccharide is a signal that stimulates biofilm formation in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U.S.A.* **109**: 20632–20636.

Karimova, G., Pidoux, J., Ullmann, A., and Ladant, D. (1998) A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci USA* **95**: 5752–5756.

Lee, K.-A., Kim, S.-H., Kim, E.-K., Ha, E.-M., You, H., Kim, B., et al. (2013) Bacterial-derived uracil as a modulator of mucosal immunity and gut-microbe homeostasis in *Drosophila*. *Cell* **153**: 797–811.

Lindenberg, S., Klauck, G., Pesavento, C., Klauck, E., and Hengge, R. (2013) The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signalling cascade in *E. coli* biofilm control. *EMBO J.* **32**: 2001–2014.

Llamas, M.A., Imperi, F., Visca, P., and Lamont, I.L. (2014) Cell-surface signaling in *Pseudomonas*: stress responses, iron transport, and pathogenicity. *FEMS Microbiol. Rev.* **38**: 569–597.

Longo, F., Motta, S., Mauri, P., Landini, P., and Rossi, E. (2016) Interplay of the modified nucleotide phosphoadenosine 5'-phosphosulfate (PAPS) with global regulatory proteins in *Escherichia coli*: modulation of cyclic AMP (cAMP)-dependent gene expression and interaction with the HupA regulatory protein. *Chem. Biol. Interact.*

Mauri, P. (2005) Identification of proteins released by pancreatic cancer cells by multidimensional protein identification technology: a strategy for identification of novel cancer markers. *The FASEB Journal*.

Mauri, P. and Dehò, G. (2008) A proteomic approach to the analysis of RNA degradosome

composition in *Escherichia coli*. *Meth. Enzymol.* **447**: 99–117.

Miller, J. (1980) The *lacI* gene: its role in *lac* operon control and its use as a genetic system. *Cold Spring Harbor Monograph Archive* **07**:.

Miller, J.H. (1972) Experiments in molecular genetics Cold Spring Harbor Laboratory.

Mills, E., Petersen, E., Kulasekara, B.R., and Miller, S.I. (2015) A direct screen for c-di-GMP modulators reveals a *Salmonella* Typhimurium periplasmic L-arginine-sensing pathway. *Sci Signal* **8**: ra57–ra57.

Monteiro, C., Saxena, I., Wang, X., Kader, A., Bokranz, W., Simm, R., et al. (2009) Characterization of cellulose production in *Escherichia coli* Nissle 1917 and its biological consequences. *Environ. Microbiol.* **11**: 1105–1116.

Perrin, C., Briandet, R., Jubelin, G., Lejeune, P., Mandrand-Berthelot, M.-A., Rodrigue, A., and Dorel, C. (2009) Nickel promotes biofilm formation by *Escherichia coli* K-12 strains that produce curli. *Appl. Environ. Microbiol.* **75**: 1723–1733.

Purcell, E.B. and Tamayo, R. (2016) Cyclic diguanylate signaling in Gram-positive bacteria. *FEMS Microbiol. Rev.* **40**: 753–773.

Righetti, F., Nuss, A.M., Twittenhoff, C., Beele, S., Urban, K., Will, S., et al. (2016) Temperature-responsive in vitro RNA structurome of *Yersinia pseudotuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* **113**: 7237–7242.

Rossi, E., Longo, F., Barbagallo, M., Peano, C., Consolandi, C., Pietrelli, A., et al. (2016) Glucose availability enhances lipopolysaccharide production and immunogenicity in the opportunistic pathogen *Acinetobacter baumannii*. *Future Microbiology* fmb.15.153–15.

Römling, U., Galperin, M.Y., and Gomelsky, M. (2013) Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* **77**: 1–52.

Römling, U., Gomelsky, M., and Galperin, M.Y. (2005) C-di-GMP: the dawning of a novel bacterial signalling system. *Mol. Microbiol.* **57**: 629–639.

Rybtke, M., Berthelsen, J., Yang, L., Høiby, N., Givskov, M., and Tolker-Nielsen, T. (2015) The LapG protein plays a role in *Pseudomonas aeruginosa* biofilm formation by controlling the presence of the CdrA adhesin on the cell surface. *MicrobiologyOpen* **4**: 917–930.

Schirmer, T. (2016) C-di-GMP Synthesis: Structural Aspects of Evolution, Catalysis and Regulation. *J. Mol. Biol.* **428**: 3683–3701.

Serra, D.O., Richter, A.M., and Hengge, R. (2013) Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. *Journal of Bacteriology* **195**: 5540–5554.

Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004) GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol. Microbiol.* **53**: 1123–1134.

Skarstad, K. and Katayama, T. (2013) Regulating DNA replication in bacteria. *Cold Spring Harb Perspect Biol* **5**: a012922–a012922.

Turnbough, C.L. and Switzer, R.L. (2008) Regulation of pyrimidine biosynthetic gene expression in bacteria: repression without repressors. *Microbiol. Mol. Biol. Rev.* **72**: 266–300– table of contents.

Ueda, A., Attila, C., Whiteley, M., and Wood, T.K. (2009) Uracil influences quorum sensing and biofilm formation in *Pseudomonas aeruginosa* and fluorouracil is an antagonist. *Microb Biotechnol* **2**: 62–74.

Vasil, M.L. and Ochsner, U.A. (1999) The response of *Pseudomonas aeruginosa* to iron: genetics, biochemistry and virulence. *Mol. Microbiol.* **34**: 399–413.

Vogel-Scheel, J., Alpert, C., Engst, W., Loh, G., and Blaut, M. (2010) Requirement of purine and pyrimidine synthesis for colonization of the mouse intestine by *Escherichia coli*. *Appl. Environ. Microbiol.* **76**: 5181–5187.

Weber, H., Pesavento, C., Possling, A., Tischendorf, G., and Hengge, R. (2006) Cyclic-di-

GMP-mediated signalling within the sigma network of *Escherichia coli*. *Mol. Microbiol.* **62**: 1014–1034.

White, A.P., Gibson, D.L., Kim, W., Kay, W.W., and Surette, M.G. (2006) Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. *Journal of Bacteriology* **188**: 3219–3227.

Wu, K., Xu, H., Zheng, Y., Wang, L., Zhang, X., and Yin, Y. (2016) CpsR, a GntR family regulator, transcriptionally regulates capsular polysaccharide biosynthesis and governs bacterial virulence in *Streptococcus pneumoniae*. *Scientific reports* **6**: 29255.

Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W., and Römling, U. (2001) The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Mol. Microbiol.* **39**: 1452–1463.

Figure legends

Figure 1. Link between pyrimidine biosynthesis and cellulose production in *E. coli* MG1655. (A) Schematic representation of *de novo* (light blue background) and salvage (light brown background) pathways for pyrimidine biosynthesis in *Escherichia coli*. The metabolic intermediate *N*-carbamoyl-aspartate and the nitrogenous base uracil, which belongs respectively to the *de novo* and salvage pathways for pyrimidine biosynthesis are highlighted in green and orange. Shared reaction between the two pathways are highlighted in a purple background. Solid lines represent a single reaction driven by an enzymatic complex encoded by genes indicated on the right. Dashed lines represent multiple consecutive reactions carried out by the products of the genes listed on the right. (B) Phenotypes of *E. coli* MG1655 and its mutant derivatives affected in pyrimidine biosynthesis or cellulose production on LB1/4 medium added with Congo red (CR) with or without 0.25 mM uracil. Infographic below the

picture summarises the ability of each strain to synthesize the metabolic intermediate *N*-carbamoyl-L-aspartate (*N*cAsp) (green=able; grey=unable). (C) Upper panels: Phenotypes on LB1/4 medium added with the fluorescent dye Calcofluor (CF) with or without 0.25 mM uracil. In the absence of added uracil, no strain showed any detectable fluorescence; *E .coli* MG1655 is shown as a representative of the CF-negative phenotype. Lower panel: CF phenotype of adherent-invasive *E. coli* (AIEC) strain LF82 with or without the addition of uracil (ura) and adenine (ADE). For both CR and CF phenotype experiments, bacterial strains were grown at 30°C for 24 hours and plates were incubated at 4°C for at least 48 hours to enhance CR and CF binding. (D) Determination of cellulose in LB1/4 medium with or without supplementation with 0.25 mM uracil.

Figure 2. DgcQ diguanylate cyclase and its interaction with components of the aspartate transcarbamylase enzyme. (A) Schematic representation of the DgcQ protein and its artificially expressed shorter variants 6xHis-DgcQ$_{cyt}$ and T18-DgcQ$_{cyt}$ comprising only the cytoplasmic domain of the protein fused with a histidine tag and the T18 subunit of *Bordetella pertussis* adenylate cyclase, respectively. (B) Bacterial Adenylate Cyclase Two-Hybrid (BACTH) assays on X-gal-supplemented plates. *E. coli* BTH101 wild-type (WT) strain, and its isogenic derivatives mutated in either the *carB* or *pyrB* genes, were transformed with a combination of two-hybrid vector plasmids (pKT25, T25; pUT18, T18), allowing expression of the indicated proteins. Only relevant combinations are presented (for the full set of interactions refer to Panel C of this figure). Blue colour is indicative of a reconstituted adenylate cyclase activity, *i.e.*, of protein-protein interaction. Strains harbouring pUT18 and pKT25 empty vectors are considered as negative control, while strain co-expressing the T25-Zip and T18-Zip fusion proteins represents the positive control. The infographic on the right highlight the ability of each strain to synthesize carbamoyl-phosphate (carbamoyl-P in the

Figure), the precursor of *N*-carbamoyl-aspartate. (C) Full set of protein interactions and their strength in BACTH assay measured as ß-galactosidase activity on overnight cultures grown in LB medium at 30°C. Results obtained in either the BTH101 strain or its *ΔpyrB* mutant, both proficient for the synthesis of *N*-carbamoyl-aspartate precursor, carbamoyl-phosphate, are highlighted by a green background; a grey background highlights results in the *ΔcarB* mutant, unable to synthesize carbamoyl-phosphate. Single experiments are represented by dots, and a minimum of three independent replicate is reported. Standard deviations are shown.

Figure 3. $DgcQ_{cyt}$ thermal denaturation temperature shift assays. In the presence of a ligand, a shift in the melting curve can be appreciated. The difference in 50% melting temperature (ΔT_m, reported in the table) is calculated at the temperature in which 50% of total fluorescence is observed. Curves and ΔT_m reported in the table are the average of three independent measurements. Standard deviations for ΔT_m are shown.

Figure 4. *In vitro* diguanylate cyclase (DGC) activity of the purified $DgcQ_{cyt}$ protein. c-di-GMP production was measured through high-pressure liquid chromatography (HPLC) analysis. For each sample, purified $DgcQ_{cyt}$ (20 µM) was incubated for 5 minutes at 30°C in the reaction buffer (see Materials and Methods) supplemented, when necessary, with either 0.1 mM UTP or 2.5 mM *N*-carbamyl-aspartate (*N*cAsp). The reaction was started by adding the substrate GTP (0.5 mM) and run for 2 hours at 30°C. c-di-GMP levels (ca. 0.8 µM) of the control reaction containing GTP only was set to 1, and relative amounts are shown. Results are mean ± standard errors of the mean of three independent experiments. The asterisks denote significant differences (*$p \leq 0.05$; **$p \leq 0.01$ Holm-Šídák multigroup analysis).

Figure 5. Schematic representation of the possible molecular network connecting pyrimidine biosynthesis, DgcQ activity, and cellulose production. Dashed lines indicate missing reactions if connecting metabolic intermediates. Green lines indicate a positive effect on the activity of the protein, while red lines indicate a negative one. IM: inner membrane; OM: outer membrane; PG: peptidoglycan; UDP-Glu: UDP-glucose.

Supplementary Figure S1. Relative expression of the (A) *yedQ* and (B) *pyrB* genes determined using Real-Time RT-PCR on RNA extracted from cultures in late exponential (Late exp., OD_{600} = 0.8) and stationary (Stat., 24h of growth) phase. 16S RNA transcript was used as reference gene. ΔCt values between the genes of interest and 16S RNA were set at 1.0 for MG1655 in Late exponential phase in LB1/4 medium without addition of uracil. Transcript levels in other growth conditions are expressed as relative values. Experiments were repeated at least three times, each time in duplicate; standard deviations are shown.

Table 1. Proteins co-eluted with $DgcQ_{cyt}$ in pull-down experiments.

Accession n.°	Protein name	Description	Score[a]		DAve[b]
			-.	uracil	
NP_417225.1	TruD	tRNA(Glu) pseudouridine(13) synthase	0	40.21	200
NP_415404.1	InfA	translation initiation factor IF-1	0	34.28	200
NP_418400.1	TrmA	tRNA m(5)U54 methyltransferase, SAM-dependent	0	30.24	200
NP_417659.1	MlaC	ABC transporter maintaining OM lipid asymmetry, periplasmic binding protein	36.26	0	-200
NP_416470.1	Dcm	DNA cytosine methyltransferase	35.24	0	-200
NP_418666.1	PyrB	aspartate carbamoyltransferase, catalytic subunit	40.3	0	-200
NP_418665.1	PyrI	aspartate carbamoyltransferase, regulatory subunit	40.31	0	-200

Cell extracts were from *E. coli* grown in LB1/4 medium (-) or LB1/4 medium supplemented with 0.25mM uracil (uracil)
Score and DAve are both algorithms of the MAProMa software (42).
[a] Score is a function of the number of uniquely identified peptides in each sample.
[b] DAve provides a relative amount ratio between the two samples. Values of ≥1 or ≤-1 indicate different relative protein amounts between two samples (positive values: higher amount in extracts from cells grown in LB1/4 + 0.25mM uracil; negative values: higher amount in extracts from cells grown in LB1/4 with no uracil supplementation); a DAve value of either 200 or –200 indicates that a given protein is exclusively present in either sample (with or without uracil, respectively).
PyrB and PyrI, the catalytic and regulatory subunits of aspartate carbamoyltransferase, are underlined.

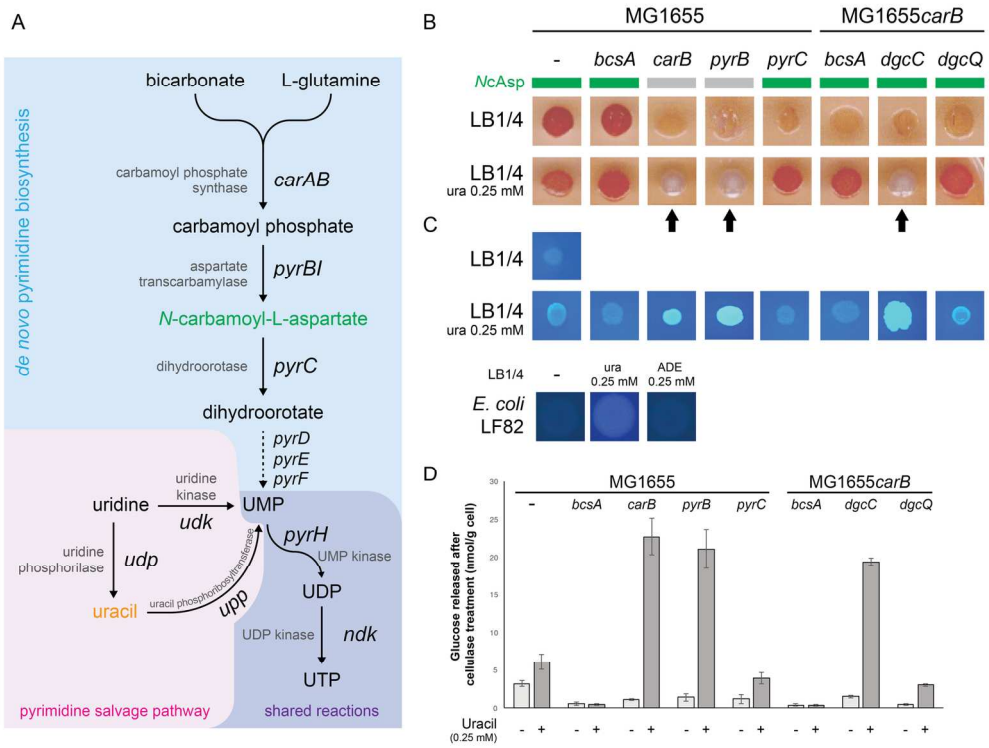


Figure 1

157x121mm (300 x 300 DPI)

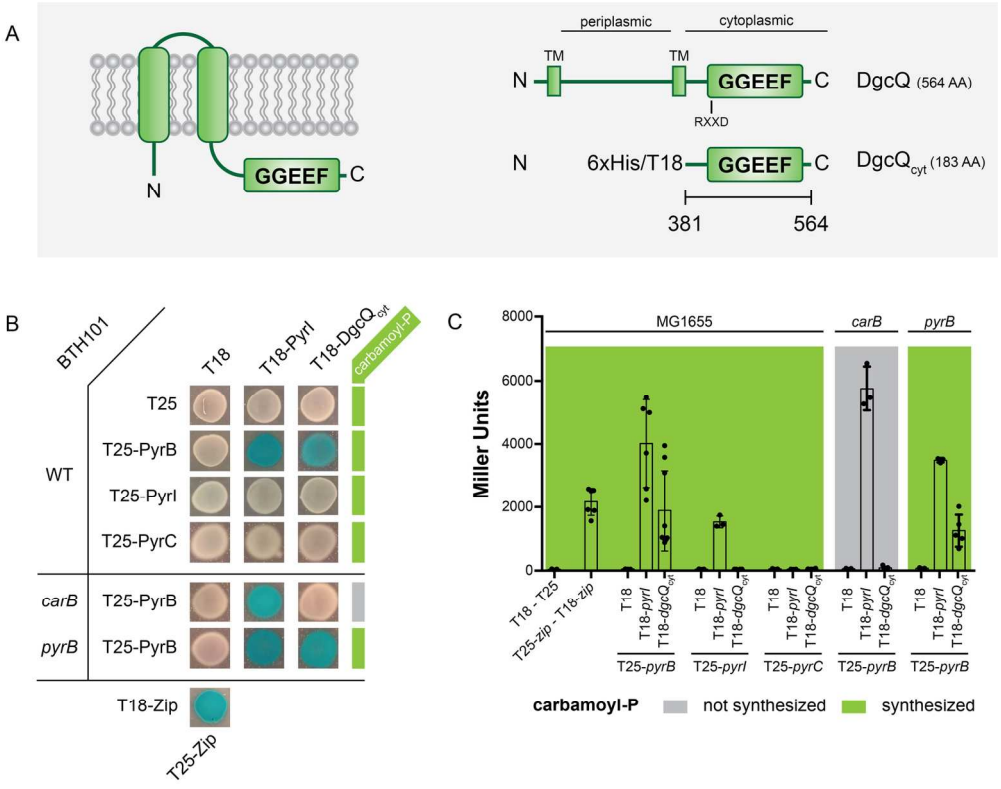
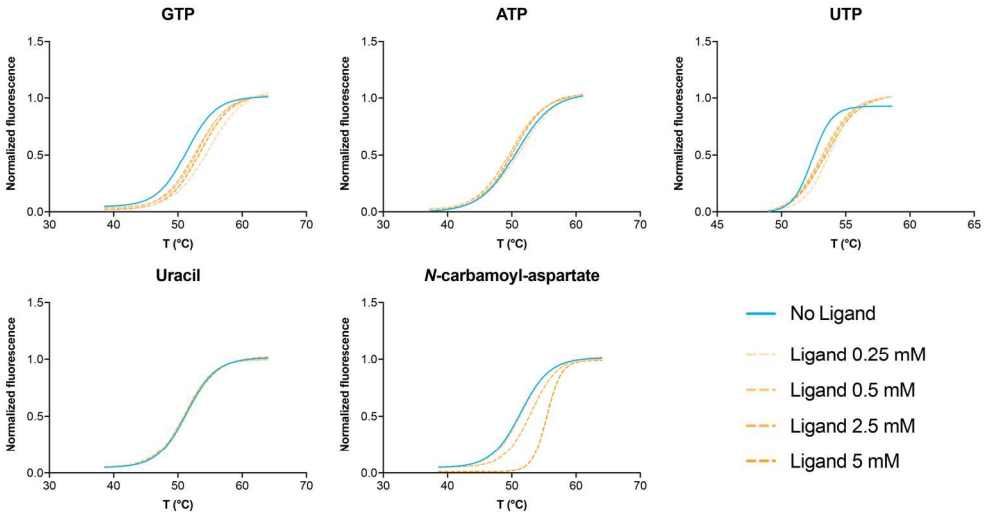


Figure 2

162x132mm (300 x 300 DPI)



ΔT_m ($T_{m\ (ligand)}$ - $T_{m\ (no\ ligand)}$)

	0.25 mM	0.5 mM	2.5 mM	5 mM
GTP	3.6 ± 1.1	1.8 ± 0.1	1.4 ± 0.1	2.1 ± 0.4
ATP	-0.5 ± 0.1	-0.6 ± 0.3	0.1 ± 0.4	0.7 ± 0.4
UTP	1.3 ± 0.5	1.1 ± 0.1	1.2 ± 0.6	1.2 ± 0.3
Uracil	0.0 ± 0.2	-0.5 ± 0.1	-0.4 ± 0.4	-0.1 ± 0.4
N-carbamoyl-aspartate	0.1 ± 0.1	0.0 ± 0.1	1.5 ± 0.1	3.9 ± 0.3

Figure 3

172x149mm (300 x 300 DPI)

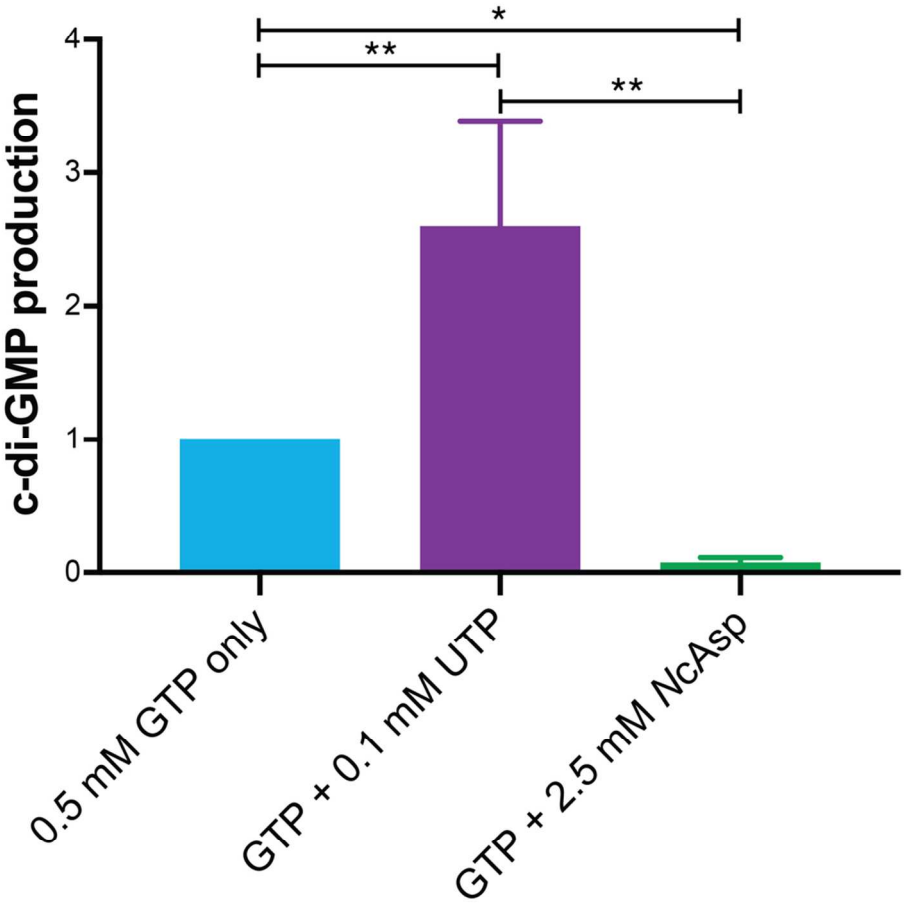


Figure 4

95x95mm (300 x 300 DPI)

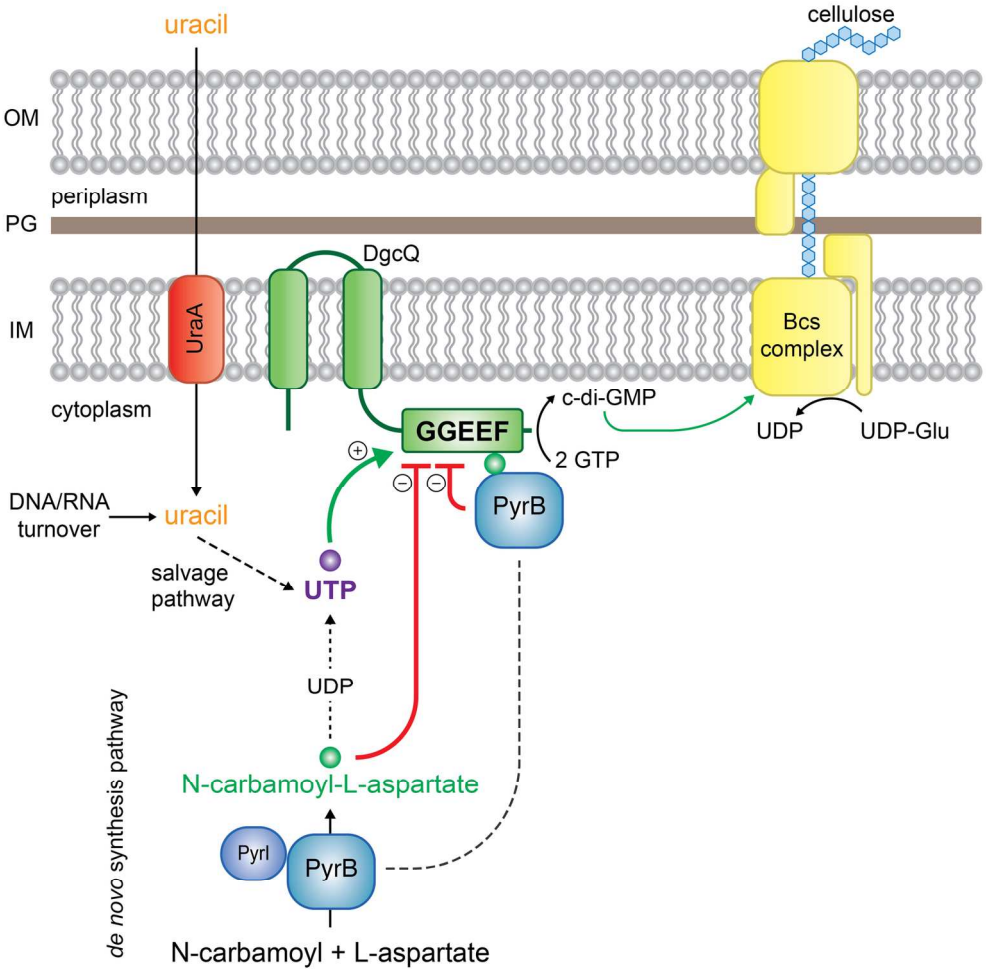


Figure 5

154x152mm (300 x 300 DPI)