Coupling stochastic behavior to metabolism: How a switch protein generates binary signaling programs in *Escherichia coli*

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel
von

Margo Corrine van Berkum

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Urs Jenal
Prof. Dr. Dirk Bumann
Prof. Dr. Matthias Heinemann

Basel, 22.06.2021

Prof. Dr. Marcel Mayor
Dekan

"(...) it is tempting to speculate that mechanisms similar to that described here, based on cyclic diguanylic acid or on related cyclic di- or oligonucleotides, may function in other organisms and other cellular processes."

Ross et al., 1987

My PhD was funded by the 'Biozentrum PhD Fellowships Program' of the University of Basel, formerly known as 'Fellowships For Excellence' (FFE). The last year of my PhD was funded by the Swiss National Science Foundation (SNSF).

Summary

The process of cellular differentiation is vital for the development of multicellular organisms and includes the use of intra- and extracellular signaling molecules that govern precise tissue patterns. In contrast, unicellular organisms like bacteria use intracellular signaling molecules to adapt their behavior and morphology to environmental changes. While cell differentiation often results from hardwired deterministic processes, bacteria can stochastically develop phenotypic variability. Cellular individuality contributes to the fitness of bacterial cultures either because it serves to delegate different functional tasks or because it serves to minimize risks in a rapidly changing environment. Although cell differentiation and stochastic behavior are well-explored on the transcriptional level, it has remained unclear how bacteria convert gradual changes of diffusible small signaling molecules into specific and robust cellular responses, and which mechanisms bacteria exploit to establish stochastic behavior of such regulatory networks. To address these questions, this study investigates c-di-GMP, a signaling molecule that is conserved across all major bacterial phyla and controls important physiological and behavioral processes like surface colonization, virulence or cell cycle progression.

C-di-GMP signaling networks can adopt highly complex architectures with multiple enzymes involved in its synthesis or degradation. In the first results chapter of my thesis, I use *E. coli* to address the question how converging input from multiple enzymes is transformed into robust and unambiguous cellular responses. Together with my collaborators, I demonstrate that *E. coli* makes use of a simple switch to convert dynamic changes of c-di-GMP into discrete binary outputs. This is mediated by an ultrasensitive switch protein, PdeL, which senses the prevailing concentration of the signaling molecule in the cell and couples this information to c-di-GMP degradation and to a transcriptional feedback loop boosting its own expression. We demonstrate that PdeL acts as a digital filter that establishes bimodal populations where individual cells exhibit either high or low c-di-GMP. The observation that PdeL effectively protects *E. coli* against specific bacteriophage predators argues that this molecular switch also serves as a bet-hedging device to minimize life-style specific risks.

The second chapter investigates a regulatory link between metabolism and c-di-GMP turnover. Our studies identified the global metabolic regulator Cra as an essential activator of *pdeL* transcription. This raised the questions if, how and why metabolic processes or growth rate modulate c-di-GMP heterogeneity. Based on my observation that PdeL-mediated c-di-GMP heterogeneity strongly depends on the available carbon source, I set out to probe the role of Cra in this process and in regulating c-di-GMP distributions in individual bacteria. I first developed reporter tools to quantify Cra activity in individual cells. With this in hand, I could demonstrate that when grown on a glycolytic source, low Cra activity results in limited *pdeL* expression and in the loss of c-di-GMP bimodality. However, c-di-GMP bimodality could be restored at least partially by boosting Cra levels or by expressing a constitutively active form of Cra. The results advocate a model in which c-di-GMP heterogeneity is tightly interlinked with metabolic processes to fine tune developmental decisions with the cells' nutrient status and growth rate.

Together, this work provides a molecular frame for how bacteria make use of simple switches to generate stochastic outcomes in signaling processes and how they tune such regulatory elements with their nutritional status. Future studies should investigate how and under which conditions the observed coupling between signaling and metabolism provides optimal fitness benefits for *E. coli*.

Table of Contents

Chapter 1

Introduction

Contents

1.1 Phenotypic heterogeneity

Genetically identical bacterial populations can display substantial phenotypic variation (Elowitz et al., 2002). These variations are proposed to be beneficial during various biological processes (Elowitz et al., 2002) and to fulfil different functional roles (Bettenworth et al., 2019; Veening et al., 2008b). Phenotypic diversification is exploited by many bacterial species, suggesting this widespread feature has evolutionary benefits (Smits et al., 2006; Thattai and Van Oudenaarden, 2004). For example, it could optimize usage of different ecosystem niches, or prepare for unexpected, unknown changes in environmental conditions. Heterogeneity in phenotypic traits and behavior could hence increase the overall population fitness, rather than the fitness of individual cells (Veening et al., 2008b).

1.1.1 Observations of phenotypic heterogeneity

Phenotypic heterogeneity can emerge in response to stress. As resources become scarce, clonal populations can display large variation in gene expression and develop distinct phenotypic traits. For example, sporulation of *Bacillus subtilis* upon entry into stationary phase is only initiated by a part of the population (Chung et al., 1994; Veening et al., 2005). Starved cells show either high or low expression levels of sporulation genes (Veening et al., 2005).

When under homogeneous conditions, some cells from an isogenic population express a certain gene while others do not, two distinct bacterial phenotypes co-exist. The two populations usually represent the two extremes: one population represents the ‘ON’ state and the other the ‘OFF’ state. In this case, plotting the population frequency as a function of the gene expression level results in a bimodal distribution (Ferrell, 2002). Many biological phenomena of bimodality have been described in both prokaryotic and eukaryotic systems (Liu et al., 2018; Paliwal et al., 2007; Shalek et al., 2013; Wang et al., 2020). Bimodal distributions have been observed in growth and gene expression (Nikel et al., 2015; Veening et al., 2005).

In the absence of direct stress, clonal bacterial populations can also display intrinsic heterogeneity. Examples of bimodality in bacteria include, but are not limited to, flagellar gene expression (Kim et al., 2020; Saini et al., 2010; Wang et al., 2020), quorum sensing (Bettenworth et al., 2019), pigmentation (Heinrich et al., 2016), metabolism (Ogura et al., 2020; Westermayer et al., 2016) and growth (Nikel et al., 2015). However, the exact mechanism giving rise to bimodal distributions has not in all cases been elucidated yet.

Heterogeneity in virulence is also a common trait among microbial pathogens (Schröter and Dersch, 2019). Recent studies revealed the occurrence of distinct bacterial subpopulations during infections that differentially influence disease progression and challenge treatment outcomes. For example, antibiotic persistence is due to a subset of dormant bacterial cells and is therefore linked to populations phenotypic heterogeneity. Persistence has been described in many bacterial species (Balaban et al., 2004; Harms et al., 2016). Antibiotic treatment kills most of the cells, but a small subpopulation survives. This subpopulation does not carry resistance markers, but is genetically identical to the population fraction killed during treatment. When antibiotic treatment is arrested, a

genetically sensitive population is restored and a bimodal persistence phenotype is again observed in a new treatment cycle (Brauner et al., 2016; Garcia-Betancur and Lopez, 2019). To date, several molecular mechanisms have been proposed to explain persistence, but controversies exist in the field (Harms et al., 2016).

1.1.2 Functions of phenotypic heterogeneity

Several functions for bimodality and phenotypic heterogeneity have been proposed. The prominent hypothesis is that the diversification in phenotype is used by bacterial populations as a form of bet-hedging, a risk-spreading strategy to increase population fitness (Bettenworth et al., 2019; Veening et al., 2008b). Bistable sporulation in *B. subtilis* was proposed as bet-hedging strategy (Veening et al., 2008a). Whereas sporulation generates highly resistant endospores, non-sporulating cells continue growth and reproduction (Veening et al., 2008a).

A second potential function of heterogeneity is division of labor, in which a more immediate population benefit is displayed (Bettenworth et al., 2019). Here, some cells deploy a certain function from which the whole population benefits via sharing, joint use or cross feeding (Schröter and Dersch, 2019). This could be an energetically costly function and hence diversification is used as a cost-efficiency strategy. Division of labor might be the strategy behind heterogeneity in the production of extracellular matrix (Veening et al., 2008b) or secondary metabolites as toxin and other bioactive compounds (García-Betancur and Lopez, 2019; Tobias and Bode, 2019).

1.1.3 Characteristics of regulatory networks

Given its wide-spread occurrence, how does phenotypic variation arise from identical genetic information? Populations naturally display cell-to-cell variability in gene expression levels. Stochastic fluctuations, also called noise, in transcription and translation are proposed to give rise to this variability (Elowitz et al., 2002; Smits et al., 2006).

Noise in gene expression can have different origins (Elowitz et al., 2002; Swain et al., 2002). First, the discrete nature of biochemical processes makes gene expression inherently noisy. Additionally, stochasticity arises from cell to cell variation in the concentration and activity of associated components such as transcriptional activators and repressors, RNA polymerase, ribosomes and degradation pathways (Swain et al., 2002). Some of these components exist only in small amounts per cell and this particularly contributes to noise via the ‘finite number effect’. The finite number effect describes that the noise in a process increases when the components of this process are present in only limited amounts.

Regulatory networks can be wired as such that noise is amplified. Through the amplification of stochastic fluctuations, two or more distinct expression states can arise within a clonal population and homogenous environment.

The regulatory systems that generate this response heterogeneity may include:

1. One or multiple feedback loops, including positive, double positive and double negative feedback loops (Smits et al., 2006; Westermayer et al., 2016);

2. One or multiple elements of non-linearity, for example protein multimerization or cooperativity in DNA binding (Nikel et al., 2015)

Feedback loops have the ability to convert a graded response into a switch-like response (Ninfa and Mayo, 2004) and to generate expression patterns characterized by two distinct states, with intermediate states rarely observed (Fig. 1).

In many cases, feedback loop regulation drives bimodal activity of key master regulators, causing bimodal expression of downstream genes (Ogura et al., 2020). A role for positive feedback has been demonstrated for the *B. subilitis* transcription factor ComK, the master competence regulator that is conserved in other Gram-positive bacteria such as *Streptococcus pneumoniae* and *Staphylococcus aureus*. ComK can bind to its own promoter region and this stimulates transcription (Hamoen et al., 1998; Van Sinderen & Venema, 1994). This positive transcriptinal feedback loop is required for competence heterogeneity (Smits et al., 2006). Additionally, ComK may be controlled by a double negative feedback loop, as ComK and a second transcription factor Rok were found to repress transcription of each other by promoter binding. The double negative feedback loop may contribute to bimodal expression of the competence system (Smits et al., 2006).

Regulation by mutual repression is also involved in bimodal flagellar gene expression in *Salmonella* species. Flagellar genes are structured in a regulatory cascade of different classes that are sequentially expressed. Bimodal expression was observed both in class II and class III genes. In *Salmonella*, a double negative feedback loop gives rise to bimodal expression of class II flagellar genes. The two proteins RflP and FliZ (in)directly mutually repress each other. As a consequence, two stable states can exist, one in which RflP is high and FliZ is low, and the opposite state in which RflP is low and consequently FliZ is high (Koirala et al., 2014; Saini et al., 2010). Bimodal expression of class III flagellar genes in *Salmonella* arises from a very sharp threshold that is set for excretion of an inhibitory regulatory protein. Variations in excretion rate on a population level could give rise to class III bimodality (Wang et al., 2020).

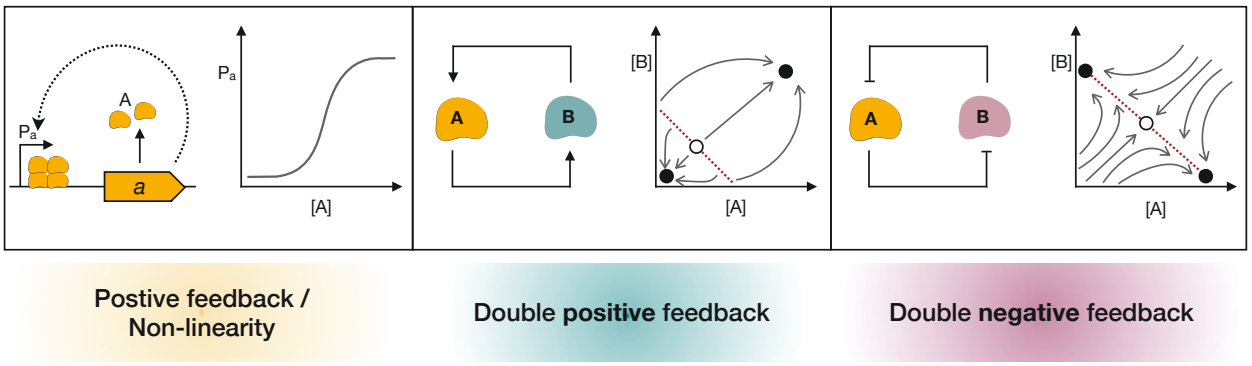


Figure 1: Characteristic regulatory elements of bimodal / bistable systems. **Left**: a bistable switch established by single positive feedback and protein multimerization. **Middle**: two components positively regulate each other, creating double positive feedback. Stable steady states exist either as both A and B off, or both A and B on (black circles). Intermediate states (white circle) are not stable. **Right**: two components negatively regulate each other. Stable steady states exist either as A on and B off, or with A off and B on (black circles). Intermediate states (white circle) are not stable. Adapted from: Ferrell et al., 2002; Smits et al., 2006.

1.2 Cyclic di-GMP

It is crucial for living cells to sense their environment and to adapt to changing conditions in order to survive (Barthe et al., 2020; You et al., 2013). From a population perspective, cells can respond in different ways including deterministically or stochastically, as discussed above (Section 1.1). For a response to take place, first the stimulus has to be transferred to the precise cellular components that in turn generate a response. Such a response could consist of (de)activation of a transcription factor or enzyme, a change in gene expression, or other. Eventually, a change in phenotype or behavior arises via downstream targets (Madigan et al., 2010).

Cells have intracellular signal transduction pathways in place to transfer stimuli and to adapt cellular processes accordingly (Madigan et al., 2010). Relay mechanisms commonly used in these pathways are phosphorylation cascades, including two-component systems and multi-step phosphorelays (Hoch, 2000; Perraud et al., 1999; Schaller et al., 2008), and signal transduction via second messengers (Jenal et al., 2017; Thompson and Malone, 2020) (Fig. 2). Bacterial second messengers are primarily purine-derived modified nucleotides (Yoon and Waters, 2021). The concentration of these nucleotide-based signaling molecules (NSMs) is controlled by enzymes that synthetize or degrade NSMs in response to a 'first' signal. NSMs enable integration and amplification of stimuli and precise control of downstream targets (Whiteley et al., 2019). NSMs can function as global regulators and induce cellular changes by binding as allosteric regulator to effector molecules (Pesavento and Hengge, 2009). Hence NSMs actively modulate cellular physiology and behavior. Their usage is widely spread among both prokaryotic and eukaryotic organisms. Different groups of NSMs can be distinguished (Fig. 2). The roles of cyclic nucleotides as cyclic AMP (cAMP) and cyclic GMP (cGMP) in signaling have been largely studied (Botsford and Harman, 1992; Daniel et al., 1998; Lucas et al., 2000; You et al., 2013). The group of cyclic dinucleotides (CDNs) consists of cyclic di-GMP (c-di-GMP) (Jenal et al., 2017; Ross et al., 1987), cyclic di-AMP (c-di-AMP) (Witte et al., 2008) and cyclic GMP-AMP (cGAMP) (Davies et al., 2012; Wu et al., 2013). NSMs also exist in linear form, such as the alarmones guanosine tetraphosphate (ppGpp) and pentaphosphate (pppGpp) (Cashel and Gallant, 1969; Steinchen et al., 2020). The related (p)ppApp is currently an emerging class of NSMs, although first detected in the late '70 (Bruhn-Olszewska et al., 2018; Nishino et al., 1979; Sobala et al., 2019; Steinchen et al., 2020). The evidence is growing for the existence of the CDNs cyclic UMP-AMP and cyclic di-UMP and for the existence of cyclic trinucleotides such as cyclic tri-AMP and cyclic AMP-AMP-GMP (Whiteley et al., 2019; Yoon and Waters, 2021). This plethora of NSMs coordinate phenotypes and cell fate across species and kingdoms. Central questions about NSM networks are: what are the precise cues and mechanisms for NSM synthesis and degradation? What phenotypes do NSMs control and how would these phenotypes contribute to bacterial adaptation (Yoon and Waters, 2021)? The focus in this work is on cyclic di-GMP, which is the most widespread bacterial CDN (Jenal et al., 2017).

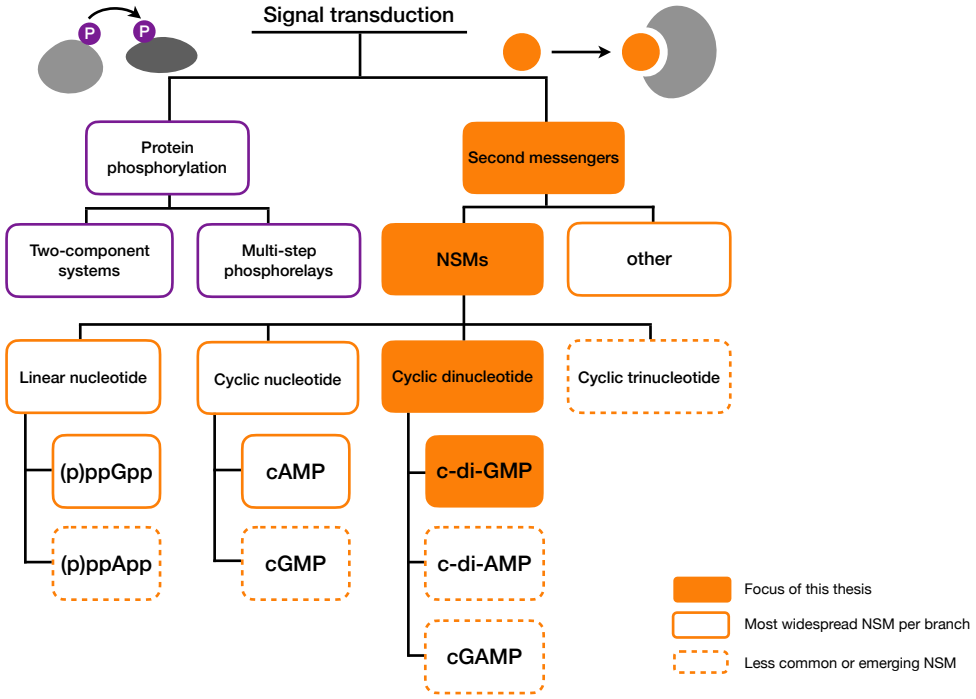


Figure 2: Diversity in intracellular signal transduction pathways. Relay mechanisms used in intracellular signal transduction pathways are phosphorylation cascades (purple boxes) and second messengers (orange boxes). Protein phosphorylation can further be distinguised in two-component systems and multi-step phosphorelays. One category of second messengers is nucleotide-based signaling molecules (NSMs). A large diversity of NSMs has so far been identified (see also text). NSMs can be futher classified in linear and cyclic mono-nucleotides; cyclic dinucleotides; and cyclic trinucleotides. Solid, orange lines: most widespread NSM of its class. Dashed lines: less common or emerging NSM. Filled, orange boxes: focus of this thesis.

1.2.1 Near-ubiquitous bacterial second messenger

The bacterial intracellular signaling molecule bis-(3'-5')-cyclic diguanylic acid (cyclic di-GMP or c-di-GMP) is conserved across all major bacterial phyla and is considered a near-ubiquitous second messenger (Conner et al., 2017; Jenal et al. 2017). It was the first CDN to be described and today considered to be the best studied (Jenal et al., 2017; Ross et al., 1987). Controlling central aspects of growth and behavior including motility, biofilm formation and virulence, c-di-GMP plays a key role in dictating bacterial phenotypes.

C-di-GMP was first identified almost 35 years ago as the specific activator of a membrane-bound cellulose synthase in the species *Komagataeibacter xylinus* (reclassified from *Gluconacetobacter xylinus*, which was reclassified from *Acetobacter xylinum*) (Ross et al 1987; Yamada et al., 1997). Using NMR and mass spectroscopy, Ross et al. provided evidence for this cyclic structure to be c-di-GMP. The authors mentioned that '*it is tempting to speculate that mechanisms similar to that described here, based on cyclic diguanylic acid or on related cyclic di- or oligonucleotides, may function in other organisms and other cellular processes*' (Ross et al., 1987). In the decades that came after, NSMs have been intensely studied. The numerous discoveries on c-di-GMP- and other NSM-dependent processes across species only confirm their writing.

The regulatory input of c-di-GMP in cellular processes is pleiotropic and extremely diverse. Intracellular c-di-GMP concentrations control growth and behavior processes including motility, exopolysaccharide production, cell morphology, quorum sensing, chemotaxis, virulence, cell cycle progression, metabolism and phage resistance (Boehm et al., 2010a; Fernandez et al., 2020; Hochstrasser and Hilbi, 2020; Karaolis et al., 2005; Liu et al., 2010; Lori et al., 2015; Mutalik et al., 2020; Nesper et al., 2017; Xu et al., 2019). C-di-GMP interferes on the transcriptional, post-transcriptional and post-translational level, as its binding effectors include transcription factors, kinases, proteases and RNA molecules (Fig. 3). The binding of one or several molecules of c-di-GMP modulates target activity or function through allosteric interaction or induction of riboswitches (Romling et al., 2013). The intensity of the effect can among others be tuned by the binding affinity of the effector molecule for c-di-GMP, indicating a multi-level network of fine-tuning mechanisms (Kunz and Graumann, 2020).

A known c-di-GMP binding domain for c-di-GMP-protein interaction is the PilZ domain (Amikam and Galperin, 2006). The structures of several PilZ domains and PilZ-domain containing proteins have been resolved from a variety of bacterial species, including but not limited to *Pseudomonas aeruginosa* (Habazettl et al., 2011), *Rhodobacter sphaeroides* (Morgan et al., 2014), *Klebsiella pneumoniae* (Schumacher and Zeng, 2016) and *Bacillus subtilis* (Subramanian et al., 2017). The elucidation of PilZ domain structures has contributed significantly to the understanding of the c-di-GMP binding mechanism and role of the conserved binding motifs. Through a comparative analysis of PilZ sequences and structures, their c-di-GMP binding motifs were recently redefined as RXXXR and [D/N]hSXXG. The same analysis allowed further classification of the PilZ domain into three branches (Galperin and Chou, 2019).

Besides PilZ domains, only few c-di-GMP binding sites can be predicted based on protein sequence, including MshEN domains (Roelofs et al., 2015; Wang et al., 2016), GIL domain (Fang et al., 2014), allosteric inhibitory sites (i-sites) and catalytic domains and their degenerate derivatives (Mouali et al., 2017; Romling et al., 2013). Probably because the flexibility of the c-di-GMP molecule allows a high number of possible interactions, many c-di-GMP binding sites cannot be predicted based on protein sequence. This makes a complete overview of c-di-GMP effector molecules challenging (Chou and Galperin, 2016; Valentini and Filloux, 2019). Complimentary methods are used to systematically extend the list of c-di-GMP binding sites (Düvel et al., 2012, 2016; Laventie et al., 2017; Nesper et al., 2012).

1.2.2 C-di-GMP synthesis and degradation

The c-di-GMP concentration is controlled by the net activity of c-di-GMP synthetizing and degrading enzymes. C-di-GMP is synthesized through the condensation and cyclization of two guanosine triphosphate (GTP) molecules catalyzed by diguanylate cyclases (DGCs). DGCs contain a catalytic GGDEF domain, named after the most common active side motif, GGDEF (Simm et al., 2004). Less common motifs in active DGCs are GGEEF (Bandekar et al., 2017; Chan et al., 2004), AGDEF (Hunter et al., 2014) and SGDEF (Romling et al., 2013). They represent circa 30% (GGEEF), 1.5.% (AGDEF) and 1.7% (SGDEF) of over 27 thousand GGDEF domains

identified in bacteria by 2011 (Romling et al., 2013). In approximately half of the catalytically active GGDEF domains, an auto-inhibitory site (i-site) is found in addition to the catalytic motif (Romling et al., 2013; Seshasayee et al., 2010). The i-site residues contain an RxxD motif that can bind c-di-GMP and allosterically inhibits cyclase activity (Dahlstrom and O'Toole, 2016; Romling et al., 2013). GGDEF domains function as homodimers (Dahlstrom and O'Toole, 2016). As both protomers bind one GMP molecule, dimerization is required for the synthesis of a c-di-GMP molecule (Paul et al., 2007; Romling et al., 2013). A c-di-GMP molecule may subsequently diffuse through the cell and/or bind a c-di-GMP effector that upon binding initiates a downstream reaction (Dahlstrom and O'Toole, 2016) (Fig. 3).

Two types of c-di-GMP specific phosphodiesterases (PDEs) exist for the degradation of the second messenger. PDEs with an EAL motif are highly abundant and degrade c-di-GMP into the linear pGpG molecule (Schmidt et al., 2005; Simm et al., 2004). Most EAL PDEs found so far form dimers or oligomers, although monomeric EAL domains show some PDE activity (Schmidt et al., 2005). Dimerization, however, appears critical for PDE activity regulation by environmental stimuli (Bai et al., 2012; Johnson et al., 2011). A second, far less abundant domain found in active PDEs is the HD-GYP domain, that degrades c-di-GMP into two GMP (Galperin et al., 1999, 2001; Ryan et al., 2009) (Fig. 3).

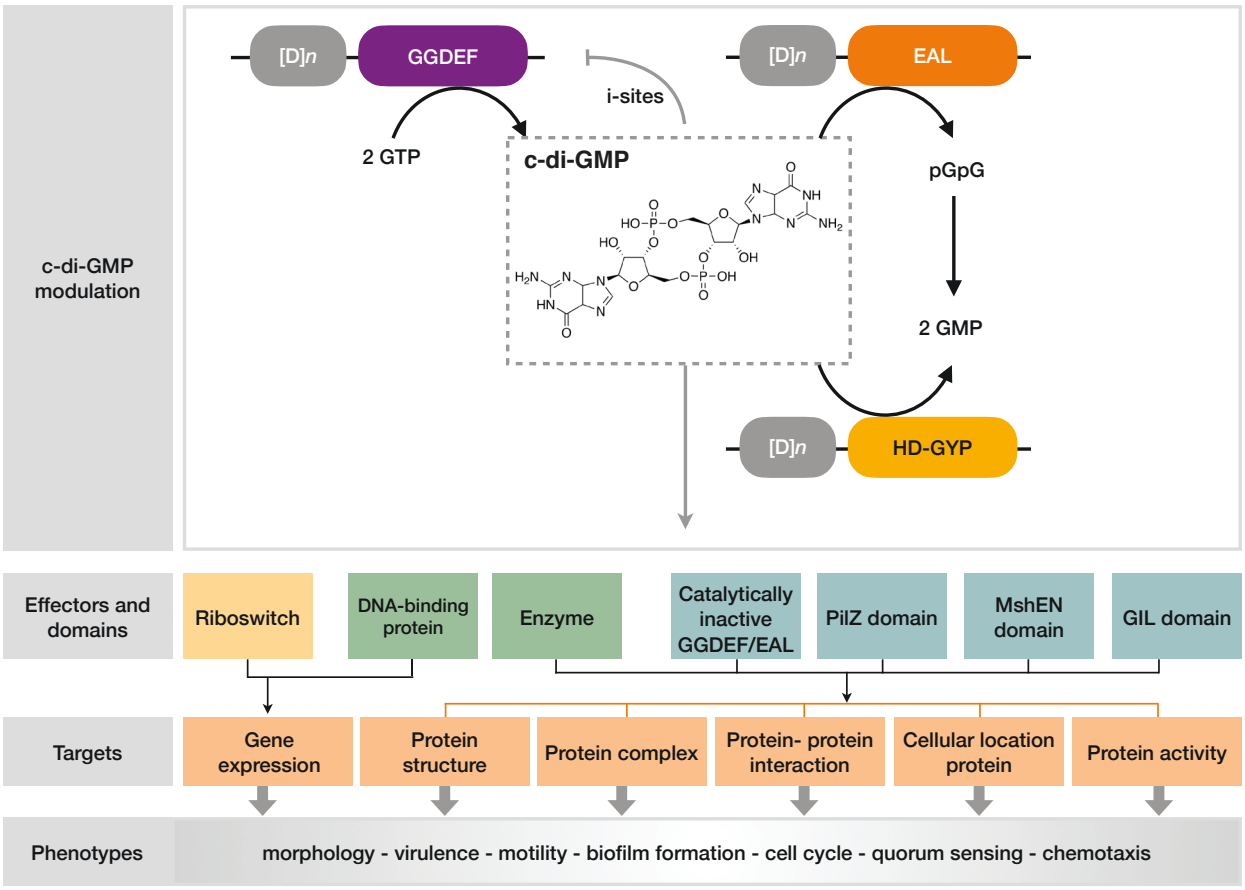


Figure 3: C-di-GMP signaling: modulators, effectors, targets and phenotypes. C-di-GMP is synthesized by diguanylate cyclases with a catalytic GGDEF domain and degraded by c-di-GMP specific phosphodiesterases characterized by an EAL (most abundant) or HD-GYP domain. Effectors of c-di-GMP include RNA and proteins with divers c-di-GMP binding domains. Binding of c-di-GMP to its effectors modulates activity of downstream targets, and thus affects bacterial phenotypes. Adapted from: Valentini and Filloux (2019)

Most bacterial genomes encode for multiple DGCs and PDEs ranging from a few to multiple dozens of catalytically active GGDEF and/or EAL domain proteins. Comparison of bacterial genome sequences reveals a high degree of conservation and almost ubiquitous presence of DGCs and PDEs (Romling et al., 2013; Chou and Galperin, 2016). The total number of GGDEF- and EAL-domain proteins per species is usually larger, as their genomes encode catalytically inactive, or degenerate, GGDEF- and EAL-proteins. These could have evolved from catalytic domains into c-di-GMP binding or other functional domains (Navarro et al., 2009; Simm et al., 2009; Whitney et al., 2012).

Active DGCs and PDEs often contain additional protein domains to the catalytic domain (Fig. 3). Taking the different bacterial phyla together in which GGDEF-, EAL- and HD-GYP proteins have been identified, the most common domain architectures include PAS-GGDEF(-EAL), GAF-GGDEF(-EAL), (REC-)REC-GGDEF, REC-EAL and REC-HD-GYP (Romling et al., 2013). Other N-terminal domains include CACHE, CHASE, CZB, GAPES or MASE sensory domains or DNA-binding domains (Dahlstrom and O'Toole, 2016; Hengge et al., 2019).

1.2.3 C-di-GMP signaling specificity

Up to several dozens of PDEs and DGCs can be transcribed from a single bacterial genome (Romling et al., 2013). The activity of these enzymes, together with a range of c-di-GMP effectors, give rise to distinct and highly coordinated phenotypes. The multiplicity of DGCs and PDEs in single species gave rise to questions about the potential impact these enzymes have on c-di-GMP effectors and eventually on c-di-GMP dependent behavior (Sommerfeldt et al., 2009): do these enzymes control a common, i.e. global cellular pool of c-di-GMP, that via diffusion modulates a common set of targets? Or can the different c-di-GMP modulating enzymes act independently and target specific effectors?

Many studies have focused on elucidating the organization and integration of c-di-GMP signals. Certain c-di-GMP modulating enzymes were reported to only control specific c-di-GMP-dependent processes. For example, the DGCs DgcE and DgcM and the PDE PdeR in *E. coli* affect biofilm formation without altering overall cellular c-di-GMP levels (Lindenberg et al., 2013; Pesavento et al., 2008; Sarenko et al., 2017; Serra and Hengge, 2019). Along with the growing evidence for signaling specificity in the c-di-GMP network, different models have been proposed for how this could be established (Fig. 4):

1. **Signaling specificity through temporal control and differential affinity in a global signaling network**

Signaling specificity could first of all be achieved through differential expression and activation of DGCs, PDEs and c-di-GMP effectors. As their transcription is controlled by different transcription factors and sigma factors, expression is shown to occur in a temporally and environmentally controlled manner (Ko and Park, 2000; Ryjenkov et al., 2006; Sommerfeldt et al., 2009). Through their additional protein domains, activity of DGCs and PDEs can subsequently depend on specific cues required for their (de)activation. Known activation mechanisms include phosphorylation and small molecule binding (e.g. cyclic AMP, oxygen, metals) at one or multiple protein domains

(Paul et al., 2007; Tuckerman et al., 2009; Vasconcelos et al., 2017; Zähringer et al., 2013). Effectors are suggested to contribute to signaling specificity through variation in their c-di-GMP binding affinity, that can differ by more than 140-fold between different effectors in the same species (Pultz et al., 2012). Thus, c-di-GMP signaling specificity is established through temporal control and variation in effector affinity.

2. **Signaling specificity through local c-di-GMP signaling**

In addition to specificity within global c-di-GMP networks, several studies support models of local c-di-GMP signaling, established through protein interactions or proximity of c-di-GMP modulating enzyme(s) and effector. These models could further explain how specificity is achieved and crosstalk avoided during simultaneous expression of c-di-GMP modulating enzymes and effector molecules. Local signaling modules have been suggested to function in different organisms including *Pseudomonas fluorescens* (Dahlstrom et al., 2015, 2016), *E. coli* (Lindenberg et al., 2013; Richter et al., 2020; Sarenko et al., 2017), *Caulobacter crescentus* (Abel et al., 2013; Lori et al., 2015) and *B. subtilis* (Kunz et al., 2020). Recently, three criteria were proposed to define local c-di-GMP signaling: i) a specific DGC/PDE knockout phenotype, ii) direct interaction between the specific PDE and/or DGC and the c-di-GMP effector or target and iii) global cellular c-di-GMP levels remain unchanged or below the binding affinity of the activated modules (Hengge, 2021). To limit crosstalk, activity of a single and strongly expressed PDE might possibly contribute to isolating local signaling modules and sustaining specificity (Jenal et al., 2017; Sarenko et al., 2017).

Global and local c-di-GMP signaling can very well be combined in a single cell (Hengge, 2021). These elements combined provide bacteria with flexibility and adaptability in complex signaling networks (Hengge, 2021).

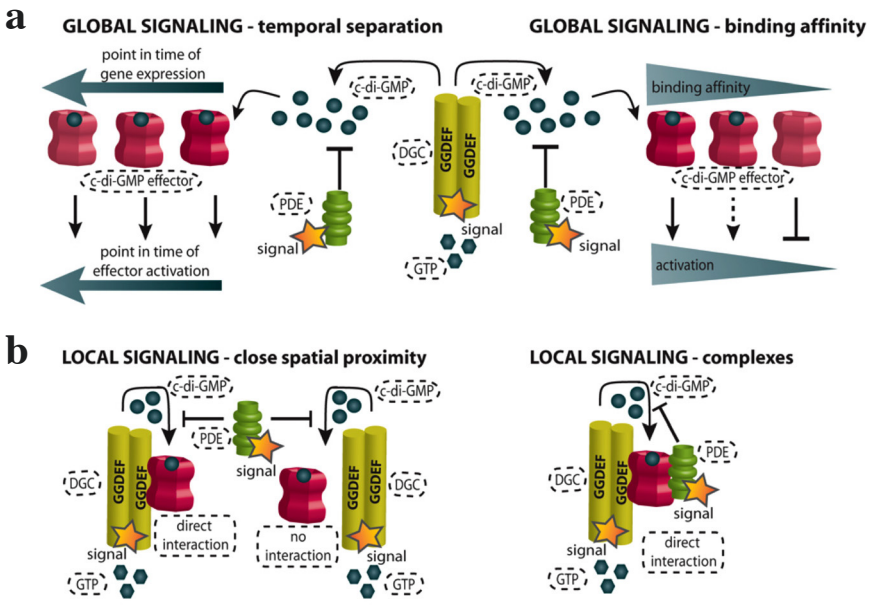


Figure 4: C-di-GMP signaling specificity explained by global and local signaling models. Spatial and temporal activitation of DGCs (yellow), PDEs (green) and effectors (magenta). **(a)** signaling specificity in global c-di-GMP networks could be established through i) temporal separation in expression and activation of PDE, DGC and effector; and ii) variation in effector binding affinity. **(b)** models of local signaling, involving local c-di-GMP pools, to explain c-di-GMP signaling specficity. Signal specificity might be achieved by close spatial proximity or protein interaction of DGC, effector and PDE. Adapted from: Kunz and Graumann (2020)

1.2.4 C-di-GMP-controlled physiology and behavior

Below I will further explain some of the major c-di-GMP-controlled processes from the bacterial kingdom, to illustrate the near-ubiquitous nature of this second messenger.

1.2.4.1 Motile-sessile transitions

Most of bacterial life on earth exists as biofilms, multicellular structures protecting populations against adverse effects including chemical and biological threats (Hall-Stoodley et al., 2004). Stress factors such as depletion of nutrients and/or oxygen can trigger the transition from a motile to sessile state (Hengge, 2020).

The first process in which c-di-GMP was identified as key regulator, cellulose synthesis (Ross et al., 1987), is a classic example of the role of c-di-GMP in these motile-sessile transitions. C-di-GMP functions as pleiotropic regulator in many bacterial species including *Escherichia*, *Pseudomonas*, *Salmonella* and *Vibrio*. Here, low c-di-GMP levels usually favor cells to be in a highly motile state, allowing expression and functioning of motility organelles. On the contrary, high cellular c-di-GMP stimulates the sessile state through activation of biofilm formation. For example, *P. aeruginosa* can produce different types of exopolysaccharide for biofilm development (Colvin et al., 2012; Ryder et al., 2007). Besides alginate, it produces polysaccharides by expressing the genes in the *pel* and *psl* locus. Transcription of *pel* and *psl* is highly regulated and inhibited by among others the transcription factor FleQ. C-di-GMP directly binds FleQ (Hickman and Harwood, 2008) and this relieves FleQ-dependent repression of *pel* and *psl*. FleQ was already identified as the master regulator of flagella gene expression in *P. aeruginosa* (Arora et al., 1997; Dasgupta et al., 2003), showing a direct reverse control. C-di-GMP also controls the motile to sessile transition in the facultative human pathogen *V. cholerae*, that is known for the cause of cholera disease. In *V. cholerae*, c-di-GMP binds and activates VpsR and VpsT. These transcription factors induce exopolysaccharide synthesis. At the same time, c-di-GMP inhibits flagellar biosynthesis on the transcriptional and post-transcriptinal level via VpsT-dependent and -independent pathways involving master regulator FlrA and additional regulators as TfoY (Srivastava et al. 2013; Pursley et al. 2018).

1.2.4.2 Cell cycle

C-di-GMP controls cell cycle progression in certain species. As example, the aquatic organism *C. crescentus* displays an asymmetric and clearly c-di-GMP-dependent division cycle. During *C. crescentus* cells division, two distinct cell types are formed. One cell pole contains a stalk and adhesive holdfast, keeping this cell attached to the surface after division. At the opposite pole of the dividing cell, the motility and chemotaxis apparatus are assembled. After division, the stalked cell commences a new DNA replication cycle (S phase) followed by cell division (G2 phase) (Fig. 5). In contrast, cell cycle is blocked into G1 phase in the motile cell until it transforms into a stalked cell. A stalked cell, as described, can undergo a new asymmetric division cycle (Abel et al., 2011, 2013; Kirkpatrick and Viollier, 2012; Lori et al., 2015). How is this process coupled to c-di-GMP?

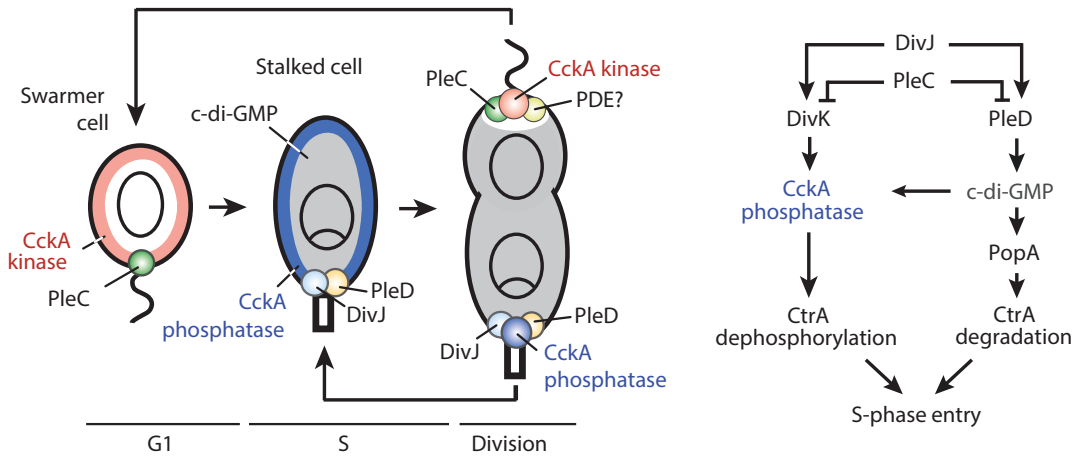


Figure 5: C-di-GMP regulates the cell cycle of *Caulobacter crescentus*. Left, polar localization c-di-GMP modulating enzymes and the c-di-GMP effector protein CckA, in different stages of *C. crescentus* cell cycle. CckA displays kinase activiy (red) at low c-di-GMP (white cytoplasm) and phosphatase activity (blue) at high c-di-GMP (grey cytoplasm). **Right**: schematic of the regulatory cascade controlling S-phase entry via the replication inhibitor CtrA. Adapted from: Lori et al., 2015

Polar localization of DGCs, PDEs and c-di-GMP effectors induces distinct cellular programs at either pole, that are decisive for replication initiation. At the stalked pole, the DGC PleD is the major generator of c-di-GMP. PleD requires phosphorylation for its catalytic activity. Here, the histidine kinase DivJ co-localizes at the stalked pole to activate PleD through phosphorylation (Abel et al., 2013; Paul et al., 2008). The CckA kinase-phosphatase switch localizes at both poles (Fig. 5). One of its targets is CtrA, a transcription factor which, when phosphorylated, blocks chromosome replication (Kirkpatrick and Viollier, 2012). When active as kinase, CckA phosphorylates CtrA. Studies have provided evidence that c-di-GMP can bind the CckA protein. CckA in complex with c-di-GMP predominantly functions as phosphatase (Dubey et al., 2016; Lori et al., 2015). The elevated c-di-GMP levels at the stalked pole therefore favor CckA phosphatase activity. CckA inactivates CtrA through dephosphorylation. Release of replication inhibition makes the cell transit into S-phase. A parallel program involving several protein modules keeps c-di-GMP levels reduced at the flagellated pole and hence maintain CckA kinase activity (Lori et al., 2015). It is proposed that asymmetry in the c-di-GMP distribution across dividing cells creates microenvironments of low and high c-di-GMP, dictating distinct CckA activity at either pole. C-di-GMP oscillates throughout the cell cycle and precisely times morphology and cell cycle (Lori et al., 2015).

1.2.4.3 Virulence

There is increasing evidence for c-di-GMP to control the expression of multiple virulence factors in pathogenic bacterial species. The relevance of c-di-GMP signaling in infections is so high, that controlling c-di-GMP by means of antimicrobial targets is investigated as treatment method to improve disease outcome. How exactly c-di-GMP controls virulence varies however widely among bacterial species. A clear trend as for motile-sessile transitions is not the case, possibly because virulence is a dynamic and multifactorial process (Valentini and Filloux, 2019). Pathogenic bacteria would need to establish colonization, invasion and immune system evasion by controlling among others the expression of secretion systems, fimbriae, adhesins or flagella. For example, c-di-GMP activates type IV pili biogenesis in *Clostridium difficile* through a riboswitch (Bordeleau et al., 2015)

and in *P. aeruginosa* through protein binding (Guzzo et al., 2009). In uropathogenic *E. coli* (UPEC), type I fimbriae contribute to initial cell adhesion and virulence. High cellular c-di-GMP levels were found to reduce type I fimbriae expression and attenuate virulence (Crépin et al., 2017), indicating the importance of c-di-GMP control in virulence traits of UPEC. Overall, precise c-di-GMP control appears essential for successful host colonization and proliferation.

1.2.4.4 Cell morphology

C-di-GMP also affects cell morphology and drives cell morphological changes. Besides *C. crescentus*, it is relevant to describe *V. cholerae* as an example because of its characteristic shape of a curved rod (Conner et al., 2017). It was recently found that high intracellular c-di-GMP levels reduce the cell's curvature, resulting in more rod-shaped *V. cholerae* cells (Fig. 6, Fernandez et al., 2020). Crva is a self-assembling periplasmic filament that generates cell curvature and promotes pathogenesis of *V. cholerae* (Bartlett et al., 2017). High c-di-GMP levels were found to reduce the *crvA* mRNA level and hence Cvra, in a VpsT-dependent manner (see also Section 1.2.4.1 and Fig. 6). Interestingly, the two morphological states were found to be beneficial for the respective behavior types, as curved cells demonstrated to be faster swimmers while rod shape cells were optimal for microcolony and biofilm formation (Fernandez et al., 2020).

1.2.5 C-di-GMP regulation in *Escherichia coli*

The work in this thesis focuses on c-di-GMP signaling in *Escherichia coli* and primarily on the regulation of the PDE PdeL and effects of PdeL activity on *E. coli* physiology. To provide a context, I'm outlining c-di-GMP signaling in *E. coli* and describing key players in c-di-GMP dependent processes.

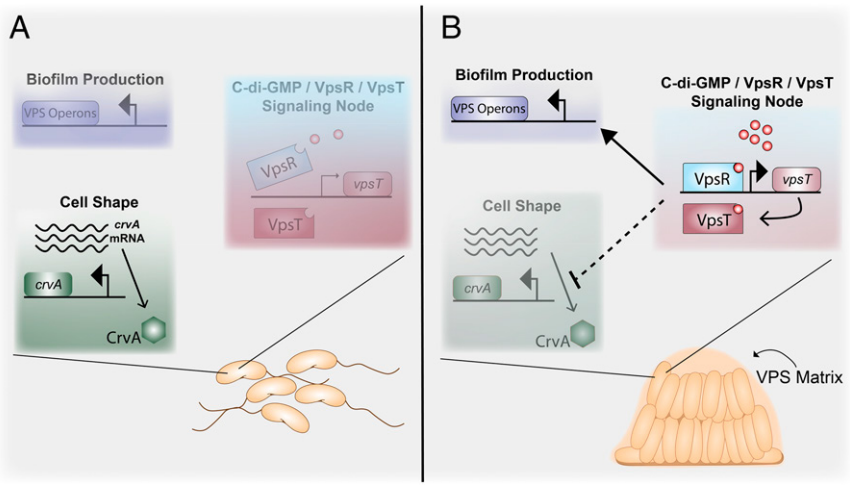


Figure 6: Cell morphology and reverse regulation of motility and matrix formation in *Vibrio cholerae*. (A) Under low c-di-GMP conditions, c-di-GMP effector protein VpsR is inactive, keeping VspT levels low and preventing expression of biofilm operons. CvrA levels are high, hence *V. cholerae* cells are flagellated and display a curved rod shape. **(B)** High c-di-GMP levels activate VspR. Expression of *vpsT* and biofilm operons is induced. *crvA* mRNA levels are reduced in a VpsT-dependent manner, resulting in cells with a straight rod shape. Adapted from: Fernandez et al., 2020

1.2.5.1 C-di-GMP synthesis and degradation in *E. coli*

As for other *Enterobacteriaceae*, c-di-GMP controls growth and behavior in *E. coli* (Romling et al., 2013). The *E. coli* K-12 MG1655 strain is a widely used laboratory strain and model organism. Its genome (Blattner et al., 1997) encodes 12 active diguanylate cyclases (DGCs), 13 active c-di-GMP specific phosphodiesterases (PDEs) and an additional four proteins with enzymatically inactive GGDEF or EAL domains (Hengge et al., 2015).

Almost all GGDEF / EAL domain proteins, except the PDE PdeH, contain an N-terminal sensory domain. The majority of these proteins is localized at the inner membrane, either through a diverse range of membrane associated sensory domains (GAF, PAS, MASE1/2/4/5, CHASE7/8, GAPES1/2/3) or through a periplasmic sensory domain (CSS; in almost half of *E. coli* K-12 PDEs) (Hengge et al., 2015).

In commensal and pathogenic *E. coli* strains, thus far two GGDEF domain proteins and four EAL domain proteins were identified which are absent in the *E. coli* K-12 genome (Branchu et al., 2013; Povolotsky and Hengge, 2016; Richter et al., 2014; Sjöström et al., 2009).

With the expansion of DGCs and PDEs identified in *E. coli* strains, a new and systematic nomenclature was proposed for all *E. coli* DGCs and PDEs. In the current nomenclature, gene names of functional DGCs and PDEs start with respectively ‘dgc’ or ‘pde’ followed by a single letter (Hengge et al., 2015).

1.2.5.2 Motility

An *E. coli* cell can produce five to six flagella, displaying a peritrichous flagellation pattern over its cell surface (Leifson, 1960; Schuhmacher et al., 2015). This in contrast to some other bacterial species, including *C. crescentus* and *P. aeruginosa*, that only assemble a single flagellum (Schuhmacher et al., 2015). Flagella enable the bacterium to swim in liquids and swarm over surfaces. Flagella can also function as adhesins and support surface attachment (Laventie and Jenal, 2020). From a structural perspective, each flagellum consists of three parts:

- the basal body, built up across the cell membranes;
- a hook, as flexible linker;
- a filament, that functions as propeller.

Both the hook and the filament are located on the outside of the cell. The basal body contains an ion-powered rotary motor, formed by a ring of stator motor complexes (MotAB) and a rotor (Apel and Surette, 2008; Santiveri et al., 2020). A high-resolution structure of the MotAB flagellar stator unit that was recently elucidated (Santiveri et al., 2020) provided insights in the role of individual residues in motor functioning. Rotation of a flagellum is required for movement and is an energy costly process. High c-di-GMP concentrations in *E. coli* inhibit flagellar rotation. C-di-GMP directly binds to the molecular brake protein YcgR. YcgR, when bound to c-di-GMP, reduces flagellar rotation through interaction with the MotAB complex (Boehm et al., 2010) (Fig. 7). With this current model of *E. coli*, c-di-GMP inhibits the rotation of fully constructed flagella. Thus,

low c-di-GMP is a prerequisite for motility on top of the expression of all flagellar components. This is in contrast with regulation of motility in species as *V. cholerae* and *P. aeruginosa*, where c-di-GMP, as discussed in Secion 1.2.4.1, inhibits flagellar gene transcription. (Arora et al., 1997; Dasgupta et al., 2003; Pursley et al., 2018; Srivastava et al., 2013).

The expression of the *E. coli* flagellum involves over fifty genes, organized in a dozen operons. Expression of these genes is highly coordinated and organized in a regulatory cascade that is very similar to that of *Salmonella* enterica serovar Typhimurium (Wada et al., 2012). Three different classes of flagellar operons can be distinguished, class I to III, which are sequentially expressed (Kalir et al., 2001; Komeda, 1982, 1986; Kutsukake et al., 1990). On top of the regulatory cascade and under the control of the only class I promoter region is master regulator $FlhD_4C_2$ (FlhDC), a heterohexameric protein complex consisting of four FlhD and two FlhC proteins (Liu and Matsumura, 1994; Wang et al., 2006). Expression of the *flhDC* operon is under tight control of various activating and inhibiting global transcription factors. Activity of these factors often depends on environmental cues, making the expression of the master regulator dependent on, among others, nutrient availability, temperature and osmolarity (Francez and Charlot et al., 2003; Lehnen et al., 2002; Shin and Park, 1995; Soutourina et al., 1999; Yakhnin et al., 2013).

Master regulator FlhDC activates class II and III genes. Class II genes encode the basal body, the hook and the flagellum-specific export apparatus. An additional class II gene is *fliA*, encoding for the flagellum-specific sigma factor σ^{28} (RpoF). This sigma factor is essential for the activation of the class III genes encoding the filament, motor, chemotaxis pathway, c-di-GMP phosphodiesterase PdeH and YcgR (Kalir et al., 2001; Ko and Park, 2000). The *fliA* gene contains two promoters, activated by RpoD and RpoF. Hence, *fliA* expression is under positive control by both FlhDC and by itself (Ikebe et al., 1999).

Within this regulatory cascade, different checkpoints and inhibitory factors were identified, ensuring class III genes are not transcribed before the hook-basal body structure is complete and functional (Chilcott and Hughes, 2000; Kim et al., 2020). As a result, the full flagellar gene cascade is not expressed per se once *flhDC* transcription is induced. Instead, expression of class II and class III genes were shown to be stochastically expressed against a background of low, constitutive *flhDC* expression (Kim et al., 2020). Flagellar gene expression is heterogeneous; only a fraction of *E. coli*

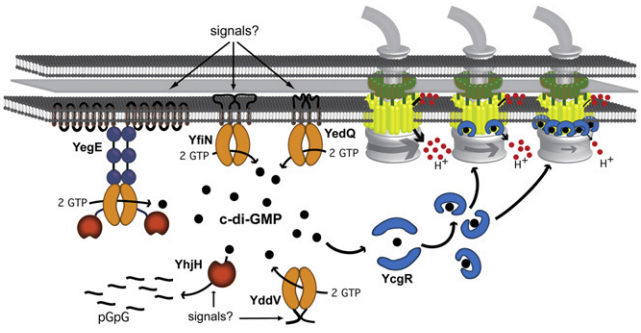


Figure 7: C-di-GMP mediated adjustment of swimming speed in *Escherichia coli*. High c-di-GMP concentrations in *E. coli* inhibit flagellar rotation. C-di-GMP directly binds to the molecular brake protein YcgR. YcgR, when bound to c-di-GMP, reduces flagellar rotation through interaction with the MotAB complex. Adapted from: Boehm et al., 2010.

cells in a population expresses class II and III genes at a given time. Also in *Salmonella*, class II and III gene expression is bimodal (Wang et al, 2020). The authors of Kim et al. 2020 described how this pulsing mechanistically works. One of the factors behind pulsing is the protein YdiV, that is described as the anti-FlhDC factor, inhibiting FlhDC activity. *ydiv* expression uncouples class I (*flhDC*) from class II gene expression. FlgM, expressed as class II gene, is an inhibitor of the flagella associated sigma factor FliA. FliA activation is dependent on excretion of FlgM through an assembled basal body. The FlgM-FliA interaction was found to introduce pulsing of class III expression as compared to class II (Kim et al., 2020). Together this shows that flagellar gene expression in *E. coli* is highly dynamic.

As mentioned above, the regulatory cascade for flagellar genes in *E. coli* is very similar to that in *Salmonella*. This concerns the structuring of different genes in the three classes. However, there is a remarkable difference in the conditions that induce flagellar expression. Whereas transcription is induced under nutrient-poor conditions in *E. coli*, *Salmonella* primarily requires rich nutrient conditions for expression of all flagellar associated genes. This difference was associated with the regulation of the anti-FlhDC factor *ydiV* (Wada et al., 2011, 2012). YdiV binds to the FlhDC complex, thereby inhibiting FlhDC functioning and induction of class II flagellar genes. The *ydiV* gene was found to be responsible for nutrient control in *Salmonella* (Wada 2011), as *ydiV* translation is enhanced in *Salmonella* under poor nutrient conditions. This mechanism of *ydiV* regulation differs from *E. coli*. In *E. coli*, *ydiV* translation is inefficient and this results in low YdiV protein level across conditions, taking away any possibility for nutrient-dependent control (Wada et al., 2012). As cyclic AMP-CRP binding is required for activation of the *flhDC* operon, *E. coli* flagellar expression is activated under conditions that prevent catabolite repression (Soutourina et al., 1999) and even reversely correlates with the nutritional value of the carbon source (Ni et al., 2020).

Several *E. coli* K-12 strains are highly motile, including MG1655 (Blattner et al., 1997), MC1000 and W3110 (Barker et al., 2004). Increased motility correlates with an insertion sequence (*IS*) element in the regulatory region of the *flhDC* operon in the *E. coli* genome. Multiple studies have reported that such a mutation upstream of the *flhDC* operon occurs in static liquid cultures or upon overnight incubation on motility agar (Barker et al., 2004; Parker et al., 2019; Wang and Wood, 2011; Zhang et al., 2017). This suggests that the benefits of increased flagellar gene expression under these conditions outweigh the energetically costly process of flagellar motility, which can account for up to several percent of a cell's energy expenses and protein synthesis (Milo et al., 2010). Since the PDE *pdeH* is a class III gene and under positive control of FlhDC and FliA, strains with the genotype *PflhDC::IS* are characterized by low c-di-GMP levels (Boehm et al., 2010; Ko and Park, 2000; Reinders et al., 2016). Because of their constitutive motility, these strains are used to study other aspects of *E. coli* motility, including the role of c-di-GMP and specific DGCs and PDEs in motility. Over the last two decades, the PDE PdeH has been pointed out as the main phosphodiesterase in *E. coli*, primarily because *pdeH* knockout mutants show poor motility during swimming assays (Boehm et al., 2010; Ko and Park, 2000; Ryjenkov et al., 2006). Hence, a Δ*pdeH* background is used to assess the contribution of individual DGCs to the cellular c-di-GMP pool by measuring the effect of DGC deletions on top of *pdeH* on swimming velocity and swarm size

(Boehm et al., 2010). A *pdeH* mutant background was also used as a trick to identify silent PDEs that are expressed, but inactive under the applied conditions (Reinders et al., 2016). By inoculation of a Δ*pdeH* strain on a motility plate followed by prolonged incubation, Reinders et al. identified motile suppressors with mutations in various PDE promoter regions and in the ORFs of *pdeL* and *pdeN*. With this, they could confirm their hypothesis that under the measured conditions, many PDEs miss the input signal required for activation or sufficient expression to measurably contribute to lowering the c-di-GMP pool (Reinders et al., 2016).

1.2.5.3 Biofilm formation

E. coli is known to produce two types of exopolysaccharide (EPS): polymeric N-acetylglucosamine (poly-GlcNAc or PGA) and cellulose. EPS and curli amyloid fibers form substantial components of the overall extracellular matrix and their abundance defines biofilm properties (Serra et al., 2013) (Fig. 8). Other extracellular matrix components include secreted protein and extracellular DNA (Flemming and Wingender, 2010).

The production and excretion of extracellular matrix components is subject to highly complex regulatory networks involving different sigma factors and nucleotide second messengers. C-di-GMP functions as positive regulator of biofilm formation. The second messenger regulates the expression of both curli and cellulose through the biofilm master regulator CsgD (Klauck et al., 2018). Expression of *csgD* is RpoS-dependent and is regulated by temperature: *csgD* is only expressed at conditions below 30°C. Elevated c-di-GMP levels activate a multicomponent signaling cascade that results in *csgD* transcription (Lindenberg et al., 2013; Pesavento et al., 2008) (Fig. 8). The MerR-like transcription factor MlrA is a direct activator of *csgD* transcription. MlrA is activated by specific interaction with the DGC DgcM. At low c-di-GMP concentrations, MlrA and DgcM are in complex with PdeR and this inhibits their activity. PdeR is both an active PDE and a c-di-GMP-sensing effector. When at elevated c-di-GMP levels, PdeR binds and catalyzes c-di-GMP, DgcM and MlrA are released. DgcM can activate MlrA and this induces cgsD expression (Lindenberg et al., 2013). CsgD is yet on top of the regulatory cascades for production and excretion of curli and cellulose.

The components required to build and export curli amyloid fibers lay in a set of curli specific genes, that are divided over two curli specific operons, *csgBAC* and *csgDEFG*. Expression of these operons is activated by the master regulator CsgD and further depends on environmental factors as temperature and osmolarity. Macrocolonies composed of curli fibers, in the absence of cellulose, are characterized by concentric rings over the structure (Serra et al., 2013).

Cellulose is synthetized and excreted by a membrane-bound protein complex consisting of BcsA and BcsB, that together form the cellulose synthase complex (Morgan et al., 2014; Whitney and Howell, 2013). BscA is allosterically activated by binding of c-di-GMP to the PilZ-domain of BcsA. DgcC was identified to be the specific DGC for c-di-GMP mediated BscA activation (Richter et al., 2020). Functional *in vivo* cellulose production and excretion requires the expression of both the *yhjR-bcsQABZC* and *bcsEFG* operon. Cellulose as a component is actually absent in biofilms formed by the laboratory strain *E. coli* K-12 and its derivatives. This is due to a TAG/TTG SNP early in the *bcsQ* coding sequence, introducing a premature stop codon (Serra et al., 2013).

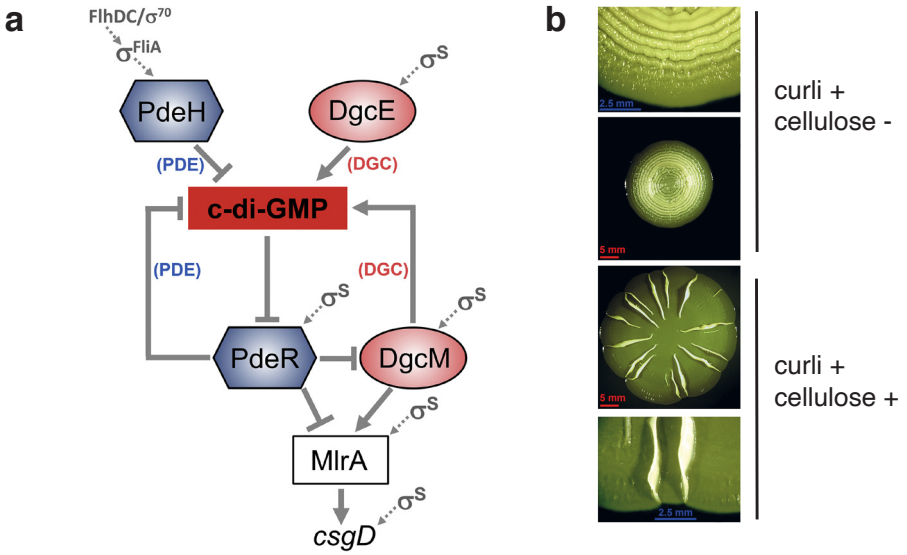


Figure 8: Regulation and properties of biofilm. **(a)** c-di-GMP dependent activation of the biofilm master regulator CsgD. **(b)** Macrocolonies of *E. coli* after five days of growth. Top: *E. coli* macrocolony expressing curli, but not cellulose, through a defect in the cellulose synthase pathway (Serra et al., 2013). Bottom: *E. coli* expressing curli and cellulose. Adapted from: Serra et al., 2013; Serra and Hengge, 2019.

Additionally, the stop codon in the *bcsQ* gene was found to affect the downstream expression of genes in the same operon. Restoring TAG back to TTG in *E. coli* K-12 *bcsQ*, in accordance with genomes of cellulose producing *E. coli* strains, drastically changes macrocolony biofilm structure (Serra et al., 2013) (Fig. 8). The presence of cellulose in addition to curli fibers results in macrocolonies featuring radial ridges and elaborate wrinkles (Serra et al., 2013), resembling the *Salmonella* 'rdar' morphotype macrocolony structure (Römling, 2005).

1.2.5.4 Carbon metabolism

Compared to other c-di-GMP regulated cellular processes, little is known about how c-di-GMP potentially controls metabolic processes in *E. coli*. Recently, one mechanism of such a process was described (Xu et al., 2019). Xu et al. found that c-di-GMP acts on central carbon metabolism via the protein deacetylase CobB and acetyl-coenzyme A synthetase (Acs). Acs activity increases upon deacetylation by CobB. C-di-GMP inhibits CobB deacetylase activity by directly binding to CobB, an interaction that is conserved in *Salmonella*. C-di-GMP-dependent CobB inhibition is part of a feedback loop, in which CobB also deacetylates the DGC DgcZ. Deacetylation of DgcZ results in enhanced stability and activity of this diguanylate cyclase. This regulatory mechanism drives a c-di-GMP dependent growth phenotype on acetate. Growth on acetate as the sole carbon source is not only impaired in a Δ*cobB* strain, but also in a *dgcZ* overexpression strain. A few CobB residues involved in c-di-GMP binding were identified. These residues are located in the N-terminal tail that is essential for dimerization. As c-di-GMP could indeed decrease CobB dimerzation, it is currently hypothesized that c-di-GMP inhibits CobB functioning by interfering with CobB dimerization (Xu et al., 2019).

1.2.5.5 Bacteriophage resistance

C-di-GMP was surprisingly identified to play a role in sensitivity of *E. coli* to bacteriophages (Mutalik et al., 2020). In three parallel high-throughput screens, Mutalik et al. quantitatively mapped host gene function to phage infection. A genome wide screen of *E. coli* genes involved in the resistance of a variety of different bacteriophages showed that in particular the resistance to N4 phage is controlled by c-di-GMP. It was observed that loss of *dgcJ* expression lowered N4 phage susceptibility and that overexpression of six phosphodiesterases including *pdeB*, *pdeC*, *pdeL* and *pdeN* also yielded increased fitness scores. The authors concluded that c-di-GMP is required for infection by phage N4. They speculated that the exact mechanism behind c-di-GMP-dependent phage susceptibility might be rather indirect, for example via cellular components that are present or absent depending on the c-di-GMP concentration. The exact mechanism remains to be elucidated (Mutalik et al., 2020).

1.2.5.6 PdeL

Out of the 25 DGCs and PDEs in *E. coli*, the focus of this thesis is on the c-di-GMP specific phosphodiesterase PdeL, previously named YahA (Hengge et al., 2015). This PDE has two domains: an N-terminal helix-turn-helix (HTH) DNA binding domain and a C-terminal catalytic EAL domain. Expression of *pdeL* is controlled by Cra, a pleiotropic regulator mainly targeting metabolic genes (Kim et al., 2018; Shimada et al., 2005, 2011). Cra activity is modulated by carbon flow and is lowest during fast glycolytic growth (Kochanowski et al., 2013; Ramseier et al., 1993, 1995). More details are provided in Section 1.3.3. A Cra binding site in the *pdeL* promoter region was detected in a genomic SELEX screen (Shimada et al., 2005) and was later confirmed by ChIP-exo (Kim et al., 2018), which is a variation of ChIP-seq (Rhee and Pugh, 2012). *In vitro* transcription assays suggested Cra to act as a *pdeL* transcriptional repressor (Shimada et al., 2005). The *pdeL* intergenic region also contains binding sites for H-NS, a transcription factor with silencing activity that binds to hundreds of regions in the *E. coli* genome (Rangarajan and Schnetz, 2018; Shimada et al., 2014; Uyar et al., 2009). Transcription of the *pdeL* gene depends on the vegetative sigma factor RpoD (also called σ^D and σ^{70}) (Shimada et al., 2014).

New attention was drawn to PdeL when several *pdeL* motile suppressor alleles restored motility in a high c-di-GMP background of an otherwise highly motile MG1655 strain (*PflhDC::IS1* Δ*pdeH*, see section 1.2.5.2) (Blattner et al., 1997; Reinders et al., 2016). Mutations in the catalytic EAL domain, but away from the active site, reduced c-di-GMP levels as compared to a Δ*pdeH* strain (Reinders et al., 2016; Sundriyal et al., 2014). Interestingly, these motile suppressor alleles also increased PdeL protein levels. Reinders et al. (2016) then compared *pdeL* promoter activity at distinct cellular c-di-GMP concentrations and observed that activity inversely correlates with cellular c-di-GMP levels. Based on these data, they argued that *pdeL* transcription is negatively controlled by c-di-GMP levels and activated when c-di-GMP levels are low. Using electrophoretic mobility shift assays (EMSA), they located a PdeL binding site in the *pdeL* intergenic region, indicating that the c-di-GMP-dependent effect could be direct through binding of PdeL to its own promoter region. A DNA binding-deficient *pdeL* mutant, as well as a Δ*pdeL* mutant, showed

constitutively low *pdeL* promoter activity independent of cellular c-di-GMP concentration. These results strongly argued that PdeL is an active PDE as well as a transcription factor, modulating its own expression in response to the prevailing c-di-GMP level. Thus, the cellular c-di-GMP concentration is an important and direct input, where low c-di-GMP is a signal for *pdeL* activation. Besides, structural elucidation of the PdeL EAL domain revealed the existence of two distinct structural conformations of PdeL-EAL dimers that depend on the presence of c-di-GMP (PDB: 4KIE and 4LJ3, Sundriyal et al., 2014).

In the search for potential further targets of PdeL as transcription factor, the DNA-binding capacities of PdeL were further explored (Yilmaz et al., 2020). Experimental data of Yilmaz et al. showed that PdeL binds to the regulatory region of the *fliFGHIJK* operon and suggested that PdeL thus represses transcription of class II flagellar genes. This description of PdeL as a repressor of motility genes (Yilmaz et al., 2020) is in contrast with observations that PdeL activity increases motility (Reinders et al., 2016). Therefore, further studies would be required to unambiguously describe the role of PdeL as transcription factor and confirm additional PdeL binding sites in the *E. coli* genome.

1.3 *Escherichia coli* metabolism

1.3.1 Central metabolic pathways

For the widely studied model organism *Escherichia coli*, the central metabolic pathways are well described. *E. coli* is a prototrophic, facultatively aerobic species (Conway and Cohen, 2015), capable of growing on a number of mono- and disaccharides (Ammar et al., 2018) and alternative carbon sources including amino acids and dicarboxylic acids (Conway and Cohen, 2015). The preferred carbon source of *E. coli* is glucose, as the highest growth rate is observed on glucose as compared to other sugars it can metabolize (Bren et al., 2016). Glucose even prevents the usage of other carbon sources through regulation of alternative catabolic pathways on the transcriptional level (Fic et al., 2009; Notley-McRobb et al., 1997). This regulatory mechanism is termed carbon catabolite repression, in which the small signaling molecule cyclic AMP (cAMP), in complex with its effector protein CRP, act as activator of many genes encoding alternative substrate degradation pathways. Activation of the sole adenylate cyclase for cAMP synthesis in *E. coli*, Cya, depends on glucose availability. As in many bacteria, the phosphotransferase system (PTS) exhibits both a transport and regulatory role (Stulke and Hillen, 1999). Extracellular glucose is transported by the glucose PTS system and phosphorylated into glucose-6-phosphate. During growth in glucose excess, the phosphorylation state of the PTS is low. Upon glucose depletion, phosphorylated PTS proteins activate Cya. Although different mechanisms are proposed for the exact Cya activation mechanism, the cAMP concentration increases due to Cya activation under glucose starvation. cAMP binds to the CRP protein and the cAMP-CRP complex activates transcription by stabilizing the RNA polymerase in the corresponding promotor regions (Green et al., 2014; Kremling et al., 2015; Stulke & Hillen, 1999).

Further repression of genes encoding alternative catabolic pathways is established through the lack of their specific inducers. This is the case for specific carbon sources such as lactose and xylose. The presence of the specific sugar is required to activate the corresponding catabolic pathway. For example, activation of the lactose catabolizing genes in the lac operon requires glucose depletion and presence of allolactose to bind the lac repressor LacI (Stulke and Hillen, 1999).

The canonical view of carbon catabolite repression describes that alternative substrate usage is prevented by glucose and as a consequence, substrates are used sequentially. When incubated in medium containing both glucose and lactose, glucose is depleted before *E. coli* switches to consuming lactose, resulting in a diauxic shift. This view of hierarchical consumption is being challenged by recent studies describing simultaneous uptake of glucose and less preferred substrates (Okano et al., 2020). Under limited glucose uptake rates, simultaneous glycerol degradation was observed. It was proposed that under conditions that the preferred substrate is limited, *E. coli* supplements its central metabolic pathways with carbon from a less preferred source.

Following substrate specific uptake and/or initial enzymatic conversions, the respective sugar, amino acid, or dicarboxylic acid is further converted through the central metabolic pathways of *E. coli*. Central metabolism in *E. coli* consists of the Embden-Meyerhof-Parnas glycolytic pathway (EMPP), the pentose phosphate pathway (PPP), the Entner-Doudoroff pathway (EDP),

the tricarboxylic acid (TCA) cycle, and fermentation pathways (see Fig. 9) (Keseler et al., 2017). Glucose is primarily metabolized by *E. coli* through the EMPP and oxidative PPP, while the EDP remains mainly inactive except during growth on gluconate (Eisenberg and Dobrogosz, 1967; Hollinshead et al., 2016). The EMPP consists of a series of ten enzymatic reactions, most of these being equilibrium reactions, and yields two pyruvate, two ATP and two NADH per molecule of glucose (Nelson and Cox, 2008). The EMPP glycolytic pathway can be divided into two stages, the preparation phase or upper glycolysis and the ATP generating phase or lower glycolysis. While the preparative phase requires ATP for the phosphorylation of glucose, ATP is generated by substrate level phosphorylation during the conversion from glyceraldehyde-3-phosphate to pyruvate. Glucose is also oxidized via the oxidative phase of the PPP generating NADPH and precursors for nucleic acid synthesis (Nelson and Cox, 2008).

During optimal glycolytic growth under aerobic conditions, *E. coli* excretes quantities of acetate. Acetate is the product of fermentation rather than respiration of the carbon source. This phenomenon is referred to as overflow metabolism and is widely observed in prokaryotes, as well as eukaryotes (referred to as the Crabtree effect in yeast and Warburg effect in proliferating mammalian cells) (Alteriis et al., 2018; Basan et al., 2015; Valgepea et al., 2010). Over time, various hypotheses have been proposed to explain overflow metabolism. It was among others hypothesized to be the result of limited respiratory capacity or cofactor recycling requirements (Majewski and Domach, 1990; Wolfe, 2005). It was also suggested that proteome allocation is more efficient in fermentation compared to respiration (Molenaar et al., 2009; O'Brien et al., 2013). A recent study could further confirm overflow metabolism in *E. coli* as a response strategy to balance the proteomic demands for energy biosynthesis and biomass to achieve rapid cell division (Basan et al., 2015). Acetate excretion was furthermore found to correlate with glycolytic growth rate (Basan et al., 2015). The conversion of glucose into acetate is strongly reduced when the glucose uptake rate is limited (Fischer and Sauer, 2003). During aerobic, but slow growth on glucose, established by glucose-chemostat cultivation, the substrate is fully oxidized by *E. coli* through parallel operation of the PEP-glyoxylate cycle and TCA cycle (Fig. 9) (Fischer and Sauer, 2003).

Furthermore, the TCA cycle is used during growth on poor carbon sources including pyruvate, acetate, and TCA cycle intermediates as succinate and fumarate (Kim et al. 2018, Nelson and Cox, 2008). Under these conditions, gluconeogenesis is performed to generate the required metabolites that can enter the PPP. Although most enzymes of the EMPP can perform enzymatic reactions in both directions and hence form the gluconeogenic pathway, some glycolytic enzymes require their gluconeogenic counterpart. These enzymes include pyruvate kinase (PykF) in the lower glycolysis, that requires replacement by phosphoenolpyruvate (PEP) synthetase to convert pyruvate into PEP; and fructose-1,6-bisphosphatase replaces 6-phosphofructokinase in the upper glycolysis to convert fructose-1,6-bisphosphate to fructose-6-phosphate (Nelson and Cox, 2008).

As the *E. coli* carbon metabolism is relatively well described, it allows rational metabolic engineering. Metabolic pathways are adapted, or introduced to extend its metabolic capacity, for alternative substrate utilization or precursor production purposes (Lim et al., 2013; McKee et al., 2012; Meng et al., 2021; Müller et al., 2015).

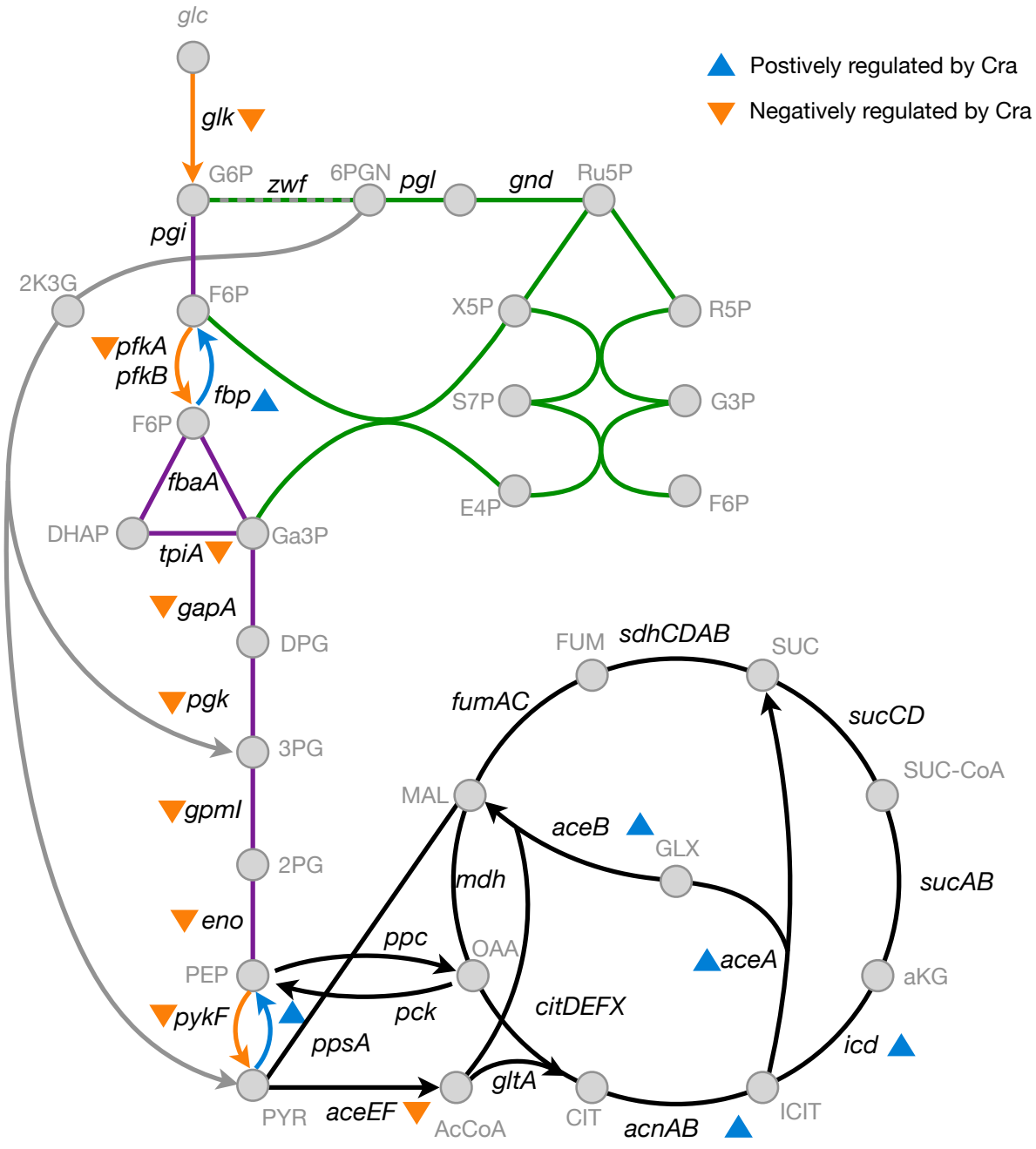


Figure 9: Central metabolic pathways in *Escherichia coli*. Central metabolic pathways in *E. coli* convert metabolites and metabolic intermediates (grey circles) via enzymatic reactions (lines). Indicated pathways are the Embden-Meyerhof-Parnas glycolytic pathway (EMPP, purple/orange), Entner-Doudoroff pathway (EDP, grey), pentose phosphate pathway (PPP, green), gluconeogenic pathway (GP, purple/blue), tricarboxylic acid cycle (TCA) and glyoxylate shunt (black). The EMPP converts glycolytic substrates to pyruvate. Many enzymatic reactions are equilibrium reactions and performed by the same enzyme during gluconeogenis (purple), whereas few enzymatic steps require separate enzymes for glycolysis (orange) or gluconeogenesis (blue). Enzymes are indicated in black italic, metabolites abbreviated in grey. Orange triangles: genes under negative control of Cra. Blue triangles: genes under positive control of Cra. Adapted from: Covert et al., 2008; Kim et al., 2018

1.3.2 Regulation of metabolism

The metabolic pathways and phenomena as described above and in Fig. 9 arise from regulatory networks that modulate carbon metabolism on the transcriptional, post-transcriptional and protein level (Kochanowski et al., 2017; Mattevi et al., 1995; Morin et al., 2016; Perrenoud and Sauer, 2005).

Metabolic transcriptional regulators including Cra, the cAMP-CRP complex, AcrA, Mlc, PdhR and IclR adjust gene expression (Covert et al., 2008; Göhler et al., 2011; Matsuoka and Shimizu, 2011; Perrenoud and Sauer, 2005).

Compared to transcriptional regulation, less is known about post-transcriptional regulation of metabolic pathways (Morin et al., 2016), but the increasing number of recent studies demonstrate it to play a pivotal role. On the post-transcriptional level, metabolic regulation is proposed through the RNA-binding protein CsrA, which is part of the carbon storage regulator (CSR) system (Babitzke and Romeo, 2007). CsrA is essential for the adequate functioning of upper glycolysis (Morin et al., 2016) and hence for glycolytic growth (Altier et al., 2000; Timmermans and Van Melderen, 2009), but is also considered to regulate additional metabolic pathways (Morin et al., 2016). The levels of multiple central metabolic enzymes are potentially regulated by CSR (Romeo et al., 1993; Sabnis et al., 1995; Wei et al., 2000; Yang et al., 1996). By binding to mRNA molecules, CsrA can prevent mRNA translation, facilitate mRNA decay (Duss et al., 2014; Seyll and Melderen, 2013; Timmermans and Melderen, 2010), or stabilize its targets and thereby specifically enhance translation (Yakhnin et al., 2013). The regulation of CsrA activity is complex and feedback controlled (Yakhnin et al., 2013). Among others, two non-coding RNA molecules, CsrB and CsrC, antagonize CsrA activity by sequestering several CsrA dimers (Liu et al., 1997; Weilbacher et al., 2003). CsrA stimulates the expression of CsrB and CsrC through the response regulator UvrY that is part of the BarA-UvrY two-component system, revealing an auto-regulating system on CsrA activity. CsrA also affects other cellular processes including motility (Wei et al., 2001), biofilm formation and dispersal (Jackson et al., 2002; Wang et al., 2005), and virulence in pathogenic strains (Bhatt et al., 2009), suggesting a role for CsrA in cellular behavior control.

Besides regulation on the transcriptional and post-transcriptional level, metabolic enzymes can be activated by allosteric regulation. For example, the metabolite fructose-1,6-bisphosphate allosterically activates pyruvate kinase (Mattevi et al., 1995; Valentini et al., 2000).

While detailed knowledge is available on the genetic regulation and biochemistry of individual enzymes in cellular metabolism and physiology, yet aspects are unclear about its overall regulation (Shimizu, 2009). Additionally, the interaction between carbon metabolism and other cellular networks is not fully understood. Open questions include: how are metabolic processes adapted by various stress responses, how do signaling molecules other than cyclic AMP affect metabolism, and does metabolism affect intracellular signaling?

1.3.3 Cra: catabolite repressor-activator

From all metabolic regulators, the focus in this work is on Cra, a transcriptional activator of the c-di-GMP phosphodiesterase *pdeL*, as explained in Section 2 of this chapter.

Cra is an acronym for catabolite repressor-activator and refers to the pleiotropic regulatory role of this transcription factor on the expression of central metabolic genes (Ramseier, 1996). The Cra protein was originally called FruR, for Fructose repressor (Ramseier et al., 1993). Its protein structure is highly similar to the *Salmonella* FruR protein (Jahreis et al., 1991; Leclerc et al., 1990; Vartak et al., 1991), which was originally identified as repressor of the *fruBKA* operon encoding fructose specific uptake and degrading enzymes (Chin et al., 1987; Geerse et al., 1986). As with *Salmonella* FruR (Chin et al., 1989), the *E. coli* Cra protein not only binds the promoter region of the fructose operon but also regulates multiple operons (Ramseier et al., 1993). Besides the *fruBKA* operon, Cra was identified to positively regulate the metabolic genes *ppsA* and *icd* and the *aceBAK* operon by binding their regulatory regions (Ramseier et al., 1993). An initial consensus sequence could be defined, that revealed numerous additional promoter regions in the *E. coli* genome with a putative Cra binding site, identifying Cra as a pleiotropic global transcriptional regulator (Ramseier et al., 1993). Subsequently, Cra was identified as repressor of the glycolytic enzyme *pykF* (Bledig et al., 1996) and confirmed as activator of the genes encoding the glyoxylate shunt (Cortay et al., 1994). Cra is not only involved in gene expression regulation across all major metabolic pathways, it consistently activates expression of enzymes involved in oxidative and gluconeogenic carbon flow while repressing glycolytic genes. Hence, Cra was proposed to modulate carbon flow direction in *E. coli* (Ramseier et al., 1995).

The current overviews of the Cra regulon were obtained by systematic, genome wide searches using SELEX, ChIP-exo and RNA-seq (Kim et al., 2018; Shimada et al., 2005, 2011). Data generated by the *in vitro* SELEX system identified 164 Cra binding sites (Shimada et al., 2011). The systematic assessment of the Cra regulon based on *in vivo* DNA binding identified 39 binding sites, 33 overlapping with SELEX data, regulating a total of 97 genes (Kim et al., 2018). While the majority of Cra targets concerns metabolic genes, there is evidence that Cra also binds promoter regions of non-metabolic genes or operons. Among others, Cra binds the promoter region of the curli encoding *csgDEFG* operon (Reshamwala and Noronha, 2011), of the c-di-GMP phosphodiesterase *pdeL* (Shimada et al., 2005), and of transcription factors including the *marRAB* operon (Shimada et al., 2011). In human pathogenic, enterohemorrhagic *E. coli* (EHEC), Cra co-regulates expression of virulence factors (Njoroge et al., 2012, 2013).

1.3.3.1 The LacI/GalR protein family

The Cra protein belongs to the LacI/GalR protein family, for which more than 1000 members have until now been identified (Swint-Kruse and Matthews, 2009; Weickert and Adhya 1992). Transcriptional regulators belonging to the LacI/GalR family are commonly found in both Gram-positive and Gram-negative bacteria, in which they act as transcriptional repressor, activator, or both (Swint-Kruse and Matthews, 2009). Many members regulate gene expression of catabolic enzymes in response to specific nutrient availability (Weickert and Adhya, 1992). Other members

control functionally distinct cellular processes including nucleotide biosynthesis in *E. coli* (Meng and Nygaard, 1990) and toxin expression in *P. aeruginosa* (Colmer and Hamood, 1998). The Cra protein is among the few members that function as global regulators (Swint-Kruse and Matthews, 2009).

LacI/GalR proteins contain a DNA-binding domain and a regulatory domain (Swint-Kruse and Matthews, 2009; Weickert and Adhya, 1992). Homodimer formation is mediated by the regulatory domain and is required for high-affinity binding to specific DNA target sequences. DNA binding is abolished upon binding of inducer molecules to the regulatory domain. Structural changes in the regulator domain upon ligand binding are propagated to the DNA binding domain and alter DNA binding affinity (Bell and Lewis, 2000; Choi et al., 1994; Kristensen et al., 1996; Schumacher et al., 2007; Weickert and Adhya, 1992).

The Cra protein contains 334 amino acid residues and has a molecular weight of 38 kDa (Artimo et al., 2012). The structure of full-length Cra has not yet been elucidated, but structures of single Cra protein domains exist. Two structures of the N-terminal helix-turn-helix motif were elucidated by NMR (PDB: 1uxc and 1uxd). A separate structure of the sugar-binding domain in apo form was elucidated by X-ray crystallography with 1.85 Å resolution (PDB: 2iks) and showed that Cra sugar binding domains dimerize. Cra proteins can also form tetramers (Shimada et al., 2005).

1.3.3.2 Regulation of Cra activity

As with other members of the LacI/GalR protein family, the DNA binding activity of Cra is modulated by an allosteric effector molecule. The metabolite Fructose-1-phosphate (F1P) is currently considered as the sole allosteric regulator of Cra (Bley Folly et al., 2018). Low millimolar concentrations of F1P displaced the Cra protein from various Cra binding sequences during *in vitro* DNA-Cra binding studies (Ramseier et al., 1993), which was confirmed by additional methods including Cra-F1P interaction (Bley Folly et al., 2018). Fructose-1,6-bisphosphate (FBP), an intermediate of the EMPP glycolytic pathway (Fig. 9), was described as a second allosteric regulator yet with a lower binding affinity (Ramseier et al., 1993). However, a recent assessment on the interaction between FBP and Cra excluded FBP as direct allosteric regulator of Cra (Bley Folly et al., 2018). Similarly, for the Cra homolog in *Pseudomonas putida*, evidence was provided that F1P functions as the sole regulator of the *P. putida* Cra protein (Chavarría et al., 2011, 2014). Yet, it remains to be elucidated how the metabolite F1P can regulate Cra activity in the absence of fructose, as until now it is assumed that in *E. coli* F1P is solely formed during fructose conversion by the FruBA phosphotransferase system (Kornberg, 2001). One suggestion for how Cra regulation across conditions could work is via the fructose-1-phosphate kinase FruK, that further phosphorylates F1P into the glycolytic intermediate FBP. It was described that FruK and Cra interact *in vivo* and that enzymatic activity of FruK can act reversely, generating F1P from FBP (Singh et al., 2017). These results together could provide a regulatory mechanism for Cra (Singh et al., 2017); however, additional research is required to support this.

1.3.3.3 Regulation of *cra* expression

Our knowledge about the regulation of *cra* expression is rather scarce. The expression of *cra* was reported to be repressed by phosphorylated PhoB and dependent on the heat shock sigma factor σ^{32} (*rpoH*) (Keseler et al., 2017).

The absolute Cra protein abundance estimation ranges from 371 copies per cell during log phase in rich medium, down to 148 copies per cell during log phase on glucose (Schmidt et al., 2016). The Cra copy number further decreases when *E. coli* is grown on glucose at lower pH or under osmotic stress, and under long-term stationary phase (one to three days) (Schmidt et al., 2016).

1.3.3.4 Physiological impact and roles of Cra

Besides the genetic and biochemical evidence for Cra as a transcriptional regulator of central metabolic genes, what is the overall *in vivo* impact of Cra on metabolic regulation and *E. coli* physiology? Cra has been proposed to modulate carbon flow direction, which primarily refers to the switch between glycolysis and gluconeogenesis (Ramseier et al., 1995). Growth on poor carbon sources including succinate, glycerol and acetate is affected in a Δ*cra* mutant strain (Kim et al., 2018). Both growth rate and lag time increase in the absence of Cra (Kim et al., 2018). Growth on galactose is also affected, but less severely. If fructose is used as sole carbon source, a Δ*cra* mutant strain displays a higher growth rate as compared to a wild type strain (Kim et al. 2018). When compared to the growth impact of a Δ*crp* mutation, metabolic regulation by Cra is more important for growth on poor carbon sources than regulation by cAMP-CRP. In contrast, Crp has a larger impact on growth on alternative sugars including fructose and galactose (Kim et al., 2018). Some glycolytic genes are antagonistically regulated by Cra and cAMP-CRP. The same holds for some genes expressing enzymes of the TCA cycle. Following the regulatory activity by Cra and cAMP-CRP on this set of genes on various sources, Cra possibly overrides the regulatory affect by CRP (Kim et al., 2018).

Not only do the cAMP-CRP and Cra regulons overlap, more regulons of transcription factors show overlap, making contributions of individual regulators to an overall response unclear (Kochanowski et al., 2017). On top of differential activity of metabolic regulators, growth rate dependent global regulation is a dominating factor in metabolic gene expression levels (Kochanowski et al., 2017). When dissecting this global, growth rate dependent transcriptional regulation from specific regulation by regulatory proteins and metabolites, a surprisingly simple system controls *E. coli* metabolism. Variation in the level of cAMP, FBP and F1P could, through their effectors Cra and CRP, explain most of the specific transcriptional control across conditions (Kochanowski et al., 2017). Even though FBP does not allosterically regulate Cra activity (Bley Folly et al., 2018), the cellular FBP concentration strongly correlates with expression of Cra-regulated genes (Kochanowski et al., 2017).

Other than metabolic adaptation through substrate specific signaling, indirect carbon source recognition was proposed as a strategy by *E. coli* to optimize nutrient sensing (Kochanowski et al. 2013, Kotte et al. 2010). Instead of responding to the availability of specific carbon sources, *E. coli* would adapt metabolism based on the total carbon flux – metabolic flux or glycolytic

flux – that arises as net result of carbon uptake and carbon flow direction. A change in substrate availability would change the metabolic flux and a system of central metabolites and transcription factors would function as flux sensors and induce metabolic adaptation as required. Furthermore, when these sensors are part of global feedback loops, it may auto-regulate metabolic pathways and pathway adaptation when required (Kotte et al., 2010). The transcription factor Cra was proposed to function as metabolic flux sensor, as Cra activity scales with the glycolytic flux measured during growth on different glycolytic sources (Kochanowski et al. 2013). The functioning of a flux sensor in *E. coli* and other organisms is still an emerging concept. Additional research could provide further support for the principle of flux sensing in microbial species.

Aim of the thesis

The intracellular signaling molecule c-di-GMP controls multiple aspects of bacterial growth and behavior. The *Escherichia coli* K-12 genome encodes 25 active diguanylate cyclases and c-di-GMP specific phosphodiesterases (PDE), but the exact cues for their activation and their impact on cellular c-di-GMP levels remain largely unknown.

In a previous study, motile suppressor alleles of the PDE PdeL restored motility in a high c-di-GMP background. Initial data suggested PdeL to function as a catalyst, c-di-GMP sensor and transcription factor (Reinders et al., 2016).

In this work, we first describe in detail the molecular mechanisms of PdeL regulation that give rise to switch-like behavior and bistable expression. Using a novel c-di-GMP biosensor developed in the Jenal lab, I aimed to elucidate the effect of stochastic *pdeL* expression on c-di-GMP distributions.

As the metabolic regulator Cra was identified as activator of *pdeL*, I investigated how Cra, hence metabolism, controls c-di-GMP distributions by controlling *pdeL* expression. Specifically, the impact of Cra activity, growth rate and variable carbon sources on *pdeL* expression and c-di-GMP distributions was assessed.

In a comparison of different *E. coli* K-12 stocks and metabolic conditions, I aimed to elucidate how PdeL couples metabolism to signaling and establishes c-di-GMP heterogeneity.

Chapter 2

A simple second messenger switch establishes binary signaling outputs to impose precise developmental transitions in *E. coli*

Alberto Reinders*, Margo van Berkum*, Benjamin Sellner*, Firas Fadel*, Andreas Kaczmarczyk, Shogo Ozaki[1], Pablo Manfredi, Matteo Sangermani[2], Alexander Harms, Tilman Schirmer, and Urs Jenal

Biozentrum of the University of Basel
Klingelbergstrasse 50/70, 4056 Basel, Switzerland

Current addresses:
[1] Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Higashi-ku, Fukuoka 812-8582, Japan
[2] Department of Circulation and Medical Imaging, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway

* Authors contributed equally
For correspondence: urs.jenal@unibas.ch, tilman.schirmer@unibas.ch

Keywords: Cellular heterogeneity, bimodality, bistability, c-di-GMP, phosphodiesterase, *Escherichia coli*, transcription, motility, biofilm, phage infection

Contents

Statement of my contributions:

The work presented in Chapter 2 is the project of A. Reinders in collaboration with members of the Schirmer lab. I acquired and analysed the data presented in Figure 5. For Figure 5b-e, I performed experiments according to the protocol used for Figure 4. For Figure 5f-h, I constructed the required strains and designed and performed the experiments. I provided a draft text for the corresponding paragraph in the Results section.

2.1 Abstract

Nucleotide-based signaling molecules (NSMs) are widespread in bacteria and eukaryotes, where they control important physiological and behavioral processes. In bacteria, NSM-based regulatory networks can be highly complex, including large numbers of enzymes involved in the synthesis or degradation of active signaling molecules. How the converging input from multiple enzymes is transformed into robust and unambiguous cellular responses has remained unclear. Here we show that *E. coli* converts dynamic changes of c-di-GMP into discrete binary outputs. This is mediated by an ultrasensitive switch protein, PdeL, which senses the prevailing concentration of the signaling molecule in the cell and couples this information to c-di-GMP degradation and to a transcriptional feedback boosting its own expression. We demonstrate that PdeL acts as a digital filter that enables precise developmental transitions, generates functional heterogeneity, confers cellular memory, and protects *E. coli* from phage predators. Based on our findings, we propose that bacteria apply simple regulatory switches to convert dynamic changes in NSMs into robust binary signaling modes.

2.2 Introduction

Biological systems need to be able to convert spatial or temporal gradients of signaling molecules into precise and robust readouts. For example, in the *Drosophila* embryo, highly accurate spatial patterning is established from an original gradient of the maternal morphogen bicoid within the first few hours of development (Gregor et al., 2007; Jaeger, 2011). Gradual changes of bicoid concentrations along the embryonic axis are converted into sharp expression patterns of its downstream target gene hunchback via switch-like, cooperative responses (Driever et al., 1989; Park et al., 2019). Similar mechanisms must exist to convert gradual changes of signaling molecules over time into specific and robust cellular responses. This is of particular relevance for bacteria, which make use of an extensive array of sensory systems for surveillance. These include receptors coupled to protein phosphorylation cascades (Bi and Sourjik, 2018; Capra and Laub, 2012; Galinier and Deutscher, 2017) or to the synthesis of small, nucleotide-based signaling molecules (NSM) (Bassler et al., 2018; da et al., 2020; Hauryliuk et al., 2015; Jenal et al., 2017; Stülke and Krüger, 2020; Zaver and Woodward, 2020). While phosphorylation cascades are generally linear and highly specific, networks relying on diffusible signaling compounds are less well defined.

NSMs are widespread in bacteria and include the linear (p)ppGpp, monocyclic compounds like cAMP, and cyclic di-nucleotides like c-di-GMP, c-di-AMP, or cGAMP. These molecules control critical bacterial processes like growth and metabolism, stress response and predator defense, as well as virulence and persistence (Bassler et al., 2018; da et al., 2020; Hauryliuk et al., 2015; Jenal et al., 2017; Stülke and Krüger, 2020; Zaver and Woodward, 2020). NSM-mediated signaling networks can adopt highly complex architectures with many bacteria harboring up to several dozens of sensors regulating the concentration of active compounds. For example, some spirochetes, alpha-proteobacteria, or mycobacteria possesses up to 30 adenylate cyclases, while many actinobacteria, cyanobacteria, or proteobacteria encode 50-100 different sensors involved in the synthesis or breakdown of c-di-GMP (Bassler et al., 2018; Galperin, 2018; Galperin et al., 2010).

This raises important questions regarding NSM control and downstream signaling. How do bacteria convert gradual temporal changes of signaling molecules into deterministic and irreversible cellular responses? And how do they absorb stochastic fluctuations of such molecules generated through random noise? This is particularly important if downstream processes are highly sensitive to concentration changes and if they occur on short time scales. For example, effector binding affinities for c-di-GMP are typically in the nanomolar range (Chou and Galperin, 2016) and some c-di-GMP-mediated processes show ultra-rapid responses on the time scale of seconds (Hug et al., 2017; Laventie et al., 2019; Nesper et al., 2017). Finally, how can small diffusible molecules regulate specific downstream processes given that concentration changes likely provoke a global cellular response? It was recently proposed that c-di-GMP can signal in spatially confined compartments with specialized sensors stimulating spatially coupled cellular processes (Richter et al., 2020). Although ‘local signaling’ relies on a direct interaction between sensors

regulating c-di-GMP synthesis and/or degradation and its downstream targets (Andrade et al., 2006; Dahlstrom et al., 2016; Giacalone et al., 2018; Lindenberg et al., 2013), molecule leakage likely occurs and needs to be absorbed to effectively isolate individual signaling modules from each other (Jenal et al., 2017; Richter et al., 2020).

One possibility to avoid detrimental fluctuations of potent signaling molecules and to convert graded inputs into switch-like, irreversible and specific responses is to sense the prevailing concentration and couple this information to a catalytic feedback. Positive and double-negative feedback loops can generate stable genetic responses when coupled to non-linear or 'ultrasensitive' behavior (Ferrell, 2002; Koshland et al., 1982). In recent years, several examples of switches were identified in bacteria generating bistable gene expression and cell fate decisions (Losick and Desplan, 2008; Norman et al., 2015; Veening et al., 2008a). However, to date, no such mechanisms are known to control the concentration of small diffusible signaling molecules.

Here, we show that *Escherichia coli* converts gradual changes of c-di-GMP into a discrete binary output, thereby generating heterogeneous populations with distinct c-di-GMP levels. This response is mediated by the phosphodiesterase PdeL that acts as a simple molecular switch establishing precise cellular levels of c-di-GMP. We show that PdeL degrades c-di-GMP and, at the same time, acts as transcription factor to stimulate its own synthesis. We demonstrate that c-di-GMP impedes PdeL activity and that catalytic and transcriptional feedbacks of PdeL generate bimodal and bistable populations with distinct behavior. Functional assays indicate that the PdeL switch sets off robust lifestyle changes that lead to effective biofilm formation and escape and that it can serve as bet-hedging device to protect *E. coli* against phage predation.

2.3 Results

2.3.1 Cra and PdeL co-regulate *pdeL* transcription

Earlier studies had shown that PdeL controls its own transcription and had suggested that PdeL autoregulation is mediated by c-di-GMP (Reinders et al., 2016). The *pdeL* promoter region (Shimada et al., 2005) contains a binding site for PdeL 675 bp upstream of the *pdeL* start codon (Reinders et al., 2016). DNA shift assays with fragments spanning the entire promoter region identified an additional binding site for PdeL overlapping a palindromic binding site for the central metabolic regulator Cra (Kochanowski et al., 2013; Shimada et al., 2011) (Fig. 1a,b). Both Cra (K_d 49 nM) and PdeL (K_d 76 nM) bind to these sites with high affinity (Fig. S1a-d), but while Cra can bind the DNA on its own, PdeL binding requires Cra (Fig. 1b,c). Hence, we termed this PdeL binding site the Cra-dependent PdeL-box (CDB) to distinguish it from the upstream Cra-independent binding site (CIB), for which PdeL showed significantly lower affinity (K_d 573 nM) (Figs. 1b; S1e-g). Although c-di-GMP negatively impacts *pdeL* transcription (Reinders et al., 2016), PdeL binding to CDB or CIB was not affected by c-di-GMP (Fig. 1b).

Transcription of *pdeL* was strongly reduced in strains lacking Cra or PdeL or in strains with mutated Cra or CDB boxes (Figs. 1d, S1h). Likewise, scrambling the right half-site of the CIB palindrome reduced *pdeL* transcription 2.5-fold (Fig. 1d, S1h). To investigate how PdeL stimulates its own transcription, we analyzed the binding of RNAP holoenzyme to the *pdeL* promoter. While RNAP alone showed poor interaction with the *pdeL* promoter region, its recruitment was strongly increased in the presence of Cra and PdeL (Fig. 1e), arguing that PdeL facilitates RNAP recruitment. The *pdeL* promoter region contains several binding sites for the transcriptional silencer H-NS (Rangarajan and Schnetz, 2018) (Fig. S1h), one of which overlaps with the left half-site of the CIB palindrome (Fig. 1a). H-NS and PdeL compete for this site *in vitro* (Fig. S1i) and *pdeL* transcription was strongly derepressed in an *hns* mutant (Fig. 1d), arguing that PdeL stimulates *pdeL* transcription by acting as an anti-silencer. Together these experiments identified Cra and PdeL as co-activators of *pdeL* transcription and proposed that Cra facilitates PdeL binding, which in turn helps recruiting RNAP to the promoter region.

2.3.2 PdeL regulates its own transcription in response to c-di-GMP

We next monitored *pdeL* transcription in response to changes of c-di-GMP. To tune c-di-GMP, we made use of the observation that *E. coli* strain CGSC 7740 lacking the phosphodiesterase PdeH has strongly increased levels of c-di-GMP (Reinders et al., 2016). By replacing the chromosomal copy of *pdeH* with a plasmid-born and IPTG inducible *pdeH*, c-di-GMP levels could be tuned from the low nanomolar (65 μM IPTG) to above 5 μM (no IPTG), while a *pdeH* wild-type strain showed intermediate c-di-GMP levels (210 nM) (Fig. 1f). The expression of a *pdeL-lacZ* reporter was maximal at low c-di-GMP concentrations but declined when c-di-GMP levels increased. Likewise, PdeL protein levels inversely scaled with c-di-GMP from 200 nM to 1.35 μM (Fig. 1f). C-di-GMP-mediated transcription of *pdeL* strictly depended on CDB and to a lesser extent on CIB, arguing that c-di-GMP impacts *pdeL* transcription via PdeL itself. A mutant lacking H-NS also

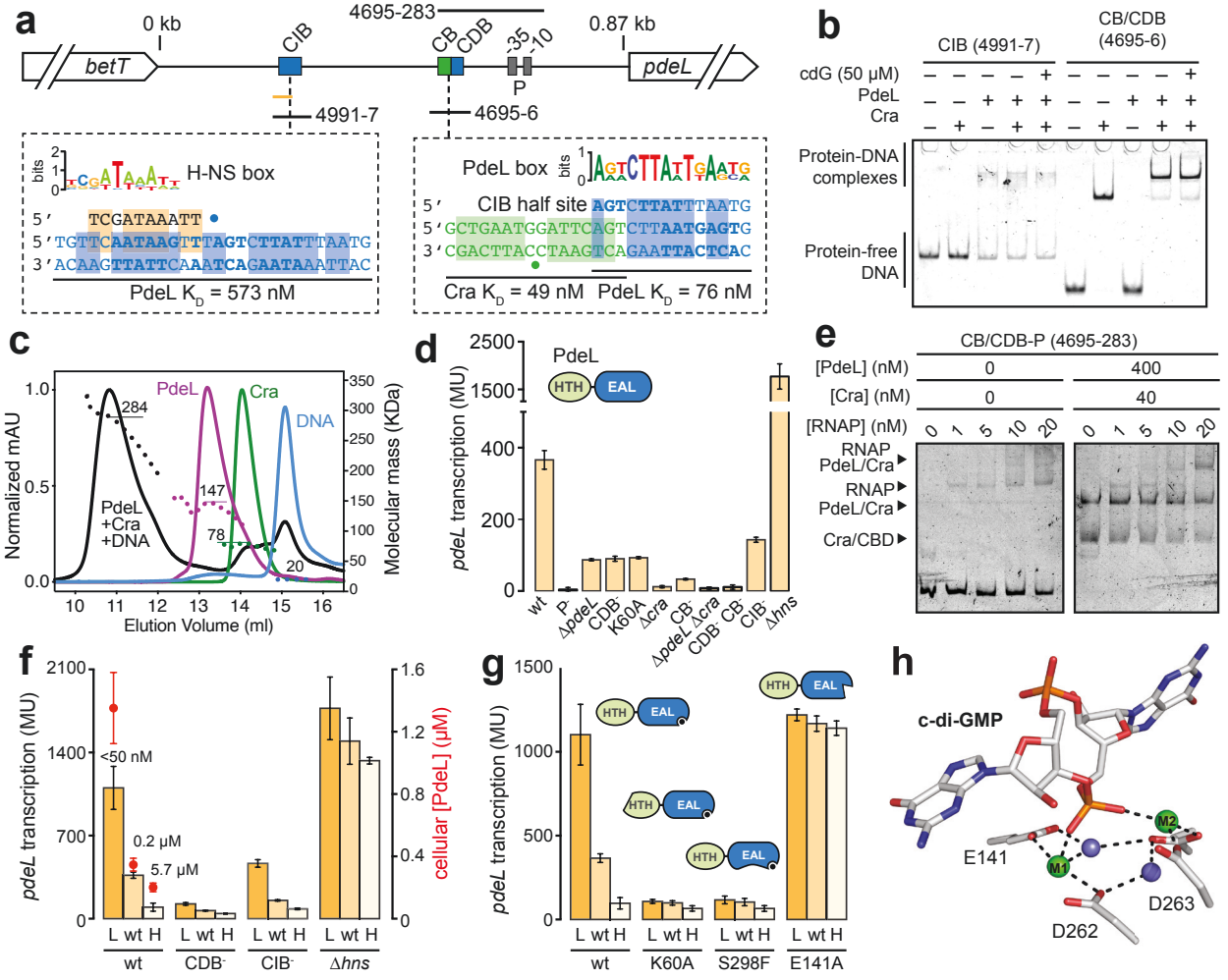


Figure 1: Regulation of *pdeL* transcription. **(a)** Schematic of *pdeL* promoter region. Binding sites for PdeL (CDB = Cra-dependent PdeL-box; CIB = Cra-independent PdeL-box), Cra (CB) and H-NS are shown in blue, green and orange, respectively. Blue and green dots mark the center of palindrome sites. Residues important for Cra and PdeL binding are highlighted in bold (see Fig. S1h) and binding affinities are indicated. Conservation of PdeL and H-NS binding sites are shown as WebLogos. **(b)** Electrophoretic mobility shift assay (EMSA) of 5' Cy3 labeled oligonucleotides and purified Cra-StrepII and PdeL-StrepII proteins. The position of the labeled oligonucleotides in the *pdeL* promoter region is indicated in **(a)**. **(c)** SEC-MALS of PdeL, Cra, and DNA fragments containing the Cra and PdeL binding sites (see: **(a)**) were analyzed individually or after mixing. The molecular masses of individual components and the complex are indicated. **(d)** Activity of *pdeL* promoter in wild type and different mutant strains. CDB^- and CB^-: sequences of binding sites were randomized to abolish TF binding (see Fig. S1h). Deletion mutants and a point mutation abolishing DNA binding of PdeL (K60A) are indicated. Inset shows domain architecture of PdeL with DNA binding domain (HTH) in green and catalytic EAL domain in blue. **(e)** Recruitment of RNA polymerase (RNAP) to the *pdeL* promoter. Increasing concentrations of RNAP were incubated with 5' Cy3 labeled target DNA (4695-283; see: **(a)**) in the absence or presence of 40 nM Cra-StrepII and 400 nM PdeL-StrepII. **(f)** Effect of c-di-GMP on *pdeL* transcription. *pdeL* transcription was assayed in strains with different c-di-GMP concentrations (L = low; wt; H = high) (see text). Levels of c-di-GMP are shown, and cellular concentrations of PdeL are indicated in red above the corresponding bars. **(g)** Transcription of *pdeL* in strains with different c-di-GMP concentrations (see **(f)**). *pdeL* mutant alleles included K60A (DNA-binding), S298F (dimerization) and E141A (c-di-GMP binding). c-di-GMP binding affinities for PdeL wt and E141A mutant were 31 nM and 686 nM, respectively. **(h)** Schematic of the PdeL active site in the T-state structure crystallized in the presence of c-di-GMP and $Ca2^+$ (PDB: 4LJ3). Bound c-di-GMP, metal (green), water (red), and residues involved in substrate binding are shown.

largely abolished c-di-GMP control (Fig. 1f). To dissect the mechanism of c-di-GMP-dependent regulation, we analyzed *pdeL* transcription in strains expressing different PdeL variants. Mutations interfering with DNA binding (K60A) (Reinders et al., 2016) or PdeL dimerization (S298F) failed to respond to c-di-GMP and showed baseline *pdeL* promoter activity (Fig. 1g). Because dimerization is critical for PdeL phosphodiesterase activity (Sundriyal et al., 2014), we

speculated that c-di-GMP-mediated transcription might be coupled to PdeL catalysis. In line with this, mutating a highly conserved active site glutamate (E141) required for c-di-GMP binding (Sundriyal et al., 2014) leads to complete de-repression of *pdeL* transcription (Fig. 1g,h). Based on these findings, we propose that PdeL senses the prevailing c-di-GMP concentration and acts as a transcription factor to boost its own expression when c-di-GMP levels are low.

2.3.3 PdeL is a c-di-GMP-mediated switch

Structural studies had indicated that the catalytic domain of PdeL adopts two distinct dimer configurations (Sundriyal et al., 2014). In the apo form, the EAL domain forms a relaxed open conformation that we termed the R-state. When bound to c-di-GMP, PdeL forms a closed or tight configuration, termed the T-state (Fig. 2a). The switch between R- and T-state is mediated via the EAL dimer interface with dimerization helices $\alpha5_A$ and $\alpha6'_B$ positioned parallel to each other in the R-state, but facing each other via their positively charged N-termini in the T-state (Fig. 2b,c). The unusual face-to-face helix conformation of the T-state is stabilized by the negatively charged residue D295 (Fig. 2c). The interchange between the two configurations entails large structural movements in loop 6, a conserved region connecting α5 with active site residues in strand β5 (Sundriyal et al., 2014) (Fig. 2b,c). This region structurally couples the dimerization interface to substrate occupancy in the active site via two highly conserved aspartate residues (D262, D263) and the anchoring glutamate E235 (Sundriyal et al., 2014) (Fig. 2b,c). E235 adopts a central hinge function by interacting either with active-site residue D263 in the R-state or, upon engagement of D263 in substrate binding, with residue T270 of the helix α5 in the T-state.

These structures provided a frame for PdeL control, in which c-di-GMP stabilizes the inert T-state while PdeL converts to the R-state at low ligand concentrations or in mutants unable to bind c-di-GMP. Accordingly, mutating residues required for T-state stabilization should lead to de-repression of *pdeL* transcription at high c-di-GMP concentrations. In line with this, mutating E235, its α5 partner T270, or the intercalating D295 lead to complete de-repression of *pdeL* transcription (Fig. 2d). Moreover, a random genetic screen for *pdeL* alleles able to restore swimming motility at high c-di-GMP concentrations (Reinders et al., 2016) (Fig. S2a) identified a range of activating substitutions in PdeL (Fig. S2b), all of which are positioned within or close to the dimerization helices α5 and α6 (Fig. S2c) and caused de-repression of *pdeL* transcription (Fig. S2d). Spontaneous mutations included T270A and D295N, arguing that all of these mutations lead to the partial or complete destabilization of the inert T-state, thereby activating *pdeL* transcription and PdeL catalysis (see below). D295, the critical residue for the stabilization of the PdeL T-state, is conserved in a large fraction of EAL phosphodiesterase domains (Fig. S2e,f). We speculate that R-to-T-state conversion represents a general feedback control mechanism through which phosphodiesterases are inhibited at high substrate concentrations.

Substrate-dependent conformational changes of the catalytic domain of PdeL were corroborated by cysteine cross-link experiments. We chose to substitute Y268 with Cys because these residues are in close proximity in the R-state but more distant in the T-state dimer (Fig. S2g,h). Oxidation

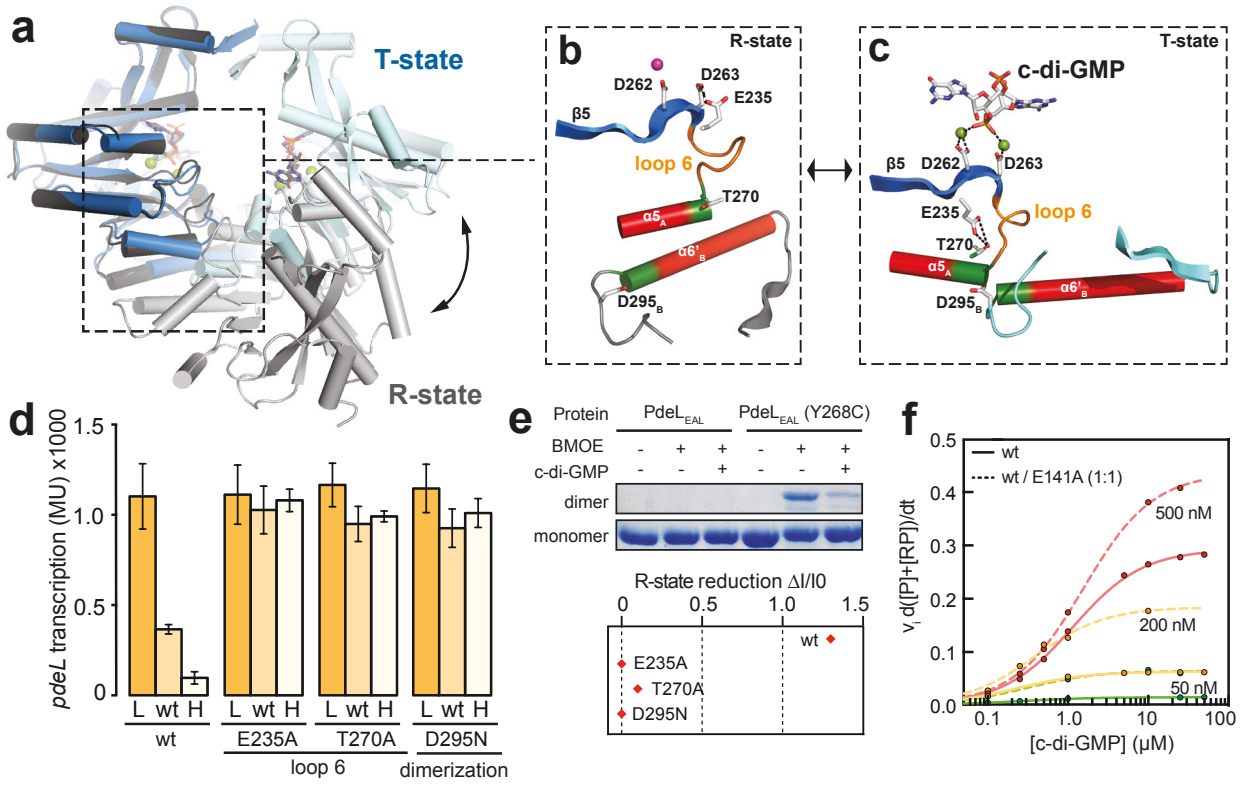


Figure 2: Distinct PdeL dimer conformations determine transcription and catalysis. **(a)** Overlay of T-state dimer (blue hues) with protomer B of the R-state configuration (grey), depicting large quaternary rearrangements between different conformations. **(b,c)** Zoom of PdeL active site, loop 6 (orange), and dimerization helices α5 and α6'$_B$. Dimerization helices are shown as magnet representation with electrostatic potential (green = positive charge, red = negative charge). Zoom of PdeL in **(b)** the canonical "open" dimer configuration (R-state) (PDB: 4LYK) and **(c)** the non-canonical "closed" dimer configuration (T-state) (PDB: 4LJ3). **(d)** *pdeL* promoter activity in strains carrying PdeL mutations in the dimerization interface and loop 6 (see **(b,c)**). L, wt, and H refer to levels of c-di-GMP as in Fig. 1f. **(e)** *In vitro* cysteine-crosslink with the purified EAL domain of PdeL (Y268C) (see Fig. S2g,h). **Top:** Crosslinks were performed without c-di-GMP, and with 50 μM c-di-GMP and proteins were separated by SDS-PAGE. PdeL wild type is shown as a negative control for crosslink specificity. **Bottom:** Quantification of *in vitro* cysteine-crosslinks of PdeL mutant variants. Band intensity differences between cysteine-crosslinks in the presence and absence of excess c-di-GMP are plotted as ΔI/I0 values. **(f)** PdeL turnover rates as a function of substrate concentration. PdeL activity (initial velocity at 37°C) was determined at different concentrations of wild-type PdeL (solid line) and a 1:1 mixture (stippled line) of PdeL wild type and E141A binding mutant. Data points were fitted with a simple Michaelis-Menten model.

of purified Y268C mutant protein revealed strong crosslinking in the absence of c-di-GMP, while cross-linking was reduced upon c-di-GMP addition. Mutating the anchoring glutamate E235, its α5 partner T270, or the intercalating D295 strongly increased cross-linking in the presence of c-di-GMP, demonstrating that mutations activating PdeL favor the R-state conformation (**Fig. 2e**).

If PdeL is forced into the inert T-state at high c-di-GMP concentrations, this may also impact its catalytic properties. One possibility is that overall activity is reduced upon substrate occupancy of both active sites of a dimer. If so, PdeL heterodimers consisting of a wild type and mutant protomer unable to bind c-di-GMP should be at least partially protected from this inhibitory effect. To test this, we compared the activity of purified wild type PdeL with a stoichiometric mix of PdeL wild type and the E141A mutant, which is unable to bind c-di-GMP and therefore catalytically inactive. Heterodimeric PdeL indeed showed increased turnover rates at higher substrate concentrations as compared to homodimeric PdeL (Fig. 2f), arguing that substrate inhibition is abolished at least partially under these conditions.

Hence, we propose that transcriptional and catalytic properties of PdeL are coupled, with high levels of c-di-GMP forcing the enzyme into an inert T-state. These results support a model in which PdeL is catalytically active at low substrate concentration but is gradually turned off at higher c-di-GMP concentrations, possibly through substrate saturation of both protomers.

2.3.4 PdeL shows strong cooperativity

Both catalytic (Sundriyal et al., 2014) and transcription activity (Fig. 1g) of PdeL depend on oligomerization. In line with this, *pdeL* promoter activity increased in a non-linear fashion (Hill-coefficient $n_h = 2.2$, Fig. 3f) with increasing PdeL concentrations. In contrast, the activity of the D295N mutant, which is unable to form a stable T-state, was constitutively high, irrespective of the protein concentration (Fig. 3a). We thus speculated that the non-linear behavior of PdeL results from oligomerization-based cooperativity and that c-di-GMP influences this equilibrium. Although full-length PdeL is a dimer in solution at low nanomolar protein concentrations (Fig. S3a), the equilibrium gradually shifted to tetramers at increasing protein concentrations (Fig. 3b). Microscale thermophoresis (MST) experiments revealed tetramerization constants in the low micromolar range (Fig. S3b). Importantly, PdeL tetramers converted back to dimers upon addition of c-di-GMP (Fig. 3c).

Crystals of full-length PdeL at 4.4 Å resolution harbor four PdeL molecules in the asymmetric unit. While the catalytic domain could be modeled successfully in the crystal packing, no density was detected for the DNA binding domain, presumably due to the flexibility caused by the inter-domain hinge. The PdeL tetramer consists of two R-state dimers (A/B and C/D) that form tetramerization interphases of 706 Å between protomers A and D and between B and C (Fig. 3d). The interface is formed through extensive contacts between N-terminal helices α1 and α2. A total of 14 residues mediate dimer-dimer contacts with reciprocal salt bridges and hydrophobic residues contributing to the stabilization of the tetramer (Fig. S3c, S4a,b). Because helices α1 and α2 are arranged in close proximity in the T-state dimer and form part of the dimerization interface, the T-state configuration is not compatible with tetramer formation (Fig. S4c). Substituting two tetramerization-specific residues, S114 (helix α1) or L168 (helix α2), with arginine generated mutants unable to form tetramers (Fig. 3e). Thus, the PdeL dimer/tetramer equilibrium is influenced by the protein concentration and c-di-GMP, forcing PdeL into the T-state conformation unable to form tetramers.

To investigate the role of tetramerization in *pdeL* transcription, we monitored *pdeL* promoter activity at varying cellular PdeL concentrations in a strain expressing *pdeL* from a tetracycline-inducible promoter (P_{tet}-*pdeL*) (Fig. S3d,e). To avoid interference from c-di-GMP, this strain also harbored an IPTG-inducible copy of *pdeH* (P_{lac}-*pdeH*), which upon maximal induction reduced c-di-GMP below the detection level (Fig. 1f). When PdeL concentrations were gradually increased, *pdeL* promoter activity increased in a highly cooperative manner ($n_h = 2.2$) with an activation constant of 230 nM PdeL (Fig. 3f). A mutant unable to form tetramers (L168R) completely abolished *pdeL* transcription activation in the physiological PdeL concentration range (Fig. 3f). Similarly, a *pdeL* promoter lacking the CIB binding site, although still induced at higher

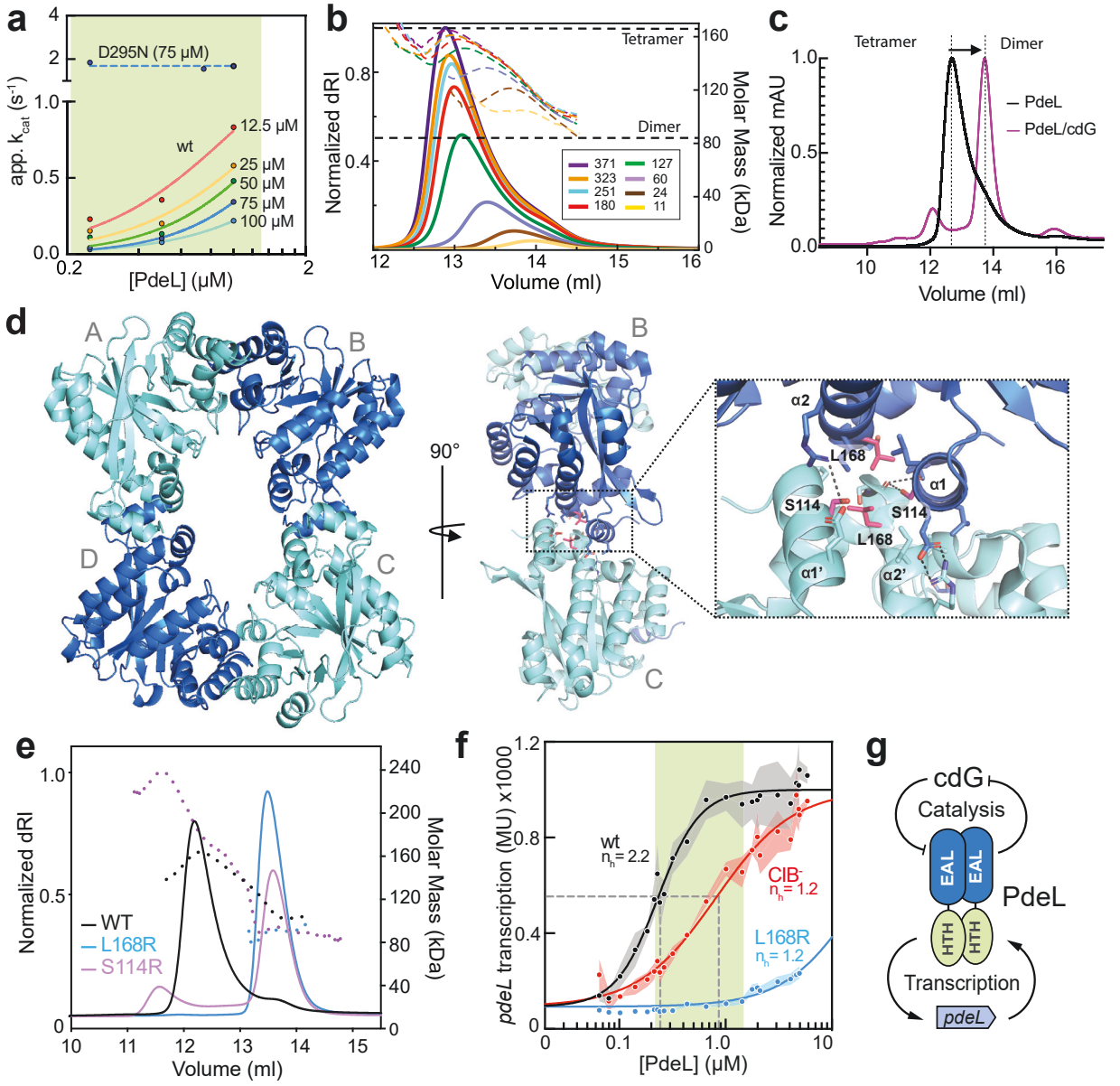


Figure 3: Cooperativity of PdeL catalysis and transcriptional activity. **(a)** Apparent k_{cat} of PdeL wild type (solid lines) and the PdeL (D295N) mutant (stippled line) as a function of increasing PdeL and c-di-GMP concentrations. Data points for D295N were fitted with a horizontal line, while PdeL wt data points were fitted with a Michaelis-Menten model with Hill-coefficient and app. k_{cat} of D295N set as v_{max} constraint. The physiologically relevant PdeL concentration range is indicated in the background in green (see Fig. 1f). **(b)** SEC-MALS analysis of PdeL at concentrations ranging from 11 to 371 μM. Colored lines correspond to the normalized refractive index with stippled lines indicating the calculated molecular mass. Black horizontal stippled lines indicate the molecular masses of PdeL dimer and tetramer. **(c)** SEC comparison of apo PdeL (black) and PdeL with c-di-GMP (purple). The arrow indicates the shift in the elution profile upon addition of c-di-GMP. **(d)** Tetramer of two R-state PdeL dimers with the four protomers (A-D) indicated. Protomers C and D were generated by applying crystallographic symmetry operations. Inset shows a zoom of the tetramerization interface between protomers A and D. Residues S114 (helix α1 and α1') and L168 (helix α2 and α2') are highlighted. **(e)** SEC-MALS of PdeL wild type (black) and mutants L168R (blue) and S114R (lila) loaded at equal concentration. **(f)** PdeL-dependent *pdeL* transcription in the absence of c-di-GMP inhibition. Strains carried an IPTG inducible copy of *pdeH* on a plasmid to reduce c-di-GMP below the detection limit (see: Fig. 1f). To tune PdeL concentrations, *pdeL* was expressed from a tetracycline-inducible promoter (P_{tet}-*pdeL*) (see: Fig. S3d,e and Materials and Methods for details). *lacZ* reporter fusions were used to assay *pdeL* promoter activity in wild type (black), CIB^- (red) and L168R mutants (tetramerization) (blue). Data points were fitted with a Michaelis-Menten fit with hill-coefficient. Kinetic values were calculated for wild type: K_{half} = 231 nM and h = 2.2; for CIB^-: K_{half} = 800 nM and h = 1.2; and for L168R: K_{half} = 18 μM and h = 1.2. The green area depicts physiologically relevant cellular PdeL concentrations (see Fig. 1f). **(g)** Schematic of PdeL feedback control determining transcription and catalysis (for details, see text).

PdeL levels, showed a non-cooperative response (n_h = 1.2) (Fig. 3f). These values are in line with PdeL binding affinity of CIB (K_D = 573 nM) and with measured PdeL levels in the promoter ON (1.35 μM) and OFF states (0.2 μM) (see Material and Methods). Thus, the CIB serves as an auxiliary element to generate switch-like, cooperative *pdeL* promoter activation in response to increasing cellular levels of PdeL.

Based on these results, we developed a simple model for PdeL regulation (Fig. 3g), in which PdeL is controlled by two connected positive feedback loops: (i) a double-negative enzymatic feedback, in which the PdeL phosphodiesterase negatively affects the c-di-GMP pool and c-di-GMP inhibits PdeL activity; and (ii) a positive transcriptional feedback where increasing PdeL levels enhance *pdeL* transcription. We propose that the observed cooperativity of PdeL depends on specific properties of the dimer-tetramer equilibrium.

2.3.5 PdeL imposes binary c-di-GMP regimes with memory

Above we showed that *pdeL* transcription inversely scales with c-di-GMP. To address how the *pdeL* promoter responds to dynamic changes of c-di-GMP, we monitored a *pdeL-lacZ* transcriptional reporter in a strain engineered to tune c-di-GMP (Fig. 1f, 4a). While gradually lowering IPTG-dependent *pdeH* expression led to a stepwise increase of c-di-GMP, *pdeL* expression remained constantly high and only turned off in the low micromolar range of c-di-GMP (Fig. 4b). Tuning c-di-GMP levels in the opposite direction led to the activation of *pdeL* transcription but exposed a large hysteresis window covering the nano- and sub-micromolar range. For example, at 1 μM c-di-GMP, *pdeL* transcription remained ON when cells had experienced low c-di-GMP before but remained OFF when cells previously experienced high c-di-GMP levels (Fig. 4b). Intriguingly, history-dependent differences in *pdeL* transcription were largest in a c-di-GMP concentration window determined for *E. coli* under similar growth conditions (Reinders et al., 2016). The hysteresis window gradually narrowed over time (Fig. 4c) and completely collapsed in mutants lacking the CIB binding (Fig. 4d). Hysteretic expression of *pdeL* appears to be mediated primarily by *pdeL* ON-kinetics, with OFF-kinetics being governed mainly by protein dilution (Figs. 4e,f). Similarly, a large hysteresis window was observed when measuring c-di-GMP levels upon tuning *pdeH* expression in both directions. Divergent c-di-GMP levels were dependent on PdeL and were only observed 4.5 h after shifting cells to the new IPTG concentrations, while the hysteresis window had collapsed 8 h after the shift (Fig. 4g,h). Thus, both *pdeL* transcription and PdeL-mediated c-di-GMP levels show bistability, indicating that PdeL is a stochastic molecular switch that confers cellular memory and robustly determines cellular levels of c-di-GMP in *E. coli*.

2.3.6 Bimodal *pdeL* expression imposes binary c-di-GMP regimes

To assay *pdeL* promoter activity at the single-cell level and at different concentrations of c-di-GMP, we engineered transcriptional fusions to different fluorophores and inserted them into strains where the chromosomal copy of *pdeH* was replaced by a tunable copy on a plasmid (Fig. 5a). While the *pdeL*::*gfp* reporter was off in all cells at high c-di-GMP and homogenously on at low c-di-GMP concentrations, bimodal expression was observed at intermediate c-di-GMP levels (Fig. 5b,c). To

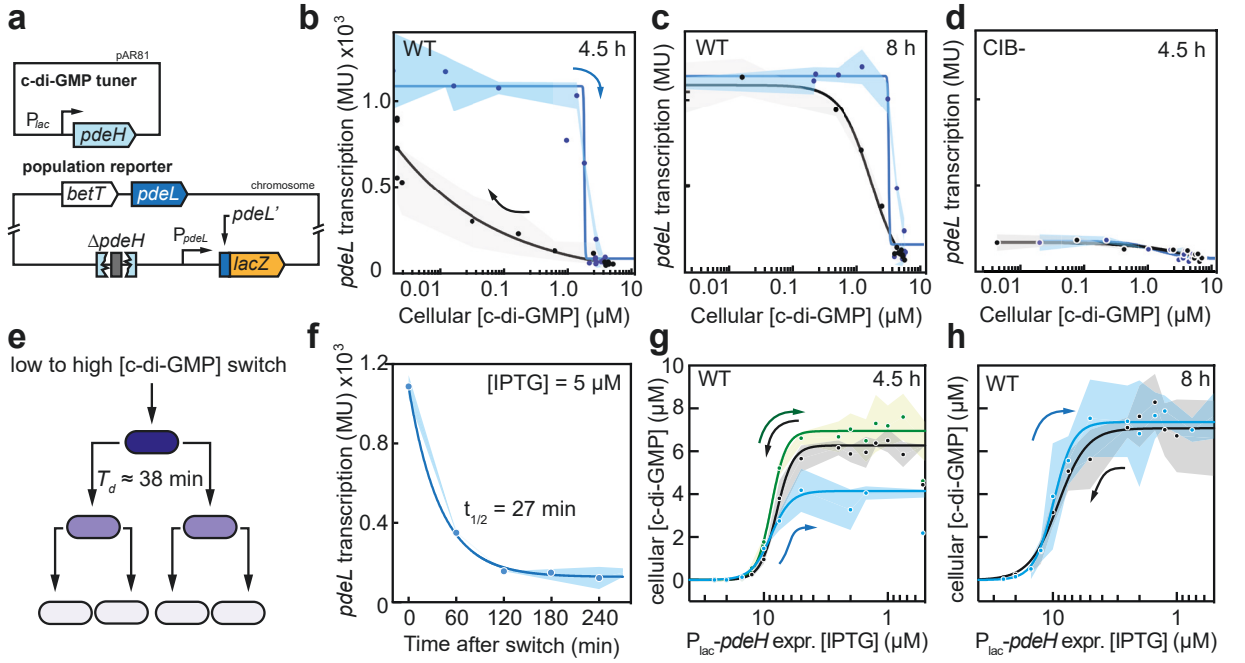


Figure 4: Expression of *pdeL* is bistable. **(a)** Schematic of reporter strains used. The *pdeL* promoter region was fused to *lacZ* by replacing the native lac promoter in the chromosome (Reinders et al., 2016). **(b,c)** Hysteresis experiments with reporter strain (see: **(a)**) containing c-di-GMP tuner plasmid to establish a range of c-di-GMP-levels. Fractions of ON and OFF cells were determined after 4.5 h **(b)** and 8 h **(c)**, respectively. Arrows indicate the direction of changes in *pdeL* transcription from pre-established c-di-GMP regimens (L = blue; H = black). Curves were fitted with a sigmoidal least square fit. **(d)** Hysteresis with the CIB$^-$ mutant with conditions as outlined in **(b)**. **(e)** Schematic of protein dilution during cell division with doubling recorded for the conditions used. **(f)** Time-dependent *pdeL* OFF-kinetics. The population reporter strain containing the c-di-GMP tuner plasmid was grown with 65 μM IPTG overnight and diluted back into fresh medium with 5 μM IPTG before *pdeL* transcription was monitored for 4.5 h. Data points were fitted with a simple decay model yielding a half-life of $t_{1/2}$ = 27 min. **(g,h)** Development of c-di-GMP concentration upon altered expression of the c-di-GMP tuner (see: **(a)**) in strains with pre-established high (black) or low (blue, green) c-di-GMP levels. Black and blue lines indicate *pdeL*+ strains, green marks the Δ*pdeL* mutant. C-di-GMP levels were measured after 4.5 h **(g)** and 8 h **(h)**, respectively.

demonstrate that bimodal *pdeL* expression generates bimodal patterns of c-di-GMP, we made use of a novel c-di-GMP sensor that was constructed by domain insertion profiling (Nadler et al., 2016) of the c-di-GMP binding protein BldD from *Streptomyces* (Tschowri et al., 2014). Fluorescence intensity of this sensor directly scales with c-di-GMP levels with maximal and minimal emission above 1 μM and below 100 nM, respectively. Combined expression of the GFP-based c-di-GMP sensor and a chromosomal *pdeL*::*mCherry* reporter in the same strain at intermediate c-di-GMP levels revealed strictly inverse fluorescence patterns indicating that cellular levels of c-di-GMP are controlled by *pdeL* expression (Fig. 5d). To corroborate this, we compared the dynamic changes of c-di-GMP in individual cells of a *pdeL*+ and a Δ*pdeL* strain when gradually reducing c-di-GMP levels using the P_{lac}-*pdeH* tuner plasmid. Both strains were grown overnight in the absence of IPTG to pre-establish high intracellular c-di-GMP concentrations and were then shifted to fresh media containing gradually increasing IPTG concentrations leading to increased expression of *pdeH*. This led to a gradual and monomodal reduction of c-di-GMP in the Δ*pdeL* strain, while in the presence of PdeL, the culture was split into two distinct subpopulations with high and low

c-di-GMP, respectively (Fig. 5e). The PdeL-mediated switch of individual cells to a low c-di-GMP regime occurred at IPTG concentrations that are significantly below those required to reduce c-di-GMP in the absence of PdeL. Thus, PdeL is a hypersensitive switch that converts gradual changes of c-di-GMP into a binary outcome.

To scrutinize the role of *pdeL* on c-di-GMP levels under more physiological conditions that do not depend on c-di-GMP tuning, we analyzed c-di-GMP in individual cells of the *E. coli* MG1655 wild type isolate CGSC 6300 (Guyer et al., 1981). Lab-adapted *E. coli* strains like CGSC 7740 that were used for motility/chemotaxis studies harbor an IS insertion in the flhCD promoter region (Barker et al., 2004; Blattner et al., 1997). This leads to the constitutive expression of flagellar genes and *pdeH*, resulting in artificially low levels of c-di-GMP. Because the ancestor strain CGSC 6300 harbors an uncompromised flhCD promoter (Barker et al., 2004, Kim et al., 2020), its c-di-GMP metabolism is not genetically altered. When grown on glucose-based minimal media, c-di-GMP levels were uniformly high (Fig. 5f). In contrast, bimodal patterns of c-di-GMP were observed when MG1655 was grown in minimal media with alternative carbon sources like glycerol, fumarate, or alpha-ketoglutarate (Fig. 5f). The addition of casamino acids to minimal glycerol media abolished c-di-GMP bimodality, phenocopying the monomodal distribution observed on glucose. We assume that the c-di-GMP increase results from the activation of a diguanylate cyclase in response to one or several amino acids, as shown previously for L-arginine in S. typhimurium (Mills et al., 2015). Importantly, c-di-GMP bimodality in glycerol media was strictly dependent on PdeL, but not on PdeH (Fig. 5h). Mutants abolishing PdeL catalysis (E141A) or tetramerization (L186R) as well as a mutant lacking the low affinity PdeL binding site in the *pdeL* promoter region (CIB) showed monomodal high c-di-GMP. In contrast, a mutation that locks PdeL in the active R-state (D295N) also abolished bimodality but generated cells with low c-di-GMP (Fig. 5h). Thus, PdeL converts natural fluctuations of c-di-GMP into a robust binary output with subpopulations maintaining either low or high levels of c-di-GMP.

2.3.7 PdeL instructs *E. coli* lifestyle and protects against phage predation

To determine if PdeL impacts the behavior of cells experiencing changes in c-di-GMP, we assayed *E. coli* biofilm formation and escape, processes that are directly controlled by c-di-GMP (Jenal et al., 2017). First, we monitored the effect of PdeL on poly-GlcNAc-dependent (PGA) biofilm formation (Steiner et al., 2013) in a strain harboring the IPTG tunable *pdeH* construct. When c-di-GMP levels were gradually increased by lowering *pdeH* expression, surface attachment was triggered only when ITPG concentrations reached levels at which *pdeL* expression is turned off. In contrast, an isogenic strain lacking PdeL showed biofilm formation already at higher *pdeH* expression levels (Fig. 6a). This indicated that PdeL provides an effective buffer against c-di-GMP noise to sustain the planktonic lifestyle of *E. coli*. Conversely, PdeL strongly promoted the escape of *E. coli* from pre-established biofilms upon lowering c-di-GMP levels. The observed large variations of escape rates are in line with stochastic induction of PdeL under these conditions (Fig. 6b).

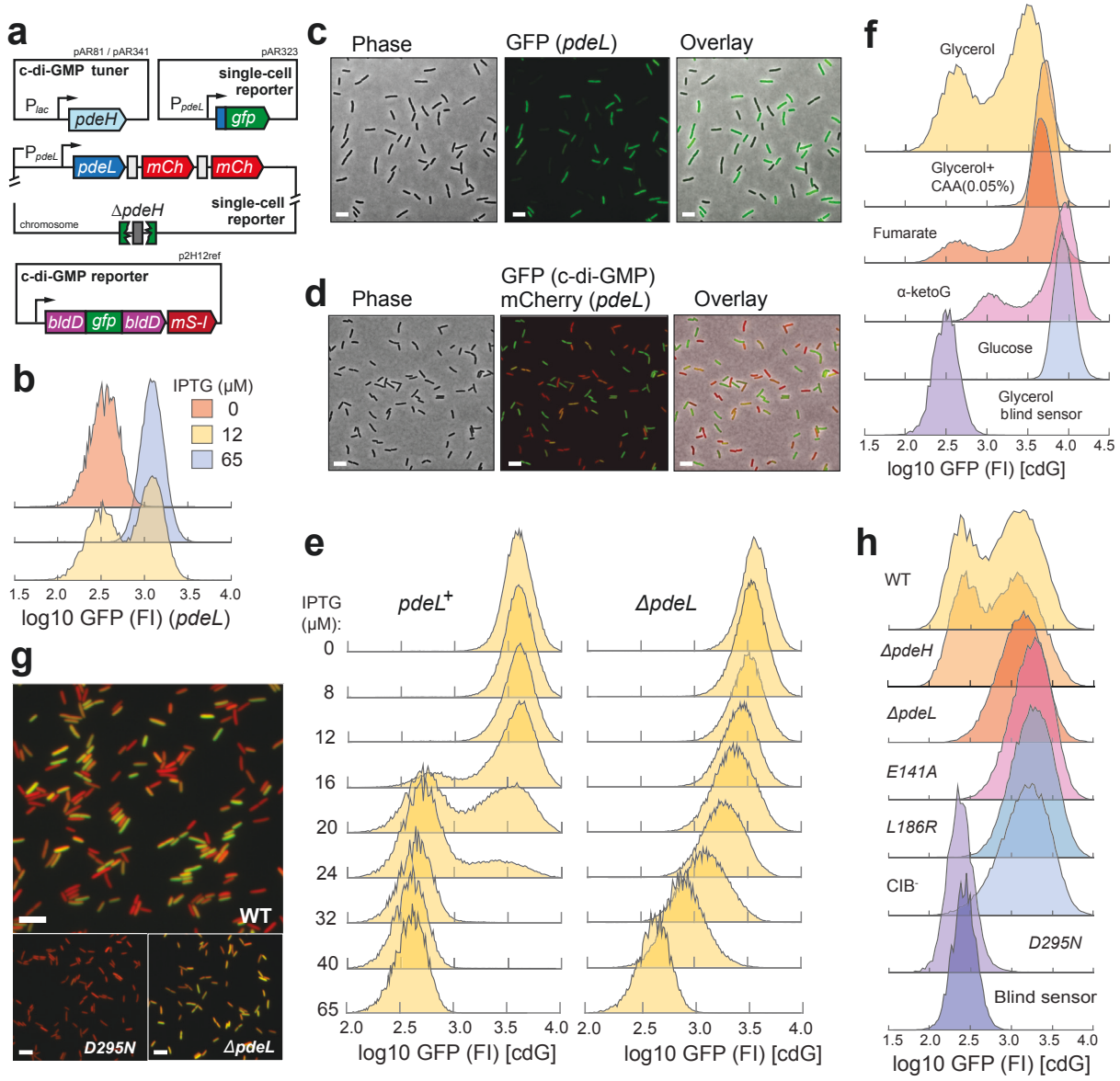


Figure 5: Stochastic expression of *pdeL* established bimodal c-di-GMP regimes. **(a)** Schematic of reporter constructs used. Single-cell *pdeL* transcription was measured with a transcriptional fusion of a tandem *mCherry* (mCh) downstream of the *pdeL* gene in the *E. coli* chromosome or with a plasmid-born *PpdeL-gfp* construct (pAR323, on the low-copy pUA66 backbone). To tune c-di-GMP levels, a *Plac*-driven copy of *pdeH* was expressed from plasmid pAR81. Expression of the operon containing the c-di-GMP reporter and *mScarlet-I* (mS-I) was induced with 200 nM anhydrotetracycline from plasmid p2H12ref. **(b)** *pdeL* promoter activity was determined in a *pdeL*+ Δ*pdeH* mutant harboring plasmids pAR81 (c-di-GMP tuner) and pAR323 (*pdeL* reporter) (see: a). Cultures grown in the absence of IPTG (high [c-di-GMP]) were shifted to media containing 0, 12, or 65 μM IPTG, grown for 6 h and analyzed by flow cytometry. **(c)** Fluorescence micrographs illustrating bimodal *pdeL* promoter activity in a strain carrying the GFP reporter plasmid pAR323 (see: a). Scale bar: 5 μM. **(d)** Fluorescence micrographs illustrating *pdeL* expression (red) and c-di-GMP levels (green) in a strain carrying a chromosomal *pdeL*::mCherry reporter and a plasmid-born copy of the c-di-GMP reporter (p2H12ref, see: a). Scale bar: 5 μm. **(e)** PdeL establishes binary c-di-GMP regimes. Pre-cultures of an *E. coli* Δ*pdeH* mutant containing a wild-type *pdeL* allele (*pdeL*+) or a *pdeL* deletion (Δ*pdeL*) and carrying plasmids pAR341 and p2H12ref (see: a) were precultured under conditions that establish high intracellular c-di-GMP levels (0 μM IPTG). Cultures were then diluted into media containing increasing levels of IPTG as indicated and grown for 8 h and analyzed by flow cytometry. **(f)** *E. coli* wild type strain carrying plasmid p2H12ref (see: a) expressing the c-di-GMP reporter was grown in minimal media with different carbon sources to mid log phase and analyzed by flow cytometry. A 'blind' sensor carrying a mutation in the c-di-GMP binding site of BldD was used as a control (Tschowri et al., 2014). **(g)** Fluorescence micrographs illustrating c-di-GMP levels in *E. coli* wild type (WT), *pdeL* (D295N), and Δ*pdeL* mutant cells carrying plasmid p2H12ref (see: a). Cultures were grown in glycerol-based minimal medium. Overlay of green: c-di-GMP (GFP), and red: constitutive *mScarlet-I* expressed from the same operon as the c-di-GMP sensor from p2H12ref. Scale bar: 5 μM. **(h)** Distribution of c-di-GMP in *E. coli* populations carrying plasmid p2H12ref (see: a) and grown in glycerol-based minimal medium. Mutants of *pdeL* included Δ*pdeL*, the c-di-GMP "blind" mutant *pdeL* (E141A), the tetramerization mutant *pdeL* (L186R), the constitutive active mutant *pdeL* (D295N), and the randomized Cra-independent PdeL-binding site (CIB) in the *pdeL* promoter region. A strain expressing a c-di-GMP binding-deficient mutant sensor was used as control as indicated in f).

Next, we considered that the c-di-GMP-mediated production of a novel surface exopolysaccharide (EPS) may serve as an entry door for bacteriophages. Specific members of these bacterial predators use EPS, capsules, or LPS as the primary receptor and have evolved virion-associated polysaccharide depolymerases to gain access to secondary cell surface receptors (Fernandes and São-José, 2018; Latka et al., 2017; Pires et al., 2016). Screening a library of *E. coli* bacteriophages for agents that specifically prey on cells with high c-di-GMP led to the identification of phage N4 (Fig. 6e), which specifically infects *E. coli* strain CGSC 6300 but not an isogenic strain constitutively expressing *pdeL* (Figs. 6c). Importantly, replacing the *pdeL* wild-type allele on the chromosome with *pdeL* (D295N) leads to complete resistance against phage N4, similar to mutations in known N4 resistance genes, including nfrA, nfrB, or wecB (McPartland and Rothman-Denes, 2008) (Figs. 6c,d). The observation that overexpression of *pdeL* (K60A) (DNA binding mutant) but not *pdeL* (E141A) (catalysis mutant) conferred phage resistance strongly indicated that PdeL phosphodiesterase activity is critical for phage protection. From this, we concluded that phage N4 selectively infects *E. coli* cells in a state where PdeL is inactive and c-di-GMP concentrations are increased. Together, these experiments emphasize the relevance of PdeL as a central regulatory player of the motile-sessile switch in *E. coli* and suggest that stochastic activation of PdeL protects *E. coli* from phage predation.

Since *E. coli* swimming motility is tightly regulated by c-di-GMP (Boehm et al., 2010) we monitored *pdeL* expression in cells moving in a chemical gradient in motility plates (Fig. S2a). Cells carrying a fluorescent reporter for *pdeL* transcription were monitored at the initial site of inoculation, at the migration front, and in a zone proximal to the visible migration front (pioneer zone) (Fig. 6f). The expression of *pdeL* showed a broad distribution at and close to the inoculation site, where cells were mostly in a quiescent, non-dividing state. At positions more distant from the inoculation site, *pdeL* expression showed clear bimodality with different ratios of ON and OFF cells at different sites (Fig. 6f), arguing that swimming and sessile cells are intermixed in this zone of the plate. In the pioneer zone, where swimming activity is expected to be most pronounced, most cells were actively expressing *pdeL*. From this, we conclude that *pdeL* expression dynamically responds to different physico-chemical conditions during chemotaxis in motility plates. Together, these experiments emphasize the relevance of PdeL as a central regulatory player of the motile-sessile switch in *E. coli* and demonstrate that stochastic activation of PdeL protects *E. coli* from phage predation.

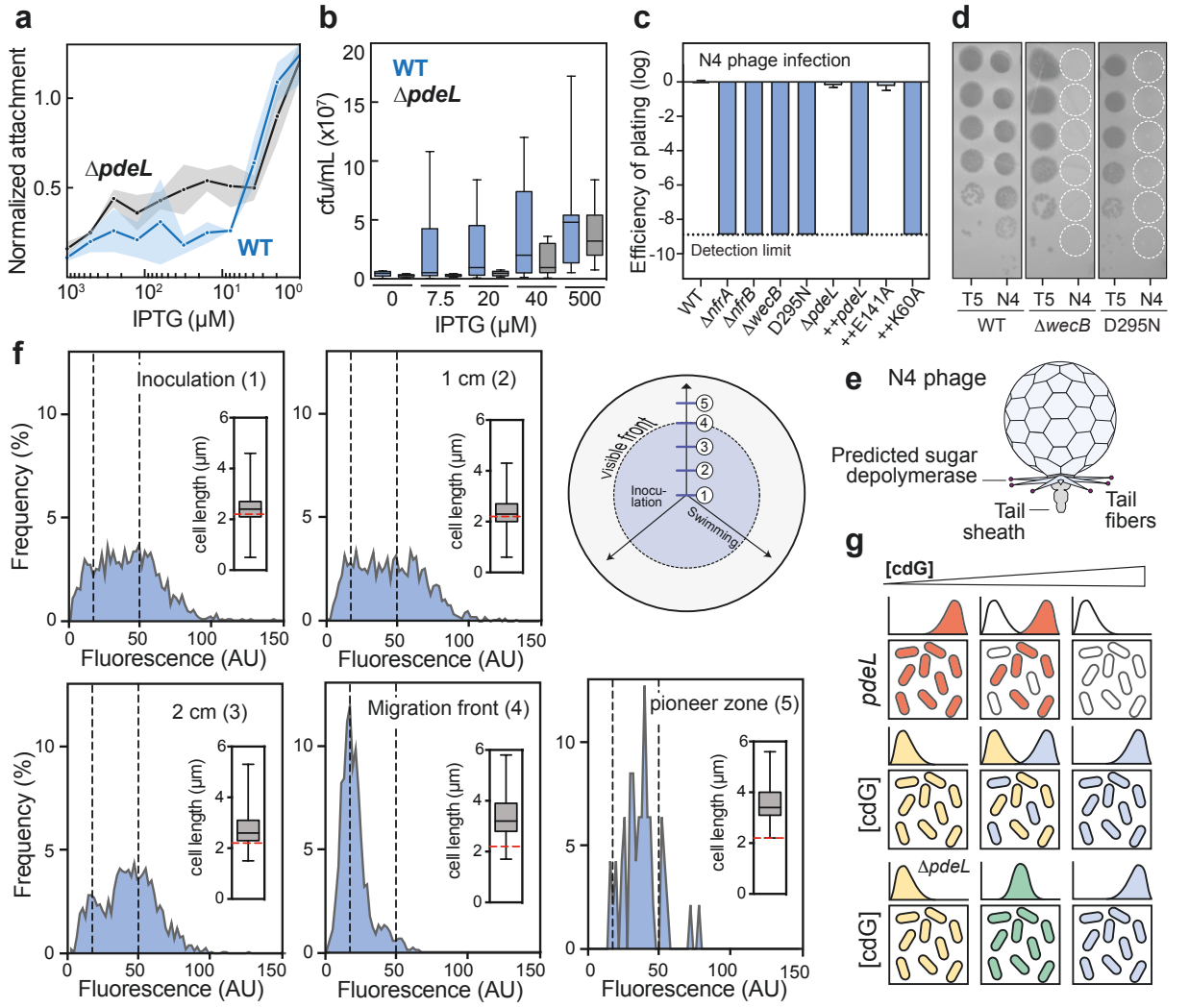


Figure 6: PdeL drives c-di-GMP-dependent biofilm formation and escape. **(a)** *E. coli* biofilm formation of a *pdeL*+ (blue) and Δ*pdeL* strain (black) at different intracellular levels of c-di-GMP. The concentration of c-di-GMP was set in the biofilm assay strain (for details see: Materials and Methods) by IPTG-mediated expression of *pdeH* from the c-di-GMP tuner plasmid (pAR81, see Fig. 4a). **(b)** Escape of *E. coli pdeL*+ (blue) and Δ*pdeL* (black) cells from pre-formed biofilm upon IPTG-mediated induction *pdeH* from the c-di-GMP tuner plasmid pAR81. Number of cells escaping to the soluble fraction was scored by serial dilution plating 3 h after fresh medium with different concentrations of IPTG was added to pre-formed biofilms. Median values are indicated with error-bars showing upper and lower quartiles. **(c)** PdeL protects against phage N4 via catalytic activity. EOP (efficiency of plating) depicts the phage susceptibility of the strain relative to wt. *pdeL* was expressed (++*pdeL*) with 100 μM IPTG from Plac on the pNDM220 plasmid. The N4 phage requires the N4 receptors NfrA, NfrB and WecB (Kiino et al., 1993) for infection and is inhibited by a PdeL (D295N) isogenic mutation. **(d)** Representative images of phage T5 and N4 plaques with 10-fold serial dilution top to bottom. **(e)** Schematic drawing of bacteriophage N4. **(f)** Transcription of *pdeL* in individual cells (Fig. 5a) travelling through a chemical gradient in TB soft agar plates. Dashed black lines mark mean *pdeL* transcription levels in liquid TB at high and low c-di-GMP, respectively (Fig: 1f). Cells were harvested 12 h after inoculation at the sites indicated in the schematic of the motility plate. The distribution of *pdeL* transcription and median cell lengths as indicator for growth state are shown. Red line shows median cell length of inoculation strain grown overnight on TB plates. **(g)** Model for PdeL control and its impact on cellular c-di-GMP concentrations. Reduction of c-di-GMP under a threshold level leads to upregulation of PdeL expression (red cells, upper panel), which in turns contributes to robustly lower cellular levels of c-di-GMP (yellow, middle panel). In contrast, a net increase in c-di-GMP levels inhibits PdeL and leads to transcriptional stalling of *pdeL* expression (white). This will contribute to robust increase of c-di-GMP in the cell (blue). Noise-induced fluctuations of c-di-GMP levels result in stochastic expression of *pdeL* establishing bimodal populations harboring distinct concentrations of c-di-GMP. PdeL activity drives bimodal distributions of cellular c-di-GMP, as a *pdeL* deletion mutant (Δ*pdeL*) displays a unimodal distribution of intermediate c-di-GMP levels (green, bottom panel).

2.4 Discussion

Binary switches generally operate on the transcriptional level to modulate the cell's genetic circuitry (Carraro et al., 2020). Here, we describe a binary switch regulating the cellular concentration of a small signaling molecule. We propose that PdeL determines c-di-GMP levels during behavioral transitions of *E. coli* by acting as a catalyst, a c-di-GMP sensor, and a transcription factor at the same time. PdeL is fully active in planktonic cells with low levels of c-di-GMP, keeping its own protein levels high and buffering against sporadic c-di-GMP bursts to avoid unwarranted activation of sessility factors. In contrast, in cells with high c-di-GMP levels, PdeL switches into its inert form, leading to a drop in PdeL levels and reduced PdeL activity. At intermediate c-di-GMP concentrations, PdeL is forced into the ON- or OFF-conformation, thereby converting gradual unimodal changes of the second messenger into robust bimodal outcomes with unambiguous activation of downstream processes.

We have shown that PdeL dynamically transits between a closed T-state dimer and an open R-state conformation with the propensity to form tetramers. The dynamic equilibrium between these two conformations is inversely influenced by the PdeL protein concentration and by c-di-GMP and may explain the cooperative behavior of PdeL both as an enzyme and as a transcription factor. Our data demonstrate that the T-to-R-state switch is mediated through structural coupling between the substrate-binding site and the dimerization interface of PdeL. Similar structural coupling was proposed for the activation of phosphodiesterases by associated sensory domains (Rao et al., 2009; Winkler et al., 2014). Likewise, EAL domains that have lost their catalytic function and have adapted a role as c-di-GMP effectors use similar mechanisms to couple c-di-GMP binding to various cellular readouts (Minasov et al., 2009; Navarro et al., 2011). Thus, tight coupling between substrate binding and dimerization could be a conserved feature of EAL domains.

PdeL is both an enzyme catalyzing c-di-GMP degradation and a sensor for its substrate. This entails a mechanism that inhibits catalysis at high substrate concentrations. Experiments with purified PdeL heterodimers composed of a catalytically active and an inert protomer, suggested that reaching a stable substrate-inhibited T-state conformation requires binding of c-di-GMP to both active sites. Based on this, we speculate that the catalytic mechanism of PdeL may be highly asymmetric, with only one protomer being active at the same time and the catalytic pocket of the neighboring protomer being kept free. Binding to both active sites would require c-di-GMP to raise above a threshold concentration, which in turn would force the dimer into the inert T-state, a conformation that is unable to transition into the highly active tetramer state. Intriguingly, the residue stabilizing the PdeL T-state, D295, is conserved in a large fraction of EAL/phosphodiesterase domains. Thus, it is possible that this represents a general feedback control mechanism through which phosphodiesterases are inhibited at high substrate concentrations.

The alternative conformations of PdeL also impact transcriptional autoregulation. Intriguingly, PdeL levels never drop below 200 nM, even at high concentrations of c-di-GMP. This concentration is high enough to engage the high-affinity binding site (CDB) in the *pdeL* promoter region and, together with Cra, maintain basal levels of PdeL. Full induction of *pdeL* transcription at low

c-di-GMP requires CIB, a low-affinity binding site further upstream in the promoter region that only comes into play when PdeL reaches higher levels. Binding to CDB and CIB may increase PdeL concentration locally, thereby shifting the equilibrium towards a highly active tetramer. Moreover, PdeL tetramers may engage both binding sites, leading to DNA bending and displacing the general gene silencer H-NS. In line with this, PdeL tetramerization, as well as an intact CIB, are critical for the transcriptional boost observed at low levels of c-di-GMP. Moreover, the region between CIB and CDB contains binding sites for IHF and Fis (Grainger et al., 2006), proteins known to promote DNA bending. How exactly PdeL interferes with H-NS and how it is able to recruit RNA polymerase to the *pdeL* promoter region remains to be shown.

Our functional analyses imply that, depending on the cells' history, PdeL acts either as a quencher or as an amplifier of behavioral transitions. At low c-di-GMP concentrations, PdeL sustains the planktonic, non-adherent state of *E. coli* by capping c-di-GMP below a threshold concentration and by effectively quenching spikes of c-di-GMP generated by the unsolicited expression or activation of diguanylate cyclases. But PdeL also strongly promoted the escape of *E. coli* from preformed biofilms. In order to escape from biofilms, bacteria need to reduce their overall c-di-GMP concentration through the activation of one or several phosphodiesterases (Rumbaugh and Sauer, 2020). Activation of PdeL under these conditions likely reinforces other phosphodiesterases, accelerating the collapse of c-di-GMP and robustly driving cells back into the motile, single-cell state. Similar mechanisms to quench or amplify concentration changes of signaling molecules may operate in other NSM networks with multiple «makers» and «breakers» (Bassler et al., 2018; Galperin, 2018; Galperin et al., 2010), in particular, if these operate on short time scales. For instance, it was shown recently that when encountering surfaces, planktonic bacteria respond to mechanical cues by increasing their c-di-GMP or cAMP levels. This blocks motility (Boehm et al., 2010; Schniederberend et al., 2019), leads to the assembly of adhesion factors (Hug et al., 2017; Laventie et al., 2019; Schniederberend et al., 2019), and to the expression of virulence genes (Persat et al., 2015b, 2015a). Molecular switches like PdeL could define an upper threshold of c-di-GMP that maintains the ability of planktonic bacteria to respond to mechanical cues in a highly sensitive manner when encountering surfaces (Lee et al., 2018). By imposing robustness to the global concentrations of signaling compounds, similar mechanisms may also limit erroneous cellular responses, thereby reducing costs resulting from noisy signaling inherent in large converging networks. Finally, PdeL may enable spatially confined signaling by keeping the global cellular pool of c-di-GMP low, thereby contributing to the «insulation» of signaling microdomains providing specificity to diguanylate cyclases that stimulate specific cellular processes locally (Richter et al., 2020). Upon switching into the off state, PdeL would abandon its neutralizing effect on diguanylate cyclases that act globally, providing a further boost to local systems. NSMs like c-di-GMP or cAMP control global and highly orchestrated cellular programs like motility, virulence, or biofilm formation (Galperin, 2018; Hengge, 2020; Jenal et al., 2017; Kalia et al., 2013). Since many of the respective cellular processes are highly sensitive to small changes in NSM concentration, bacteria may need mechanisms to convert

gradual unimodal changes of signaling compounds into stable binary outputs. In contrast, NSMs like (p)ppGpp, which often relies on a single combined synthase and hydrolase catalyst, may solve this problem by directly controlling their synthesis/hydrolysis activities in a switch-like fashion (Tamman et al., 2020).

Our experiments demonstrated that *pdeL* is expressed in a bimodal and bistable fashion, conferring *E. coli* with memory about its c-di-GMP state. During transitions from low to high c-di-GMP, memory is likely stored in the prevailing concentration of PdeL itself, which eventually decreases by dilution during growth after its de novo synthesis is suspended. In contrast, bacteria experiencing a reduction of c-di-GMP may retain memory in the form of a hyper-stable inert T-state of PdeL. Collective survival strategies such as bet-hedging and division of labor lead to increased stability and fitness of bacterial populations and are used by bacteria to minimize risks inherent in specific lifestyles or environmental conditions (Diard et al., 2013; Dubnau and Losick, 2006). Similarly, bistable c-di-GMP regimes likely enable *E. coli* to maintain subpopulations with distinct phenotypes and behavior during precarious and costly lifestyle transitions. For instance, maintaining a subpopulation of motile cells during surface colonization may promote the dissemination, as shown previously for the opportunistic human pathogen *P. aeruginosa* (Laventie et al., 2019). Likewise, the co-existence of different c-di-GMP programs may effectively protect against predators like bacteriophages or amoebal grazers, which generally prey on unprotected single bacteria, while bacteria in biofilms are protected by an extracellular matrix (Labrie et al., 2010; Simmons et al., 2020; Vidakovic et al., 2017). In contrast, some bacteriophages also infect their prey by using specific surface-exposed EPS as primary receptors (Pires et al., 2016). Given the prominent role of c-di-GMP in stimulating EPS synthesis and secretion (Krasteva et al., 2010; Merighi et al., 2007; Steiner et al., 2013; Thongsomboon et al., 2018; Whitney et al., 2012), high c-di-GMP states may impose a specific predator-mediated burden and predispose bacteria for phage infections. We show here that this is indeed the case for *E. coli*, which is effectively protected from infections by phage N4 when PdeL is active and c-di-GMP levels are low. N4 infection requires NfrB, an inner membrane protein with homology to glycosyltransferases and to a MshEN-like c-di-GMP binding domain (McPartland and Rothman-Denes, 2008; Wang et al., 2016). We speculate that phage N4 uses an as yet unidentified, c-di-GMP dependent EPS as the primary receptor to infect *E. coli* and that the stochastic regulation of PdeL effectively protects part of the *E. coli* population from phage attack.

This study raises several important future questions. How widespread are mechanisms that impose cellular heterogeneity to NSM-based networks? Are molecular mechanisms converting graded changes of NSMs into binary outputs maybe a necessary consequence of complex signaling architectures of networks operating with small signaling molecules like cAMP, c-di-GMP, and others? Also, what is the exact role of PdeL and similar digital converters in spatiotemporal control of NSM networks? E.g., does the PdeL transcription factor exclusively control its own expression, or does this regulon expand to additional *E. coli* genes? Finally, why is *pdeL* autoregulation strictly coupled to the activity of Cra, a sensor of the metabolic flux through the central carbon metabolism (Bley Folly et al., 2018; Kochanowski et al., 2013; Ramseier, 1996)? Cra activity

is highest under gluconeogenic conditions and at low growth rates, indicating that c-di-GMP heterogeneity may contribute to population fitness primarily under such growth conditions. The observations that both Cra and c-di-GMP play important roles in regulating virulence of pathogenic *E. coli* (Branchu et al., 2013; Hu et al., 2013; Njoroge et al., 2013; Richter et al., 2014) indicate that important behavioral processes are coordinated with growth and metabolic activity in this organism. How this contributes to the successful adaptation of pathogenic *E. coli* to the human host remains to be determined.

2.5 Materials & Methods

Bacterial strains and growth media

The bacterial strains and plasmids used in this study are listed in Table S1 and S2 respectively. *E. coli* K-12 MG1655 (Blattner et al., 1997) and its derivatives were grown as indicated in the dedicated methods sections. *E. coli* K-12 MG1655 CGSC 6300 was directly obtained from the Coli Genetic Stock Center (CGSC). For strain construction and pre-cultures, LB (Luria Bertani) medium was used. Physiology experiments were performed either in TB (Tryptone Broth; 10 g/l tryptone, 5 g/l NaCl) or M9 minimal medium (Gerosa et al., 2013) to which carbon sources were added from concentrated stock solutions. P1 phage lysate preparation and transduction were carried out as described in (Boehm et al., 2010).

Gene deletions and l-RED-mediated recombineering

Chromosomal gene deletions and modifications: Gene deletions were essentially carried out either as described (Datsenko and Wanner, 2000) or with the use of a comprehensive mutant library (Keio collection (Baba et al., 2006)) and P1 mediated transduction. Chromosomal 3xflag-tagging of genes was carried out according to the published method (Uzzau et al., 2001). For unmodified strains, AB330 (see strain list Table S1) or pKD46 was used. pKD46-mediated recombineering was used for construction of strains already harboring chromosomal modifications. Selection markers were removed by site-specific recombination using pCP20 (Datsenko and Wanner, 2000).

Construction of promoter-lacZ fusions: Construction of chromosomal promoter-lacZ fusions were carried out via λ-RED-mediated recombination as described above. AB989 (see strain list Table S1) was used as a recipient strain. The donor PCR fragment harboring the promoter of interest was designed to site-specifically excide *Prha-ccdB* and integrate upstream of the native lacZ ORF to generate a merodiploid translational fusion. Successful integration events were selected through growth on rhamnose minimal plates.

Electrophoretic mobility shift assay (EMSA)

5’ Cy3-labeled input DNA was generated either via oligonucleotide annealing or PCR. For oligonucleotides used see Table S3. 10 nM of the input DNA and purified proteins were incubated for 10 min at room temperature in buffer consisting of 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM MgCl2, 10 % Glycerol, 1 mM DTT, 0.01 % Triton X-100, 0.1 mg/mL BSA and 25 μg/mL λ-DNA. As indicated in the Figures, samples were incubated in the presence or absence of 2 mM CaCl2 and 50 μM c-di-GMP. Samples were run on 8 % polyacrylamide gel. DNA-protein complexes were analyzed using Typhoon FLA 7000 (GE Healthcare).

β-galactosidase reporter assay

Strains were grown in TB medium o/n at 37°C. The next day cultures were diluted back 1:500 into fresh medium and grown at 37°C until desired OD_{600}. 500 μL of the culture were mixed with 380 μL Z-buffer (75 mM Na_2HPO_4, 40 mM NaH_2PO_4, 1 mM KCl, 1 mM $MgSO_4$) supplemented

with 100 μL 0.1 % SDS and 20 μL chloroform. Samples were vortexed for 10 sec and left on the bench for 15 min. 200 μL sample were transferred into a clear 96-well plate. As substrate 25 μL 4 mg/mL 2-nitrophenyl- β-D-galactopyranoside (o-NPG) solution (dissolved in Z-buffer) were added. The initial velocity of the color reaction was determined at a wavelength of 420 nm.

Protein purification

$PdeL_{EAL}$ variants were purified by single StrepII-tag or His-tag affinity purifications, whereas for full-length PdeL and other transcription factors, a heparin purification step was added.

StrepII-tag purification: All proteins were cloned into a pET28a vector (Novagen) between NcoI and NotI restriction sites. Proteins were overexpressed in BL21 (AI) cells grown at 30°C in 2 L LB medium. For overexpression of mutant protein variants, the corresponding wild-type version of the gene was deleted in the overexpression strain. At an OD_{600} of 0.6 the culture was induced with 0.1 % L-arabinose. Cells were harvested 4 h post-induction by centrifugation at 6000 g for 30 min at 4°C. The cell pellet was resuspended in 7 mL Buffer A (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT) including a tablet of c0mplete mini EDTA-free protease inhibitor and a spatula tip of DNaseI. Cells were lysed by 4 passages of French press, and the lysate cleared at 4°C in a table-top centrifuge set at full speed for 40 min. The cleared supernatant was loaded on 1 mL StrepTactin Superflow Plus resin. The supernatant was reloaded another two times before washing with a total of 50 mL Buffer A. The column was washed with 10 mL Buffer B (100 mM Tris-HCl pH 8.0, 50 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT). 500 μL aliquots of proteins were eluted with Buffer B supplemented with 2.5 mM d-Desthiobiotin.

His-tag purification: His-tagged proteins were applied to a 5 ml Ni-chelating column equilibrated with 50 mM Tris pH 8, 1M NaCl, 5mM $MgCl_2$. Proteins were eluted with 50 mM Tris pH 8, 1 M NaCl, 5 mM $MgCl_2$, 25 mM TCEP and 250 mM imidazole. Eluted fractions were pooled and concentrated, and further purified by size exclusion chromatography using a Superdex 16/60 S200 column equilibrated in 20 mM Tris pH 8, 200 mM NaCl and 5 mM MgCl2. The protein was quantified by recording absorbance at 280 nm and stored at −20 °C at ~9 mg/ml concentration.

Heparin Purification: A 1 mL HiTrap Heparin HP was washed with 10 mL $H_2O_{dest.}$, followed by equilibration with 10 mL Buffer B. The eluate from the StrepII-tag affinity purification was loaded three times. After loading, the column was washed with 10 mL Buffer A followed by a washing step with 10 mL Buffer C (100 mM Tris-HCl pH 8.0, 350 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT). The protein was eluted in 500 μL fractions with a total of 10 mL Buffer D (100 mM Tris-HCl pH 8.0, 2 M NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT). The fractions containing the highest protein concentration were pooled and dialyzed o/n at 4°C against 1.5 L Dialysis Buffer (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA,

1 mM DTT). PdeLEAL variants used for cysteine crosslink assays were dialyzed against CXA Buffer (100 mM Tris-HCl pH 7.2, 250 mM NaCl, 2 mM EDTA). The final protein concentration was recorded at 280 nm, and the content of co-purified nucleotide contaminants determined as a ratio of 260/280 nm.

Microscale thermophoresis

Experiments were carried out with a Monolith NT.115 device using Premium capillaries (both from NanoTemper Technologies). Experiments were performed in PBS supplemented with 5mM DTT, and 0.05% Tween 20 in Premium treated capillaries with 40% LED power and 80% IR-Laser power at 22°C. Laser on and off times was set to 30 and 5 seconds, respectively. In all assays, PdeL-His was labeled using the Monolith NT His-Tag Labeling Kit RED-tris-NTA. Labeled PdeL-His was kept at a constant final concentration of 50 nM. For binding assays, 16 two-fold serial dilutions of an unlabeled partner were used.

For Protein-DNA interaction assays, 5'-3' and 3'-5' oligos corresponding to CIB and CDB DNA fragments were annealed slowly from 95°C to 20°C with varying DNA concentrations. Starting concentrations were 0.5 μM (CIB) and 6 μM (CDB) in the absence of Cra and 0.2 μM in the presence of 50 nM Cra. Probes were incubated for 1h before loading into the capillaries. Experiments were carried out in triplicates with three independent preparations.

To determine the K_d of PdeL tetramerization, unlabeled PdeL-Strep was titrated starting from 0.21 μM. Measurements started 16 min after mixing labeled and unlabeled PdeL and repeated every 18 min for 3 h 30 min. In a parallel experiment, PdeL-Strep was titrated starting from 0.4 μM and incubated for 4 h. Finally, an MST experiment was performed with PdeL-Strep initially being titrated as indicated above before the concentration of PdeL-Strep was doubled in all samples of the serial dilutions, while PdeL-His was kept at 50 nM.

K_d values were calculated using MO Affinity Analysis software (NanoTemper). Assimilated curves were analyzed with GraphPad using non-linear regression and one site-specific binding.

SEC-MALS analysis

PdeL samples (100 μl) varying from 11 to 371 μM were loaded onto a Superdex 200 (10/300) column (GE Healthcare) at constant flow (0.5 ml/min) in 20 mM Tris pH 8.0, 200 mM NaCl, 5 mM $MgCl_2$, 5 mM DTT. The SEC instrument was coupled to an in-line multi-angle light-scattering and differential refractive index detectors (Wyatt Heleos 8+ and Optilab rEX) to measure the apparent mass values (MW_{app}) during elution. The inter-detector delay volumes and band broadening, the light-scattering detector normalization, and the instrumental calibration coefficient were calibrated using a standard 2 mg/ml of BSA solution (Thermo Pierce) run in the same buffer, on the same day, according to standard Wyatt protocols. SEC-MALS experiments with PdeL in the presence of the CIB or CDB DNA fragments or Cra protein were performed with the same column and device. In the loaded samples, PdeL concentration was at 94 μM, Cra 94 μM, the DNA fragments CIB and CDB at 120 μM. SEC experiments were performed with the same column using 120 μM PdeL and 240 μM c-di-GMP in the presence of 5 mM $CaCl_2$.

Crystallization

Protein purification: Wild-type and mutant *pdeL* alleles were cloned in pET28a (Novagen) with C terminal His- or Strep-tags. Proteins were expressed in *E. coli* BL21 (DE3) (Novagen) grown in LB medium containing kanamycin (50 μg/ml) with IPTG (500 μM) added at an OD_{600} of 0.6. Cultures were further incubated overnight at 16°C before cells were harvested by centrifugation and disrupted using a French pressure cell. The lysate was clarified by centrifugation, and His-tagged proteins were loaded on a Ni column equilibrated with buffer A (50 mM Tris pH 8, 1M NaCl, 5mM $MgCl_2$) and eluted with buffer B (50 mM Tris pH 8, 1M NaCl, 5mM $MgCl_2$, 25mM TCEP, and 250 mM imidazole). StrepII-tagged PdeL was purified via StrepTactin Superflow Plus resin pre-equilibrated with buffer A and eluted with Buffer A supplemented with 2.5 mM d-Desthiobiotin. Eluted fractions were pooled and concentrated and further purified by size exclusion chromatography using a Superdex 16/60 S200 column equilibrated in 20mM Tris pH 8, 200 mM NaCl, 5mM MgCl2, 25mM TCEP.

Crystallization: PdeL protein and CIB DNA fragments were mixed with a 2:1 stoichiometry to a final concentration of 100 μM. Crystals were obtained using the sitting-drop vapor-diffusion method after optimizing F7 condition in clear strategy screen (molecular dimension). 20% glycerol as cryoprotectant and crystals were flash-frozen in liquid nitrogen.

X-ray diffraction data collection, molecular replacement, and refinement

X-ray diffraction datasets were collected at the Swiss Light Source (SLS) synchrotron on the PXI beamline. Diffraction datasets were processed using XDS (Kabsch, 2010) and the resulting intensities were scaled using SCALA from the CCP4/CCP4i2 suite (Potterton et al., 2018) (Collaborative Computational Project, Number 4, 1994). The structure was solved with phaser (McCoy et al., 2007) using the coordinate of PdeL-EAL (PDB: 4LYK) monomer as the search model. Refinements were carried out using REFMAC5. Model building was performed with Coot (Emsley and Cowtan, 2004) and the quality of the model was assessed using MolProbity (Williams et al., 2018). Structural figures were prepared using PyMOL (http://www.pymol.org).

Mass photometry

The Refeyn OneMP mass photometer was used to determine PdeL oligomerization state at low protein concentrations. 18 μl buffer (20 mM Tris pH 8.0, 200 mM NaCl, 5 mM $MgCl_2$) were pre-loaded into a silicone well and mixed with 2 μl of protein (final concentration of 25 nM) prior to data acquisition. 6000 frames were collected using default instrument parameters. DiscoverMP software provided by Refeyn was used for data analysis using default parameters for event extraction and fitting.

Immunoblotting

Cells were grown in TB medium at 37°C until desired OD_{600}. An equivalent of 1 mL of an OD_{600} of 1.0 was pelleted and resuspended in 100 μL SDS Laemmli buffer. Cells were lysed by boiling the sample at 98°C for 10 min. 8 μL of total cell extracts were loaded onto a 12.5 %

SDS-polyacrylamide gel, and proteins transferred using a wet blot system. Proteins with 3xFlag-tag were detected with a 1:10.000 dilution of monoclonal mouse α-Flag monoclonal antibody and a 1:10.000 dilution of polyclonal rabbit α-mouse horseradish-peroxidase (HRP) secondary antibody. Proteins were visualized with enhanced chemiluminescence (ECL) detection reagent and imaged in a gel imager (GE ImageQuant LAS 4000).

Phosphodiesterase enzyme assay

Phosphate Sensor assay: Phosphate sensor assay was essentially performed as described in (Reinders et al., 2016). Briefly: Conversion of c-di-GMP into pGpG was measured indirectly by a coupled alkaline phosphatase (AP)/phosphate sensor online assay. The terminal phosphate of the pGpG product is cleaved by the coupling enzyme AP (20 U/μl, Roche), and the phosphate concentration is determined from the fluorescence increase through binding of phosphate to the phosphate sensor (0.5 μM; Thermo Fisher). PdeL and c-di-GMP concentrations were used, as shown in the individual experiments. Fluorescence increase was detected by excitation at 430 nm and emission at 468 nm.

FPLC assay: Assay was performed as described in (Sundriyal et al., 2014). Enzymatic activity was assayed offline by FPLC-based steady-state nucleotide quantification following incubation for varying durations. Enzymatic reactions were carried out at 20°C in 100 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EDTA, 1 mM DTT and 50 M thiamine pyrophosphate as FPLC standard. PdeL and c-di-GMP concentrations were used as described in the result section. The reaction was started by addition of enzyme to a total reaction volume of 600 μl. Samples volumes of 100 μl were withdrawn and the reaction was stopped at different time points by addition of 10 μl of 100 mM $CaCl_2$ and subsequent heating at 98°C for 10 min.

The samples were then analyzed using ion-exchange chromatography (1-mL Resource-Q column) after addition of 890 μL 5 mM ammonium bicarbonate (NH_4CO_3) to increase the volume to 1 mL. 500 μL of this was then loaded onto the column. The column was washed thoroughly and the bound nucleotides were eluted with a linear NH_4CO_3 gradient (5 mM to 1 M) over 17 column volumes. The amount of pGpG product was determined by integration of the corresponding absorption (253 nm) peak after normalization of the data with respect to the internal thiamine pyrophosphate standard.

Cysteine crosslink assay

10 μM $PdeL_{EAL}$-3xFlag-StrepII variants purified in CXA Buffer (100 mM Tris-HCl pH 7.2, 250 mM NaCl, 2 mM EDTA) were incubated for 10 min at room temperature in the presence of 10 mM $CaCl_2$, either in presence or absence of 50 μM c-di-GMP. Proteins were crosslinked for 1 h at room temperature with an 8-fold molar excess (80 μM) of bismaleimidoethane (BMOE). Crosslink reaction was quenched for 15 min at room temperature by addition of 50 mM DTT. Samples were supplemented with SDS Laemmli buffer and proteins denatured by heating at 98°C for 5 min. Samples were loaded on a 1.5 mm thick 12.5 % SDS-polyacrylamide gel and detected by staining with Coomassie Brilliant Blue G according to the staining protocol from Sigma (product

number B2025). In brief: Gel was fixed for 30 min in Fixing Solution and stained over night at room-temperature in staining solution. Gel was destained with 10 % acetic acid (v/v), 25 % (v/v) methanol for 60 sec with shaking. Gel was rinsed with 25 % methanol and destained for 24 h with fresh 25 % methanol.

Microscopy

Cells were placed on a PBS pad solidified with 1% agarose. Fluorescence and differential interference contrast (DIC) microscopy were performed on a DeltaVision Core (Applied Prescision, USA) microscope equipped with an Olympus 100X/1.30 Oil objective and an EDGE/sCMOS CCD camera. Exposure time for microscopy picture was 0.05 sec for bright field (POL), 0.1 sec for GFP for the c-di-GMP sensor, 0.2 sec for *pdeL-gfp* and 0.4 or 0.5 sec for *pdeL-mCherry*. For all settings, the ND filter was set to 100% transmission. For strains grown on M9 glycerol, microscopy was performed using an inverted microscope (Eclipse Ti2, Nikon Instruments Europe B.V.) with 100x oil immersion objective (CFI Plan Apo λDM100x Oil, Nikon). Exposure time was set to 0.05 sec for phase contrast, 0.2 sec for GFP, and 0.2 sec for mCherry. Images analysis was performed with the open source software-package Oufti (Paintdakhi et al., 2015) to determine cell outlines. Outlines were used to compute mean single-cell fluorescence using WHISIT (Sprecher et al., 2017).

C-di-GMP measurements

C-di-GMP measurements were performed according to the published procedure (Spangler et al., 2010). In brief: *E. coli* cells were grown in 24 mL TB medium at 37°C to an OD_{600} of 0.5. Cells were pelleted and washed in 300 μL ice-cold distilled water. After washing, the cell pellet was resuspended in 300 μL ice-cold extraction solvent (acetonitrile/methanol/$H_2O_{dest.}$, 40/20/20 v/v/v). After pelleting, the supernatant was transferred to a 2 mL safe-lock tube, and the extraction procedure repeated twice with 200 μL extraction solvent. Biological triplicates were performed and analyzed by HPLC-MS/MS. Measured values were mathematically converted into cellular c-di-GMP concentration. Constants of *E. coli* cell volume and cfu/mL needed for calculation were experimentally determined. A standard curve correlating *pdeH* induction and c-di-GMP levels was used to interpolate c-di-GMP values below the detection limit of around 10 nM c-di-GMP. The threshold was set to a minimum of 1 nM c-di-GMP per cell. To establish precise intracellular c-di-GMP levels, an *E. coli* Δ*pdeH* mutant was engineered with a plasmid-encoded IPTG inducible copy of *pdeH* (*Plac*::*pdeH*, pAR81). Engineered strains also harboring different reporter constructs were grown overnight in the presence of 1 mM IPTG or without IPTG. After diluting precultures into fresh medium containing variable concentrations of IPTG, cells were grown to an OD_{600} of 0.5 and harvested for c-di-GMP measurements.

Absolute protein concentration determination via selected reaction-monitoring (SRM) LC-MS analysis

600 μL of *E. coli* cultures grown in TB to an OD_{600} of ca. 0.5 were pelleted and washed twice with 1 mL ice-cold PBS. Cells were lysed in 1 % sodium deoxycholate (SDC), 100 mM Tris pH 8.5 by sonication. Proteins were denatured by heating at 95°C for 10 min. Protein alkylation was performed with chloracetamide. Protein digestion was performed by subsequent treatment with Lys-C (enzyme/protein ratio 1:200) and trypsin (enzyme/protein ratio 1:50). Peptides were acidified with TFA and desalted using PreOmics (ThermoFisher) cartridges.

An aliquot of heavy reference peptide mix was spiked into each sample at a concentration of 20 fmol of heavy reference peptides per 1 μg of total endogenous protein mass. The heavy peptide mix contained 10 chemically synthesized proteotypic peptides (JPT Peptide Technologies GmbH) of the two target proteins (5 peptides each) that showed the highest MS1 responses in a previous large-scale study (Schmidt et al., 2016). In a first step, selected reaction-monitoring (SRM) assays (Peterson et al., 2012) were generated from a mixture containing 500 fmol of each reference peptide and shotgun LC-MS/MS analysis on a Q-Exactive HF platform. The setup of the µRPLC-MS system was as described previously (Ahrné et al., 2016). Chromatographic separation of peptides was carried out using an EASY nano-LC 1000 system (Thermo Fisher Scientific), equipped with a heated RP-HPLC column (75 µm x 37 cm) packed in-house with 1.9 µm C18 resin (Reprosil-AQ Pur, Dr. Maisch). Peptides were analyzed per LC-MS/MS run using a linear gradient ranging from 95% solvent A (0.1% formic acid) and 5% solvent B (99.9% acetonitrile, 0.1% formic acid) to 45% solvent B over 60 minutes at a flow rate of 200 nL/min. Mass spectrometry analysis was performed on Q-Exactive HF mass spectrometer equipped with a nanoelectrospray ion source (both Thermo Fisher Scientific). Each MS1 scan was followed by high-collision-dissociation (HCD) of the 10 most abundant precursor ions with dynamic exclusion for 20 seconds. Total cycle time was approximately 1 second. For MS1, 3e6 ions were accumulated in the Orbitrap cell over a maximum time of 100 ms and scanned at a resolution of 120,000 FWHM (at 200 m/z). MS2 scans were acquired at a target setting of 1e5 ions, accumulation time of 50 ms, and a resolution of 30,000 FWHM (at 200 m/z). Singly charged ions and ions with unassigned charge state were excluded from triggering MS2 events. The normalized collision energy was set to 27%, the mass isolation window was set to 1.4 m/z and one microscan was acquired for each spectrum.

The acquired raw files were searched against a decoy database using the MaxQuant software (Version 1.0.13.13) containing standard and reverse sequences of the predicted SwissProt entries of *E. coli* (www.ebi.ac.uk, release date 2016/05/02), retention time standard peptides, and commonly observed contaminants (in total 10402 sequences) generated using the SequenceReverser tool from the MaxQuant software (Version 1.0.13.13). The precursor ion tolerance was set to 10 ppm, and fragment ion tolerance was set to 0.02 Da. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues unless followed by proline), 3 missed cleavages were allowed, carbamidomethylation (C) was set as fixed modification and arginine (+10 Da), lysine (+8 Da) and oxidation (M) were set as a variable modification. The resulting msms.txt file was converted to a spectral library panel with the 5 to 10 best

transitions for each peptide using an in-house software tool. This was then imported into the SpectroDive program (Version 7.5, Biognosys, Schlieren, Switzerland), and a transition list for quantitative SRM analysis was generated. Here, all samples were analyzed on a TSQ-Vantage triple-quadrupole mass spectrometer coupled to an Easy-nLC (Thermo Fisher, Scientific). In each injection, an equivalent of 1.5 µg of peptides including heavy peptide references was loaded onto a custom-made main column (Reprosil C18 AQ, 3 µm diameter, 100 Å pore, 0.75 × 300 mm) and separated using the same gradient mentioned above. The mass spectrometer was operated in the positive ion mode using ESI with a capillary temperature of 275 °C, a spray voltage of +2200 V. All of the measurements were performed in an unscheduled mode and a cycle time of 2 sec. A 0.7 FWHM resolution window for both Q1 and Q3 was set for parent- and product-ion isolation. Fragmentation of parent-ions was performed in Q2 at 1.2 mTorr, using collision energies calculated with the SpectroDive software (version 7.5). Each condition was analyzed in biological quadruplicates. All raw files were imported into SpectroDive for absolute peptide and protein quantification.

Biofilm assays

Attachment assays were carried out as described previously (Boehm et al., 2009). Briefly: 200 μL TB medium provided in a clear 96-well microtiter plate were inoculated 1:40 with an o/n culture grown at 37°C. The plate was incubated statically at 37°C for 24 h unless indicated differently. After recording the OD_{600} of the total biomass, the planktonic phase of the culture was discarded and the wells washed with $H_2O_{dest.}$ from a hose. The remaining attached biomass was stained with 200 μL 0.3 % crystal violet (0.3 % (w/v) in 5 % (v/v) 2-propanol, 5 % (v/v) methanol) for 20 min. The plate was washed with $H_2O_{dest.}$ from a hose and the stained biofilm dissolved in 20 % acetic acid for 20 min. The intensity of crystal violet stain was quantified at 600 nm and normalized to the initially measured total biomass.

For biofilm escape assays, cells harboring pAR81 were allowed to attach to plastic surfaces in 96-well microtiter plates for 7 hrs at 30°C in TB medium supplemented with 30 μg/ml ampicillin. Plates were gently washed with deionized water, dried, and incubated for 3 h at 30°C after adding fresh TB media was supplemented with IPTG to induce plasmid-borne *Plac-pdeH* and. After incubation, 10 μL of the planktonic phase in the center of the well were isolated to determine cfu/mL by spotting serial dilutions in LA plates supplemented with ampicillin. A pipetting robot (Tecan freedom evo) that allows for sensing of the liquid surface was used to improve reproducibility.

Motility assay

A single colony was picked onto a TB swarmer plate (0.3 % agar). Plates were incubated at 37°C for 3-4 h. Swarm halos were recorded with a NIKON Coolpix990 and swarm radius quantified via ImageJ (NIH, USA).

Bacteriophage N4 propagation and infection assay

Phage N4 was propagated on *E. coli* MG1655 (CGSC 6300) and stored at 4°C in SM(G) buffer containing 0.1 M NaCl, 10 mM $MgSO_4$, 0.05 M Tris-HCl pH 7.5 and 0.01% gelatin. Phage titer was determined by spotting 2.5 µl of a 10-fold serial dilution on a lawn of *E. coli* MG1655 via top agar overlay method as described before (Kropinski et al., 2009) with slight modifications. In brief, 100 µl of *E. coli* overnight culture was mixed with 3 ml top agar (0.4% LB-agar supplemented with 20 mM $MgSO_4$ and 5 mM $CaCl_2$) and pipetted on an LB agar plate pre-warmed to 60°C. 2.5 µl of a serial dilution of phage N4 was spotted after top-agar solidified.

Flow cytometry

For flow cytometry of TB cultures, TB medium supplemented with 100 µg/ml ampicillin, 30 µg/ml kanamycin, and 0 or 65 µM IPTG was inoculated from single colonies. For strains carrying the c-di-GMP sensor, the medium was additionally supplemented with 200 nM anhydrotetracycline (aTc). TB cultures were grown overnight at 37°C and shaking. The next day cultures were diluted 1:500 in fresh TB medium supplemented with 40 µg/ml ampicillin, 20 µg/ml kanamycin, various IPTG concentrations, and where necessary, 200 nM aTc and grown at 37°C for 4.5 h. At this point, cell density was between OD_{600} 0.3-0.5 for all cultures. Cultures were diluted 1:200 in fresh TB medium with identical supplements and grown at 37°C for another 1.5 to 3.5 h up to a total incubation time of 6 to 8 h. Samples of max 800 µl were taken, kept on ice and diluted into 1x PBS just before analysis.

Flow cytometry of M9 minimal medium cultures started with day cultures in LB supplemented with 100 µg/ml ampicillin, which were inoculated from single colonies and grown for 6-7 h at 37°C. Cells were washed once in M9 medium without carbon source and diluted to an OD_{600} of 0.01 or lower in 5 ml M9 medium supplemented with 0.4% glycerol, 0.4% glycerol and 0.05% casamino acids, 0.5% glucose, 0.2% fumarate or 0.2% a-ketoglutarate. M9 cultures were incubated at 37°C with shaking at 180 rpm (Multitron, INFORS HT), for at least 17 h before measurements. At this point, cell density was between OD_{600} 0.1-0.4 for all cultures. To ensure exponentially growing cultures at the moment of sampling, M9 overnight cultures were diluted 1:500 in fresh M9 minimal medium with identical carbon source once in between where necessary. Samples were kept on ice and diluted into 1x PBS just before analysis.

Cells were analyzed on a BD LSRFortessa flow cytometer at medium flow rate and a maximum event rate of 10’000 events/s. Per sample, 100’000 events were recorded. Parameters measured were forward scatter (FSC-H), and side scatter (SSC-H and SSC-W). GFP was excited at 488/20 and detected with a 512/20 emission filter. Where applicable, mScarlet-I was excited using the pre-set excitation/emission ‘mCherry’. Data was collected using the Diva (BD Biosciences) software. FlowJoTM Software (version 10.6.1 for Windows) was used for the import and gating of raw data. The forward-scatter (FSC-H) and side-scatter (SSC-H and SSC-W) were used to separate cells from background particles. For analysis of cultures carrying the c-di-GMP sensor,

a third gate was applied in which we used the ‘mCherry’ channel to gate for cells expressing mScarlet-I. Only mScarlet-I-positive cells were included in the analysis of the GFP signal coming from the c-di-GMP sensor. Gated populations were exported as ‘Scale Values’ in csv-files, and GFP distributions were visualized using MATLAB version R2019b (MathWorks) scripts.

2.6 Supplementary material

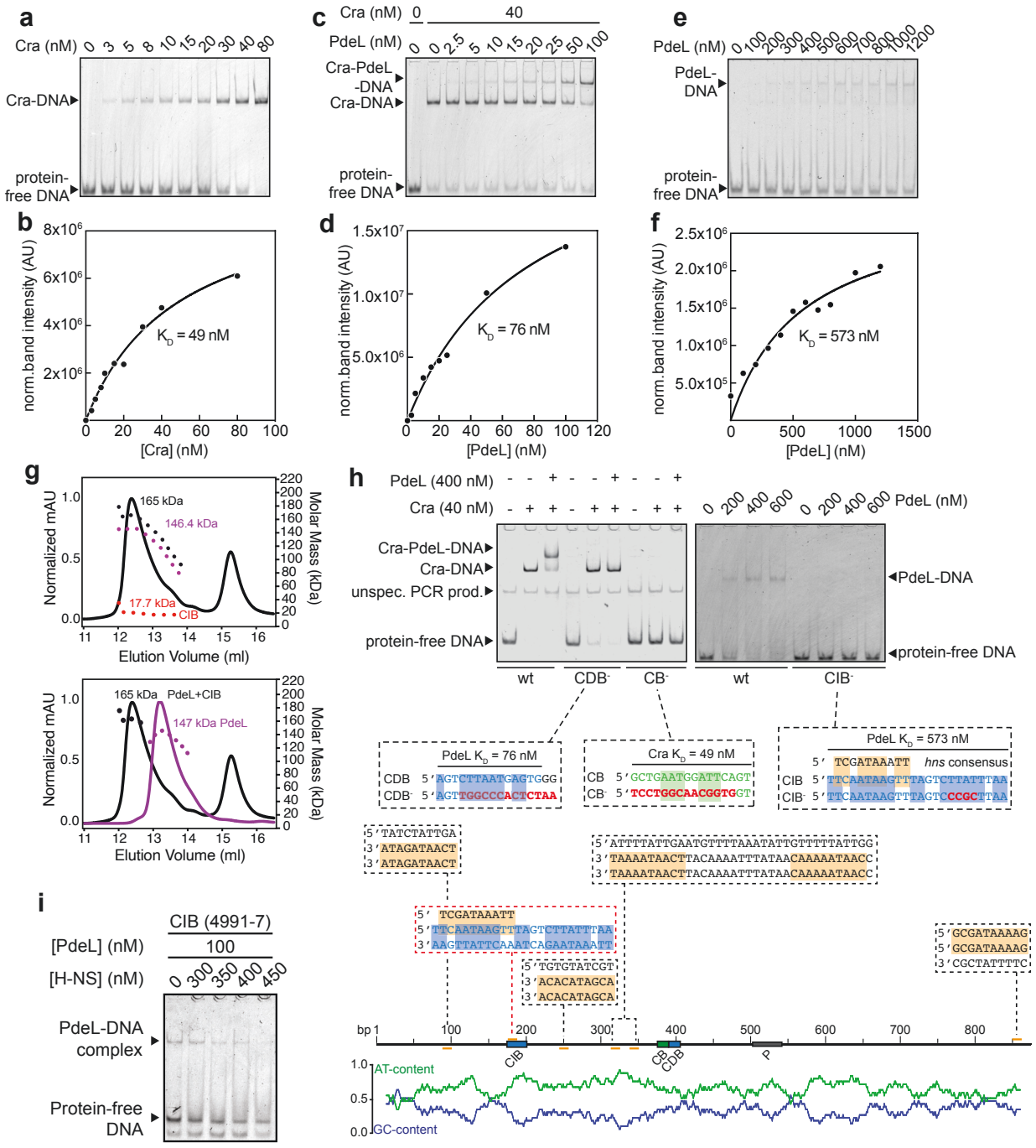


Figure S1: Binding specificity and affinities of Cra and PdeL to *pdeL* promoter region. **(a)** Binding of purified Cra-StrepII to the *pdeL* intergenic region as tested by electrophoretic mobility shift assay (EMSA). Binding was assayed using 10 nM of Cy3-labeled oligonucleotide spanning the Cra-box (CB) and 10 nM of Cy3-labeled oligonucleotide spanning the Cra-dependent PdeL-box (CDB) (see: Fig. 1a). Cra concentrations used are indicated. **(b)** Saturation binding fit of band intensities from **(a)**. **(c)** EMSA and **(d)** binding affinity of purified PdeL-StrepII in the presence of 40 nM Cra-StrepII using the same oligos as in **(a)**. The PdeL binding constant was calculated from band intensities of the super-shift (Cra-PdeL-DNA-complex). **(e)** Binding of purified PdeL-StrepII to the Cra-independent PdeL-box (CIB). The labeled oligonucleotide included the CIB region and 10 bp up- and 33 bp downstream of CIB. **(f)** Saturation binding fit of band intensities from **(e)**. **(g)** SEC-MALS of PdeL DNA containing CIB (see: Fig. 1a) analyzed individually (bottom) and after mixing (top). The molecular masses of individual components and the complex are indicated. **(h)** Binding specificity of Cra and PdeL to the *pdeL* promoter region using oligos with mutated binding sites. Left panel: binding of Cra (40 nM) and PdeL (400 nM) to wild type (CB, CDB) and scrambled (CDB^-, CB^-) binding sites. Right panel: binding of PdeL (200 - 600 nM) to wild type and scrambled (CIB^-) CIB binding site. Sequences below the graphs indicate Cra and PdeL binding sites (upper line) with scrambling mutations highlighted in red in the line below. Note that mutations abolishing binding of PdeL to CIB were chosen outside of the putative H-NS consensus sequence. The graph on the bottom shows the putative H-NS binding boxes within the *pdeL* promoter region. PdeL and Cra binding sites are indicated in blue and green, respectively. H-NS binding sites (orange) were identified using the Virtual Footprint website http://www.prodoric.de/vfp/ (Münch et al., 2005). H-NS binding to CIB (red box) was experimentally verified in **(i)**. AT- and GC-content of *pdeL* intergenic region with a binning of 5 bp is shown at the bottom. **(i)** PdeL and H-NS compete for CIB binding. EMSA assay with labeled DNA covering CIB (see Fig. 1a) and concentrations of purified proteins as indicated.

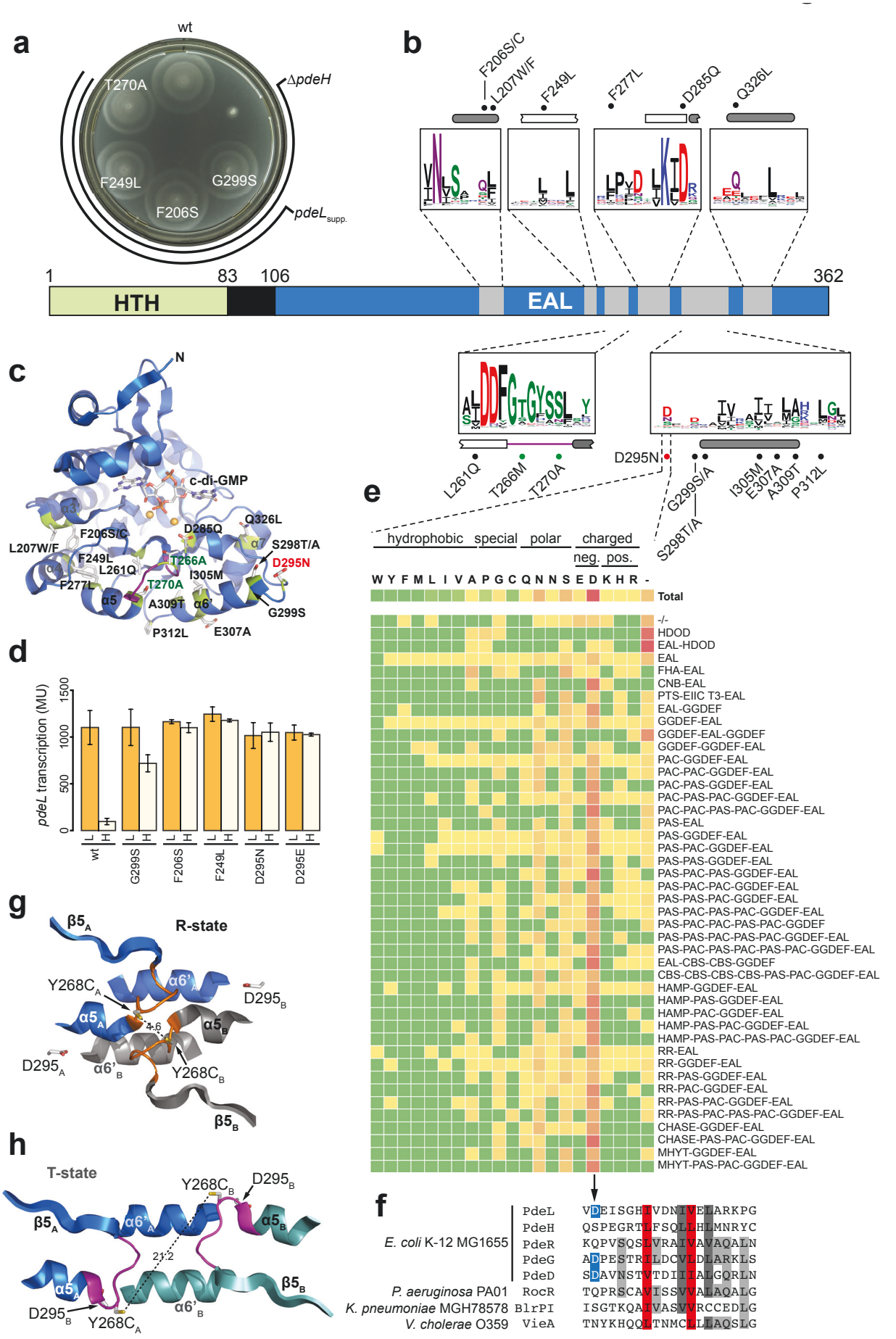

a
wt
ΔpdeH
T270A
F249L
G299S
F206S
pdeL supp.
b
F206S/C
L207W/F
F249L
F277L
D285Q
Q326L
1
83
106
362
HTH
EAL
c
N
c-di-GMP
α3'
F206S/C
D285Q
Q326L
L207W/F
F249L
T266A
α7
S298T/A
α4
L261Q
D295N
F277L
T270A
I305M
α5
A309T
α6'
G299S
P312L
E307A
d
pdeL transcription (MU)
0
500
1000
1500
L H
wt
G299S
F206S
F249L
D295N
D295E
e
L261Q
T266M
T270A
D295N
G299S/A
S298T/A
I305M
E307A
A309T
P312L
hydrophobic
special
polar
charged
neg.
pos.
W Y F M L I V A P G C Q N N S E D K H R -
Total
-/-
HDOD
EAL-HDOD
EAL
FHA-EAL
CNB-EAL
PTS-EIIC T3-EAL
EAL-GGDEF
GGDEF-EAL
GGDEF-EAL-GGDEF
GGDEF-GGDEF-EAL
PAC-GGDEF-EAL
PAC-PAC-GGDEF-EAL
PAC-PAS-GGDEF-EAL
PAC-PAS-PAC-GGDEF-EAL
PAC-PAC-PAS-PAC-GGDEF-EAL
PAS-EAL
PAS-GGDEF-EAL
PAS-PAC-GGDEF-EAL
PAS-PAS-GGDEF-EAL
PAS-PAC-PAS-GGDEF-EAL
PAS-PAC-PAC-GGDEF-EAL
PAS-PAS-PAC-GGDEF-EAL
PAS-PAC-PAS-PAC-GGDEF-EAL
PAS-PAC-PAC-PAS-PAC-GGDEF
PAS-PAS-PAC-PAS-PAC-GGDEF-EAL
PAS-PAC-PAS-PAC-PAS-PAC-GGDEF-EAL
EAL-CBS-CBS-GGDEF
CBS-CBS-CBS-CBS-PAS-PAC-GGDEF-EAL
HAMP-GGDEF-EAL
HAMP-PAS-GGDEF-EAL
HAMP-PAC-GGDEF-EAL
HAMP-PAS-PAC-GGDEF-EAL
HAMP-PAS-PAC-PAS-PAC-GGDEF-EAL
RR-EAL
RR-GGDEF-EAL
RR-PAS-GGDEF-EAL
RR-PAC-GGDEF-EAL
RR-PAS-PAC-GGDEF-EAL
RR-PAS-PAC-PAS-PAC-GGDEF-EAL
CHASE-GGDEF-EAL
CHASE-PAS-PAC-GGDEF-EAL
MHYT-GGDEF-EAL
MHYT-PAS-PAC-GGDEF-EAL
f
E. coli K-12 MG1655
P. aeruginosa PA01
K. pneumoniae MGH78578
V. cholerae O359
PdeL VDEISGHIVDNIVELARKPG
PdeH QSPEGRTLFSQLLHLMNRYC
PdeR KQPVSQSLVRAIVAVAQALN
PdeG ADPESTRILDCVLDLARKLS
PdeD SDAVNSTVTDIIIALGQRLN
RocR TQPRSCAVISSVVALAQALG
BlrPI ISGTKQAIVASVVRCCEDLG
VieA TNYKHQQLTNMCLLLAQSLG
g
R-state
β5A
Y268CA
α6'
D295B
α5A
4.6
α5B
D295A
α6'B
Y268CB
β5B
h
T-state
Y268CB
D295B
β5A
α6'A
α5B
21.2
α5A
β5B
D295B
α6'B
Y268CA

Figure S2: Location and properties of activating *pdeL* alleles. **(a)** Motility plate with *pdeL* suppressors restoring motility in a Δ*pdeH* mutant. **(b)** Domain architecture of PdeL with the DNA binding domain (HTH) in green and the catalytic EAL domain in blue. Suppressor mutations in the highly conserved loop 6 (green dots), in the R-state stabilizer Asp295 (red dot), or elsewhere in the *pdeL* coding sequence (black dots) are indicated. Alpha-helices (rounded grey bars), β-sheets (blank rectangles) and unstructured regions (line) are marked. Conservation of regions containing motile suppressors is shown as WebLogos of an alignment of 500 non-redundant EAL-domain proteins. **(c)** Crystal structure of EAL domain in its T-state conformation with Ca^{2+} (orange), c-di-GMP, loop 6 (purple) and positions of suppressor mutations (sticks and light green). **(d)** Activity of a *pdeL-lacZ* transcriptional reporter introduced into a selection of *pdeL* suppressor strains from **(b)** and **(c)**. Low **(L)** and high **(H)** levels of c-di-GMP were established as indicated in Fig. 1f. **(e)** Conservation scores of the aspartic acid residue corresponding to D295 of PdeL in 500 non-redundant EAL-domain proteins with different domain architectures (see Material and Methods). Green and red colors indicate low and high occurrence, respectively. Amino acids are classified according to their chemical properties. EAL = phosphodiesterase; GGDEF = diguanylate cyclase. **(f)** Alignment of the PdeL region containing D295 with a selection of PDEs. All PDEs of *E. coli* K-12 with a conserved Asp (blue) at this position are listed. **(g)** Close-up of loop 6 (orange) and dimerization helices in the R-state conformation (Fig. 2b). Protomers are colored in marine and grey. The distance of Y268C substitutions between two protomers is shown in Å and indicated by a stippled black line. **(h)** Loop 6 (magenta) and dimerization helices in the T-state configuration (Fig. 2c). Protomers are shown in marine and light-teal. Note the almost 5-fold longer distance between Y268C residues as compared to **(g)**.

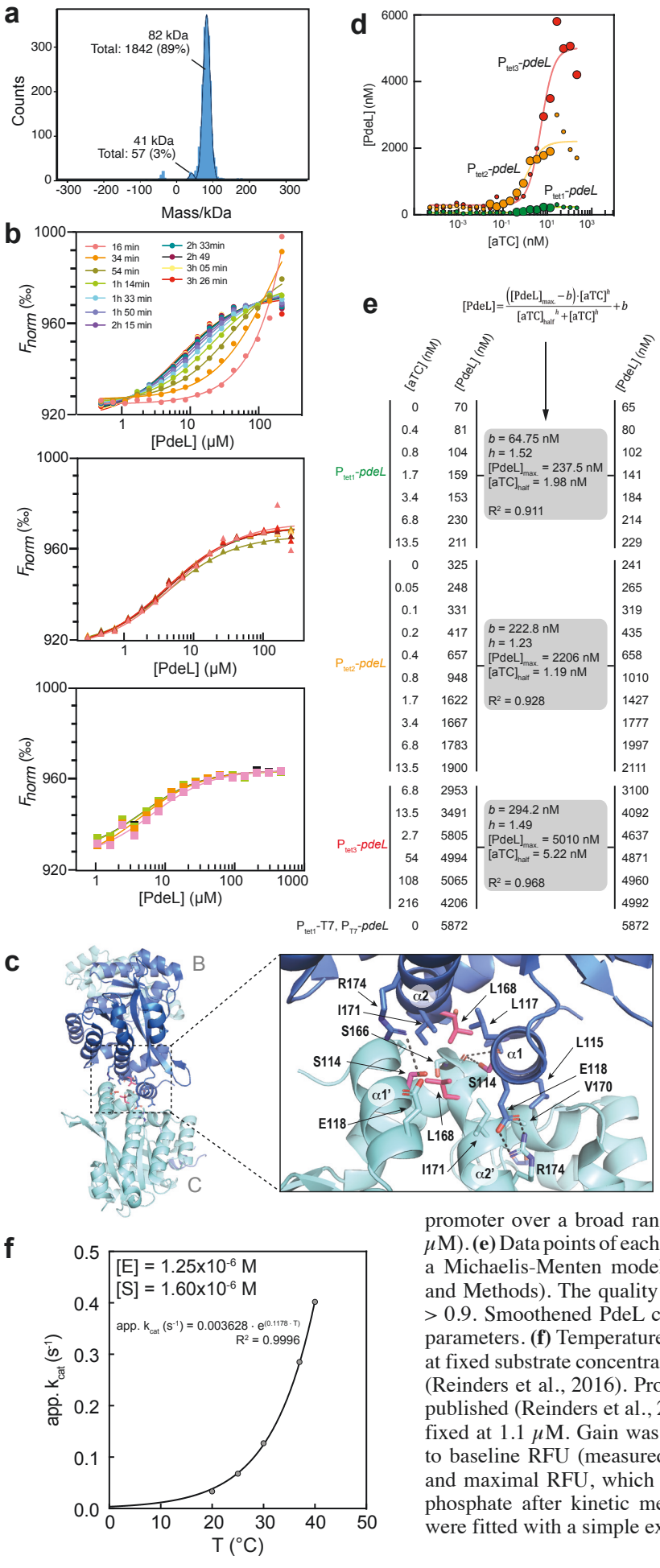


Figure S3: PdeL oligomerization and tuning of cellular levels of PdeL. **(a)** PdeL is a dimer at a low protein concentration. Refeyn analysis with mass histograms shown for 1899 steps of 25 nM PdeL. A Gaussian model is shown in black and the distribution of monomer (41 kDa) and dimers (82 kDa) are indicated. **(b)** Redistribution analysis of PdeL by MST. PdeL dimer/tetramer exchange determined thermophoretic mobility upon titration of purified Strep-tagged PdeL to a constant pool (50 nM) of fluorescently labeled His-tagged PdeL (**top**). Measurements were carried out in roughly 18 min intervals as indicated until reaching equilibrium after 3.5 h. **Middle panel:** Labeled and unlabeled PdeL were mixed and equilibrated for 4 hrs. Thermophoretic mobility was determined after 3-fold dilution of the mixture. **Bottom panel:** Labeled and unlabeled PdeL were equilibrated as in **(b)** and thermophoretic mobility was determined after addition of a concentrated solution of unlabeled PdeL. Tetramerization Kds were determined to be between 3.2 and 6.4 μM. **(c)** Dimer-dimer interface of R-state tetramer with two protomers B and C visible. Inset shows a zoom of the tetramerization interface with residues of helices α1 and α2 highlighted, which are involved in dimer-dimer interaction. **(d)** PdeL protein concentration as measured by SRM. PdeL expression was tuned with three different Ptet-*pdeL* constructs harboring weak (green), intermediate (orange), and strong RBS (red) sequences providing different translational activities. For very high *pdeL* expression, a plasmid-based construct was used, in which P_{tet1}-driven T7 polymerase drives the expression of *pdeL* under control of the T7 promoter. The large dots depict conditions chosen to analyze the activity of the *pdeL* promoter over a broad range of PdeL concentrations (0.65 – 5.9 μM). **(e)** Data points of each P_{tet} construct used in **(b)** were fitted with a Michaelis-Menten model with Hill-coefficients (see Materials and Methods). The quality of individual fits is indicated with R2 > 0.9. Smoothened PdeL concentrations were calculated from fit parameters. **(f)** Temperature-dependent activity of PdeL (1.25 μM) at fixed substrate concentration (1.6 μM) determined as outlined in (Reinders et al., 2016). Progress curves were fitted with script as published (Reinders et al., 2015). For all conditions the $K_{D\,[ES]}$ was fixed at 1.1 μM. Gain was adjusted for each condition according to baseline RFU (measured before start of kinetic measurement) and maximal RFU, which was determined by addition of excess phosphate after kinetic measurement. Resulting app. k_{cat} values were fitted with a simple exponential curve.

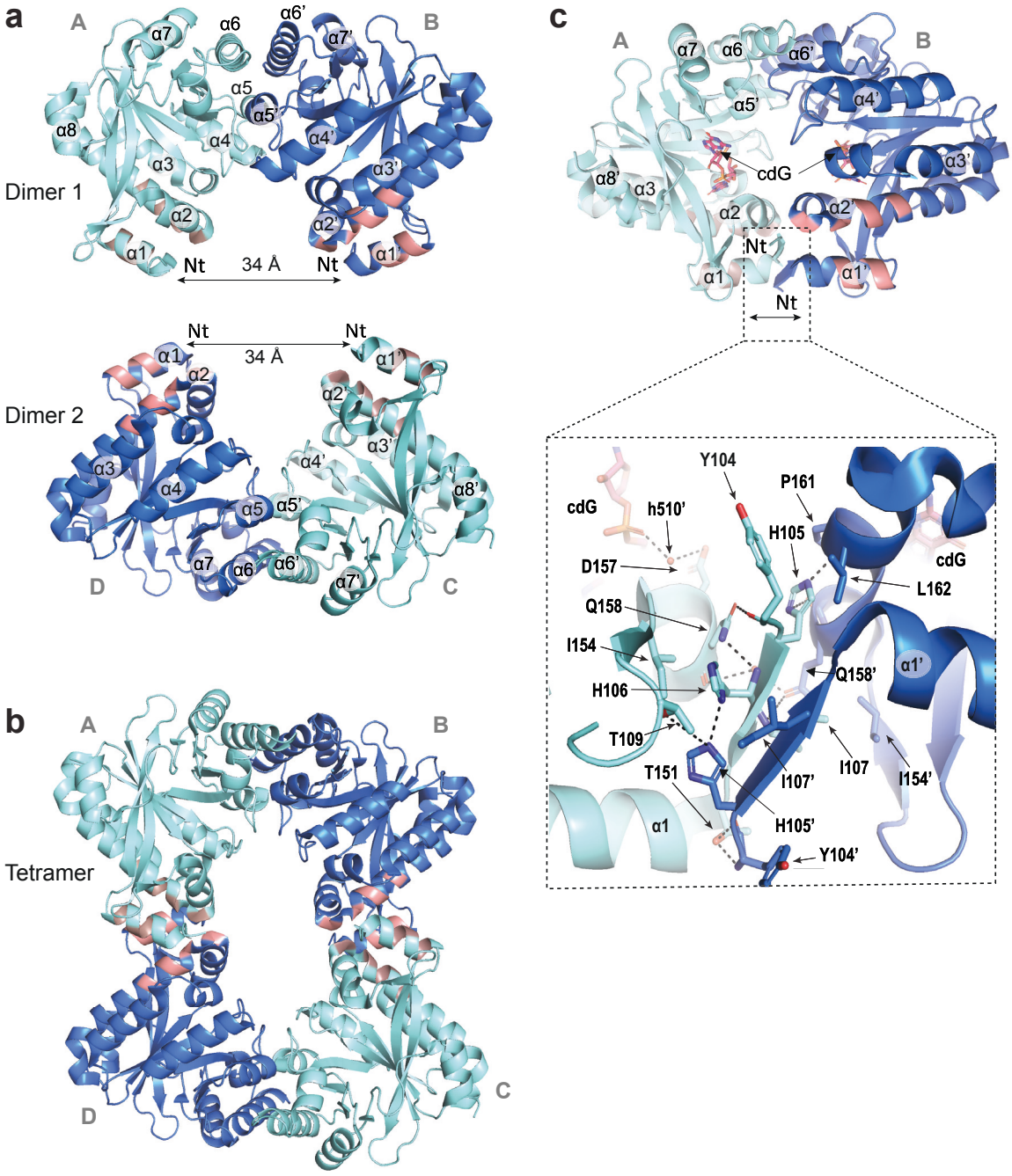


Figure S4: PdeL R- and T-state dimer configurations. **(a)** Two PdeL dimers in the R-state configuration with monomers A and C in cyan and B and D in blue. Residues of helices α1 and α2 forming the tetramerization interface are colored in light brown (monomers A and C) and pink (monomers B and D), respectively. The distance between the N-termini (Nt) of two protomers is indicated. **(b)** The PdeL tetramer is formed through interactions of the colored residues indicated. **(c)** PdeL T-state dimer configuration (PDB: 4lj3) with c-di-GMP shown in sticks **(top)**. The structure illustrates the reduced distance between the N-termini as compared to the R-state configuration. Direct monomer-monomer contacts in the Nt region prevents interactions involved in tetramerization from occurring. **Bottom:** zoom-in on the N-terminal interactions of the PdeL T-state with residues involved shown as sticks and electrostatic interactions indicated as dashed lines. Monomer-monomer contacts lead to the formation of an a/b strand in the N-terminus of each monomer that stabilizes this conformation.

Table S1: Strains used in this study

Strain	Genotype	Source
MG1655	*E. coli K-12*	Blattner et al., 1997
BL21 (AI)	*F^- ompT $hsdS_B$ (r_B^- m_B^-) gal dcm araB::T7 RNAP-tetA*	Life techn.
AB330	*λ cI857 Δ(cro-bioA)*	A. Böhm
AB607	*ΔpdeH::frt*	Boehm et al., 2010
AB989	*λ cI857 Δ(cro-bioA) kan::P_{rha}-ccdB-lacZ*	A. Böhm
AB2137	*ΔpdeH::frt pdeL (G299S)-3xflag::frt*	Reinders et al., 2016
AB2202	*ΔpdeH::frt pdeL (F206S)-3xflag::frt*	Reinders et al., 2016
AB2203	*ΔpdeH::frt pdeL (F249L)-3xflag::frt*	Reinders et al., 2016
AB2271	*ΔpdeH::frt ΔpdeL::frt*	this study
AB2377	*ΔpdeH::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2378	*ΔpdeH::frt pdeL (G299S)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2400	*kan::$P_{pdeL\ CB-}$-lacZ (merodiploid translational fusion)*	this study
AB2401	*kan::$P_{pdeL\ CB-\ \&\ CDB-}$-lacZ (merodiploid translational fusion)*	this study
AB2402	*kan::$P_{pdeL\ CDB-}$-lacZ (merodiploid translational fusion)*	this study
AB2519	*pdeL (E235A)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2520	*pdeL (D263N)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2521	*pdeL (S298F)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2535	*pdeL (G299S)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2569	*ΔpdeL::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	Reinders et al., 2016
AB2571	*ΔpdeH::frt ΔpdeL::frt*	this study
AB2609	*pdeL (K60A)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	Reinders et al., 2016
AB2727	*kan::$P_{pdeL\ P1-}$-lacZ (merodiploid translational fusion)*	this study
AB2731	*pdeL (E141A)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2789	*ΔpdeL::frt Δcra::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2806	*Δcra::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2830	*Δhns::frt pdeL-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2846	*ΔpdeH::frt pdeL (D263N)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2847	*ΔpdeH::frt pdeL (K60A)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2848	*ΔpdeH::frt pdeL (S298F)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2849	*ΔpdeH::frt pdeL (E141A)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2851	*ΔpdeH::frt pdeL (E235A)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2905	*pdeL (K283R)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2907	*ΔpdeH::frt pdeL (K283R)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2923	*ΔpdeH::frt Δhns::frt pdeL-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2937	*pdeL (D263N) (K283R)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2939	*ΔpdeH::frt pdeL (D263N) (K283R)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2940	*pdeL (F206S)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2942	*ΔpdeH::frt pdeL (F206S)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2943	*pdeL (F249L)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study

(continued) Table S1: strains used in this study

Strain	*Genotype*	Source
AB2945	*ΔpdeH::frt pdeL (F249L)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB2986	*kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB2996	*ΔpdeH::frt csrA::Tn5Δ(kan)::frt*	this study
AB2997	*ΔpdeH::frt ΔpdeL::frt csrA::Tn5Δ(kan)::frt*	this study
AB3292	*kan::$P_{pdeL\ CIB-}$-lacZ (merodiploid translational fusion)*	this study
AB3299	*ΔpdeH::frt frt::$P_{const.\ weak.}$-pgaA-D*	this study
AB3302	*ΔpdeH::frt ΔpdeL::frt frt::$P_{const.\ weak.}$-pgaA-D*	this study
AB3335	*ΔpdeH::frt $P_{pdeL\ CIB-}$-pdeL-3xflag::frt kan::PpdeL CIB--lacZ (merodiploid translational fusion)*	this study
AB3340	*ΔpdeH::frt kan::$P_{pdeL\ CDB-}$-lacZ (merodiploid translational fusion)*	this study
AB3368	*ΔpdeH::frt pdeL (T270A)-3xflag::frt*	this study
AB3381	*pdeL (T270A)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB3383	*ΔpdeH::frt pdeL (T270A)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB3431	*ΔpdeH::frt pdeL-[$RBS_{synth.}$-mCherry]$_2$::frt*	this study
AB3447	*pdeL (D295N)-3xflag::frt kan::P_{pdeL}-lacZ (merodiploid translational fusion)*	this study
AB3449	*ΔpdeH::frt pdeL (D295N)-3xflag::frt kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB3455	*ΔpdeH::frt frt::P_{tet}-tetR-pdeL kan::PpdeL-lacZ (merodiploid translational fusion)*	this study
AB3496	*ΔpdeH::frt pdeL (D263N) (K283R)-[$RBS_{synth.}$-mCherry]$_2$::frt*	this study
AB3499	*ΔpdeH::frt $P_{pdeL\ CIB-}$-pdeL-[$RBS_{synth.}$-mCherry]$_2$::frt*	this study
AB3501	*PpdeL (CIB-)-pdeL::frt*	this study
AB3508	*ΔpdeH::frt frt::P_{tet}-tetR-pdeL kan::PpdeL CIB--lacZ (merodiploid translational fusion)*	this study
AB3673	*pdeL (E141A)::frt*	this study
AB3718	*pdeL (D295N)::frt*	this study
AB3812	*E. coli K-12 MG1655 CGSC 6300*	Coli Genetic Stock Center
AB3871	*ΔpdeH::frt frt::Ptet-RBS (synth. strong)-tetR-RBS (synth. weak)-pdeL (L168R)*	this study
AB4490	*CGSC 6300 ΔpdeH::frt*	this study
AB4491	*CGSC 6300 ΔpdeL::frt*	this study
AB4513	*pdeL (E141A)::frt-kan-frt*	this study
AB4514	*pdeL (D295N)::frt-kan-frt*	this study
AB4519	*CGSC 6300 pdeL (E141A)::frt*	this study
AB4520	*CGSC 6300 pdeL (D295N)::frt*	this study
AB4532	*PpdeL (CIB-)-pdeL::frt-kan-frt*	this study
AB4533	*ΔpdeH::frt frt::Ptet-RBS (synth. strong)-tetR-RBS (synth. weak)-pdeL (L168R)::Frt-kan-Frt*	this study
AB4534	*CGSC 6300 Δ(PpdeL-pdeL)::frt-cat-frt*	this study
AB4536	*CGSC 6300 PpdeL (CIB-)-pdeL::frt*	this study
AB4540	*CGSC 6300 pdeL (L186R)::frt*	this study

Table S2: Plasmids used in this study

Plasmid	Genotype (resistance)	Source
pET28a	pBR332 lacI P_{T7} (*kan*) 6xHis expression vector	Novagen
pKD46	λ RED^+ (*amp*)	Datsenko and Wanner, 2000
pCP20	FLP^+ (*amp*)	Cherepanov et al., 1995
pNDM220	*repA parR parM* P_{lac} (amp)	Gotfredsen and Gerdes, 1998
pAR1	P_{lac}-*pdeL-strepII* in pET28a (*kan*)	Reinders et al., 2016
pAR3	P_{lac}-*pdeL (D263N)-strepII* in pET28a (*kan*)	this study
pAR19	P_{lac}-*cra-strepII* in pET28a (*kan*)	this study
pAR28	P_{lac}-*hns-strepII* in pET28a (*kan*)	this study
pAR52	P_{lac}-*pdeL (K283R)-strepII* in pET28a (*kan*)	this study
pAR62	P_{lac}-*pdeL (D263N) (K283R)-strepII* in pET28a (*kan*)	this study
pAR81	P_{lac}-$RBS_{synth.}$-*pdeH-3xflag* in pNDM220 (*amp*)	this study
pAR201	P_{lac}-*pdeL (D295N)-strepII* in pET28a (*kan*)	this study
pAR202	P_{lac}-$pdeL_{EAL}$-*3xflag-strepII* in pET28a (*kan*)	this study
pAR205	P_{lac}-$pdeL_{EAL}$ *(Y268C)-3xflag-strepII* in pET28a (*kan*)	this study
pAR210	P_{lac}-$pdeL_{EAL}$ *(Y268C) (T270A)-3xflag-strepII* in pET28a (kan)	this study
pAR211	P_{lac}-$pdeL_{EAL}$ *(Y268C) (E235A)-3xflag-strepII* in pET28a (kan)	this study
pAR212	P_{lac}-$pdeL_{EAL}$ *(Y268C) (D295N)-3xflag-strepII* in pET28a (kan)	this study
pAR226	*pdeL-3xFlag::kan* in pUC19 *(amp)*	this study
pAR231	*pdeL::kan* in pUC19 (*amp*)	this study
pAR323	P_{pdeL}-*gfpmut2* in pUA66 (*kan*)	this study
pAR341	P_{lac}-$RBS_{synth.}$-*pdeH-3xflag* in pNDM220 (*kan*)	this study
p2H12ref	P_{tet}-*sensor-mScarlet-I* in pBR322 (*amp*)	UJ11206, Kaczmarczyk and Jenal, unpublished
p2H12ref-blind	P_{tet}-*sensor*-mScarlet-I* in pBR322 (*amp*)	UJ11207, Kaczmarczyk and Jenal, unpublished

Table S3: Primers used for EMSA in this study

Primer number	Description	Sequence
4991	CIB_fwd	5' Cy3-GTTGCGAATGTTCAATAAGTTTAG
4997	CIB_ref	5' Cy3-ATGCGTCATTTCAAATGATCAGC
4695	CB-CDB_fwd	5' Cy3-TGCTGAATGGATTCAGTCTTAATGAGTGGG
4696	CB-CDB_rev	5' Cy3-CCCACTCATTAAGACTGAATCCATTCAGCA
4283	RNAP_bind_pdeL_rev	5' Cy3-GAGCAAAGGCGCATTATATG

Chapter 3

Genetic variations between common *E. coli* MG1655 laboratory stocks impose differences in signaling and adaptation

Margo van Berkum, Théo Gervais, Urs Jenal

Biozentrum of the University of Basel
Klingelbergstrasse 50/70, 4056 Basel, Switzerland

Contents

Statement of my contributions:

Data presented in Figure S2 has been generated by Théo Gervais during his rotation in the Jenal lab. I constructed all strains and designed and performed all other experiments for which data is presented in Chapter 3. I analyzed all data and I wrote the complete chapter.

3.1 Abstract

The *Escherichia coli* K-12 MG1655 strain is a widely used laboratory strain and model organism. Although genetic diversification is unwanted for laboratory strains, variations in the sequences of MG1655 stocks were previously identified. These differences include the presence or absence of insertion sequence (*IS1*) elements in the *flhDC* regulatory region and *crl* gene. How these genetic differences result in different phenotypes is still largely unexplored.

Here, we compare aspects of signaling and growth for the MG1655 stocks CGSC 6300 and CGSC 7740. We show that the P_{flhDC}*::IS1 crl*::*IS1* genotype of the CGSC 7740 stock drives a growth phenotype of reduced lag time after sudden substrate switches. A difference in lag time is conserved among very diverse pre-shift conditions including glucose minimal medium and rich tryptone broth. Whole proteome comparison of the two MG1655 stocks showed a distinct protein signature with more than 100 proteins significantly up- or downregulated, that was remarkably well conserved among different pre-shift growth conditions.

We also describe a third *IS1* element located in the *dgcJ* gene encoding a putative diguanylate cyclase. In contrast to the *IS1* element in P_{flhDC}, the *IS1* element in *dgcJ* did not affect motility nor did it change cytosolic cyclic-di-GMP levels under the conditions analyzed, suggesting DgcJ activity to be highly specific.

Our results demonstrate how differently *E. coli* K-12 MG1655 stocks can respond and behave under standard laboratory conditions as the result of few genetic changes. Sequence annotation updates as well as the sequencing of laboratory strains are examples to keep the research community informed about the genotypes of widely used model organisms.

3.2 Introduction

The *E. coli* laboratory strain K-12 MG1655 is a widely used strain in research on *E. coli* genetics and physiology. Other common K-12 strains include W3110 and BW25113, the parental strain of the Keio collection (Baba et al., 2006). Although they are all K-12 strains, genetic differences exist and come in the occurrence of nucleotide deletions and insertions, base substitutions and Insertion Sequence (*IS*) element abundance (Hayashi et al., 2006; Riley et al., 2006). BW25113 was derived from K-12 following a series of generalized transduction and allele replacements (Baba et al., 2006; Datsenko and Wanner, 2000). The full genomes of MG1655, W3110 and BW25113 have been sequenced and annotated (Blattner et al., 1997; Grenier et al., 2014; Hayashi et al., 2006). Genome sequence information however should be seen as a snapshot in time (Riley et al., 2006), as the availability of detailed sequence analysis and experimental evidence on gene functions are still expanding. Accurate sequences and annotation are of importance, because it provides fundamental information for the *E. coli* research community and beyond.

Strain collections exist across the world and from here, bacterial stocks can be obtained upon request. The Coli Genetic Stock Center (CGSC) at Yale University (New Haven, CT, USA) holds almost ten thousand different non-pathogenic *E. coli* strains, primarily K-12 derivatives including MG1655, W3110 and most strains from the Keio collection. The center's online database (https://cgsc.biology.yale.edu/) has three entries for the wild type strain *E. coli* K-12 MG1655: CGSC 6300, CGSC 7740 and CGSC 8237. They are all indicated as the wild type background 'F-, λ-, rph-1'. From these three entries, CGSC 6300 was deposited first, by M.S. Guyer in 1981 (Guyer et al., 1981) when its full sequence was not known yet. It was this strain entry CGSC 6300 that was requested by the Blattner lab, who sequenced the strain and thus published the first whole genome sequence of *E. coli* K-12 in the mid-90s (Blattner et al., 1997). Importantly, Blattner et al. sent an MG1655 culture back to the CGSC after sequencing. This sequenced 'subculture of MG1655' was added to the CGSC as the new database entry CGSC 7740 and with the addition of 'seq': MG1655(seq).

Over time it has become evident that the stocks CGSC 6300 and CGSC 7740 carry genetic differences in the occurrence of *IS* element insertions, single nucleotide insertions and deletions as well as SNPs (Barker et al., 2004; Freddolino et al., 2012). Compared to CGSC 6300, the genome of CGSC 7740 carries among others an *IS1* element in the regulatory region of the *flhDC* operon (Barker et al., 2004), an *IS1* element in the open reading frame of the *crl* gene, an *IS5* element in the regulatory region of the *oppABCDF* operon and loss of function mutations in the *glpR* and *gatC* gene (Freddolino et al., 2012).

The *E. coli* K-12 MG1655 genome contains 45 *IS* elements of different kinds (Blattner et al., 1997). *IS* elements are mobile genetic elements that can duplicate and reintegrate themselves elsewhere in genomes, forming an element of genome plasticity (Mahillon and Chandler, 1998; Naas et al., 1994). Mutations in bacterial genomes arise as a result of *IS* element translocation. *IS* elements usually only encode information required for their motility, but can activate neighboring genes by forming new promoter sites. Integration of an *IS* element inside an open reading frame is associated with

loss of function of that specific gene (Barker et al., 2004; Freddolino et al., 2012; Mahillon and Chandler, 1998). Different families of *IS* elements exist. *IS1* elements are 768 bp long and prefer AT-rich DNA regions for insertion. With integration of an *IS1* element, commonly 9 bp of target DNA are duplicated at the site of insertion (Calos et al., 1978; Grindley, 1978).

What would be the consequences of these *IS1*-driven genetic differences on MG1655 phenotypes? The same reports describing these genetic differences, predict or provide the first evidence for distinct phenotypes under specific conditions (Barker et al., 2004; Freddolino et al., 2012). The most eminent phenotype so far that distinguishes the two MG1655 stocks is motility. CGSC 6300 is described as 'poorly motile', whereas CGSC 7740 is described as 'highly motile' (Barker et al., 2004). This motility phenotype is assigned to the *IS1* element in P_{flhDC} (P_{flhDC}*::IS1*). The *flhDC* operon encodes the regulatory proteins FlhD and FlhC that function in the heterohexameric assembly $FlhD_4C_2$ (further indicated FlhDC) as the master regulator of flagellar biosynthesis (Liu and Matsumura, 1994; Wang et al., 2006). Expression of this master regulator is under tight regulation of several inhibitors and activators responding to environmental and nutritional cues. Among others, *flhDC* expression is subject to catabolite repression as binding of the cAMP CRP complex to the *flhDC* promoter region is essential for transcriptional activation (Soutourina et al., 1999; Yokota & Gots, 1970). The *IS1* element in the *flhDC* regulatory region is positioned in the same direction as the *flhDC* operon and disrupts a binding site of an *flhDC* repressor. It was proposed that P_{flhDC}*::IS1* results in high motility through reduced *flhDC* repression (Barker et al., 2004). In a single cell analysis of CGSC 6300, it was recently observed that flagellar genes are stochastically activated. Pulsing of flagellar gene expression is a consequence of its regulatory mechanism and leads to bimodal expression patterns (Kim et al., 2020). In contrast, promoter activity of flagellar genes was unimodally high in CGSC 7740, confirming the constitutive expression of flagellar genes in this background (Kim et al., 2020). The *E. coli* flagellar genes are encoded in a regulatory cascade divided in class I, class II and class III genes, resulting in sequential expression of structural components of the flagellar organelle and regulatory factors (see also chapter 1). The *flhDC* operon is the only class I operon, that activates class II operons encoding among others the hook-basal body complex and the flagella-associated sigma factor *fliA*. FliA is the required sigma factor for expression of class III genes including *fliC* encoding flagellin and the c-di-GMP phosphodiesterase *pdeH*. The level of the nucleotide second messenger c-di-GMP needs to be low to enable flagella rotation, as c-di-GMP in complex with the molecular brake protein YcgR (encoded as class III gene) reduces flagellar motor activity (Boehm et al., 2010; Paul et al., 2010). Besides flagellar genes, FlhD alone or in complex as FlhDC controls other cellular processes including cell division (Prüss and Matsumura, 1996; Prüss et al., 1997), anaerobic respiration and central carbon metabolism (Prüss et al., 2003).

The second *IS1* element, in the *crl* gene, is likely to affect growth of *E. coli* under laboratory conditions (Freddolino et al., 2012). As the *IS1* element is inserted in the *crl* open reading frame, the *crl*::*IS1* genotype most likely causes a loss of function of *crl* (Freddolino et al., 2012). The Crl protein functions as activator of the stress sigma factor σ^S or RpoS (Pratt and Silhavy, 1998; Typas et al., 2007). RpoS in complex with the core RNA polymerase (RNAP) stimulates the

expression of numerous genes involved in the general stress response, long-term survival, stationary phase and biofilm formation (Hammar et al., 1995; Lange and Hengge-Aronis, 1991; Loewen and Hengge-Aronis, 1994; Olsen et al., 1993; Tanaka et al., 1993; Weber et al., 2005). Crl can diarectly bind to RpoS and this increases the affinity of RpoS for RNAP (Bougdour et al., 2004; England et al., 2008). Hence, Crl facilitates RpoS-RNAP complex formation ($E\sigma^S$). Overall, Crl modulates the balance between RpoS and other sigma factors as the vegetative RpoD (σ^{70}) and flagella-associated FliA (σ^{28}), and thereby the balance between growth and stress resilience (Hengge, 2020; Nystrom, 2004; Pesavento et al., 2008; Typas et al., 2007).

In this study, we are providing evidence for the existence of an additional *IS1* element that is located in the *dgcJ* gene and distinguishes the two MG1655 stocks CGSC 6300 and CGSC 7740. Despite its annotation as putative diguanylate cyclase, we did not find evidence for *dgcJ* to affect motility or cytosolic cyclic-di-GMP concentrations, that are however primarily altered by the *IS1* element in the *flhDC* intergenic region. The *IS1* elements in P_{flhDC} and *crl* in the MG1655 stock CGSC 7740 drive a growth phenotype of reduced lag time after sudden substrate switches including but not limited to the switch from glucose to fumarate. We set one additional step towards unraveling the mechanism behind this adaptation advantage by whole proteome analysis of both MG1655 stocks. These results expand our knowledge on physiological differences between MG1655 stocks and how they arise from few genetic changes in primarily regulatory proteins. This knowledge can further support deliberate strain choices in the context of experimental studies on specific phenotypes.

3.3 Results

3.3.1 Three *IS1* elements genetically distinguish MG1655 stocks CGSC 6300 and CGSC 7740

In order to verify the respective absence and presence of described *IS1* elements in the *E. coli* K-12 MG1655 stocks CGSC 6300 and CGSC 7740 (Barker et al., 2004; Freddolino et al., 2012), we PCR amplified the P_{flhDC}-*flhDC* and *crl* region from the CGSC 6300 and CGSC 7740 genome upon arrival of these strains from the Coli Genetic Stock Center (CGSC). Indeed, PCR results indicated ~700 bp long DNA inserts in CGSC 7740 (Fig. 1a). Sequencing of the PCR fragments confirmed that in both regions the DNA insert encodes an *IS1* element.

Through whole genome sequencing of CGSC 7740 and subsequent alignment to the *E. coli* K-12 MG1655 sequence NC_000913.3 from the NCBI database, an *IS1* element insertion in the *dgcJ* open reading frame was detected. This gene is currently annotated as intact *dgcJ* in NC_000913.3. We wondered whether this *IS1* element could be a third, yet unannotated *IS1* element that genetically distinguishes MG1655 stocks CGSC 6300 and CGSC 7740. PCR amplification and sequencing of the *dgcJ* region indicated that the CGSC 6300 genome contains an intact *dgcJ* open reading frame, while the genotype of CGSC 7740 is *dgcJ::IS1* (Fig.1a). This result suggests

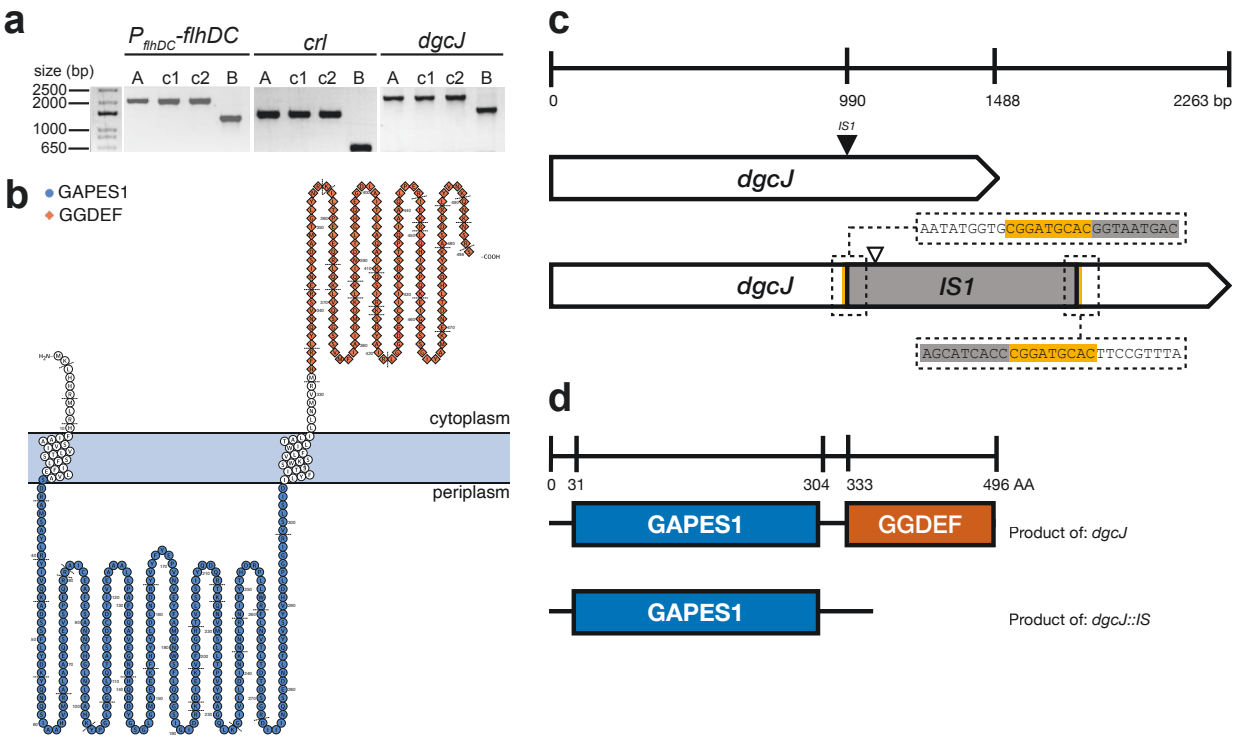


Figure 1: Identification of *IS1* elements in MG1655 stock CGSC 7740. **(a)** PCR amplification of the region including the *flhDC* promoter and open reading frame (P_{flhDC}-*flhDC*; left), the *crl* open reading frame (middle) and *dgcJ* open reading frame (right). Strains ordered from the CGSC are indicated with A (CGSC 7740) and B (CGSC 6300). MG1655 c1 and c2 are control strains where c1 is the MG1655 reference strain in the laboratory of U. Jenal and c2 a MG1655 stock received from the laboratory of M. Heinemann. **(b)** Visualization of the predicted DgcJ protein domain localization. **(c)** Graphical summary of *dgcJ* and *dgcJ::IS1* gene. Filled triangle: point of insertion of *IS1* element in *dgcJ* gene. Open triangle: stop codon introduced by *IS1* element insertion. Sequences in the boxes show the *dgcJ* sequence (white), 9-bp duplication of *dgcJ* (yellow) and *IS1* (grey). **(d)** Graphical summary of the *dgcJ* and suggested *dgcJ::IS1* gene product. The full-length DgcJ protein contains 496 amino acid residues and has two domains: an N-terminal gammaproteobacterial periplasmic sensor domain (GAPES1 domain) and a C-terminal GGDEF domain.

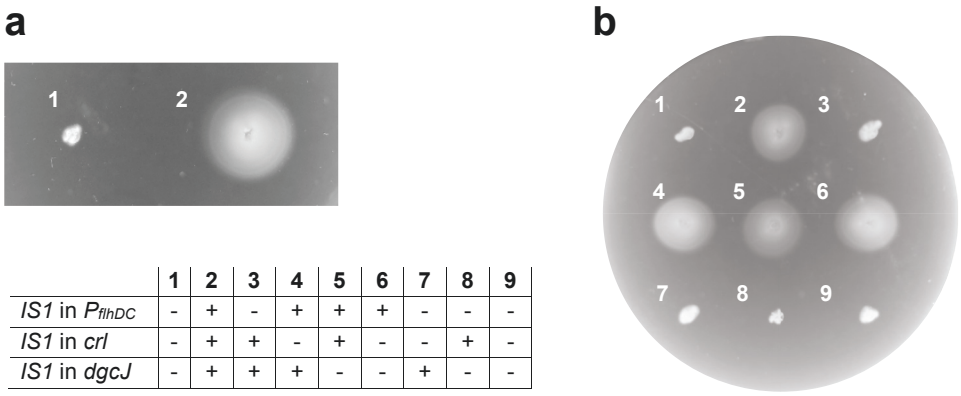
that the presence of *dgcJ::IS1* coincides with *crl*::*IS1* and P_{flhDC}*::IS1*. To corroborate on this, we verified the presence of all three *IS1* elements in the motile MG1655 strain originating from the laboratory of F. Blattner (University of Wisconsin) (Blattner et al., 1997) and in a motile MG1655 strain received from the laboratory of M. Heinemann (University of Groningen). We confirmed the presence of all three *IS1* elements in both strains (Fig. 1a). From this we conclude that a third *IS1* element genetically distinguishes stocks CGSC 6300 and CGSC 7740.

The *dgcJ* gene encodes a putative diguanylate cyclase with an N-terminal gammaproteobacterial periplasmic sensor domain (GAPES1 domain), a C-terminal catalytic GGDEF domain and two transmembrane domains (Fig. 1b,d) (Blum et al., 2020). Whereas the sensor domain is predicted to locate in the periplasm, the catalytic domain is predicted to locate in the cytoplasm (Fig. 1b) (Blum et al., 2020). DgcJ is one of the 12 diguanylate cyclases in *E. coli* K-12 that synthesize the second messenger c-di-GMP (Hengge et al., 2016; Pesavento et al., 2008). In order to determine how the *IS1* element in *dgcJ* affects *dgcJ* expression and functioning, we defined the direction of the *IS1* element and the exact location of its insertion based on sequencing data (see Supplementary material). The *IS1* element in *dgcJ* is inserted in the reverse direction compared to the *dgcJ* gene alignment. The *IS1* element is inserted 990 bp downstream of the start of the *dgcJ* open reading frame and is thus located in between the regions encoding the GAPES1 and GGDEF domain (Fig. 1c,d). As a consequence of its sequence alignment and point of insertion, the *IS1* element sequence introduces a stop codon in *dgcJ* 90bp away from the point of *IS1* insertion (Fig. 1c). Because of the reversed direction of the *IS1*, there is most likely no promotor region driving expression of the separate GGDEF domain. Based on this we predict that the genotype *dgcJ::IS1* generates a non-functional cyclase, as the catalytic GGDEF domain is uncoupled by the premature stop codon (Fig.1 c,d). Since DgcJ has been reported to interfere with motility at 37°C (Pesavento et al., 2008), we were wondering whether the disruption of *dgcJ* contributes to the distinct motility phenotypes of CGSC 6300 and CGSC 7740 (Barker et al., 2004).

3.3.2 *flhDC*, but neither *dgcJ* nor *crl*, modulates *E. coli* motility

The swarm rates of stocks CGSC 6300 and CGSC 7740 largely deviate, with CGSC 7740 having an over six fold higher swarm rate on motility agar (Barker et al., 2004). We confirmed the motility phenotype with a motility assay. Indeed, CGSC 7740 was highly motile as it showed a clear halo around the point of inoculation on the motility agar after six hours of incubation at 37°C (Fig. 2a). In contrast, CGSC 6300 is poorly motile as no halo was observed (Fig. 2a).

Mutations occurring in the *flhDC* promoter region of non-motile strains are associated with increased motility (Barker et al., 2004; Fahrner and Berg, 2015; Zhang et al., 2017). The deletion of both *dgcJ* and *rpoS* can increase motility via different mechanisms (Dong and Schellhorn, 2009; Pesavento et al., 2008). Therefore, we aimed to verify if *crl*::*IS1* and *dgcJ::IS1* also contribute to the motility phenotype. In order to dissect the contribution of each of the three *IS1* elements to the observed motility phenotype, we first removed the individual *IS1* elements from the CGSC 7740 genome by replacing the affected regions with the sequence from CGSC 6300 (for details see Materials and Methods). This resulted in nine strains, including the wild type CGSC 7740 and



	1	2	3	4	5	6	7	8	9
IS1 in P_{flhDC}	-	+	-	+	+	+	-	-	-
IS1 in *crl*	-	+	+	-	+	-	-	+	-
IS1 in *dgcJ*	-	+	+	+	-	-	+	-	-

Figure 2: *IS1* element in P_{flhDC} affects motility. Outgrowth of *E. coli* cells on motility agar plate (TB + 0.3% agar) after six hours of incubation at 37°C. **(a)** Motility plate incubated with wt CGSC 6300 **(1)** and wt CGSC 7740 **(2)**. **(b)** Motility plate inoculated with wt CGSC 6300 **(1)**, wt CGC 7740 **(2)** and derivatives of CGSC 7740 with *IS1* elements removed from P_{flhDC}, *dgcJ* and/or *crl* (3-9) as indicated in the table.

CGSC 6300, covering all possible combinations of one, two or all three of the *IS1* elements considered. Next, we assessed the motility of all strains on motility agar plates. As expected, restoring the *flhDC* regulatory region (CGSC 7740 P_{flhDC}) completely abolished motility to the level as observed for CGSC 6300 (Fig. 2b). In contrast, restoration of neither the *dgcJ* nor the *crl* gene altered motility in CGSC 7740 (Fig. 2b). Restoring P_{flhDC} in addition to *dgcJ* and/or *crl* then fully abolished motility (Fig. 2b). This was expected based on the observed effect on motility of the P_{flhDC} restoration alone. Overall, we concluded that the *IS1* element in P_{flhDC} drives increased motility in CGSC 7740 at 37°C.

3.3.3 P_{flhDC}*::IS1* affects global c-di-GMP signaling

The three operons affected by *IS1* elements in CGSC 7740 are all directly or indirectly involved in the regulation of the second messenger c-di-GMP. It is assumed that P_{flhDC}*::IS1* increases *pdeH* expression (Barker et al., 2004; Ko and Park, 2000; Zhao et al., 2007). *crl*::*IS1* could well affect the expression of several diguanylate cyclases under RpoS control (Pesavento et al., 2008; Weber et al., 2005, 2006). A non-functional DgcJ is the direct consequence of *dgcJ::IS1*. We investigated how cellular c-di-GMP levels are affected by these three *IS1* elements using a novel single cell biosensor for c-di-GMP (Kaczmarczyk and Jenal, unpublished). The sensor was expressed in strains with all possible *IS1* element combinations during growth on M9 glycerol and c-di-GMP distributions were measured. A culture of the original CGSC 6300 stock displayed a bimodal c-di-GMP distribution (Fig. 3a), as previously reported for this stock when grown on gluconeogenic sources (see Chapter 2). In contrast, the c-di-GMP distribution in CGSC 7740 was unimodal and c-di-GMP levels were low, as the distribution overlapped with that obtained from the c-di-GMP binding-deficient control sensor (Fig. 3a). Restoring P_{flhDC} in the CGSC 7740 background restored bimodal c-di-GMP (Fig. 3a). Removing the other two *IS1* elements in addition to P_{flhDC} did not have any additive effect. Besides, restoring the *crl* and *dgcJ* individually or combined in CGSC 7740 did not change the c-di-GMP pattern during growth on glycerol as compared to CGSC 7740 wt (Fig. 3a).

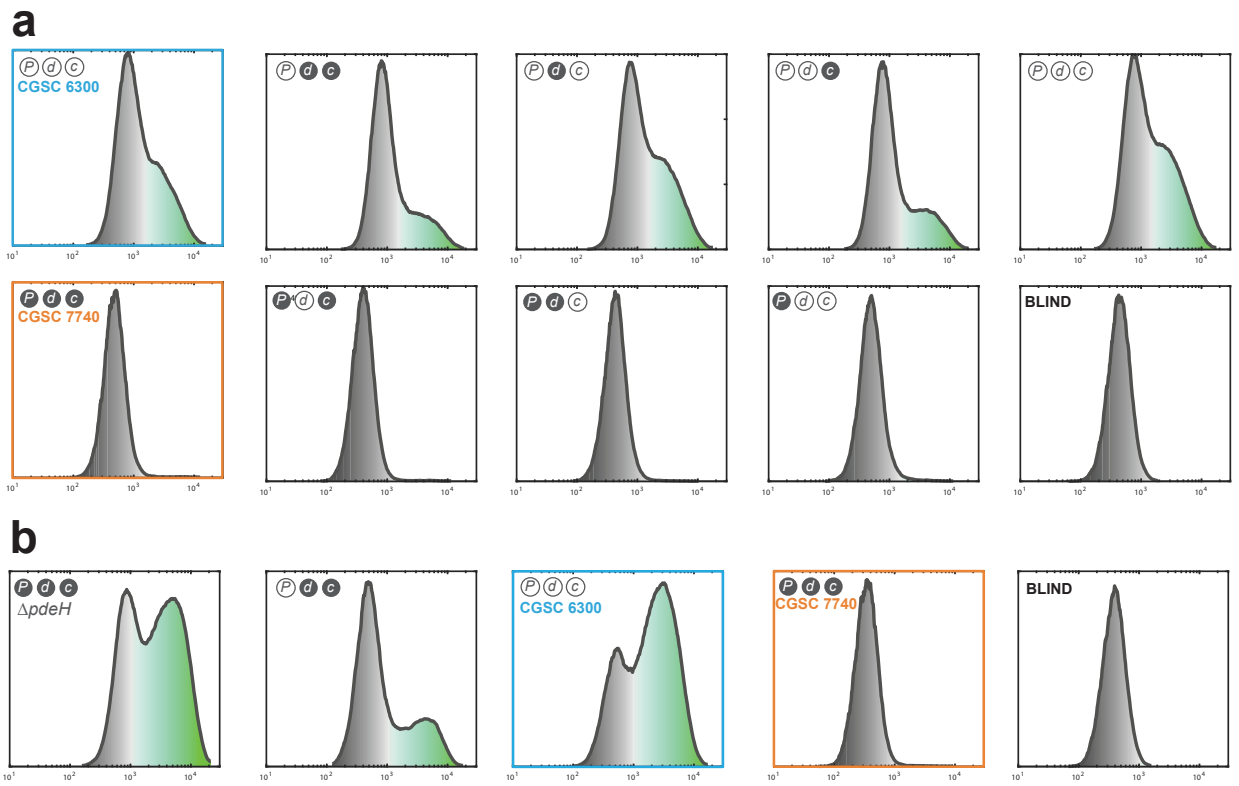


Figure 3: *IS1* element in P_{flhDC} affects single-cell c-di-GMP distributions via *pdeH*. Single cell measurements of the cytosolic c-di-GMP level of *E. coli* cells transformed with the plasmid-based c-di-GMP biosensor (Kaczmarczyk and Jenal, unpublished). C-di-GMP distributions in background CGSC 6300 (blue frame), CGSC 7740 (orange frame) and CGSC 7740 mutants lacking one or several of the *IS1* elements. Open/white circles: no *IS1* element in the specific region. Filled, grey circles: *IS1* element present. P: P_{flhDC}, d: *dgcJ*, c: *crl*. 'BLIND': representative measurement of CGSC 7740 expressing a c-di-GMP-binding deficient biosensor (Kaczmarczyk and Jenal, unpublished). **(a)** c-di-GMP distributions in background CGSC 6300, CGSC 7740 and CGSC 7740 mutants lacking one or several of the *IS1* elements were determined during exponential growth on glycerol minimal medium. LB pre-cultures were washed and diluted in M9 glycerol supplemented with 200 nM anhydrotetracycline to induce c-di-GMP sensor expression and grown until exponential phase. **(b)** c-di-GMP distributions as in **(a)** for the Δ*pdeH* strain and control strains.

The FlhDC transcriptional regulator directly activates *pdeH* expression. Therefore, we wondered whether increased expression of *pdeH* alone is responsible for the unimodal low distribution of c-di-GMP. We expressed the c-di-GMP biosensor in CGSC 7740 *pdeH* deletion background and observed that *pdeH* deletion is sufficient to restore bimodal c-di-GMP distribution (Fig. 3b). From this we concluded that altered *pdeH* expression in CGSC 7740 is the driver of cytoplasmic changes in c-di-GMP concentrations.

As the *flhDC* operon requires the cAMP-CRP complex as an essential activator of transcription (Soutourina et al., 1999), we wondered whether the observed differences in c-di-GMP distributions are maintained under low cAMP. Hence, we expressed the c-di-GMP sensor in CGSC 6300 and CGSC 7740 during growth on glucose as the sole energy and carbon source. On glucose, the c-di-GMP distribution in CGSC 6300 populations was no longer bimodal, but instead displayed a unimodal high c-di-GMP distribution (Fig. S1 and chapter 2). The c-di-GMP distribution of a CGSC 7740 population is unimodal low on glucose (Fig. S1). So, also at low cAMP levels, the two MG1655 stocks show distinct c-di-GMP patterns.

3.3.4 *IS1* elements shorten lag time after substrate shift in a cumulative manner

RpoS is involved in setting the SPANC balance, a trade-off between stress resistance and nutritional competence (Ferenci, 2005). While RpoS induces the general stress response, attenuating RpoS activity seems to represent an evolutionary important strategy to increase the cell's capability to grow on a range of alternative carbon sources including fumarate and acetate (Chiang et al., 2011; King et al., 2004). With Crl being the activator of RpoS (Bougdour et al., 2004; Typas et al., 2007), we hypothesized RpoS activity to be reduced in a *crl*::*IS1* background and wondered how this affects growth, in particular under dynamic conditions with changes in substrate availability. Therefore, we analyzed the MG1655 growth phenotype after sudden substrate switches, events which are usually accompanied by a substantial lag time before growth on the new substrate is observed (Barthe et al., 2020; Basan et al., 2020; Kotte et al., 2014). First, we compared the growth phenotypes of the CGSC 6300 and CGSC 7740 wt stocks after a substrate switch from glucose to fumarate. Exponentially growing cultures on glucose minimal medium were switched to minimal medium with fumarate as the sole carbon source. Here, we observed that the lag phase of CGSC 6300 is twice as long as that of CGSC 7740 (Fig. 4a,d). In addition, the maximum OD_{600} of 0.4 of CGSC 7740 was 20% lower than that of CGSC 6300, which reached a maximum OD_{600} of 0.5 on fumarate (Fig. 4a,e). A difference in lag time was also observed when switching from glucose to acetate and was also conserved for a shift from rich Tryptone Broth (TB) to fumarate (Fig. 4b,c). These observations suggest that the difference in lag time is not specific for switching from glycolysis to gluconeogenesis. A more fundamental growth difference between the two MG1655 stocks would allow for such a consistent pattern for a diverse substrate range.

Next, we investigated how *crl* affects the lag time after substrate shift by repeating the shift from glucose to fumarate and including the CGSC 7740 strain with restored *crl* (CGSC 7740 *crl*). Indeed, *crl* affected the lag time, as CGSC 7740 *crl* showed an intermediate lag phase that is 70% longer than the CGSC 7740 wt, but still 20% shorter than CGSC 6300 (Fig. 4d). We concluded that *crl* expression alone cannot explain the difference in lag time between CGSC 7740 and CGSC 6300. We went on and asked how the other two known *IS1* elements, P_{flhDC}*::IS1* and *dgcJ::IS1* affect lag time. Restoring P_{flhDC} in addition to *crl* had an additive effect on the lag time (Fig. 4d). The lag time of this CGSC 7740 *crl* P_{flhDC} strain was identical to the CGSC 6300 stock. Restoring only P_{flhDC} in CGSC 7740 showed an intermediate lag time 33% shorter than CGSC 6300. Besides, the maximum OD_{600} value reached by CGSC 6300 during growth on fumarate was also observed in a CGSC 7740 background as soon as the P_{flhDC} was restored (Fig. 4e), indicating *flhDC* expression may affect growth dynamics or yield.

Remarkably, restored *dgcJ* in the CGSC 7740 background (CGSC 7740 *dgcJ*) did not significantly change lag phase, but had an additive effect on lag time when restored in a CGSC 7740 P_{flhDC} background (CGSC 7740 P_{flhDC} *dgcJ*, Fig. 4d). Expression of *dgcJ* might increase in the CGSC 7740 P_{flhDC} background due to a shift in sigma factor balance in favor of *dgcJ*, whose expression is highest during exponential growth (Sommerfeldt et al., 2009).

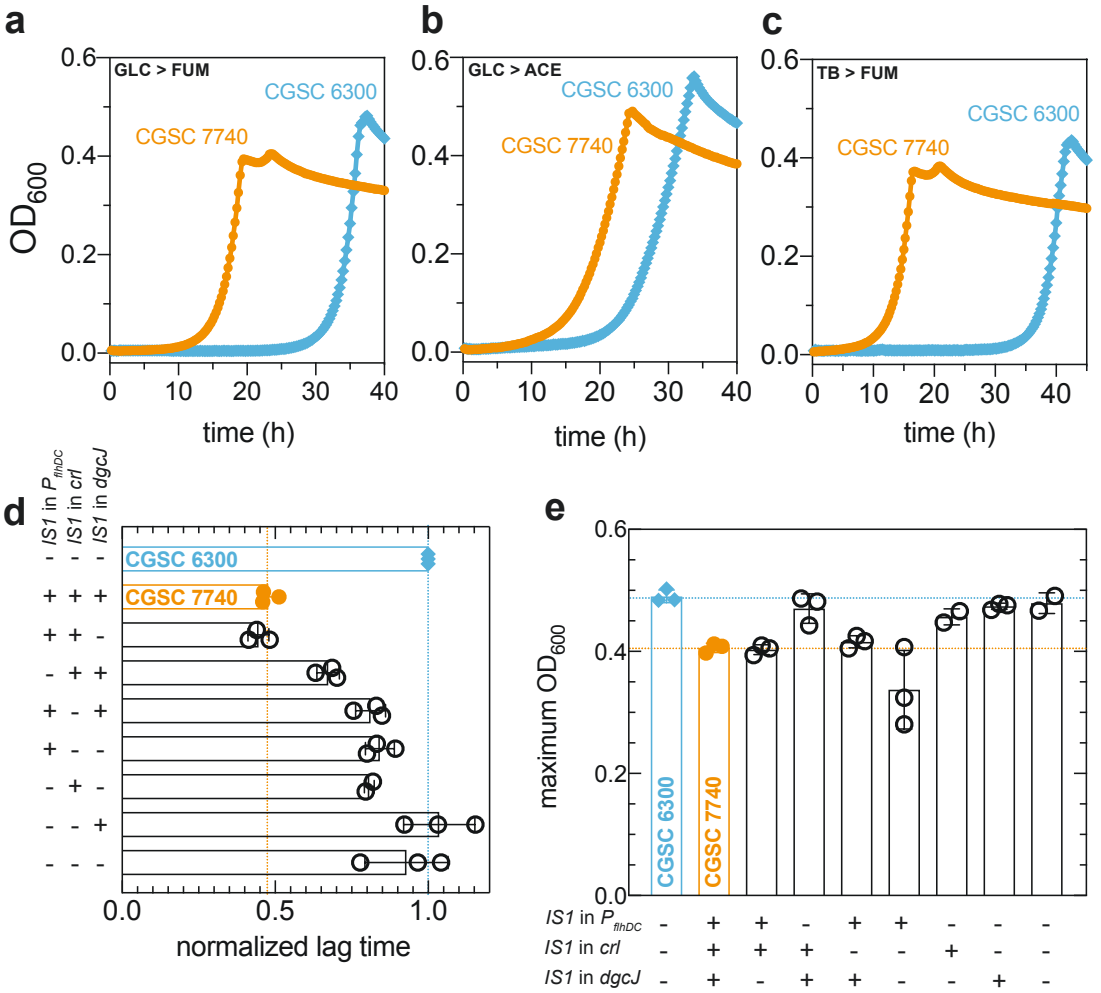


Figure 4: *IS1* elements affect growth after substrate switches. Growth, normalized lag time and maximum OD_{600} for CGSC 6300 (blue), CGSC 7740 (orange) and CGSC 7740 derivatives lacking one or several *IS1* elements. (a,b,c) Representative growth curves of CGSC 6300 and CGSC 7740 after a substrate switch from **(a)** glucose to fumarate, **(b)** glucose to acetate and **(c)** TB medium to fumarate as the sole carbon source. Incubation in medium with the new substrate started at time = 0. Each datapoint represents the average of at least three technical replicates. GLC, glucose; FUM, fumarate; ACE, acetate; TB, tryptone broth. **(d)** Normalized lag time after substrate switch from glucose to fumarate for CGSC 6300, CGSC 7740 and CGSC 7740 derivatives lacking one or several *IS1* elements. Growth curves were obtained as in **(a)**. Lag time was defined as the incubation time required to reach an OD_{600} of 0.05. Lag times were then normalized to average lag time of CGSC 6300. Each point represents a biological replicate consisting of at least three technical replicates. Orange line: average normalized lag time of CGSC 7740. Blue line: average normalized lag time of CGSC 6300. **(e)** Maximum OD_{600} reached by strains during growth on fumarate after a substrate switch from glucose to fumarate. Maximum OD_{600} was measured for CGSC 6300, CGSC 7740 and CGSC 7740 derivatives lacking one or several *IS1* elements. Orange line: average maximum OD_{600} of CGSC 7740. Blue line: average OD_{600} of CGSC 6300.

Could the difference in lag phase between CGSC 6300 and CGSC 7740 be the result of altered c-di-GMP levels during the pre-shift condition? The three *IS1* elements interfere with c-di-GMP regulation (Section 3.3.3) and the two MG1655 stocks show distinct cellular c-di-GMP levels on glucose (Fig. S1). For example, if the effect of P_{flhDC}*::IS1* on lag phase would be through c-di-GMP, a Δ*pdeH* mutant in a CGSC 7740 background would then prolong the lag phase similarly to restoring P_{flhDC} in CGSC 7740 (Fig. 4d). To test this hypothesis, we compared growth of wt with a Δ*pdeH* mutant, both in a CGSC 7740 background (Fig. S2). A Δ*pdeH* strain switched from glucose to fumarate resumed growth as early as its ancestor strain (Fig. S2a). Similarly, the lag time of a Δ*pdeH* mutant was not prolonged as compared to wild type after a glucose to acetate switch (Fig. S2b). We did however observe a difference in the maximum OD_{600} reached during growth on acetate (Fig. S2b). The affected culture density as observed is consistent with

previous observations that elevated c-di-GMP levels impair growth on acetate (Xu et al., 2019). Here, deacetylation of acetyl-coA synthetase (Acs), required for efficient acetate metabolism, was proposed to be reduced through a c-di-GMP-dependent reduction in deacetylase activity of CobB (Xu et al., 2019). By comparing wt to Δ*pdeH*, no c-di-GMP dependent effect on lag time was observed under the conditions tested. In conclusion, the *IS1* elements in *crl* and P_{flhDC} together can explain the difference in lag phase after substrate shift. As both Crl and the FlhDC have regulatory functions, we propose the differential expression of their downstream targets to drive the lag time difference.

3.3.5 Proteome analysis in rich medium and glucose minimal medium

We hypothesized that the difference in lag time of the two MG1655 stocks after substrate shift could be due to a difference in protein abundance of a set of (metabolic) proteins. Already during the pre-shift condition in glucose minimal medium or TB medium, a difference in protein expression would then set the conditions allowing for a faster growth resumption on the new substrate. To identify which proteins are differently expressed under pre-shift conditions, we performed comparative whole proteome analysis of cells of CGSC 7740 and CGSC 6300 grown in glucose minimal medium and TB medium (Fig. 5). To assess the impact of the *IS1* elements on protein levels, we additionally verified for all detected proteins whether they have been identified to be part of the FlhDC regulon or the FliA or RpoS sigmulon (Fig. 5) (Santos-Zavaleta et al., 2018).

Between the two growth conditions, we found a remarkable similarity in the list of upregulated (Table 1) as well as downregulated proteins (Table 2). In cells grown on glucose minimal medium, 2119 proteins were detected (Fig.5a) and a total of 2120 proteins were detected in TB

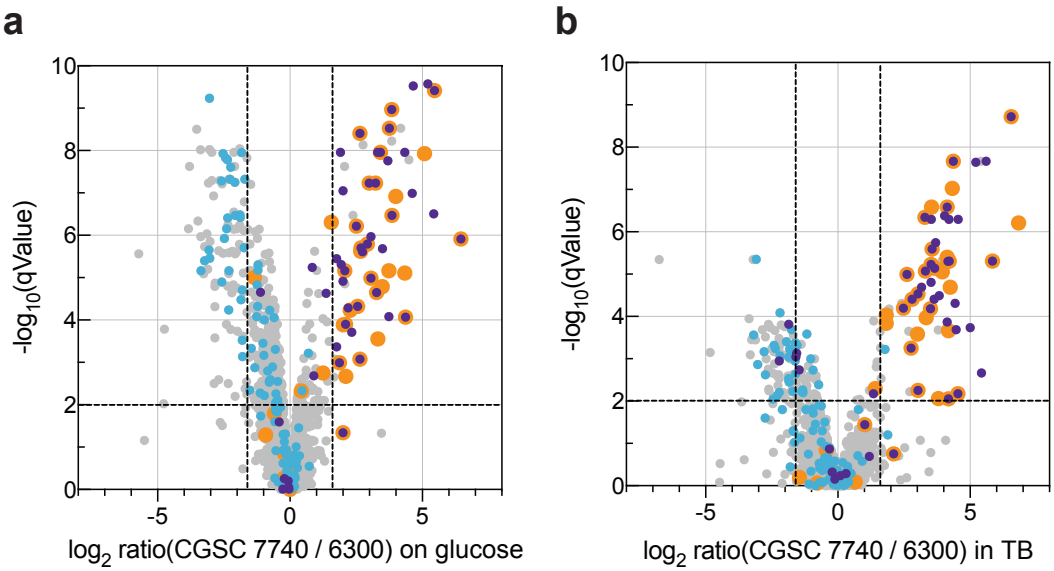


Figure 5: Comparative whole proteome analysis of CGSC 6300 and CGSC 7740 during growth on glucose and TB. Protein abundance in CGSC 7740 as compared to CGSC 6300, indicated as the log2 of the protein abundance ratio. Positive log2 values indicate a protein upregulation in CGSC 7740; proteins with a negative log2 values were downregulated in CGSC 7740. Significance of fold changes is indicated as -log(qValue), where qValue represents the adjusted p-Value. Dashed lines: three-fold change in relative protein abundance (vertical) and significance cut-off (horizontal). Highlighted proteins are known to require the RpoS sigma factor (blue), FliA sigma factor (purple) and/or FlhDC transcriptional regulator (orange) for their expression (Santos-Zavaleta et al., 2018). Proteome comparison during growth on **(a)** glucose minimal medium and **(b)** TB medium.

medium (Fig.5b). In both conditions, we found 59 significantly* upregulated proteins with a fold change of at least three (Table 1). 85% of these proteins are either part of the FlhDC regulon, the FliA sigmulon or both (Table 1). Their upregulation could therefore be explained by the P_{flhDC}*::IS1* genotype. The FliA protein itself was upregulated 14-fold and 10-fold on glucose and TB respectively. The FlhD and FlhC protein were not detected, so no assessment could be made about changes in their protein levels.

The c-di-GMP phosphodiesterase PdeH, expressed as a class III gene under FlhDC and FliA control (Ko and Park, 2000; Zhao et al., 2007), was upregulated eight-fold in glucose and 23-fold in TB (Table 1). The c-di-GMP phosphodiesterase PdeL, whose transcription is repressed at high c-di-GMP concentrations, was seven-fold upregulated in the CGSC 7740 background during growth on TB (Table 1), most likely as a consequence of PdeH activity. PdeL expression requires activation by metabolic regulator Cra (see chapter 2), a pleiotropic regulator with low activity under glycolytic conditions (Ramseier, 1996). This probably explains why the level of PdeL protein was not significantly upregulated during growth on glucose (Table 1). FliZ, expressed as a class II flagellar gene, is upregulated at least six-fold (Table 1). Given its inhibitory function on the transcription of RpoS-controlled genes (Pesavento and Hengge, 2012; Pesavento et al., 2008), elevated FliZ levels may downregulate RpoS activity in addition to the *crl* gene loss of function.

In glucose minimal medium, 93 proteins were significantly downregulated at least three-fold in CGSC 7740 compared to CGSC 6300 (Table 2). A total of 81 proteins were significantly downregulated at least three-fold during growth on TB medium (Table 2). In both conditions, expression of 30% of the downregulated proteins requires RpoS, according to the current overview of RpoS-controlled operons (Santos-Zavaleta et al., 2018). This indicated that 70% of strongly (at least three-fold) and significantly* downregulated proteins could not directly be linked to reduced RpoS activity in CGSC 7740. They could be newly identified members of the RpoS sigmulon, or proteins could be altered indirectly. Among the proteins with the highest fold change were Crl and the glycerol-3-phosphate repressor GlpR (Table 2). We explained the downregulation of Crl to be the effect of *crl*::*IS1*. The difference in GlpR level could be explained by a single nucleotide deletion present in the *glpR* gene in CGSC 7740, but not in CGSC 6300, leading to *glpR* gene loss of function (Freddolino et al., 2012). This in turn could explain the increase in protein levels of GlpA, GlpB, GlpC, GlpD, GlpF, GlpK and GlpT, all under negative GlpR control (Santos-Zavaleta et al., 2018) (Table 1). The downregulation in GatC and GatD levels are also due to additional genetic differences between the MG1655 stocks. The two proteins are required for galactitol uptake and utilization (Nobelmann and Lengeler, 1996). A frameshift in the *gatC* ORF in CGSC 7740 results in a truncated *gatC* product. Expression of *gatD* was suggested to be affected due to polar effects (Freddolino et al., 2012).

Increased activity of the metabolic regulator Cra prior to a substrate shift shortens lag phase on the new substrate (Kotte et al., 2014). Based on our proteome data, we tried to infer whether there is a difference in Cra activity between CGSC 7740 and CGSC 6300. This did not give a clear indication, as we found a downregulation of proteins under positive, as well as a downregulation of proteins under negative Cra control. Downregulation of proteins under positive Cra control included

*significant: -log10(qValue) > 2.0

AceA, AceB, AcnA, BetB and PoxB. Downregulation of proteins for which Cra functions as repressor included Epd, FbaA, FbaB, ManX, ManY, ManZ and TpiA (Santos-Zavaleta et al., 2018). FbaA and FbaB are aldolases in the upper glycolysis that catalyze the reversible conversion of Fructose-1,6-Bisphospate to glyceraldehyde 3-phosphate and dihydroxyacetone phosphate (DHAP) (Keseler et al., 2017). It is so far unclear how downregulation of these proteins affects metabolic flux distributions.

Table 1: Proteins upregulated in CGSC 7740 during growth on glucose minimal medium and TB. Protein description based on *E. coli* proteome from Uniprot. Regulon and sigmulon annotation based on (Santos-Zavaleta et al., 2018). n.s.: not significant.

Protein	Protein description	Fold change GLC	Fold change TB	FlhDC regulon	FliA/RpoS sigmulon
FliD	Flagellar hook-associated protein	87.9	57.3	FlhDC	FliA
FliC	Flagellin	43.9	93.9	FlhDC	FliA
YnjH	Uncharacterized protein	43.1	43.2		FliA
Tar	Methyl-accepting chemotaxis protein II	36.9	37.2		FliA
FlgH	Flagellar L-ring protein	34.0	20.1	FlhDC	
CheZ	Protein phosphatase	25.2	48.7		FliA
CheW	Chemotaxis protein	24.5	32.0		FliA
PdxK	Pyridoxine/pyridoxal/pyridoxamine kinase	22.5	42.3		
FliO	Flagellar protein	20.6	17.6	FlhDC	FliA
FlgD	Basal-body rod modification protein	20.2	114.0	FlhDC	
CheY	Chemotaxis protein	20.1	16.4		FliA
GlpD	Aerobic glycerol-3-phosphate dehydrogenase	18.2	3.0		
FlhA	Flagellar biosynthesis protein	16.0	17.6	FlhDC	
FliA	RNA polymerase sigma factor	14.5	9.8	FlhDC	FliA
YcgR	Flagellar brake protein	14.4	20.7	FlhDC	FliA
YjcZ	Uncharacterized protein YjcZ	14.4	15.2		
FlgM	Negative regulator of flagellin synthesis	13.5	11.5	FlhDC	FliA
CheR	Chemotaxis protein methyltransferase	13.3	17.5		FliA
FlgE	Flagellar hook protein	13.3	18.2	FlhDC	
MotA	Motility protein	13.0	11.5		FliA
CheA	Chemotaxis protein	11.2	22.2		FliA
FlgI	Flagellar P-ring protein	11.2	8.1	FlhDC	
FliF	Flagellar M-ring protein	10.7	11.8	FlhDC	FliA
FlgA	Flagella basal body P-ring formation protein	10.0	19.0	FlhDC	
Tap	Methyl-accepting chemotaxis protein IV	9.9	14.2		FliA
FlgN	Flagella synthesis protein	9.7	12.6	FlhDC	FliA
FliL	Flagellar protein	9.4	5.6	FlhDC	FliA
PdeH	Cyclic di-GMP phosphodiesterase	8.3	23.2	FlhDC	FliA
Tsr	Methyl-accepting chemotaxis protein I	8.3	23.4		FliA
FliM	Flagellar motor switch protein	7.9	18.5	FlhDC	FliA
FliH	Flagellar assembly protein	7.6	11.4	FlhDC	FliA
YeaH	UPF0229 protein	6.8	13.8		
FlxA	Protein	6.7	21.7		FliA
FlgF	Flagellar basal-body rod protein	6.4	2.7	FlhDC	
FliG	Flagellar motor switch protein	6.4	8.2	FlhDC	FliA
FliK	Flagellar hook-length control protein	6.2	18.1	FlhDC	FliA
FliZ	Regulator of sigma S factor	6.2	9.9	FlhDC	FliA
FliS	Flagellar secretion chaperone	5.9	n.s.	FlhDC	FliA
FliN	Flagellar motor switch protein	5.7	17.9	FlhDC	FliA
GlpF	Glycerol uptake facilitator protein	5.2	6.4		
ModC	Molybdenum import ATP-binding protein	5.0	12.5		FliA
YecR	Uncharacterized protein	4.8	15.4	FlhDC	
OppB	Oligopeptide transport system permease protein	4.6	-4.7		FliA
FlgG	Flagellar basal-body rod protein	4.3	10.1	FlhDC	
FliI	Flagellum-specific ATP synthase	4.3	6.8	FlhDC	FliA
ComR	HTH-type transcriptional repressor	4.2	n.s.		
FlgK	Flagellar hook-associated protein	4.2	7.0	FlhDC	FliA
GlpK	Glycerol kinase	4.2	3.8		
CheB	Chemotaxis response regulator protein-glutamate methylesterase	4.0	11.6		FliA

Table 1 (continued)

Protein	Protein description	Fold change GLC	Fold change TB	FlhDC regulon	FliA/RpoS sigmulon
FlgC	Flagellar basal-body rod protein	4.0	13.9	FlhDC	
MotB	Motility protein	4.0	9.0		FliA
Trg	Methyl-accepting chemotaxis protein III	3.8	18.4		FliA
FlgL	Flagellar hook-associated protein	3.7	6.1	FlhDC	FliA
OppD	Oligopeptide transport ATP-binding protein	3.7	-2.9		FliA
OppC	Oligopeptide transport system permease protein	3.4	-3.6		FliA
OppF	Oligopeptide transport ATP-binding protein	3.4	-2.8		FliA
PotB	Spermidine/putrescine transport system permease protein	3.3	n.s.		
YeeD	Putative sulfur carrier protein	3.2	n.s.		
GlpB	Anaerobic glycerol-3-phosphate dehydrogenase subunit B	3.0	3.6	FlhDC	
GlpA	Anaerobic glycerol-3-phosphate dehydrogenase subunit A	2.4	3.6	FlhDC	
GlpT	Glycerol-3-phosphate transporter	2.1	3.1		
YmgG	UPF0757 protein YmgG	2.0	3.1		
Aer	Aerotaxis receptor	1.8	13.0	FliA	
EmrA	Multidrug export protein	1.7	5.3		
GlpC	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	1.3	11.6	FlhDC	
FliE	Flagellar hook-basal body complex protein	n.s.	8.2	FlhDC	FliA
PdeL	Cyclic di-GMP phosphodiesterase	n.s.	6.6		
EnvC	Murein hydrolase activator	n.s.	3.4		RpoS

Table 2: Proteins downregulated in CGSC 7740 during growth on glucose minimal medium and TB. Protein description based on *E. coli* proteome from Uniprot. Regulon and sigmulon annotation based on (Santos-Zavaleta et al., 2018). n.s.: not significant.

Protein	Protein description	Fold change GLC	Fold change TB	FliA/RpoS regulon
Crl	Sigma factor-binding protein	-52.4	-108.0	
YbgS	Uncharacterized protein	-27.4	n.s.	
YodD	Uncharacterized protein	-26.8	-12.5	
GlpR	Glycerol-3-phosphate regulon repressor	-14.2	-7.2	
HdeA	Acid stress chaperone	-13.9	-4.7	
GadA	Glutamate decarboxylase alpha	-11.5	-4.8	
HdeB	Acid stress chaperone	-11.1	-5.4	
YebV	Uncharacterized protein	-10.4	-9.9	
GadB	Glutamate decarboxylase beta	-10.3	n.s.	
MdtE	Multidrug resistance protein	-10.3	-4.9	RpoS
HdeD	Protein	-9.6	n.s.	
YciF	Protein	-9.4	-5.4	RpoS
YebF	Protein	-9.3	-10.3	
YahO	Uncharacterized protein	-8.8	-4.7	
PsiF	Phosphate starvation-inducible protein	-8.6	n.s.	
Bfr	Bacterioferritin	-8.4	-6.7	
GlsA1	Glutaminase 1	-8.3	-4.2	
MsyB	Acidic protein	-8.2	-9.1	RpoS
SodC	Superoxide dismutase [Cu-Zn]	-8.2	-3.6	RpoS
YhbO	Protein/nucleic acid deglycase 2	-8.2	-6.1	
YjbJ	UPF0337 protein	-8.1	-5.6	
YdhS	Uncharacterized protein	-7.7	-3.1	
YgiW	Protein	-7.7	-9.2	
YcaC	Probable hydrolase	-7.4	-8.6	
YchH	Uncharacterized protein	-7.3	n.s.	
AidB	Putative acyl-CoA dehydrogenase	-7.2	n.s.	
GatC	PTS system galactitol-specific EIIC component	-7.1	-28.4	
YedP	Mannosyl-3-phosphoglycerate phosphatase	-7.1	-5.0	
YghA	Uncharacterized oxidoreductase	-7.1	-4.2	
PaaC	1,2-phenylacetyl-CoA epoxidase, subunit C	-6.4	n.s.	
YgaM	Uncharacterized protein	-6.4	-5.2	
Slp	Outer membrane protein	-6.3	-6.1	
PfkB	ATP-dependent 6-phosphofructokinase isozyme 2	-6.0	-8.5	RpoS
YdcS	Putative ABC transporter periplasmic-binding protein	-6.0	-6.9	RpoS
OsmY	Osmotically-inducible protein Y	-5.9	-4.3	
HchA	Protein/nucleic acid deglycase 1	-5.8	-3.6	RpoS
YniA	Putative kinase	-5.8	n.s.	
ElaB	Protein	-5.6	n.s.	RpoS
KatE	Catalase HPII	-5.4	-5.3	RpoS
YehZ	Glycine betaine-binding protein	-5.3	n.s.	
EcnB	Entericidin B	-5.2	-8.2	RpoS
FadE	Acyl-coenzyme A dehydrogenase	-5.1	n.s.	
GatD	Galactitol 1-phosphate 5-dehydrogenase	-5.1	-4.0	
Sra	Stationary-phase-induced ribosome-associated protein	-5.1	n.s.	RpoS
TktB	Transketolase 2	-5.1	-3.2	RpoS
FadM	Long-chain acyl-CoA thioesterase	-5.0	n.s.	
GadC	Probable glutamate/gamma-aminobutyrate antiporter	-5.0	-4.8	
OtsB	Trehalose-6-phosphate phosphatase	-5.0	n.s.	RpoS
UspC	Universal stress protein C	-5.0	n.s.	
YahK	Aldehyde reductase	-5.0	-3.2	
YhfG	Uncharacterized protein	-5.0	-3.6	RpoS
GabD	Succinate-semialdehyde dehydrogenase [NADP(+)]	-4.8	-3.1	RpoS
Tam	Trans-aconitate 2-methyltransferase	-4.7	n.s.	RpoS
YcgJ	Uncharacterized protein	-4.7	n.s.	
YhhA	Uncharacterized protein	-4.6	n.s.	
ChaB	Putative cation transport regulator	-4.5	-4.4	
YgdI	Uncharacterized lipoprotein	-4.5	-2.8	
YqjD	Uncharacterized protein	-4.4	-3.4	
OsmE	Osmotically-inducible putative lipoprotein	-4.3	n.s.	

Table 2 (continued)

Protein	Protein description	Fold change GLC	Fold change TB	FliA/RpoS regulon
PatA	Putrescine aminotransferase	-4.3	n.s.	
YdiZ	Uncharacterized protein	-4.3	-6.2	
YeaG	Uncharacterized protein	-4.3	-6.2	
FbaB	Fructose-bisphosphate aldolase class 1	-4.2	-5.1	RpoS
YbbJ	Inner membrane protein	-4.2	n.s.	
YccJ	Uncharacterized protein	-4.2	-4.6	RpoS
YnhG	Probable L,D-transpeptidase	-4.2	-2.3	
Ahr	Aldehyde reductase	-4.1	-3.7	
MscS	Small-conductance mechanosensitive channel	-4.0	n.s.	
YmdF	Uncharacterized protein	-4.0	n.s.	
FimA	Type-1 fimbrial protein, A chain	-3.9	n.s.	
AmyA	Cytoplasmic alpha-amylase	-3.7	-4.7	
TalA	Transaldolase A	-3.7	-3.5	RpoS
TreA	Periplasmic trehalase	-3.7	-6.7	RpoS
YbaY	Uncharacterized lipoprotein	-3.7	-4.3	
YegP	UPF0339 protein	-3.7	-2.9	
AldB	Aldehyde dehydrogenase B	-3.6	-4.6	RpoS
OtsA	Trehalose-6-phosphate synthase	-3.6	-3.9	RpoS
Blc	Outer membrane lipoprotein	-3.5	n.s.	RpoS
PuuB	Gamma-glutamylputrescine oxidoreductase	-3.5	n.s.	RpoS
WrbA	NAD(P)H dehydrogenase (quinone)	-3.5	-4.5	RpoS
FadH	2,4-dienoyl-CoA reductase	-3.4	n.s.	
YbdK	Putative glutamate--cysteine ligase 2	-3.4	-3.0	
YsgA	Putative carboxymethylenebutenolidase	-3.4	-3.0	
AstC	Succinylornithine transaminase	-3.3	-3.0	RpoS
Dps	DNA protection during starvation protein	-3.3	-3.3	
GabT	4-aminobutyrate aminotransferase	-3.3	-3.6	RpoS
PoxB	Pyruvate dehydrogenase	-3.3	-5.9	
XylF	D-xylose-binding periplasmic protein	-3.2	n.s.	
PdxI	Pyridoxine 4-dehydrogenase	-3.1	n.s.	
YjgR	Uncharacterized protein	-3.1	-2.3	
YqjC	Protein	-3.1	n.s.	
YqjE	Inner membrane protein	-3.1	n.s.	
BtuE	Thioredoxin/glutathione peroxidase	-3.0	-2.0	
YgaU	Uncharacterized protein	-2.9	-3.6	
AstE	Succinylglutamate desuccinylase	-2.8	-3.4	RpoS
YqjG	Glutathionyl-hydroquinone reductase	-2.8	-4.4	
YbhB	UPF0098 protein	-2.5	-4.0	
YegS	Lipid kinase	-2.5	-4.6	
AstD	N-succinylglutamate 5-semialdehyde dehydrogenase	-2.4	-3.0	RpoS
PuuE	4-aminobutyrate aminotransferase	-2.3	-3.2	RpoS
YaiA	Uncharacterized protein	-2.3	-3.5	
YibT	Uncharacterized protein	-2.3	-3.1	
YohF	Uncharacterized oxidoreductase	-2.1	-5.2	
BcsG	Cellulose biosynthesis protein	-2.0	-3.1	
DgcM	Diguanylate cyclase	-1.7	-5.6	
GatB	PTS system galactitol-specific EIIB component	-1.7	-3.0	
RssB	Regulator of RpoS	-1.7	-3.4	RpoS
YedF	Putative sulfur carrier protein	-1.5	-3.6	
OppA	Periplasmic oligopeptide-binding protein	2.6	-3.0	FliA
OppC	Oligopeptide transport system permease	3.4	-3.6	FliA
OppB	Oligopeptide transport system permease protein	4.6	-4.7	FliA
YdcH	Uncharacterized protein	n.s.	-9.6	
Prr	Gamma-aminobutyraldehyde dehydrogenase	n.s.	-6.5	
YedE	UPF0394 inner membrane protein	n.s.	-5.4	
YdcJ	Uncharacterized protein	n.s.	-4.7	
Ivy	Inhibitor of vertebrate lysozyme	n.s.	-3.3	
DppC	Dipeptide transport system permease protein	n.s.	-3.0	

3.4 Discussion

Bacteria are highly adaptive by nature. Their short division time combined with genome plasticity facilitates the acquisition of mutations. Mutations that improve adaptation or survival are enriched under the prevailing conditions including laboratory conditions (Naas et al., 1994; Parker et al., 2019). Unintended evolution of *E. coli* strains should however be avoided for synthetic biology applications and to secure the required experimental reproducibility (Umenhoffer et al., 2010).

We reported about phenotypic differences in growth and signaling between two laboratory stocks of the commonly used *E. coli* K 12 MG1655 strain, that arise from genetic differences in the form of *IS1* elements. The *E. coli* K-12 MG1655 genome contains 45 *IS* elements of different kinds (Blattner et al., 1997). MG1655 stock CGSC 7740 has at least three additional *IS1* elements as compared to stock CGSC 6300. These *IS1* elements locate in the *flhDC* promoter region and the *crl* and *dgcJ* open reading frame (Barker et al., 2004; Freddolino et al., 2012). The P_{flhDC}*::IS1* genotype in CGSC 7740 resulted in constitutively low levels of the second messenger c-di-GMP, via activation of the c-di-GMP specific phosphodiesterase PdeH (Fig. 3, Fig. 6). Furthermore our data showed that P_{flhDC}*::IS1* and *crl*::*IS1* drive a growth phenotype of reduced lag time upon substrate shift (Fig. 4, Fig. 6). A whole proteome comparison of the two MG1655 stocks showed that the levels of several hundreds of proteins were significantly altered. The proteome comparison provides new insight in how differently these two MG1655 backgrounds can be under standard laboratory conditions and is a first step in identifying the key players responsible for reducing lag time in MG1655 stock CGSC 7740.

3.4.1 Minor impact of *dgcJ* on motility and cytosolic c-di-GMP levels

In addition to the previously identified *IS1* elements in P_{flhDC} (Barker et al., 2004; Blattner et al., 1997) and the *crl* gene (Freddolino et al., 2012), we described a third *IS1* element, located in the *dgcJ* gene, that is present in CGSC 7740 but absent in CGSC 6300. The genotype of stock CGSC 7740 should therefore be P_{flhDC}*::IS1 crl*::*IS1 dgcJ::IS1*. In contrast to P_{flhDC}*::IS1* and *crl*::*IS1*, the *dgcJ* gene is misannotated in the current online reference genome of *E. coli* K-12 MG1655, as RefSeq accession no. NC_000913.3 currently contains the intact *dgcJ* and not *dgcJ::IS1*. Because the *dgcJ* gene encodes a putative diguanylate cyclase (Hengge et al., 2016; Pesavento et al., 2008) and *E. coli* cells adjust their swimming velocity in a c-di-GMP concentration-dependent manner (Boehm et al., 2010), we investigated the impact of DgcJ on the motility phenotype of CGSC 7740. DgcJ was previously reported to reduce motility in K-12 W3110 in assays performed at 37°C, indicating that c-di-GMP synthesis by DgcJ contributes to a common c-di-GMP pool (Pesavento et al., 2008). In our study, restoring the *dgcJ* gene in background MG1655 CGSC 7740 did not affect motility at 37°C (Fig. 2b). There are however differences between the motility assay in this study and one previously described (Pesavento et al., 2008). In the set-up of Pesavento et al. (2008), the highly motile strain K-12 W3110 (with genotype P_{flhDC}*::IS5* (Barker et al., 2004)) was rendered non-motile through deletion of *pdeH*. Subsequently, it was assessed which diguanylate cyclases affect motility by deletion of individual or combinations of diguanylate cyclases.

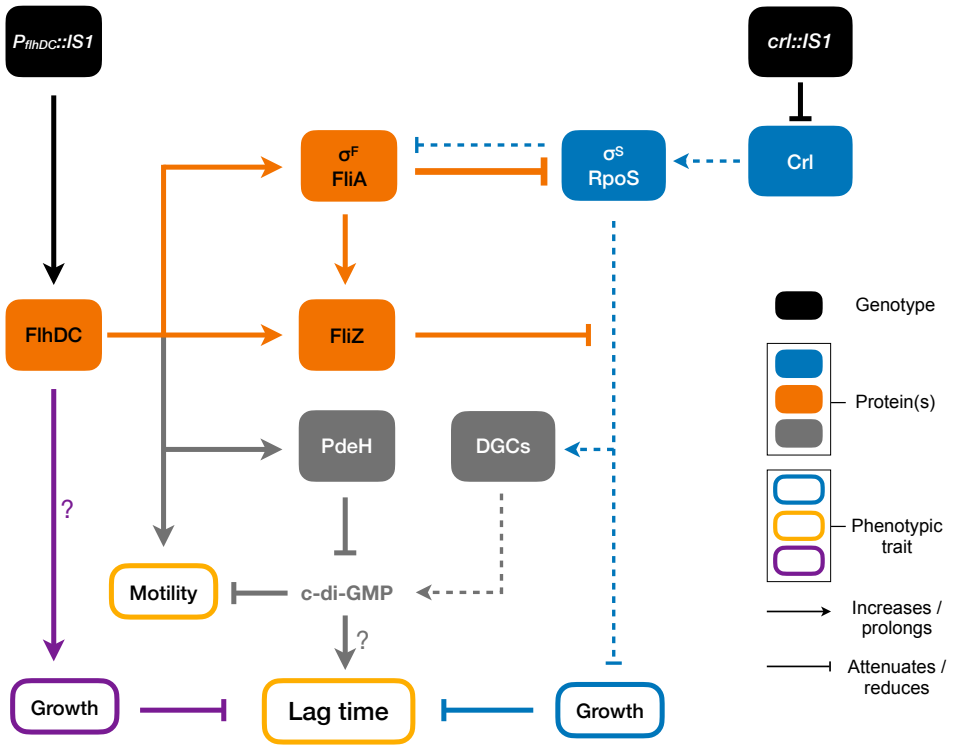


Figure 6: Proposed working model of c-di-GMP signaling and growth adaptation mediated by genetic variations in the occurrence of *IS1* elements. Overview of the genetic variations and the phenotypes they affect as investigated in this study. Solid lines: the regulatory effect is amplified in CSGC 7740 background. Dashed lines: the regulatory effect is attenuated in the CGSC 7740 background. The *IS1* element in the *flhDC* intergenic region (P_{flhDC}*::IS1*) leads to constitutive *flhDC* expression in CGSC 7740. Through constitutive activation of the flagellar genes downstream of FlhDC, CGSC 7740 is highly motile. The *IS1* element in the *crl* open reading frame is considered a *crl* loss of function mutation. For the reduced lag time of CGSC 7740 after sudden substrate switches, several mechanisms are proposed. Non-flagellar, downstream targets of FlhDC could stimulate growth and reduce lag time after a sudden substrate (purple). Alternatively, the contribution of P_{flhDC}*::IS1* to reduced lag time could primarily be by increasing levels of FliA and FliZ (orange). Together with reduced Crl activity, FliA and FliZ would attenuate RpoS activity by an increase in alternative sigma factor abundance (FliA) and inhibiting RpoS-controlled expression (FliZ). Attenuated RpoS activity would improve growth and shorten lag time after sudden substrate switch (blue). RpoS and FlhDC/FliA also stimulate expression of, respectively, diguanylate cyclases and c-di-GMP specific phosphodiesterase (grey). C-di-GMP levels are reduced in CGSC 7740. We could however not provide evidence in this study that c-di-GMP levels affect the lag time after substrate switches.

Pesavento et al. (2008) described that deletion of *dgcJ* exhibited most pronounced suppression of the W3110 Δ*pdeH* motility phenotype. In our study, we verified a possible reduction in motility by restoring *dgcJ* in a MG1655 CGSC 7740 background with *pdeH* still expressed. A possible increase in c-di-GMP synthesis by restoring *dgcJ* in CGSC 7740 could have been masked by constitutive *pdeH* expression. Therefore, the effect of *dgcJ* on motility should be assessed in a CGSC 7740 Δ*pdeH* mutant at 37°C, before drawing further conclusions. It is furthermore worth mentioning that the detection of the *dgcJ::IS1* genotype in CGSC 7740 explains earlier findings about the role of diguanylate cyclases in motility. Boehm et al. (2010) defined a set of four diguanylate cyclases in K 12 MG1655 that adjust swimming velocity through their activity. In contrast to the findings in K-12 W3110 by Pesavento et al. (2008), DgcJ was not among these four DGCs. Boehm et al. used the highly motile MG1655 strain (Blattner et al., 1997), that must have had genotype P_{flhDC}*::IS1 crl::IS1 dgcJ::IS1*. Non-functional DgcJ would explain why a *dgcJ* deletion mutant does not alter c-di-GMP levels.

Besides motility, we assessed the cytosolic c-di-GMP concentrations on the single-cell level with a c-di-GMP sensor. We choose M9 glycerol as the growth condition, as CGSC 6300 displays a bimodal c-di-GMP distribution in this condition (Chapter 2 and Fig. 3). We observed no difference between a *dgcJ* or *dgcJ::IS1* genotype (Fig. 3). These observations raise questions about the specificity of *dgcJ* expression and activity, as well as the conditions that would select for a *dgcJ::IS1* genotype. As we have not specifically looked at *dgcJ* expression in M9 glycerol, one possibility could be that required signals for *dgcJ* transcription are missing under the conditions used in this study. Besides, all *E. coli* K-12 diguanylate cyclases contain one or several sensory domains, that might regulate catalytic activity (Blum et al., 2020; Hengge et al., 2016). In the case of DgcJ, specific cues for activation via its GAPES1 sensor domain (Fig. 1) could be missing. Indeed, we recently found that DgcJ requires a compound for its activation that is present in complex medium (Luria-Bertani medium (LB), which contains tryptone and yeast extract). The exact cue for DgcJ activation is so far unknown, but this could explain why *dgcJ* does not alter cytosolic c-di-GMP distributions when assessed on M9 based buffer supplemented with glycerol (Fig. 3). Apart from the DgcJ activating mechanism, another remaining question is what conditions selected the *dgcJ::IS1* genotype. Out of the 25 c-di-GMP regulating enzymes in *E. coli* K-12 (Hengge et al., 2016), what would be the advantage of disrupting *dgcJ* specifically? The field of bacterial signaling is postulating that PDEs and DGCs could be involved in local c-di-GMP signaling. Criteria for a diguanylate cyclase to be part of a local signaling module include a specific phenotype and direct interactions with a target system (Hengge, 2021). It was proposed that a specific diguanylate cyclase, c-di-GMP phosphodiesterase and c-di-GMP effector together can form a module through their close proximity in space (Hengge, 2021; Richter et al., 2020). DgcJ, with two predicted transmembrane domains, periplasmic sensory domain and a cytosolic catalytic domain (Fig.1) could thus be precisely localized in the cell. In this case, DgcJ would affect very specific c-di-GMP-controlled processes and affect a local c-di-GMP pool. What could then be the local c-di-GMP effector? What cellular processes could *dgcJ* be regulating as such that an MG1655 population is enriched for a disrupted *dgcJ* gene, *dgcJ::IS1*? C-di-GMP and in particular *dgcJ* expression was only recently demonstrated to correlate with bacteriophage resistance (Mutalik et al., 2020). Deletion of specifically *dgcJ* resulted in resistance against the N4 bacteriophage (Mutalik et al., 2020). Indeed, we recently observed that the bacteriophage N4 infected *E. coli* at high c-di-GMP levels when established through catalytic activity of DgcJ in complex medium. *E. coli* cells carrying a nonfunctional *dgcJ*, e.g. *dgcJ::IS1*, were resistant against N4 . We observed that a second specific DGC, DgcQ, activated by extracellular arginine, contributed to c-di-GMP-dependent N4 infection. The precise target system that mediates N4 phage susceptibility through DgcJ and DgcQ activity remains yet to be elucidated.

3.4.2 The *flhDC* upstream region is a mutation hotspot

We observed that from the three *IS1* elements analyzed in this study, only P_{flhDC}*::IS1* affected motility at 37°C (Fig. 2b). Increased motility through an *IS* element insertion in the *flhDC* regulatory region is not unique for *E. coli* K-12 MG1655 CGSC 7740. The K-12 derivatives MC1000, RP437,

W3110 and YK410 carry an *IS5* element in the *flhDC* intergenic region of their genomes and these strains were also reported to be highly motile (Barker et al., 2004). Moreover, insertions as found in these K-12 stocks are highly reproducible in conditions allowing flagellar rotation. Multiple studies have reported on spontaneous mutations occurring upstream of the *flhDC* operon in static liquid cultures or upon overnight incubation on motility agar, rendering this region a mutation hotspot (Barker et al., 2004; Parker et al., 2019; Wang and Wood, 2011; Zhang et al., 2017). This could among others suggest that the MG1655 ancestral stock CGSC 6300 has been exposed to those conditions that select and enrich for constitutive *flhDC* expression.

Our data showed that constitutive *flhDC* expression allowed more rapid adaptation to new substrates. This was not limited to a switch from a glycolytic to a gluconeogenic substrate, but was also observed when we switched from TB medium to fumarate as the sole carbon source (Fig. 4). From the many downstream targets of FlhDC and FlhD, which genes would drive the observed reduction in lag time? It was already described that in addition to flagellar genes, the products of the *flhDC* operon activate metabolic genes involved in anaerobic respiration and the Entner Doudoroff pathway (Prüss et al., 2003). However, it was not previously shown that activation of *flhDC* results in a shorter lag time under dynamic substrate conditions (Fig. 4). Remarkably, FlhD also has a regulatory role in cell division (Prüss and Matsumura, 1996). During growth on tryptone broth, a Δ*flhD* mutant reaches higher cell densities than a K-12 YK410 wt strain (with genotype P_{flhDC}*::IS5*) (Prüss and Matsumura, 1996). Since *flhDC* expression in CGSC 6300 is low (Kim et al., 2020), it could well mimic a Δ*flhD* mutant. A Δ*cadA* mutant strain has a similar growth phenotype as a Δ*flhD* mutant. Expression of *cadA* is stimulated by FlhD, although not via direct transcriptional activation (Prüss et al., 1997). The *cad* operon is under the control of multiple regulators and CadA activity is inhibited by ppGpp (Kanjee et al., 2011; Santos-Zavaleta et al., 2018). FlhD-stimulated *cadA* expression could explain the observed difference in maximum optical cell density (Fig.4). Additional experiments in a P_{flhDC}*::IS1* Δ*cadA* or P_{flhDC}*::IS1* Δ*flhD* background should be performed to provide more insight in their potential contribution to lag time.

An alternative mechanism to consider for FlhDC-dependent lag time reduction involves the flagella-associated sigma factor FliA and the RpoS inhibiting factor FliZ (Pesavento and Hengge, 2012) (Fig. 6). Both FliA and FliZ are directly activated by FlhDC (Santos-Zavaleta et al., 2018). Sigma factors, including FliA and RpoS, compete for limited availability of RNA Polymerase (RNAP) (Pesavento et al., 2008; Typas et al., 2007). Cellular transitions are driven by a shift in sigma factor activity (Pesavento et al., 2008) and the sigma factor balance as such represents a trade-off between growth and stress resilience (Ferenci, 2005; Hengge, 2021; Nystrom, 2004). Attenuated RpoS activity is widely associated with a growth advantage on alternative substrates (King et al., 2004; Nystrom, 2004). We postulate that nutrient limiting conditions as described above select for *flhDC*-activating mutations, because it attenuates RpoS activity via FliA and FliZ. In addition to loss of function of *crl*, increased amounts of alternative sigma factor (FliA) and RpoS inhibiting factor (FliZ, Pesavento and Hengge, 2012) would further reduce RpoS activity in CGSC 7740 (Fig. 6).

Overall, the adaptation advantage of the P_{flhDC}*::IS1* genotype is sufficiently large to outweigh the energetically costly process of flagellar motility, which can account for up to several percent of a cell's energy expenses and protein synthesis (Milo et al., 2010). In the studies that observed increased *flhDC* expression through spontaneous mutation, it remains unclear whether it is primarily the motility phenotype or this metabolic phenotype that provides a benefit to otherwise natively low *flhDC* expressing cells.

3.4.3 Mechanism behind lag time differences after substrate shift

Cells display a lag phase when shifting between substrates. A diauxic shift is observed, when a preferred substrate in a mixture is depleted and adaptation to catabolism of the remaining substrate is required (Chu and Barnes, 2016; Monod, 1949). A sudden switch, in which the prevailing substrate is instantly depleted and replaced by an other, less preferred substrate, is often accompanied by a substantial lag phase (Basan et al., 2020; Kotte et al., 2014). The underlying mechanisms of this phenomenon are extensively studied in *E. coli* and other bacterial species (Barthe et al., 2020; Basan et al., 2020; Kotte et al., 2014; Nikel et al., 2015). Based on observations that the lag time on the new substrate is shortened when the growth rate on the pre-shift substrate is reduced, a growth adaptation trade-off was proposed (Basan et al., 2020). In our study, the growth rate of CGSC 6300 and CGSC 7740 on glucose was comparable. A difference in pre-shift growth rate to possibly explain lag time did therefore not exist.

Alternatively, increased expression of the enzymes required for specific uptake of the new substrate can shorten lag phase. This was shown for substrate switches to fumarate, glycerol, xylose and lactose (Barthe et al., 2020; Kaiser et al., 2018; Kotte et al., 2014; Nikel et al., 2015). Increased expression of gluconeogenic enzymes under glycolytic conditions was also found to shorten lag phase (Basan et al., 2020). Fore example, lag time after a shift from a glycolytic source to acetate was shortened when expression of the *aceAB* operon, encoding enzymes in the glyoxylate shunt, was induced during the pre-shift glycolytic growth phase (Basan et al., 2020). However, in our proteome analysis, *aceA* and *aceB* were not significantly upregulated in CGSC 7740, and even slightly downregulated on glucose. Further analysis of the proteomes of the two MG1655 stocks would be required for quantitative conclusions about enzyme abundance in central metabolic pathways. Comparing metabolic flux distributions of CGSC 6300 and CGSC 7740 during pre-shift conditions is also recommended. This could present differences in carbon flux that possibly underlay the lag time difference.

Most studies on lag time so far can be seen as bottom-up approaches, where genetic (Barthe et al., 2020; Kotte et al., 2014) or epigenetic (Basan et al., 2020) changes were applied to a reference strain after which the effect on lag time was assessed. In contrast, the observed difference in lag time between MG1655 stocks CGSC 6300 and CGSC 7740 is a top-down case. It is certainly recommended to further investigate the underlying mechanism, as it would increase our knowledge of how bacteria naturally rewire their networks to shorten lag time and improve nutrient adaptation. The difference in lag time may be a matter of global resource allocation

or signaling, or lay in the levels of only a few key enzymes. It also remains to be elucidated whether the observed population lag arises from a difference in lag time of individual cells, or a difference in the fraction of cells resuming growth. Therefore, additional single cell observations are recommended during nutrient shift experiments.

3.4.4 Strain choice can largely impact experimental outcome

Our results indicate that the choice of *E. coli* K-12 stock can largely impact experimental outcome. It is unknown how often the choice for the stock is either arbitrary or well considered. Studies requiring a specific phenotype for the reference strain make the choice more deliberate. For example, in studies on *E. coli* motility a highly motile reference strain was chosen (Boehm et al., 2010; Pesavento et al., 2008; Reinders et al., 2016). In the work on flagellar gene expression in CGSC 6300 by Kim et al. (2020), the authors explicitly discussed how their results would have looked in the other MG1655 background, CGSC 7740 (Kim et al., 2020). These discussions contribute to our understanding, as they provide insights into how certain *E. coli* stocks, in this case the K-12 MG1655 CGSC 7740, will display different expression patterns (Kim et al., 2020). We would advocate describing bacterial strains as precisely as possible, including besides strain also the stock and source. This will allow more precise comparison between studies. In the case of the *E. coli* MG1655 stocks CGSC 6300 and CGSC 7740, we advocate increasing awareness about their genetic differences in the occurrence of *IS1* elements and single nucleotide insertions and deletions (Barker et al., 2004; Freddolino et al. 2012). Because *IS1* elements can easily be detected based on DNA fragment lengths of PCR products (Barker et al., 2004; Tamar et al., 2016), one could easily determine which *E. coli* MG1655 stock is used the lab.

3.5 Materials and Methods

Strains and plasmids

An overview of the strains used in this study is provided in Table S1 and plasmids are listed in Table S2. *E. coli* K-12 MG1655 stocks CGSC 6300 and CGSC 7740 were requested from the Coli Genetic Stock Center (CGSC) (Yale University, New Haven, CT, USA). MG1655 reference strain 'c1' (Fig. 1a) is the MG1655 reference strain in the lab of U. Jenal and directly originates from the laboratory of F. Blattner (Blattner et al., 1997). MG1655 reference strain 'c2' (AB3746, see Table S1) (Fig. 1a) was kindly shared by the lab of M. Heinemann (University of Groningen, Groningen, The Netherlands). The c-di-GMP sensor plasmids were designed and provided by A. Kaczmarczyk from the lab of U. Jenal.

Chromosomal engineering

Removal of single *IS1* elements from CGSC 7740 was performed in parallel, generating strains CGSC 7740 *dgcJ* (AB3864), CGSC 7740 P_{flhDC} (AB3870) and CGSC 7740 *crl* (AB3882). *IS1* elements were removed from strain CGSC 7740 via two-step scarless λ-Red recombineering (Datsenko and Wanner, 2000). Each *IS1* element was first replaced by a sequence containing a kanamycin resistance cassette and a rhamnose-inducible toxin encoding *ccdB* gene (kan-Prha-ccdB). To restore the *IS1*-affected chromosomal region, the kan-Prha-ccdB sequence was subsequently replaced by the required sequence amplified from the CGSC 6300.

The insert kan-Prha-ccdB was PCR amplified from plasmid pKD45 with 50-bp overhang of recombineering sites upstream and downstream of each *IS1* element, using primers listed in Table S3 (_fwd1, _rev1). A genomic DNA preparation of strain CGSC 6300 (GenElute Bacterial Genomic DNA kit (Sigma-Aldrich)) was used as template to generate linear DNA fragments of P_{flhDC}, *crl* and *dgcJ* with sufficiently long overhangs for recombineering (primers _fwd2, _rev2). Correct intermediate and final constructs were checked based on PCR fragment length and sequencing using primers '_seq' primers.

Strain CGSC 7740 was transformed with plasmid pKD46 and λ-Red recombineering was performed as described previously (Datsenko and Wanner, 2000). In short, a single colony of pKD46-bearing CGSC 7740 was inoculated in 5 ml LB supplemented with ampicillin and incubated overnight at 30°C. The grown culture was diluted 1:100 the next day in 50ml LB without NaCl supplemented with 100 μg/ml ampicillin and incubated in a pre-warmed incubator (Multitron, INFORS HT) at 30°C and 160-200 rpm. At an OD_{600} of 0.3-0.4, induction of λ-Red genes was initiated by addition of 600 μl of 0.2% L-arabinose and incubation was continued until OD_{600} 0.6-0.8. The culture was kept on ice for at least 30 minutes and washed four times in 20 ml 10 % ice cold glycerol. Linear DNA fragments were inserted via electroporation. After transformation with a linear kan-Prha-ccdB fragment with overhangs, cell suspensions were diluted in LB, incubated

to allow phenotypic expression, plated on LB agar supplemented with kanamycin and incubated at 30°C. Selection for scarless integration of linear fragments P_{flhDC}, *crl* and *dgcJ* was performed by plating on MMA rhamnose plates (1x MMA, 1mM MgSO4, 0.2% Rhamnose and 1.4% agarose) and incubation at 30°C. Strains were cured from plasmid pKD46 by growth at 42°C.

For the removal of a second and third *IS1* element from CGSC 7740, P1 transduction was used. P1 phage lysate preparation and transduction were carried out as described in (Miller, 1992). In short, P1 lysates of the intermediate strains CGSC 7740 *dgcJ*::kan-Prha-ccdB (AB3863), CGSC 7740 P_{flhDC}::kan-Prha-ccdB (AB3865) and CGSC 7740 *crl*::kan-Prha-ccdB (AB3882) were prepared. CGSC 7740 derivatives were transduced with P1 lysate and positive clones were selected on LB agar supplemented with kanamycin and sodium citrate. Correct insertion of the kan-Prha-ccdB fragment was checked based on band length of PCR amplification over the region. Then, the second step of scarless recombineering was repeated as described above, replacing the kan-Prha-ccdB fragment.

Growth media

LB (Luria-Bertani) medium was used for growth of pre-cultures and for strain construction, unless stated otherwise. LB Agar plates consisted of LB medium supplemented with 1.5% agar.

TB (Tryptone Broth) medium consisted of 10 g/l tryptone (Bacto) and 5 g/l NaCl. Motility plates consisted of TB medium supplemented with 0.3 % agar (Agar Bacteriology Grade – PanReac Applichem). Minimal medium consisted of M9-based minimal medium (Gerosa et al., 2013). Minimal medium was supplemented with sterile concentrated stock solutions of carbon sources glucose (10%), glycerol (10%), or sodium fumarate (4%) and used at final concentrations of 0.5% glucose, 0.4% glycerol and 0.2% fumarate minimal medium.

Cultivation conditions

M9-based minimal medium was always inoculated with liquid pre-cultures grown in LB medium. Single colonies from LB Agar plates were used to inoculate 5 ml aliquots of LB medium. LB cultures were grown for 6-7 h at 37°C.

For growth in glucose minimal medium, grown LB cultures were diluted directly 1:200 in M9-based minimal medium supplemented with 0.5% glucose. Glucose cultures were incubated overnight at 37°C and again diluted 1:500 in fresh medium and grown for 6-7h to obtain exponentially growing cultures at an OD_{600} 0.3-0.5.

For growth in glycerol minimal medium, cells grown in LB were pelleted and resuspended in M9 medium without carbon source. Cell suspensions were diluted to an OD_{600} of 0.01 or lower in 5 ml M9 medium supplemented with 0.4% glycerol. M9 cultures were incubated at 37°C with shaking at 180 rpm (Multitron, INFORS HT), for at least 17 h to a final cell density between OD_{600} 0.1-0.4 for all cultures.

TB medium was inoculated with single colonies from LB Agar plates and cultures were incubated overnight at 37°C. Overnight cultures were diluted 1:500 in fresh TB medium and grown for 4 h to an OD_{600} of 0.4-0.5.

PCR amplification and sequencing

All primers were ordered from Sigma Aldrich unless stated otherwise. For primer sequences, see Table S3. The genomic region covering the *flhDC* operon and *flhDC* intergenic region was amplified using primers 12519 and 12521; the *crl* open reading frame was amplified using primers 13067 and 13068; the *dgcJ* open reading frame was amplified using primers 12553 and 12554. After clean-up of PCR products (PCR clean-up kit, Macherey-Nagel), 10 μl of product together with 2 μl of 10 mM primer was sequenced using Sanger sequencing (Microsynth AG, Balgach, Switzerland). To translate the DNA sequence of *dgcJ::IS1* into protein sequence, the SIB Expasy translate tool was used (web.expasy.org/translate/ (Artimo et al., 2012)).

DgcJ protein domain localization

The predicted DgcJ protein domain localization (Fig. 1b) was visualized using Protter software (https://bit.ly/3qto3sg (Omasits et al., 2014)).

Motility assays

Motility assays were performed similar to Parker et al. (2019). Motility plates (TB + 0.3 % agar) were poured one day before the motility assay. Single colonies from LB Agar plates were inoculated in LB and incubated at 37 °C the night before the motility assay. The next morning, all cultures were diluted 1:100 in 5 ml fresh LB and grown into exponential phase at 37 °C for 1.5 h until an OD_{600} of 0.3. Pre-warmed motility plates were inoculated with 3 μl of cell culture grown to an OD_{600} of 0.3. Plates were incubated at 37 °C for 6 h in a closed box that also contained a beaker with demi-water, to keep the plates moisture and dark. Pictures of motility plates were taken on a custom-made stage that illuminated the plates from below.

Flow cytometry

Flow cytometry of M9 minimal medium cultures started with day cultures in LB supplemented with 100 μg/ml ampicillin, which were inoculated from single colonies and grown for 6-7 h at 37°C. Cells were washed once in M9 medium without carbon source and diluted to an OD_{600} of 0.01 or lower in 5 ml M9 medium supplemented with 0.4% glycerol. Glycerol cultures were incubated at 37°C with shaking at 180 rpm (Multitron, INFORS HT), for at least 17 h before measurements. At this point, cell density was between OD_{600} 0.1-0.4 for all cultures. To ensure exponentially growing cultures at the moment of sampling, M9 overnight cultures were diluted 1:500 in fresh M9 minimal medium with identical carbon source once in between where necessary. Samples were kept on ice and diluted into 1x PBS just before analysis.

Cells were analysed on a BD LSRFortessa flow cytometer at medium flow rate and a maximum event rate of 10’000/s. Per sample, 100’000 events were recorded. Parameters measured were forward scatter (FSC-H), side scatter (SSC-H and SSC-W), ‘mCherry-H’ for mScarlet-I and ‘GFP-H’ for GFP. Data was collected using the Diva (BD Biosciences) software. FlowJo software (Tree Star, Inc.) was used for import and gating of raw data. The forward-scatter (FSC-H) and side-scatter (SSC-H and SSC-W) were used to separate cells from background particles. For

analysis of cultures carrying the c-di-GMP sensor, a third gate was applied in which we used the 'mCherry' channel to gate for cells expressing mScarlet-I. Only mScarlet-I-positive cells were included in the analysis of the GFP signal coming from the c-di-GMP sensor. Gated populations were exported from FlowJo as 'Scale Values' in csv-files and GFP distributions were visualized using MATLAB (MathWorks) scripts.

Substrate switching experiments

Cultures were grown in glucose minimal medium to an OD_{600} of 0.3-0.5. Per culture, a volume of 1.4 ml was sampled, pelleted and washed twice in 800 μl M9-based medium without carbon source. The pellet was resuspended in 1000 μl M9-based medium without carbon source. 800 μl was used to measure the OD_{600} of the cell suspension and the remaining volume was diluted to and OD_{600} of 0.1. Cell suspensions were further diluted to and OD_{600} of 0.01 in 180 μl of medium in a 96-well plate. The rows A and H and columns 1 and 12 of 96-well plates were always filled with ddH_2O and never used for cultivation because of possible evaporation during long-term incubation. Growth / optical density in 96-well plates was followed during incubation using a plate reader (Epoch2 or Synergy2, Biotek Instruments Inc.) and Gen5 (version 3.10, Biotek Instruments Inc.) software. Cultivation settings: plate reader pre-heated to 37°C; fast, continuous shaking; OD_{600} was measured every 15 minutes.

Determination of lag time

Data captured during substrate switching experiments was corrected for background signal by subtracting average signal from wells filled with water from each data point of inoculated wells. As a semi-quantitative approach, the lag time was defined as the time required for cultures to reach an OD_{600} of 0.05 after background correction. The average lag time of strain CGSC 6300 was set to 100% and lag time (%) for all other strains calculated relative to CGSC 6300 average lag time. Maximum OD_{600} values reached during growth were retrieved with a custom-made MATLAB script that searched for the maximum value in each well, using data after background subtraction.

Proteome analysis

For each strain, the proteome was analyzed from three biological replicates. A volume of circa 1 ml was sampled from *E. coli* cultures grown in glucose minimal medium to an OD_{600} of around 0.3 and circa 600 μL was sampled from of *E. coli* cultures grown in TB to an OD_{600} of around 0.5. Samples were directly put on ice, pelleted and washed with 1 mL ice-cold Phosphate-buffered saline (PBS) (8 g/l NaCl, 0.2 g/l KCl, 1.44 g/l $Na_2HPO_4.2H_2O$, 0.24 g/l KH_2PO_4, adjusted to pH 7.4). Pellets were flash-frozen in liquid nitrogen and stored at -80°C until further processing.

Protein extraction and digestion. Cells were lysed in 50 µl lysis buffer (1% sodium deoxycholate, 10 mM TCEP, 100 mM Tris base pH 8.5, supplemented with 15 mM chloroacetamide). Samples were sonicated for 20 minutes using a 30 seconds on / 30 seconds off program (Bioruptor Pico sonication device, Diagenode) and heated for 10 minutes at 95°C and 300 rpm (ThermoMixer C, Eppendorf). Porcine trypsin was added to a final enzyme:protein ratio of 1:50 and samples were digested overnight at 37°C and 300 rpm (ThermoMixer C, Eppendorf).

Solid phase extraction. Digested protein samples were spun down, mixed with 50 µl 5% TFA (solution in water) and 100 µl 1% TFA in 2-propanol. Samples were loaded on PR-Sulfonate Cartridges (SDB-RPS, PreOmics) and spun down for 4 minutes at 2000 rcf (Eppendorf Centrifuge). Cartridges were washed twice with 200 µl 1% TFA in 2-propanol and spun down as above. Cartridges were then washed twice with 200 µl 0.2% TFA in water and spun down as above. Peptides were eluted in two centrifugation rounds with 100 µl Elution buffer (1% (v/v) ammonium hydroxide, 19% water and 80% acetonitrile) each, for 2 minutes at 2000 rcf. Peptide solutions were dry eluted in a vacuum concentrator. Dried peptides were dissolved via ultrasonification and shaking in LC buffer (0.15% Formic acid, 2% acetonitrile) to a final peptide concentration of 0.5 µg/µl. Peptide concentration was verified with a SpectroStar Nanodrop analyzer. Samples were stored at -20°C until measurement.

Relative protein concentration determination using LC-MS. Samples were analyzed by LC-MS on an Orbitrap Elite system (Thermo Scientific). Proteins were identified and quantified over all samples by MASCOT search against the *E. coli* K-12 MG1655 reference proteome UP000000625 (Uniprot (Bateman et al., 2020)). Protein abundances were quantified by MS1-based Label Free quantitation and protein fold changes between conditions were statistically analyzed by SafeQuant (version 2.3.4). The qValue represents adjusted pValue for multiple testing using Benjamini-Hochberg method.

3.6 Supplementary material

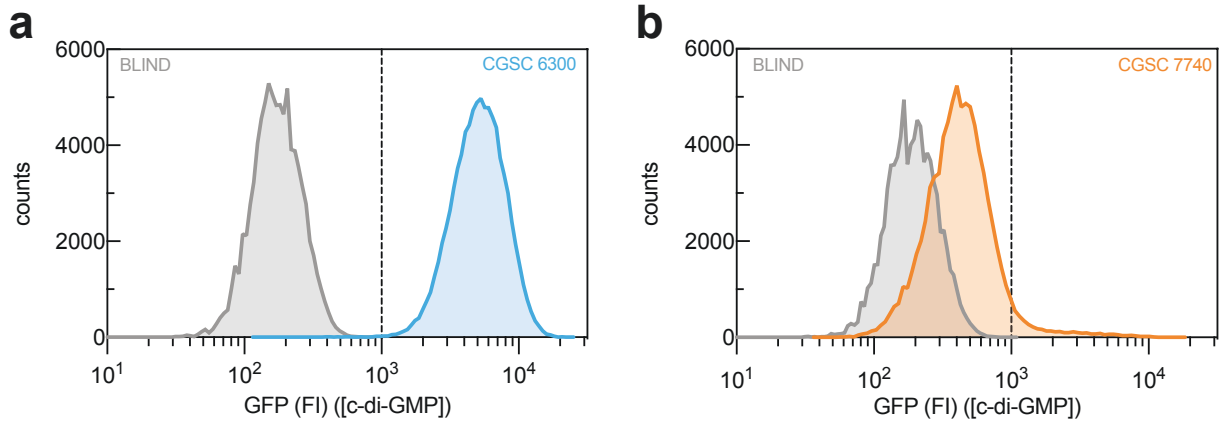


Figure S1: Distinct c-di-GMP levels in CGSC 6300 and CGSC 7740 during growth on glucose. Single cell c-di-GMP levels during exponential growth on glucose minimal medium were measured with a c-di-GMP biosensor expressed from plasmid p2H12ref (Kaczmarczyk and Jenal, unpublished) in **(a)** CGSC 6300 and **(b)** CGSC 7740. The biosensor signal (GFP) positively correlates with [c-di-GMP]. The biosensor signal distributions are shown in comparison to the GFP signal from a mutated biosensor that cannot bind c-di-GMP (p2H12ref-blind (Kaczmarczyk and Jenal, unpublished), in grey). Biosensor expression was induced with 200 nM anhydrotetracycline (aTc). Every graph shows the result after gating of 100'000 events.

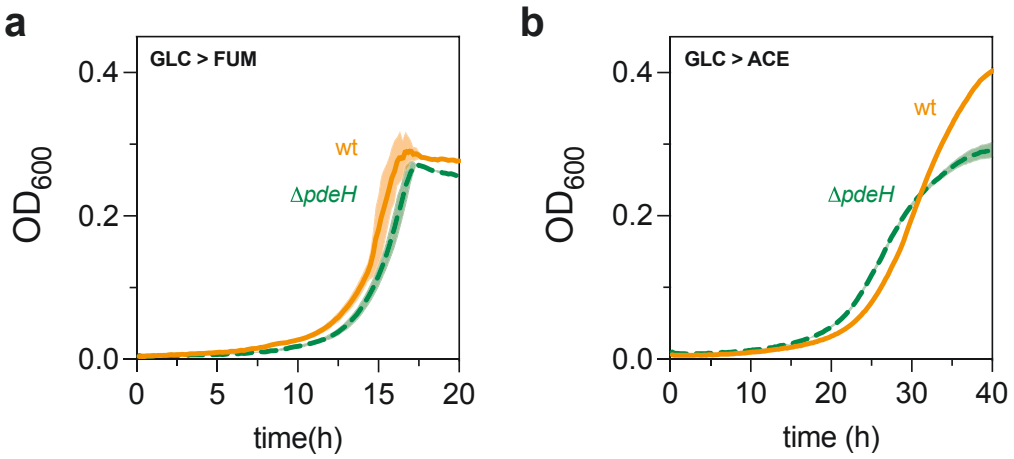


Figure S2: : Growth of CGSC 7740 and CGSC 7740 *ΔpdeH* after substrate shift. Growth curves (optical density (OD_{600}) over time) of a CGSC 7740 genotype (MG1655 (Blattner et al., 1997)) wt (orange) and Δ*pdeH* (green, broken line) after a substrate switch from **(a)** glucose to fumarate and **(b)** glucose to acetate as the sole carbon source. Incubation in medium with the new substrate started at time = 0. Each datapoint represents the average of three technical replicates. GLC: glucose, FUM: fumarate, ACE: acetate.

3.6.1 DNA sequence of *dgcJ::IS1*

Sequencing data of *dgcJ::IS1* of strain CGSC 7740. Indicated are *dgcJ* start codon (bold), *dgcJ* stop codon (black underlined), *IS1* sequence (grey underlined) and duplicated *dgcJ* sequence as a result of *IS1* insertion (orange italic).

ATGAAATTGCACCATAGAATGCTCCGGCATTTTATCGCCGCAAGTGTCATTGTGC
TGACATCTTCCTTCCTTATTTTTGAACTTGTCGCCAGCGACAGAGCAATGAGTGC
CTATCTGCGCTATATCGTGCAGAAAGCAGACTCCTCCTTTCTTTATGATAAGTATC
AGAATCAGAGTATTGCCGCGCATGTGATGCGCGCTCTCGCTGCTGAGCAGTCGG
AAGTGTCGCCAGAACAGCGGCGCGCCATCTGCGAGGCTTTTGAGTCTGCCAATA
ACACCCATGGCTTAAACCTGACTGCCCATAAATACCCGGGCTTACGCGGCACACT
ACAAACCGCATCCACTGACTGCGACACAATTGTGGAAGCTGCAGCACTATTACC
CGCTTTTGATCAGGCAGTGGAAGGCAACCGCCACCAGGATGATTACGGTTCAGG
TCTTGGGATGGCCGAAGAGAAATTTCACTATTATCTCGATCTCAATGACCGCTAT
GTCTATTTTTATGAGCCGGTTAATGTTGAATACTTTGCGATGAATAACTGGTCCTT
CCTGCAGTCAGGAAGTATTGGCATCGATCGCAAAGATATTGAAAAGGTATTTAC
CGGGCGTACCGTATTGTCGAGCATTTACCAGGATCAGCGTACTAAACAGAACGT
GATGAGTTTGCTGACGCCGGTATATGTCGCAGGGCAGCTAAAAGGGATTGTGCT
GCTGGATATTAACAAAAACAATCTGCGGAATATCTTTTATACTCATGACCGCCCT
CTCCTCTGGCGTTTTCTCAATGTCACGCTAACCGATACCGATTCGGGGCGCGACA
TTATCATCAACCAGAGCGAAGATAATCTGTTCCAGTATGTCAGTTACGTCCATGA
CTTACCGGGCGGCATTCGTGTCTCGTTATCCATTGATATTCTTTACTTTATCACGT
CTTCGTGGAAAAGCGTTCTGTTCTGGATTTTGACGGCGTTAATTTTGCTGAATATG
GTG***CGGATGCAC***GGTAATGACTCCAACTTATTGATAGTGTTTTATGTTCAGAT
AATGCCCGATGACTTTGTCATGCAGCTCCACCGATTTTGAGAACGACAGCGA
CTTCCGTCCCAGCCGTGCCAGGTGCTGCCTCAGATTCAGGTTATGCCGCTCA
ATTCGCTGCGTATATCGCTTGCTGATTACGTGCAGCTTTCCCTTCAGGCGGG
ATTCATACAGCGGCCAGCCATCCGTCATCCATATCACCACGTCAAAGGGTGA
CAGCAGGCTCATAAGACGCCCCAGCGTCGCCATAGTGCGTTCACCGAATAC
GTGCGCAACAACCGTCTTCCGGAGACTGTCATACGCGTAAAACAGCCAGCG
CTGGCGCGATTTAGCCCCGACATAGCCCCACTGTTCGTCCATTTCCGCGCAG
ACGATGACGTCACTGCCCGGCTGTATGCGCGAGGTTACCGACTGCGGCCTG
AGTTTTTTAAGTGACGTAAAATCGTGTTGAGGCCAACGCCCATAATGCGGGC
TGTTGCCCGGCATCCAACGCCATTCATGGCCATATCAATGATTTTCTGGTGC
GTACCGGGTTGAGAAGCGGTGTAAGTGAACTGCAGTTGCCATGTTTTACGG
CAGTGAGAGCAGAGATAGCGCTGATGTCCGGCGGTGCTTTTGCCGTTACGC
ACCACCCCGTCAGTAGCTGAACAGGAGGGACAGCTGATAGAAACAGAAGCC
ACTGGAGCACCTCAAAAACACCATCATACACTAAATCAGTAAGTTGGCAGCA
TCACC***CGGATGCAC***TTCCGTTTATACCAAAATGTGTCGCGAGAAAATATTAGTGA
TGCGATGACTGGACTGTATAATCGCAAAATTTTAACCCCTGAACTGGAGCAGCG
GTTGCAGAAACTGGTGCAATCCGGTTCTTCGGTGATGTTTATTGCTATTGACATG
GACAAGTTAAAGCAAATAAATGACACCCTCGGTCATCAGGAGGGGGATTTAGCG
ATTACGTTATTAGCTCAGGCGATTAAACAGTCGATTCGTAAAAGTGATTATGCCA
TCCGACTCGGTGGCGATGAATTCTGCATCATTCTTGTCGATTCGACGCCGCAAAT
TGCAGCACAACTGCCTGAACGTATCGAAAAACGTCTGCAACATATCGCGCCGCA
GAAAGAGATCGGCTTCTCTTCCGGTATTTACGCGATGAAAGAAAACGATACGTT
ACATGATGCGTATAAAGCTTCCGATGAGCGTTTATATGTCAATAAGCAGAACAA
AAACAGCCGTTCATGA

Table S1: Strains used in this study

Strain	Genotype	Source
MG1655	F- λ- rph-1 (PflhDC::IS1 crl::IS1 dgcJ::IS1)	Blattner et al., 1997
CSGC 6300	F- λ- rph-1	Coli Genetic Stock Center
CGSC 7740	F- λ- rph-1 (PflhDC::IS1 crl::IS1 dgcJ::IS1)	Coli Genetic Stock Center
AB3746	F- λ- rph-1 (PflhDC::IS1 crl::IS1 dgcJ::IS1)	M. Heinemann
AB607	MG1655 ΔpdeH::frt	Boehm et al., 2010
AB3863	CGSC 7740 ΔdgcJ::kan Prha-ccdB	This study
AB3865	CGSC 7740 ΔPflhDC::kan Prha-ccdB	This study
AB3874	CGSC 7740 Δcrl::kan Prha-ccdB	This study
AB3882	CGSC 7740 crl	This study
AB3864	CGSC 7740 dgcJ	This study
AB3870	CGSC 7740 PflhDC	This study
AB3883	CGSC 7740 crl dgcJ	This study
AB3884	CGSC 7740 PflhDC crl	This study
AB3897	CGSC 7740 PflhDC dgcJ	This study
AB3885	CGSC 7740 PflhDC crl dgcJ	This study

Table S2: Plasmids used in this study

Plasmid	Genotype (*resistence*)	Source
pKD46	λ RED$^+$ (*ampicillin*)	Datsenko and Wanner, 2000
pKD45	kan-Prha-ccdB (*ampicillin, kanamycin*)	Datsenko and Wanner, 2000
pCP20	FLP$^+$ (*ampicillin*)	(Cherepanov and Wackernagel, 1995)
p2H12ref	P_{tet}-sensor-mScarlet-I (*ampicillin*)	UJ11206, Kaczmarczyk & Jenal, *unpublished*
p2H12ref-blind	P_{tet}-sensor*-mScarlet-I (*ampicillin*)	UJ11207, Kaczmarczyk & Jenal, *unpublished*

*mutation in the c-di-GMP binding domain of the c-di-GMP sensor

Table S3: Primers used for analysis and removal of IS1 elements

primer number	primer name	template	sequence (bold overlaps with template)
13973	crl_fwd1	pKD45	GCTAAATTTTGCCAATTTGGTAA AACAGTTGCATCACAACAGGAGATAGCA **TCAGAAGAACTCGTCAAGAAG**
13974	crl_rev1	pKD45	GAATATTAATAATGATGATATTG TCGCGGTTGGCATGACGTATCAGTTTTAATG **GGATATTATCGTGAGGATG**
13975	crl_fwd2	CGSC6300	**GCTAAATTTTGCCAATTTGGT AAAACAGTTGCATCACAACAGGAGATAGCA**
13976	crl_rev2	CGSC6300	**GAATATTAATAATGATGATAT TGTCGCGGTTGGCATGACGTATCAGTTTTAATG**
13977	crl_seq_fwd	CGSC 7740	CGGTAAATTATTAGAGATCC
13068	crl_seq_rev	CGSC 7740	CGACGTCGGTGCTACGTATT
13708	dgcJ_fwd1	pKD45	GAACTTTATCTGGTTTTCTCGTT TCACTAACCGAAGGAGTGCCATTTATC **TCAGAAGAACTCGTCAAGAAG**
13709	dgcJ_rev1	pKD45	CAGTCATTTCACGCTCCTGAGAT TACAAGCAAACAACCACAGAAGGTTA **GGATATTATCGTGAGGATG**
13710	dgcJ_fwd2	CGSC 6300	GAACTTTATCTGGTTTTCTCGTT TCACTAACCGAAGGAGTGCCATTTATC
13711	dgcJ_rev2	CGSC 6300	CAGTCATTTCACGCTCCTGAGAT TACAAGCAAACAACCACAGAAGGTTA
13712	dgcJ_seq_fwd	CGSC 7740	GTGATTATTCATCATATTTAAGTGCAAAAATATTC
13713	dgcJ_seq_rev	CGSC 7740	CCTAGAGCGGTGCCACGAATTTCTGACACC
12820	PflhDC_fwd1	pKD45	GATTTTCAATAATGCGTGATGCA GATCACACAAAACACTCAATTACTTAAC **TCAGAAGAACTCGTCAAGAAG**
12821	PflhDC_rev1	pKD45	CTTTGTATTTAATTAGTTTGTTG TGCGGTAAGTGTCTGTTTAAAAATAGC **GGATATTATCGTGAGGATG**
12822	PflhDC_fwd2	CGSC 6300	GAAGTGACAAACCAGTTGATTG
12823	PflhDC_rev2	CGSC 6300	CTTTGCGTTTCTTCATGCATC
12519	PflhDC_seq_fwd	Var*	GGAATGTTGCGCCTCACCG
12521	PflhDC_seq_rev	Var	CCTGTTTCATTTTTGCTTGCTAGC
13067	crl_fwd_fig1	Var	CAGGAAATCACCGACTGGAT
13068	crl_seq_rev	Var	CGACGTCGGTGCTACGTATT
12553	dgcJ_fwd_fig1	var	CAACAAAGTGATTATTCATC
12554	dgcJ_rev_fig1	Var	CGATTAACCGCTGATGCGTC

*Various templates

Chapter 4

A role for metabolic flux sensor Cra in second messenger heterogeneity

Margo van Berkum, Shogo Ozaki[1], Raphael Dias Teixeira, Alberto Reinders, Urs Jenal

Biozentrum of the University of Basel
Klingelbergstrasse 50/70, 4056 Basel, Switzerland

Current address:
[1] Department of Molecular Biology, Graduate School of Pharmaceutical Sciences, Kyushu University, Higashi-ku, Fukuoka 812-8582, Japan

Contents

Statement of my contributions:

The data presented in Chapter 4 reflect primarily my work. Mutant *cra* (R197A) was designed and constructed by S. Ozaki and A. Reinders. EMSA data presented in Figure 3d,f is the work of S. Ozaki. R. Dias Teixeira performed the protein superimposition as shown in Figure 3b and S1a. I designed and performed all other experiments for which data is presented in Chapter 4. I analysed the data and I wrote the chapter.

4.1 Abstract

Networks of small intracellular signaling molecules like c-di-GMP or cyclic AMP globally control bacterial physiology and behavior. It is generally envisioned that the cellular concentration of these signaling molecules is controlled deterministically in response to environmental cues. We have recently reported that *E. coli* excludes intermediate concentrations of c-di-GMP by making use of a regulatory switch that converts graded changes of c-di-GMP into binary states. Bimodal regimes of c-di-GMP are established by the phosphodiesterase PdeL and *pdeL* expression depends on the available carbon source. While *pdeL* transcription requires the metabolic flux sensor Cra, it has remained unclear how carbon metabolism or growth rates affect c-di-GMP in individual cells and if Cra contributes to signaling heterogeneity. Here, we probe the role of Cra in regulating c-di-GMP distributions in individual bacteria. We demonstrate that when grown in glucose minimal media, Cra activity is low resulting in the loss of *pdeL* expression and in the loss of c-di-GMP bimodality. In contrast, a Cra mutant that is unable to bind its allosteric regulator Fructose-1-phosphate (F1P) displayed increased Cra activity on glucose, increased *pdeL* promoter activity and also affected protein levels of other Cra targets. We demonstrate that Cra activity and *pdeL* expression are also increased when the overall glucose uptake is limited through the controlled expression of the glucose transporter *ptsG*. Initial experiments indicated that increasing Cra activity cannot restore c-di-GMP bimodality under conditions where Cra is normally limited by the metabolic conditions. We have proposed that the PdeL switch allows *E. coli* to hedge its bets when undergoing fundamental behavioral transitions between motile and sessile states. The results presented here, indicate that c-di-GMP bet-hedging is tightly interlinked with the *E. coli* metabolism to fine tune developmental decisions with the cells' nutrient status and growth rate. Further investigations are required to elucidate under which physiological conditions the regulatory interplay between Cra and PdeL provides optimal fitness benefits to *E. coli*.

4.2 Introduction

Genetically identical bacterial populations can display substantial phenotypic variations (Elowitz et al., 2002). These variations are proposed to be beneficial during various biological processes (Elowitz et al., 2002) and to fulfil different functional roles including bet-hedging and division of labor (Bettenworth et al., 2019; Veening et al., 2008b). Phenotypic diversification is broadly exploited by many bacterial species, suggesting it to be a principle with evolutionary benefits (Smits et al., 2006; Thattai and Van Oudenaarden, 2004).

Heterogeneity is often observed on the level of gene expression, driving distinct phenotypic traits (Smits et al., 2006). Recently we however observed heterogeneity in the cellular levels of the small signaling molecule c-di-GMP (chapter 2), that contols growth and behavior phenotypes including motility, biofilm formation and virulence across all major bacterial phyla (Jenal et al., 2017, Romling et al., 2013). We observed that *E. coli* populations display bimodal c-di-GMP distributions and identified the c-di-GMP specific phosphodiesterase *pdeL* as driver of this heterogeneity (chapter 2).

PdeL establishes bimodal c-di-GMP distributions during growth on gluconeogenic substrates including glycerol. Cells of a *pdeL* deletion strain all display high c-di-GMP levels under these conditions (chapter 2 and Fig. 1). During growth on glucose bimodal c-di-GMP distributions are also lost, thus growth on glucose phenocopies a *pdeL* deletion strain (Fig. 1).

Establishing these binary signaling outputs through *pdeL* depends on cells' metabolic state, as the metabolic regulator Cra is an essential activator of *pdeL* transcription (chapter 2). Although Cra was initially reported to repress *in vitro pdeL* transcription (Shimada et al., 2005), binding of Cra to the *pdeL* promoter region is required for *in vivo pdeL* expression (chapter 2). Remarkably, the *pdeL* promoter contains one of the tightest binding sites for Cra (K_d of 2.4 μM) (Shimada et al., 2005). The promoter of the *fruBKA* operon, on which Cra acts as repressor (hence the original name of Cra was FruR, Fructose Repressor) has the highest Cra binding affinity of 1.9 μM (Shimada et al., 2005). Promoter regions of Cra-activated genes such as *aceB* and *ppsA* have a ten-fold higher K_d for Cra binding as compared to fruBKA (Shimada et al. 2005). Cra bound to P_{pdeL} facilitates binding of a PdeL dimer to a palindromic sequence that overlaps with the Cra binding site (chapter 2). Hence Cra and PdeL act as co-activators of *pdeL* expression (chapter 2).

The activity of Cra inversely scales with what is called the glycolytic flux (Kochanowski et al., 2013), i.e. the amount of sugar compounds catabolized through the glycolysis pathway per time. The metabolite Fructose-1-phosphate is currently considered the sole allosteric regulator of Cra, that modulates Cra activity through binding to the LacI-type sugar binding domain (Bley Folly et al., 2018; Ramseier et al., 1993; Weickert & Adhya, 1992). It is however still unknown how activity of Cra on its targets is modulated during growth on other substrates than fructose. Consistently with its activity, however, Cra activates genes encoding enzymes of the TCA cycle and enzymes required for gluconeogenesis, and represses transcription of glycolytic enzymes (Ramseier et al., 1993, 1995). PdeL, as well as a set of virulence genes, are among the non-metabolic Cra targets (Njoroge et al., 2012, 2013).

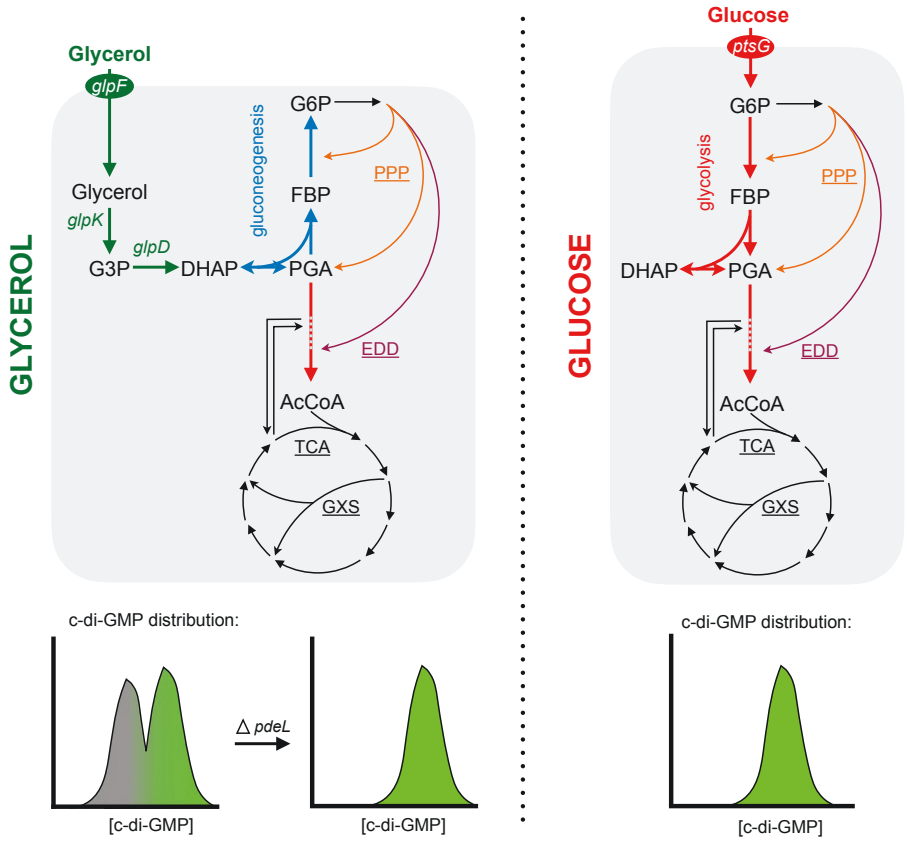


Figure 1: Metabolism of glycerol and glucose and corresponding c-di-GMP distributions. Schematic of metabolism and corresponding c-di-GMP distributions on glycerol (left) and glucose (right) in *E. coli*. Left: the glycerol specific dissimilation pathway (green arrows) converts glycerol via G3P into the glycolytic intermediate DHAP. Glycerol passes the inner membrane through facilitated diffusion (glpF). Under aerobic conditions, glycerol is phosphorylated to G3P at the expense of ATP by Glycerol kinase (glpK) in the cytosol. G3P is oxidized to DHAP by the aerobic G3P dehydrogenase (glpD). From here, lower glycolysis (red arrow) takes place. Gluconeogenesis (blue arrows) is required to generate G6P and feed the PPP. During growth on glycerol as the sole energy and carbon source, *E. coli* populations displayed a bimodal c-di-GMP distribution (grey: low c-di-GMP, green: higher c-di-GMP). Upon deletion of *pdeL*, only the higher c-di-GMP peak was observed (green, chapter I). Right: uptake of glucose through the glucose transporter (ptsG) phosphorylates glucose into G6P, which is further catabolized in the upper and lower glycolysis pathway (red arrows). During growth on glucose as the sole energy and carbon source, *E. coli* populations displayed a unimodal c-di-GMP distribution corresponding to high c-di-GMP levels (green, chapter I). Abbreviations: glpF, gene encoding glycerol facilitator; glpK, encoding glycerol kinase; glpD, encoding glycerol 3-phoshpate dehydrogenase; G3P, sn-glycerol 3-phosphate; DHAP, dihydroxyacetone phosphate; G6P, glucose 6-phosphate; FBP, fructose-1,6-bisphosphate; PGA, glyceraldehyde 3-phosphate; PPP, pentose phosphate pathway; EDD, Entner-Doudoroff pathway; TCA, tricarboxylic acid cycle; GXS, glyoxylate shunt. Orange, purple and broken red line represent multiple enzymatic reactions and metabolic intermediates that are not shown for simplicity. Also, multiple other metabolic reactions taking place during glucose and glycerol metabolism are not shown for simplicity.

We characterized PdeL *in vivo* during growth on tryptone or glycerol as the sole carbon and energy source (chapter 2). During these experiments, we assumed unimodal high Cra activity and focused on the c-di-GMP dependent feedback loops on *pdeL* expression. Bimodal c-di-GMP distribution on glycerol is lost in a catalytically inactive *pdeL* strain and growth on glucose phenocopies this outcome (chapter 2). As Cra activity on glycerol is higher compared to growth on glucose (Lehning et al., 2017), how does reduced Cra activity on glucose contribute to the loss of c-di-GMP bimodality? What is the role of Cra in setting bimodal c-di-GMP distributions through *pdeL*? Can bimodal c-di-GMP be established on glucose by increasing Cra-activity? We

hypothesized that the c-di-GMP phenotype on glucose could be the result of reduced Cra activity that acts through *pdeL*. This hypothesis assumes that *pdeL* expression on glucose is limited by Cra and that bimodal c-di-GMP distributions could be established by increasing Cra activity on glucose.

Here, we further investigate the role of Cra in bimodal c-di-GMP distributions. We explore different tools to increase Cra activity during growth on glucose and assess their effect on c-di-GMP distributions. We verify that Cra activity is increased by induction of *cra* expression, in a *cra* mutant that lacks F1P binding capability and by lowering the glycolytic flux via the *E. coli ptsG* glucose transporter. Increasing Cra activity during growth on glucose did however not result in bimodal c-di-GMP distributions as seen during gluconeogenesis. From our data we suggest that additional regulatory mechanisms affect the c-di-GMP pool. These c-di-GMP regulating systems are predicted to display differential activity during growth on glucose compared to gluconeogenic growth, resulting in the observed substrate-dependent c-di-GMP distributions.

4.3 Results

4.3.1 Assembly of a Cra activity reporter to distinguish metabolic states of individual cells

Our observation of bimodal distributions of c-di-GMP within *E. coli* populations (chapter 2) and the potentially important role of Cra in mediating these unicellular patterns through the control of *pdeL*, emphasize the importance of measuring the physiological state of individual cells rather than population averages. To investigate the exact role of Cra in establishing bimodal c-di-GMP distributions, Cra activity measurements on the single cell level are required. Cra activity was defined as the normalized promoter activity of a Cra target gene for which Cra is the sole known regulator (Kochanowski et al., 2013). This is the case for the *pykF* and *ppsA* promoters (Bledig et al., 1996; Geerse et al., 1989; Keseler et al., 2017; Nègre et al., 1998). The promoter activity readout of these Cra-controlled genes is normalized by the activity of the same promoter but with a scrambled, i.e., randomized Cra binding site (P*). This would leave the promoter activity condition-dependent but Cra-independent (Kochanowski et al., 2013). Fluorescence-based Cra activity measurements have previously been performed on a population level based on the *pykF* promoter (P_{pykF}-*gfp* and P^*_{pykF}-*gfp*) (Kochanowski et al., 2013) and later on a single cell level based on the *ppsA* promoter (P_{ppsA}-*gfp* and P^*_{ppsA}-*rfp*) (Lehning et al., 2017).

Here, we used a single cell Cra activity reporter based on the Cra-repressed *pykF* promoter and two fluorescent proteins of distinct color (Fig. 2a). Expression of a *cfp* was under control of *pykF* and expression of *yfp* was controlled by P^*_{pykF}. Both fluorescent proteins were expressed from the same plasmid (Fig. 2a). Cra activity was defined as the single cell YFP/CFP ratio, where a higher ratio corresponds to higher Cra activity.

The functioning of this Cra activity reporter was then validated by comparing its readout during growth on different substrates that should infer different Cra activities (Kochanowski et al., 2013, Lehning et al., 2017), as well as by verifying differences in readout between a *cra* and Δ*cra* during growth on glucose. Cultures carrying the Cra activity reporter plasmid were grown up to mid-exponential phase on different substrates and analyzed by capturing YFP and CFP intensities of individual cells by microscopy (Fig. 2b). We found that single cell Cra activity increased when reporter cells were grown from glucose to glycerol, galactose, and casamino acids (Fig. 2c). Hence, we could distinguish different Cra activities on different substrates. Although Cra activity was lowest on glucose, ratios dropped to even lower levels in a Δ*cra* mutant (Fig. 2d), indicating that the reporter is able to read out low levels of Cra activity and distinguish it from fluorescence levels observed in the absence of Cra. These data indicated that the novel tool indeed allows to quantify Cra activities in individual cells in a physiologically relevant window.

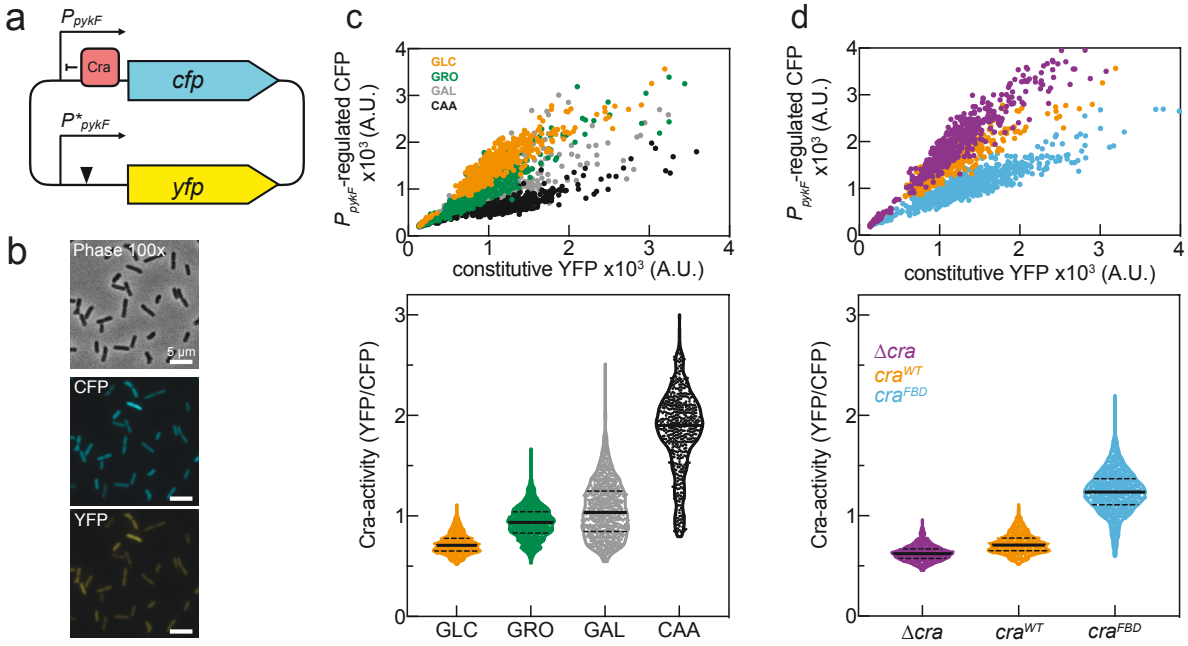


Figure 2: A fluorescence-based Cra activity reporter distinguishes metabolic states of single cells across conditions. **(a)** Graphical lay-out of the single cell Cra activity reporter plasmid. Expression of *cfp* is under control of the *pykF* promoter. Cra functions as repressor of P_{pykF}. A *pykF* promoter sequence with a scrambled Cra binding site, P* (Kochanowski et al. 2013), drives expression of *yfp* from the same plasmid. **(b)** Section of a snapshot of *E. coli* cells carrying the Cra activity reporter plasmid. Cultures were grown into exponential phase on glucose and spotted on agar pads. Phase contrast, as well as CFP and YFP intensity (all at 100x magnification) were captured by microscopy. Scale bar: 5 μm. **(c)** Single cell Cra activity of *E. coli* wild type cells grown on the substrates glucose (GLC, orange), glycerol (GRO, green), galactose (GAL, grey) and casamino acids (CAA, black). Every data point represents a single cell. **Top:** Single cell P_{pykF}-controlled CFP intensity as a function of Cra-independent YFP intensity. **Bottom:** Cra activity (defined as the YFP:CFP intensity ratio) calculated for the single cells displayed in the top graph. **(d)** Single cell measurements of the Cra activity reporter in a *Δcra* (purple), *cra* (orange) and cra^{FBD} (blue) strain during exponential growth on glucose. **Top:** Single cell P_{pykF}-controlled CFP intensity as a function of Cra-independent YFP intensity. **Bottom:** Cra activity (defined as the YFP:CFP intensity ratio) as calculated for the data points displayed in the top graph.

4.3.2 Genetic uncoupling of Cra activity from metabolic fluxes

To be able to better define the entire range of Cra activities, we next thought to generate a constitutive Cra mutant that no longer responds to metabolic signals. Cra activity is allosterically regulated by the metabolite Fructose-1-phosphate (F1P) that binds to the Cra protein thereby reducing the affinity for its DNA binding sites (Bley Folly et al., 2018; Ramseier et al., 1993). To define mutations that abolish F1P binding to *E. coli* Cra, we used information provided by the crystal structures of *E. coli* Cra in its apo form (Cra^{Ec}) and of its ligand bound homologs. Cra contains an N-terminal helix-turn-helix (HTH) DNA binding domain and a C-terminal LacI-type sugar binding domain (Fig. 3a). Since the structure of the Cra^{Ec} sugar binding domain (PDB: 2iks) was solved without its ligand, we used ligand bound structures of homologs of Cra^{Ec} to define its ligand binding pocket. The sugar binding domain of Cra from Pseudomonas putida (Cra^{Pp}) (Chavarría et al., 2014) shows 63.9% sequence similarity with the C-terminal domain of Cra^{Ec} (Fig. S1). Not only are residues involved in binding F1P in Cra^{Pp} (PDB: 3o75) conserved in Cra^{Ec} (Fig. S1), but superposition of the crystal structures of the sugar binding domains of Cra^{Ec} and Cra^{Pp} confirmed conservation of the structural positioning of these residues (Fig. 3b, Fig. S1). Several of the residues coordinating F1P binding in Cra^{Pp} are also conserved in LacI (Fig. S1), where they form direct hydrogen bonds with the inducer IPTG (Daber et al., 2007). Replacement

of one of these residues, R197, was shown to result in the complete loss of LacI response to its inducer (Kleina and Miller, 1990; Markiewicz et al., 1994). Based on these structural homology studies, we choose to replace residue R197 in Cra^{Ec} with a nonpolar Alanine residue to generate the *cra* (R197A) mutant (Fig. 3a).

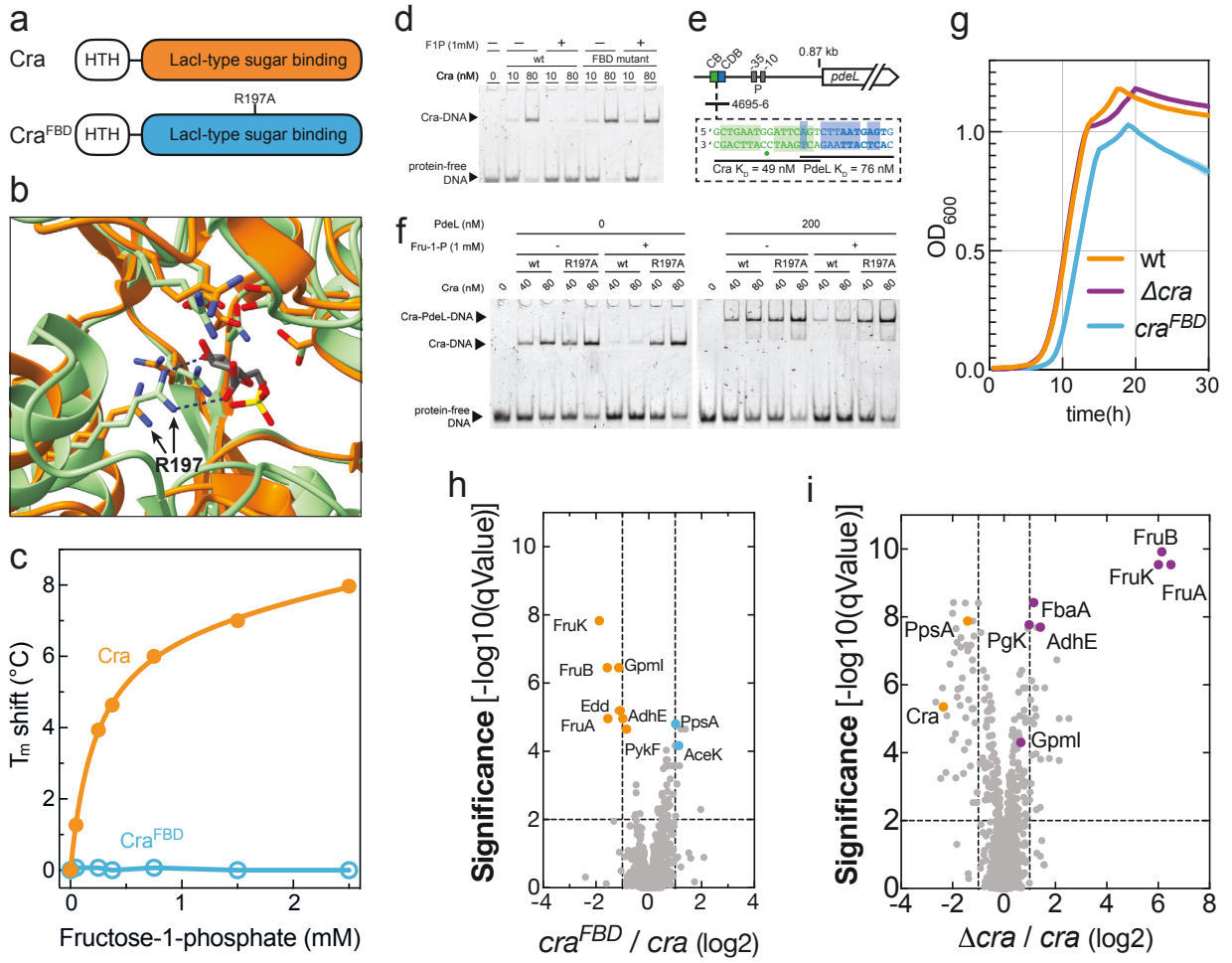


Figure 3: An F1P binding deficient Cra mutant displays increased Cra activity *in vitro* and *in vivo*. **(a)** Schematic lay-out of the protein domains of Cra and Cra^{FBD}. The Cra protein consists of an N-terminal Helix-turn-helix motive (HTH) and a C-terminal LacI-type sugar binding domain. The R197A mutation in Cra^{FBD} is located in the sugar binding domain. **(b)** Zoom-in of the superimposition of the sugar binding domain of *P. putida* Cra (Cra^{Pp}, light green, PDB: 3o75) bound to Fructose-1-phosphate (F1P, grey) and *E. coli* Cra (Cra^{Ec}, orange, PDB: 2iks). Cra^{Pp} residues shown in sticks interact with F1P. The F1P-interacting residues that are conserved between Cra^{Pp} and Cra^{Ec} are also shown in sticks in the Cra^{Ec} structure. Residue R197 is indicated for both structures with black arrows. Hydrogen bonds between F1P and R197 of Cra^{Pp} are indicated with dark broken lines. Color-coded atoms are nitrogen (blue), oxygen (red) and phosphate (yellow). **(c)** Thermal Shift Assay (TSA) of Cra (orange) and Cra^{FBD} (blue) in the presence of F1P. The increase in protein denaturation temperature (T_m shift (°C)) of purified Cra-strepII and Cra^{FBD}-strepII was measured at increasing F1P concentrations. The reference T_m was defined as the T_m in the absence of F1P. **(d)** Electrophoretic Mobility Shift Assay (EMSA) of 5' Cy3 labeled oligonucleotide covering the Cra binding and Cra-dependent PdeL binding site and 0, 10 and 80 nM of purified Cra-strepII (wt) and Cra^{FBD}-strepII. DNA band shifts were assessed in the presence (+) or absence (-) of 1 mM F1P. **(e)** graphical overview of Cra and PdeL DNA binding sequences used in the EMSA in **(d)** and **(f)**. **(f)** EMSA of 5' Cy3 labeled oligonucleotide covering the Cra binding and Cra-dependent PdeL binding site and purified Cra-StrepII ('wt'), Cra^{FBD}-strepII ('R197A') and PdeL-StrepII proteins. The position of the labeled oligonucleotides in the *pdeL* promoter region is indicated in **(e)**. DNA band shifts were assessed with (200 nM) or without (0) PdeL-strepII and in the presence (+) or absence (-) of 1 mM F1P. **(g)** Growth curves (optical density (OD_{600}) over time) of *cra* (orange), *cra*^{FBD} (blue) and *Δcra* (purple) cultures grown on glucose. **(h,i)** Whole proteome analysis of strains *cra*, *cra*^{FBD} and *Δcra* during exponential growth on minimal medium supplemented with glucose as the sole carbon source. Significance of the log2 fold change of the relative protein abundance *cra*^{FBD}/cra **(h)** and *Δcra*/cra **(i)**. Cra targets downregulated in *cra*^{FBD} and *Δcra* are indicated in orange, as is the Cra protein in **(i)**. Cra targets upregulated in *cra*^{FBD} as compared to *cra* are indicated in blue. Cra targets upregulated in a *Δcra* background as compared to *cra* are indicated in purple.

To demonstrate that Cra (R197A) is unable to bind F1P, we purified the mutant and wild type protein and assessed ligand binding by means of thermal shift assays. While binding of F1P to wild type Cra was readily observed based on large thermal shifts of the protein upon addition of increasing concentrations of F1P, binding to Cra^{FBD} was completely abolished (Fig. 3c). We thus refer to this mutant as 'F1P binding deficient Cra' (Cra^{FBD}). To assess the potential changes in Cra-DNA binding, we used the Cra binding site of the *pdeL* promoter region (Shimada et al., 2011) for DNA shift assays with purified Cra or Cra^{FBD}. Cra-DNA complex formation was identical for Cra and Cra^{FBD} in the absence of F1P (Fig. 3d). However, while the Cra-DNA complex was not observed in the presence of 1 mM F1P, Cra^{FBD}-DNA complex was observed in the presence of the ligand (Fig. 3d), indicating that DNA binding of this mutant is no longer influenced by this concentration of the metabolite. We have shown recently that both Cra and PdeL bind to the *pdeL* promoter with high affinity and that Cra is responsible for recruiting PdeL to this site (Reinders et al., 2016; Shimada et al., 2005). In line with this and with the data above, both Cra and Cra^{FBD} were able to recruit PdeL to the *pdeL* promoter sequence, but while addition of F1P strongly reduced PdeL binding in the presence of Cra, both proteins remained fully bound to DNA in the presence of Cra^{FBD} (Fig. 3f).

To probe the properties of Cra^{FBD} *in vivo*, we replaced the *cra* gene on the *E. coli* chromosome with the cra^{FBD} allele. The resulting cra^{FBD} mutant strain had a slight growth defect on glucose (Fig. 3g), suggesting that tuning of Cra activity by F1P is crucial for optimal growth on glucose. We did not observe growth of the cra^{FBD} strain on fructose (Fig. S2). This strong *in vivo* effect of Cra^{FBD} on fructose metabolism likely results from the repression of the *fruBKA* operon. The *fruBKA* promoter region contains two tight Cra binding sites (Ramseier et al., 1993; Shimada et al., 2005). Remarkably, cra^{FBD} also imposed a slight growth defect on gluconeogenic substrates including glycerol, fumarate and alpha-ketoglutarate. Plate reader-based growth measurements showed a lower maximum optical density in a cra^{FBD} background for each substrate (Fig. S2).

Next, we used the Cra activity reporter to show that the cra^{FBD} strain indeed has increased Cra activity as compared to the wild type strain when grown in glucose minimal media (Fig. 2d). To investigate if increased Cra^{FBD} activity also affects protein levels of its downstream targets, we performed a whole proteome analysis of the cra^{FBD} strain grown on glucose and compared it to *E. coli* wild type and to the Δ*cra* mutant. Seven proteins were found to be downregulated at least two-fold in the cra^{FBD} strain (-log10(qValue) > 2 and log2(cra^{FBD}/*cra*) < -1) (Fig. 3h, Table 1). Six of these proteins were previously shown to be repressed by Cra: AdhE, Edd, GpmI, FruB, FruA and FruK (Santos-Zavaleta et al., 2018). The Cra-repressed PykF protein level was 1.8-fold reduced in a cra^{FBD} background (Fig. 3h, Table 1). Nine proteins were upregulated at least two-fold in a cra^{FBD} strain (-log10(qValue) > 2 and log2(cra^{FBD}/*cra*) >1) (Fig. 3h, Table 1), of which two are known Cra targets: PpsA and AceK (Santos-Zavaleta et al., 2018). The other seven upregulated proteins could be so far unidentified Cra targets or could be upregulated indirectly.

Table 1: Differential protein expression in a *cra^{FBD}* and *Δcra* background during growth on glucose. Relative protein abundance in a *cra^{FBD}* strain compared to wt *cra*, indicated as the protein abundance fold change ('fold change *cra^{FBD}*/*cra*'). Listed are all proteins with a significant fold change (-log(qValue)>2). Protein description according to UniProt KB. Additionally, the corresponding *Δcra*/*cra* fold change is indicated including significance (-log(qValue)). Cra regulon: Cra is a known regulator of expression of proteins marked 'V' (Santos-Zalaveta et al., 2018). n.s., not significant.

Protein name	Protein Description	fold change *cra^{FBD}*/*cra*	-log(qValue) *cra^{FBD}*/*cra*	fold change *Δcra*/*cra*	-log(qValue) *Δcra*/*cra*	Cra regulon
FruK	1-phosphofructokinase	-3.7	7.8	64.5	9.5	V
FruB	Multiphosphoryl transfer protein	-3.0	6.5	70.7	9.9	V
FruA	PTS system fructose-specific EIIB'BC component	-2.9	5.0	90.1	9.5	V
Mqo	Malate:quinone oxidoreductase	-2.3	5.2	n.s.	<2.0	-
GpmI	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	-2.2	6.5	1.6	4.3	V
Edd	Phosphogluconate dehydratase	-2.1	5.2	n.s.	<2.0	V
AdhE	Aldehyde-alcohol dehydrogenase	-2.0	5.0	2.7	7.7	V
PykF	Pyruvate kinase I	-1.8	4.7	1.4	2.9	V
Eno	Enolase	-1.4	2.1	1.3	4.7	V
PtsI	Phosphoenolpyruvate-protein phosphotransferase	-1.4	2.2	n.s.	<2.0	V
GapA	Glyceraldehyde-3-phosphate dehydrogenase A	-1.4	3.0	1.4	2.9	V
HtpX	Protease	-1.4	2.8	n.s.	<2.0	-
PurE	N5-carboxyaminoimidazole ribonucleotide mutase	-1.3	2.2	n.s.	<2.0	-
Agp	Glucose-1-phosphatase	1.3	2.3	n.s.	<2.0	-
SucB	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	1.3	2.3	-1.4	3.1	-
SucD	Succinate--CoA ligase [ADP-forming] subunit alpha	1.3	2.2	-1.5	5.3	-
PoxB	Pyruvate dehydrogenase	1.4	2.2	1.3	4.1	V
FumC	Fumarate hydratase class II	1.4	2.3	n.s.	<2.0	-
GabD	Succinate-semialdehyde dehydrogenase	1.4	2.8	n.s.	<2.0	-
Icd	Isocitrate dehydrogenase	1.4	2.7	-1.4	5.0	V
YaeQ	Uncharacterized protein	1.4	2.0	n.s.	<2.0	-
PckA	Phosphoenolpyruvate carboxykinase	1.4	2.1	-1.2	2.0	V
GltA	Citrate synthase	1.5	3.5	-1.4	5.3	-
DkgA	2,5-diketo-D-gluconic acid reductase A	1.5	2.5	n.s.	<2.0	-
Trg	Methyl-accepting chemotaxis protein III	1.5	2.3	-3.3	5.4	-
ProX	Glycine betaine/proline betaine-binding periplasmic protein	1.6	2.5	n.s.	<2.0	-
Fbp	Fructose-1,6-bisphosphatase class 1	1.6	3.6	-1.3	3.1	-
TktB	Transketolase 2	1.6	3.8	2.4	7.7	-
AceB	Malate synthase A	1.6	4.0	-1.4	3.1	V
ArgT	Lysine/arginine/ornithine-binding periplasmic protein	1.6	3.0	n.s.	<2.0	-
AceA	Isocitrate lyase	1.6	3.6	-1.6	4.8	V
TalA	Transaldolase A	1.6	3.8	1.2	2.2	-
AldA	Lactaldehyde dehydrogenase	1.6	2.8	-1.7	6.4	-
GlpF	Glycerol uptake facilitator protein	1.6	2.4	-2.1	2.8	-
UgpB	sn-glycerol-3-phosphate-binding periplasmic protein	1.7	2.1	n.s.	<2.0	-
YhbO	Protein/nucleic acid deglycase 2	1.7	2.6	2.4	5.9	-
FadE	Acyl-coenzyme A dehydrogenase	1.7	2.4	n.s.	<2.0	-
OsmY	Osmotically-inducible protein Y	1.8	3.2	1.7	2.7	-
Acs	Acetyl-coenzyme A synthetase	1.8	3.6	1.7	5.9	-
YgaM	Uncharacterized protein YgaM	1.9	2.7	1.4	2.0	-
YehZ	Glycine betaine-binding protein YehZ	2.0	3.6	n.s.	<2.0	-
MglA	Galactose/methyl galactoside import ATP-binding protein	2.0	4.2	-1.2	2.7	-
YedP	Mannosyl-3-phosphoglycerate phosphatase	2.0	2.1	1.5	2.0	-
PpsA	Phosphoenolpyruvate synthase	2.0	4.8	-2.7	7.9	V
FadB	Fatty acid oxidation complex subunit alpha	2.1	3.0	n.s.	<2.0	-
AceK	Isocitrate dehydrogenase kinase/phosphatase	2.2	4.2	-1.4	2.0	V
MglB	D-galactose-binding periplasmic protein	2.2	3.6	-1.6	2.8	-
FadA	3-ketoacyl-CoA thiolase	2.3	4.7	n.s.	<2.0	-
ActP	Cation/acetate symporter ActP	2.6	4.7	n.s.	<2.0	-

We analyzed the current annotations for their expression regulation and found that expression is either positively regulated by cAMP-CRP or ppGpp, or inhibited by phosphorylated DNA regulator AcrA (Keseler et al., 2017). An increase in Cra activity may enhance cAMP, ppGpp or phosphorylated AcrA levels and hence indirectly result in protein upregulation.

We also took a look at the 27 proteins that are significantly upregulated in *cra*^{FBD} as compared to wt *cra*, but with a lower fold change, between 1.3- and 1.9-fold (log2(*cra*^{FBD}/*cra*) = 0.34-0.92). Five of these proteins are known to be under Cra control: PoxB, Icd, PckA, AceA and AceB (Santos-Zavaleta et al., 2018). From the remaining 22 proteins, 77% are either reported to be stimulated by cAMP-CRP or ppGpp or inhibited by phosphorylated ArcA (Keseler et al., 2017). Further studies are required to verify if this correlation is direct, and more general what mechanisms upregulated non-Cra targets in the *cra*^{FBD} background. Cra and Cra^{FBD} protein levels were not significantly different, ruling out an effect of increased Cra activity because of increased Cra levels. Consistently, a fraction of the proteins upregulated in the *cra*^{FBD} strain showed a reverse pattern in the proteome comparison of *E. coli* wild type and the Δ*cra* mutant (Fig. 3i, Table 1).

Altogether, these studies demonstrated that Cra activity can be uncoupled from sensing metabolic fluxes by mutagenizing its F1P binding site.

4.3.3 Expression of *pdeL* correlates with Cra levels and activity

Based on our earlier observations, we hypothesized that Cra controls c-di-GMP distributions by controlling *pdeL* transcription (chapter 2). To test this, we first compared *pdeL* expression levels in tryptone broth (TB) and glucose minimal medium (MMG) using a transcriptional P_{pdeL}-lacZ reporter. While *pdeL* expression reached intermediate levels during growth of *E. coli* in TB (Fig. 4a), a basal expression level comparable to a Δ*cra* mutant was observed on glucose (Fig. 4b). To test if this was due to limiting Cra protein levels, *cra* and *cra*^{FBD} were expressed from an IPTG-inducible promoter on a plasmid (P_{lac}-*cra* / P_{lac}-*cra*^{FBD}). Increasing inducer concentrations lead to gradually increasing concentrations of Cra both on TB and on glucose (Fig. S3). In parallel, *pdeL* expression gradually increased in both media to similar levels observed in TB (Fig. 4a,b). This suggested that during growth on glucose, *pdeL* expression is limited by reduced Cra levels hence activity of Cra. In contrast, increasing Cra protein above wild type levels did not boost *pdeL* expression in TB (Fig. 4b).

The expression of *pdeL* is jointly controlled by Cra and PdeL, the latter of which responds to changing concentrations of c-di-GMP. We had shown earlier that in complex (TB) medium, the *pdeL* promoter is strongly activated by PdeL itself when c-di-GMP concentrations are very low or by PdeL mutant variants with abolished c-di-GMP control (chapter 2). In line with this, reducing the c-di-GMP levels (overexpression of *pdeH*) or expressing *pdeL* mutants strongly increased *pdeL* promoter activity in TB (Fig. 4c). In contrast, when cells were grown in MMG, *pdeL* promoter activity was only increased upon lowering c-di-GMP levels or when using a strain expressing the PdeL D295N variant, which has constitutive catalytic activity. The catalytically inactive E141A

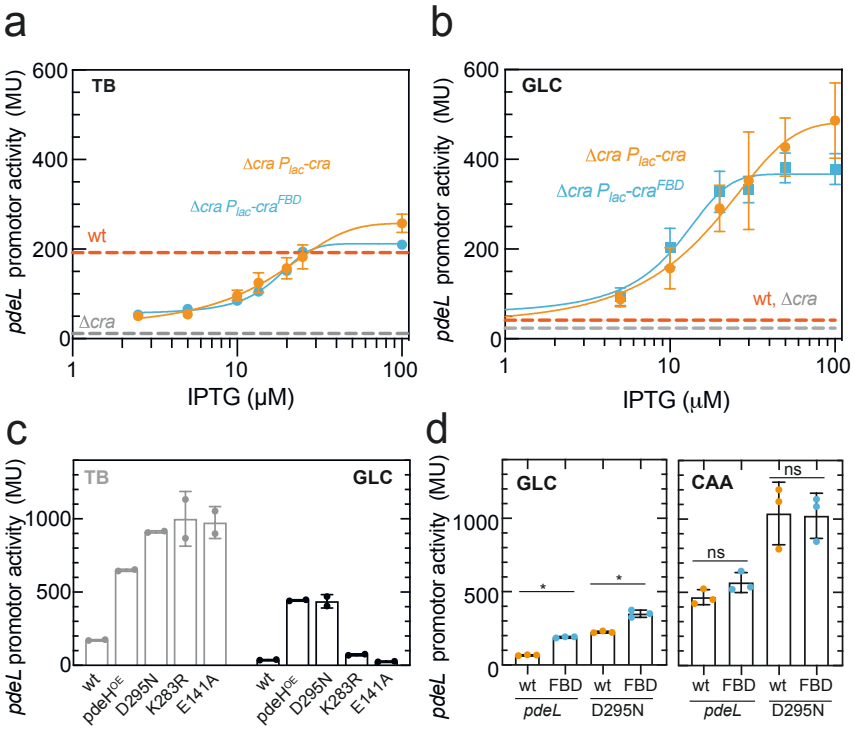


Figure 4: Low Cra levels and activity limit *pdeL* expression in glucose minimal medium. **(a,b)** *pdeL* promoter activity at increasing IPTG concentration during growth on **(a)** tryptone broth (TB) and **(b)** glucose minimal medium. Promotor activity was measured with a population P_{pdeL}-*lacZ* reporter in a wild type background (dashed red line), *Δcra* (dashed grey line) and in a *Δcra* background transduced with a low copy plasmid carrying the P_{lac}-*cra* construct (pAR218) (orange) or P_{lac}-cra^{FBD} (pAR213) (blue). **(c)** *pdeL* promoter activity of several *pdeL* mutants during growth on TB (in grey, see also chapter 2) and glucose minimal medium (black). *pdeL* expression was compared for the *pdeL* wild type background (wt), *pdeH* overexpression strain (*pdeH*^{OE}: Δ*pdeH* + P_{lac}-*pdeH* in medium supplemented with ampicillin and 100 μM IPTG), constitutively active *pdeL* mutant (D295N), *pdeL* with a mutated catalytic base that is responsible for the activation of the water molecule driving hydrolysis of c-di-GMP (K283R) and c-di-GMP binding deficient *pdeL* mutant (E141A). **(d)** *pdeL* promoter activity during growth on minimal medium supplemented with glucose (GLC) or casamino acids (CAA) as the sole carbon source. Expression of *pdeL* wt (*pdeL*) and constitutively active *pdeL* (D295N) in a *cra* (wt, orange) and *cra*^{FBD} (FBD, blue) background. Comparisons indicated with (*) are significantly different ($P < 0.05$). ns: not significant.

and K283R mutants were unable to increase *pdeL* expression on glucose (Fig. 4c). Thus, while a drop in c-di-GMP can boost *pdeL* promoter activity even under conditions where Cra levels and activity are low, this requires the full activation potential of PdeL. Apparently, this is provided by the D295N mutant, but not by the E141A and K283R variants.

Besides Cra protein levels, we investigated if the prevailing Cra DNA-affinity is limiting *pdeL* expression in MMG. As shown in Fig. 4d, *pdeL* promoter activity increased almost three-fold in a *cra*^{FBD} background as compared to a *cra* wild type strain. In contrast, Cra^{FBD} showed only a small stimulatory effect on *pdeL* expression in a complex medium containing casamino acids (CAA) (Fig. 4d). This can be explained by the higher activity of Cra under these conditions (Fig. 2c). The *pdeL* promoter activity was further increased in strains expressing the constitutive PdeL variant D295N, both in minimal and complex media. Together, this reinforced the notion that the *pdeL* promoter is jointly controlled by Cra and PdeL in response to metabolic cues and c-di-GMP, respectively.

4.3.4 Cra activity and *pdeL* expression inversely scale with *E. coli* growth rates

Cra was proposed to function as a metabolic flux sensor with Cra activity inversely correlating with glycolytic flux (Kochanowski et al., 2013). Hence, we set out to investigate if Cra activity and *pdeL* expression could be boosted by lowering the glycolytic flux. To modulate the glycolytic flux, we generated two strains that allowed tuning of the glucose uptake via the controlled expression of the glucose transporter PtsG (Basan et al., 2015; Litsios et al., 2018). For population level measurements, we engineered a Δ*ptsG* mutant expressing the *ptsG* gene from an IPTG-inducible promoter on a low copy number plasmid (pAR350) (P_{lac}-*ptsG*) (Fig. 5a). The doubling time on glucose of this strain inversely scaled with the IPTG concentration added to the medium both in the *cra* wild type and cra^{FBD} mutant background (Fig. 5b). For single cell measurements, we engineered a chromosomal induction system replacing the *ptsG* promoter region in the native locus with the anhydrotetracycline-inducible (aTc) P_{tet} promoter (Fig. 5a). Tuning *ptsG* expression with aTc showed similar inducer-dependent changes of growth rates, reaching a plateau at a slightly higher doubling time than wild type *E. coli* (Fig. 5c). Moreover, Cra activity scales with *ptsG* induction (Fig. 5d) and with overall growth rates (Fig. 5e) indicating that under glycolytic conditions Cra activity is a direct readout of growth rate. While Cra activity distributions at different *ptsG* induction levels were largely unimodal, the window between individual cells displaying maximal and minimal Cra activities gradually widened at declining growth rates (Fig. 5d).

To investigate how *pdeL* expression was modulated at different growth rates, we introduced the chromosomal *lacZ* reporter for the *pdeL* promoter activity into the Δ*ptsG* mutant harboring the P_{lac}-*ptsG* plasmid. When assaying *pdeL* promoter activity at different doubling times in MMG, we found that it gradually increased as growth rates declined (Fig. 5f). While *pdeL* expression was minimal, and similar to *E. coli* wild type, at maximal growth rates, *pdeL* promoter strength gradually increased under growth limited conditions eventually approaching levels similar to those observed in a cra^{FBD} background. In contrast, *pdeL* promoter activity did not change at different growth rates in cells expressing the constitutive Cra variant, Cra^{FBD}, with the exception of the highest induction level of *ptsG* at which *pdeL* expression also showed a slight reduction (Fig. 5f, Fig. S4). Altogether, these data demonstrated that *pdeL* expression is modulated by Cra as a function of the *E. coli* growth rate.

4.3.5 Cra determines binary distribution of c-di-GMP in *E. coli* populations

The experiments described above demonstrated that Cra is a key regulator of *pdeL*, limiting its expression during rapid growth on glucose but strongly activating the *pdeL* promoter upon reduced glucose availability. Given that *pdeL* expression is binary and that this stochastic control establishes binary c-di-GMP concentrations in populations of *E. coli* (chapter 2), we next wanted to investigate how Cra-mediated metabolic control influences c-di-GMP bimodality at the single cell level. To assess c-di-GMP concentrations in individual *E. coli* cells during growth on glucose, we expressed a GFP-based c-di-GMP biosensor together with a c-di-GMP-insensitive mScarlet-I from a single aTc-inducible operon (Fig. 6a). GFP and RFP intensities of individual cells from exponential phase were measured using flow cytometry during a set of preliminary experiments.

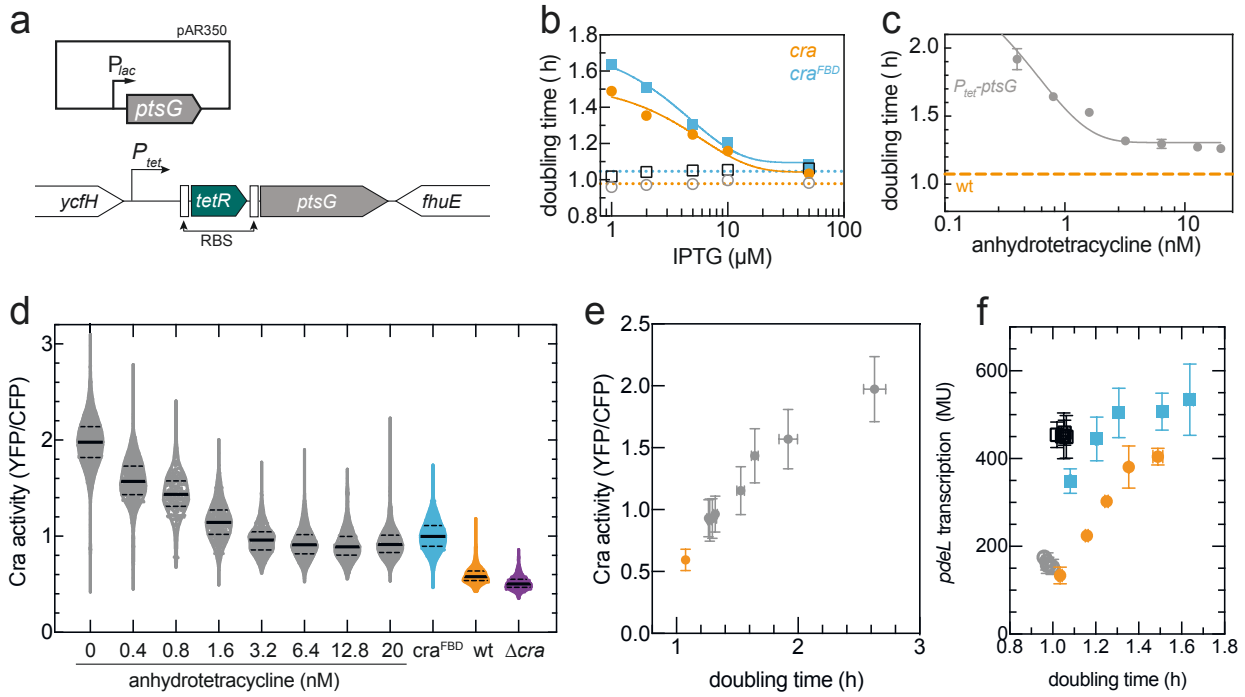


Figure 5: Cra activity and *pdeL* expression scale with glucose uptake flux. **(a)** Graphical lay-out of the constructs for tuning of *ptsG* expression. **Top:** expression of the *ptsG* gene from an IPTG-inducible promoter on a low copy number plasmid (pAR350) (P_{lac}-*ptsG*). **Bottom:** chromosomal P_{tet}-*ptsG* construct. The complete *ptsG* intergenic region was replaced by the anhydrotetracyline (aTc) inducible P_{tet} promoter, followed by a strong ribosomal binding site (RBS), the gene encoding the tetracycline repressor protein (*tetR*, green), and a weak RBS for the *ptsG* gene. **(b)** Growth rate on glucose minimal medium at increasing IPTG levels. The glucose transporter *ptsG* was expressed from a P_{lac}-*ptsG* construct on a low copy plasmid (pAR350) in a Δ*ptsG* (orange circles) and *cra*^{FBD} Δ*ptsG* (blue squares) strain background. Reference growth rates were determined for *E. coli* wt + EV (grey circles) and *cra*^{FBD} + EV (black squares). The average reference growth rate is indicated in broken lines for wt + EV (orange) and *cra*^{FBD} + EV (blue). EV: pNDM220 empty vector. **(c)** Growth rate (doubling time in h) as measured for P_{tet}-*ptsG* (grey) at increasing aTc concentration during growth on glucose. Orange broken line: growth rate of wt *E. coli* on glucose. **(d)** Cra activity distributions measured with the single cell Cra activity reporter plasmid (see Fig. 2a) for different *ptsG* induction levels in a P_{tet}-*ptsG* background (grey, construct as in **(a)**) and in a wt (orange), *cra*^{FBD} (blue) and *Δcra* (purple) background. Cultures were sampled during exponential growth on glucose and YFP and CFP fluorescence intensities were captured by microscopy. **(e)** Data from **(c)** and **(d)** combined showing average Cra activity as a function of the population growth rate (doubling time in h) on glucose for wt (orange) and the P_{tet}-*ptsG* strain induced at different aTc concentrations (grey). **(f)** *pdeL* promoter activity at increasing doubling time in glucose minimal media. *pdeL* promoter activity was measured in strains Δ*ptsG* + P_{lac}-*ptsG* (orange), *cra*^{FBD} Δ*ptsG* + P_{lac}-*ptsG* (blue), wt + EV (grey) and *cra*^{FBD} + EV (black) at increasing IPTG levels. The IPTG-growth rate correlation was derived from **(b)** and see also Fig. S4 for P_{pdeL}-*lacZ* – IPTG correlations.

As we had observed an over five-fold increase in *pdeL* promoter activity on glucose during *cra* expression induction (Fig. 4b), we first assessed c-di-GMP distributions in a Δ*cra* strain bearing the plasmid-based P_{lac}-*cra* construct. Preliminary data showed that without IPTG the c-di-GMP distribution was unimodal and c-di-GMP levels were high (Fig. 6b, Fig. S5). At increasing *cra* expression (Δ*cra* P_{lac}-*cra*, 100 μM IPTG), we observed that a small subpopulation of cells (0.4%) had shifted to a low c-di-GMP signal while the mScarlet-I signal was unaltered (Fig. 6b, Fig. S5). When comparing c-di-GMP distributions in a *cra* wild type and the *cra*^{FBD} background during growth on glucose, we observed a subpopulation (1.2%) that had shifted to low c-di-GMP in the latter (Fig. 6c, Fig. S6).

To measure the c-di-GMP distribution under increasing induction of the P_{tet}-*ptsG* construct, we modified the c-di-GMP biosensor expression system to avoid induction interference. While the expression of the original version was controlled by P_{tet}, in the new construct the biosensor

operon was inserted behind a constitutive promoter (Fig. 6a). When comparing preliminary data of c-di-GMP levels under conditions of low (0 nM) and high (20 nM) anhydrotetracycline induction representing different growth rates (Fig. 5d), we observed a small shift of the GFP signal towards lower GFP intensities but no change from a mono- to a bimodal distribution of c-di-GMP (Fig. 6d, Fig. S7). Even though the mScarlet-I signal did not change over the P_{tet}-*ptsG* induction range (Fig. S7), additional controls are required to exclude a growth rate dependent effect on the obtained signal from the biosensor. Besides, these data should be compared with c-di-GMP distributions in a Δ*pdeL* P_{tet}-*ptsG* background and complemented with single cell *pdeL* expression under increasing P_{tet}-*ptsG* induction.

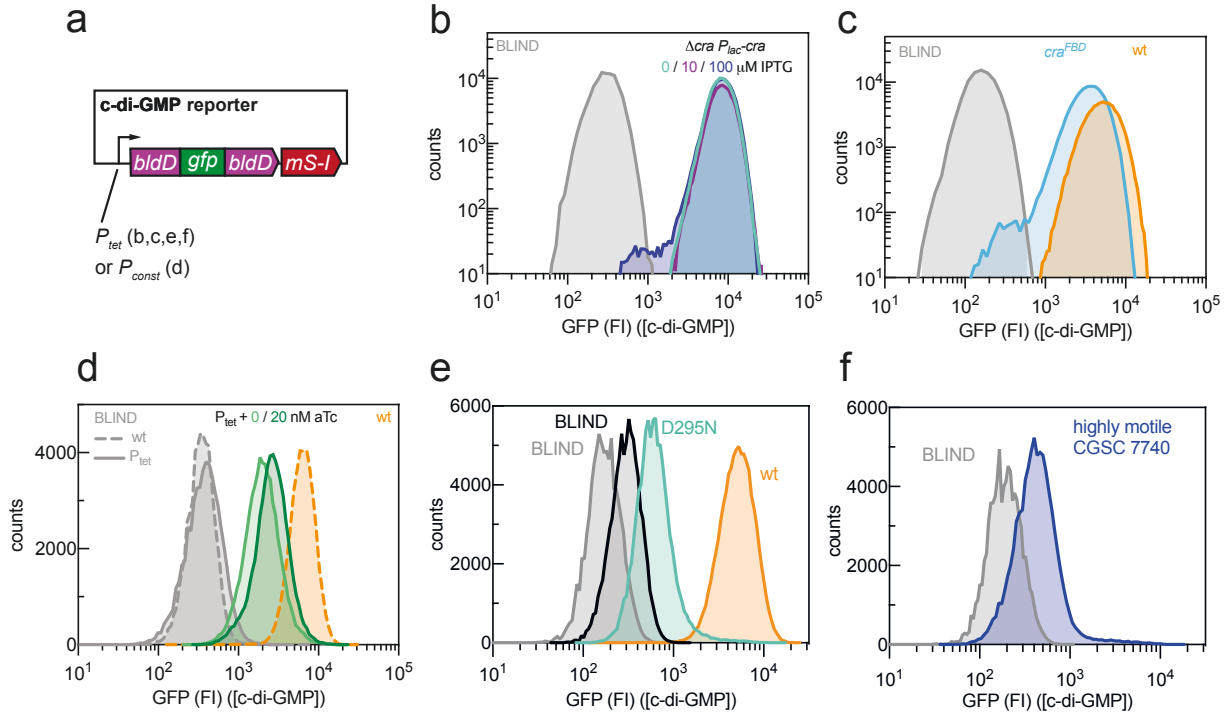


Figure 6: Increasing Cra levels or activity mediates bimodal c-di-GMP distribution during growth on glucose. Single cell c-di-GMP levels during exponential growth on glucose minimal medium were measured with a c-di-GMP biosensor expressed from a plasmid (p2H12ref), see **(a)**. The biosensor signal (GFP) positively correlates with [c-di-GMP]. The biosensor signal distributions are always shown in comparison to the GFP signal from a mutated biosensor that cannot bind c-di-GMP (p2H12ref-blind, in grey). Every graph shows the result after gating of 100'000 events. **(a)** Schematic of the c-di-GMP biosensor plasmid. Biosensor expression was induced with 200nM anhydrotetracycline (aTc), except for the data presented in figure **(d)**, for which a constitutively expressed sensor operon was used. **(b)** c-di-GMP distributions in a *Δcra* strain with *cra* expression induced from a P_{lac}-*cra* construct on a low copy plasmid (similar as in Fig. 4b). *cra* expression was either non-induced (green) or induced by supplementing the medium with 10 μM (purple) or 100 μM (dark blue) IPTG. **(c)** c-di-GMP distributions in a cra^{FBD} background (blue) as compared to wt *cra* (orange) and wt *cra* expressing p2H12ref-blind (grey). **(d)** c-di-GMP distributions at various *ptsG* expression levels. Expression of *ptsG* was induced from a chromosomal P_{tet}-*ptsG* construct (Fig. 5a) with aTc concentrations ranging from 0 (light green) to 20 (dark green) nM (see also Fig. S7 for intermediate aTc concentrations). The c-di-GMP biosensor was expressed from a constitutive promoter (construct P_{const}-p2H12ref) and the GFP signal was compared to wt *E. coli* expressing P_{const}-p2H12ref (wt, orange). Negative controls included expression of the c-di-GMP binding deficient biosensor (P_{const}-p2H12ref-blind) in strains P_{tet}-ptsG (0 nM aTc, solid grey line) and wt strain (broken grey line). **(e,f)** c-di-GMP distributions of wt (orange), *pdeL* (D295N) (green) and the *E. coli* K-12 MG1655 CGSC 7740 (blue) that contains additional *IS1* elements in the genome as compared to wt (Barker et al. 2004), resulting among others in constitutive expression of the c-di-GMP phosphodiesterase *pdeH*. For each graph, the mutated biosensor p2H12ref-blind was expressed in the same strain background as the functional c-di-GMP biosensor ("BLIND", grey (wt/ CGC7740) / black (*pdeL* (D295N)).

We have shown above that even at maximum growth rate in glucose minimal medium, conditions under which Cra activity is clearly limiting *pdeL* expression, lowering c-di-GMP levels or uncoupling PdeL from c-di-GMP control can boost the *pdeL* promoter significantly (Fig. 4b). To address if this also affects population distribution of c-di-GMP, we compared the c-di-GMP biosensor output in *E. coli* wild type with a strain harboring the *pdeL* (D295N) allele. Strikingly, while c-di-GMP distribution was monomodal and high in wild type, it switched to a monomodal low distribution in the *pdeL* (D295N) background (Fig. 6e, Fig. S6). The latter is similar to the monomodal c-di-GMP distribution observed in the hyper-motile MG1655 strain CGSC 7740, which harbors an *IS1* insertion in the *flhDC* promoter region causing constitutive expression of the PdeH phosphodiesterase (Fig. 6f, Fig. S6) (see chapter 3). In conclusion, although *pdeL* expression is limited by low Cra activity if *E. coli* grows on glucose, it can still be boosted by conditions that drive c-di-GMP concentrations low enough to fully activate PdeL (Fig. 4c).

4.4 Discussion

Expression of the c-di-GMP specific phosphodiesterase (PDE) PdeL is co-activated by Cra and PdeL and boosted under low c-di-GMP levels. Based on earlier observations, we hypothesized that the metabolic regulator Cra controls c-di-GMP distributions by controlling *pdeL* expression.

Here, we showed Cra to be a key regulator of *pdeL*. Cra limits *pdeL* expression during fast growth on glucose through limited Cra levels (Fig. 4b) and reduced Cra activity (Fig. 4d). Expression of *pdeL* could be activated:

1. in a Cra-dependent manner, by increasing Cra levels and Cra activity on glucose. C-di-GMP distributions remained however largely unaltered;
2. in a Cra-independent manner, by lowering c-di-GMP levels or abolishing negative c-di-GMP feedback control in *pdeL* (D295N). PDE activity of PdeL (D295N) was sufficient to strongly reduce c-di-GMP levels in *E. coli* populations grown on glucose.

Following these observations, a remaining key question about Cra is its role in the activation of *pdeL*. Binding of Cra to the *pdeL* promoter region increases the affinity for a PdeL dimer to bind the Cra-dependent PdeL binding box (CDB, see chapter 2). At low c-di-GMP levels, as mimicked by *pdeL* (D295N), is Cra required for full activation of *pdeL*? Would similar *pdeL* (D295N) expression (Fig. 4c) and c-di-GMP levels (Fig. 6e) be obtained in the absence of Cra, or is low, but residual Cra activity required? This could be investigated by measuring *pdeL* (D295N) transcription and c-di-GMP distributions in a Δ*cra pdeL* (D295N) mutant strain. The maximum *pdeL* expression, i.e. *pdeL* (D295N) expression, in the absence of Cra binding, is still largely unexplored. We propose two scenarios for possible experimental outcomes and a corresponding role of Cra in *pdeL* expression:

1. **Cra as prerequisite for *pdeL* expression**

In a first hypothesis, Cra activity would be a prerequisite for PdeL to be fully active and to drive binary signaling programs. If c-di-GMP measurements in a *pdeL* (D295N) Δ*cra* on glucose show a unimodal high distribution, this would indicate that residual Cra activity is required and it would support this hypothesis. During growth on gluconeogenic sources, a scrambled Cra binding site in the *pdeL* promoter region (CB^-, see chapter 2) could be used instead of a Δ*cra* mutant background, to miminize interference in metabolic processes. Also, here, when bimodal c-di-GMP patterns are lost in the absence of Cra activation during gluconeogenic growth, *pdeL* expression would rely on Cra first, independently of c-di-GMP.

2. **Cra as a dynamic element in *pdeL* activation**

In a second hypothesis, Cra and c-di-GMP may influence *pdeL* expression simultaneously and the sensitivity of PdeL as a switch protein may therefore be the result of their joint input. Cra activity, dependent on the carbon flow, may set the sensitivity of *pdeL* to stochastically define c-di-GMP levels. Moreover, it cannot be excluded that Cra may actively contribute to generating

bimodal c-di-GMP distributions, by cell to cell variation in its activity on the *pdeL* promoter. Since Cra binding largely affects PdeL binding to the *pdeL* promoter (chapter 2), fluctuations in Cra activity could subsequently be amplified through the PdeL- and c-di-GMP feedback loops working on *pdeL* expression.

Although still largely unimodal, we observed a shift in the c-di-GMP distribution when lowering the glycolytic flux (Fig. 6d). Besides, we measured a small subpopulation of low c-di-GMP in the cra^{FBD} background. It remains to be investigated whether PdeL is involved in these Cra- and flux-dependent effects on c-di-GMP distributions (Fig. 6b-d). C-di-GMP distribution measurements under elevated Cra activity and reduced growth should be repeated in a *pdeL* deletion mutant background to provide further insight.

4.4.1 C-di-GMP distributions on glucose

In this study, we made a first investigation how increased Cra activity during growth on glucose affects the c-di-GMP level in individual cells (Fig. 6a-c). Bimodal c-di-GMP distributions as otherwise seen on glycerol and gluconeogenic substrates did not appear. Instead, we measured a subpopulation of maximum 1% with a reduced GFP signal when increasing Cra activity directly through *cra* titration or replacing *cra* by cra^{FBD}. GFP distributions were unimodal, but with reduced fluorescence intensity as compared to wild type under reduced glycolytic flux through *ptsG* expression control. Further experiments and additional controls are required to draw conclusions about the mechanisms behind these observations. It is recommended to reproduce the flow cytometry measurements (Fig. 6b-d) and to complement these with fluorescence microscopy measurements of the c-di-GMP sensor. Microscopy data allow for a better normalization of the GFP (c-di-GMP) signal intensity by constitutive mScarlet-I intensity. For the *cra* induction and cra^{FBD} expression data, it should be verified whether cells with lower GFP intensity are characterized by lower mScarlet-I intensity or not. For c-di-GMP distributions in the P_{tet}-*ptsG* background, microscopy data would allow better normalization for the cell size. A reduction in growth rate also affects cell size (Volkmer and Heinemann, 2011). It should be verified that the shift in GFP signal is not a direct result of smaller cell size.

To address the role of Cra through *pdeL* on c-di-GMP, it would be crucial to include c-di-GMP measurements at increased Cra activity in a *pdeL* deletion background. In particular for the small c-di-GMP subpopulations and the shift in GFP intensity (Fig. 6a-c), measuring c-di-GMP in Δ*pdeL* cells (i.e. Δ*pdeL* Δ*cra* P_{lac}-cra; Δ*pdeL* cra^{FBD}; Δ*pdeL* P_{tet}-*ptsG*) will provide insight into whether this effect is *pdeL* dependent.

4.4.2 Cra protein levels regulate Cra activity

We showed that by inducing *cra* or cra^{FBD} expression, *pdeL* promoter activity increased up to five-fold on glucose (Fig. 4b). From this observation we concluded that *pdeL* expression is enhanced under elevated Cra protein levels. It was expected that Cra binding to the *pdeL* promoter

region is limited during growth on glucose. Even though the Cra binding site in the *pdeL* promoter region is among the tightest in the known Cra regulon (Shimada et al., 2005), it is expected that Cra targets are regulated in a condition-dependent manner and growth on glucose is a condition with relatively low Cra activity (Kochanowski et al., 2013).

Induction of *cra* and cra^{FBD} expression increased *pdeL* promoter activity similarly (Fig. 4a,b). These data emphasize the impact of the absolute Cra protein concentration, in addition to F1P modulation, on the overall Cra activity. Based on these findings, we hypothesize that the regulation of the Cra protein concentration could be an important, yet scarcely studied aspect of Cra activity regulation that deserves more attention in the future. So far, only little is known about the regulation of Cra protein levels: the expression of *cra* was reported to be under negative control of phosphorylated transcriptional regulator PhoB and dependent on the heat shock sigma factor σ^{32} (*rpoH*) (Keseler et al., 2017). We calculated Cra protein concentrations under various conditions based on published data on Cra copy numbers (Schmidt et al., 2016) and condition-dependent *E. coli* cell size (Volkmer and Heinemann, 2011). During growth in the absence of amino acids, the Cra protein concentration (Cra copies per cell volume) reversely correlates with growth rate (μ, (h^{-1})).

Since we induced *cra* expression in a wild type *pdeL* background (Fig. 4a,b) and c-di-GMP inhibits *pdeL* expression (chapter 2), it cannot be excluded that part of the increment of *pdeL* promoter activity is due to a reduction in the level of c-di-GMP. Intriguingly, *pdeL* promoter activity on glucose at increasing *cra* expression plateaus at an activity level similar to the level reached under low c-di-GMP or in a constitutively active *pdeL* (D295N) mutant background (Fig. 4b,c). In order to dissect the contributions of Cra and c-di-GMP on *pdeL* promoter activity (Fig. 4a,b), additional measurements are advised in a *pdeL* (D295N) background under increasing *cra* expression.

4.4.3 A single cell Cra activity readout

Because Cra-DNA affinity is modulated by the allosteric regulator F1P (Bley Folly et al., 2018; Ramseier et al., 1993), measuring the promoter activity of a Cra-regulated gene to estimate Cra activity is superior to using the promoter activity from the *cra* gene itself. We assessed Cra activity on the single cell level using a similar reporter system as described before (Lehning et al., 2017). The functioning of the single cell Cra activity reporters is based on the same principles (Kochanowski et al., 2013), but we also noted differences in the final construct and measurement approach. In this study, the Cra activity reporter plasmid was based on the Cra-repressed *pykF* promoter (identical to Kochanowski et al., 2013) that controlled *cfp* (P_{pykF}-cfp) and yfp (P^*_{pykF}-yfp) expression (Fig. 2a). In contrast, the single cell reporter by Lehning et al. (2017) is based on the Cra-activated *ppsA* promoter controlling *gfp* (P_{ppsA}-*gfp*) and *rfp* (P^*_{ppsA}-rfp) expression from a single plasmid. Lehning et al. found that P_{ppsA}-controlled *gfp* expression showed a higher dynamic GFP intensity range between different carbon sources as compared to P_{pykF}-*gfp*. Hence the *ppsA* promoter was preferred over *pykF*. Importantly, Lehning et al. measured fluorescence intensity by flow cytometry. This is an understandable choice, since they used the

Cra activity reporter as a high-throughput solution to substitute metabolic flux analysis. Lehning et al. applied the reporter to screen and select for mutants with adapted metabolic properties, hence a high-throughput method as flow cytometry would indeed be required. In this study, we used the single cell Cra activity for analysis rather than for screening and selection. We used fluorescence microscopy to assess Cra activity distributions within isogenic populations and compared Cra activity between different strains and conditions (Fig. 2, Fig. 5d). We made a test run analyzing the CFP and YFP intensities by flow cytometry, as this would allow for a high throughput assessment. However, flow cytometry was found to be insufficiently sensitive to measure differences in the signal across conditions or strains (data not shown). By means of fluorescence microscopy, we could well measure and distinguish condition-dependent differences in the Cra activity reporter readout (Fig 2, Fig. 5d). Hence, all our Cra activity reporter data have been acquired by means of fluorescence microscopy.

4.4.4 Cra (R197A) is a F1P binding deficient Cra mutant

We rationally designed and constructed a Cra mutant that lacks the ability to bind its allosteric regulator Fructose-1-phosphate (Fig. 3a-d; Bley-Folly et al., 2018; Ramseier et al., 1993). Mutation of a single residue, R197A, that was predicted to be one of the F1P interacting residues (Fig. 3b), was sufficient to abolish F1P binding *in vitro* at relevant F1P concentrations. We compared the Cra and Cra^{FBD} denaturation temperatures and DNA band shift assays in the presence of low millimolar F1P concentrations, a range similar to other *in vitro* studies on Cra (Bley-Folly et al., 2018; Ramseier et al., 1993,).

The assessment of the interaction between Cra and fructose-1,6-bisphosphate (FBP) (Bley-Folly et al., 2018) was published at the time we performed our *in vitro* Cra^{FBD} studies. For several decades it was generally considered that FBP also functions as physiological effector of Cra in *E. coli* (Kim et al., 2018; Kochanowski et al., 2013, 2017; Kotte et al., 2010; Lehning et al., 2017; Ramseier et al., 1993, 1995; Wei et al., 2016). Allosteric regulation of Cra was again addressed by Bley Folly et al. (2018), who used an experimental set-up with various techniques to exclude FBP as a direct allosteric regulator of Cra. This simplified our studies as such, that we no longer included FBP binding studies of Cra and Cra^{FBD}. Although it remains to be elucidated how the metabolite F1P is produced in the absence of Fructose, the evidence on *in vitro* binding of F1P to the *E. coli* Cra protein is clear (Bley Folly et al., 2018; Ramseier et al., 1993). Our studies of Cra^{FBD} clearly show that regulation of Cra by F1P is relevant *in vivo*. Loss of F1P binding enhanced the activity of this pleiotropic regulator during growth on glucose, as could be observed on the transcription level and protein level of Cra targets (Fig 2d, 3h).

There is previous work on semi-rational design of Cra mutants (Wei et al., 2016), but with the opposite aim to increase the ligand affinity (for FBP, as the research was performed prior to the FBP binding assessment by Bley-Folly et al. (2018)). In this study, we provided evidence for the loss of F1P binding of the Cra mutant Cra (R197A). To our best knowledge, a Cra protein mutant unable to respond to its physiological regulator was not constructed before. In this Cra^{FBD} mutant, Cra activity is uncoupled from its allosteric regulator. It would certainly be interesting to

follow up on the *in vivo* data presented here by performing a metabolic flux analysis in the *cra*^{FBD} background on different carbon sources. Together with more detailed measurements on growth yield, flux analysis may shine further light on why the *cra*^{FBD} strain shows a growth effect both on glucose as well as on gluconeogenic substrates (Fig. 3g, Fig. S2). In future studies on metabolism, Cra^{FBD} could be used to interfere with native *E. coli* metabolic regulation and among others be of value in further studies on *cra* expression regulation.

4.4.5 Control over glycolytic flux via glucose uptake system

To reduce the glycolytic flux during growth on glucose, we decided to control the expression of the glucose uptake system, in this case the glucose transporter *ptsG*. By varying the inducer concentration (IPTG for P_{lac}-*ptsG* and anhydrotetracycline for P_{tet}-*ptsG*), we modulated the glucose influx, hence glycolytic flux (Basan et al., 2015; Litsios et al., 2018). Inducible expression of *ptsG* and other genes encoding components for specific substrate uptake systems was previously used to establish reduced uptake fluxes (Basan et al., 2015; Okano et al., 2020). An alternative set-up would be chemostat cultivation of strains containing their wild type substrate uptake regulation. During chemostat cultivation, the growth rate is set by the dilution rate. Instead of mimicking reduced substrate availability, the glycolytic flux would be limited because of an actual limitation of the amount of glucose available for uptake per time. A chemostat set-up is widely used to study *E. coli* metabolism and physiology including system-wide proteome allocation under reduced substrate or energy availability (Kotte et al., 2014; Nanchen et al., 2008; Renilla et al., 2012; Rijsewijk et al., 2011; Schmidt et al., 2016; Taymaz-Nikerel et al., 2009). Glucose-limited chemostat cultivation was previously also used to draw the correlation between Cra activity and glycolytic flux (Kochanowski et al., 2013). Depending on the experimental requirements, an inducible substrate uptake system or substrate limited continuous cultivation could be preferred. As we only modulated glucose uptake, but under multiple conditions, the use of *ptsG*-inducible strains met our experimental requirements well. Because we aimed to measure single cell Cra activity under inducible *ptsG*, measurements were however more sensitive to differences in *ptsG* induction between cells as compared to measuring population averages. To avoid differences in *ptsG* induction between cells as much as possible, we therefore combined the Cra activity reporter plasmid with a chromosomal P_{tet}-*ptsG* induction construct. It would however still be interesting to measure *pdeL* expression and c-di-GMP distributions under limitation of glucose or other substrates in a chemostat and to compare the results with the data presented here.

4.4.6 Choice of MG1655 strain backgrounds in the context of c-di-GMP regulation and glycerol metabolism

A note has to be made on the use of the *E. coli* MG1655 strain backgrounds in this study. The validation of the different tools to increase Cra activity including all Cra activity reporter measurements was performed in the highly motile *E. coli* K-12 MG1655 background (Blattner et al., 1997), that is genetically different from the MG1655 stock we used to assess

c-di-GMP distributions (Fig. 6a-e). As c-di-GMP distributions were observed to be bimodal on gluconeogenic sources in the CGSC 6300 stock only (chapter 2), or at least require the absence of an *IS1* element upstream of the *flhDC* operon (chapter 3), the CGSC 6300 background was used for the c-di-GMP distribution measurements under increased Cra activity (Fig. 6a-e).

Among others, genetic regulation of the glycerol specific dissimilation pathway (Fig. 1) is different between the two MG1655 stocks. This is due to a loss of function mutation in the *glpR* gene (Freddolino et al., 2012) in the highly motile strain and CGSC 7740 stock (Blattner et al., 1997). GlpR functions as the specific repressor of glycerol catabolism operons including *glpFKX* and *glpD* (Fig.1). As a result, the products of these operons were found to be elevated in the absence of glycerol (chapter 3). It should be taken into account that the loss of *glpR* function might affect carbon metabolism and possibly Cra activity. In *Pseudomonas putida*, expression of the glycerol catabolizing pathway is similar as in *E. coli* (Kim et al., 2013). The regulation of GlpR and genetic wiring of catabolic enzymes causes metabolic bimodality and persister formation in *P. putida* (Nikel et al., 2015). In *E. coli*, bimodal c-di-GMP distributions on glycerol were restored in CGSC 7740 upon restoring *flhDC* or *pdeH* deletion, suggesting that the c-di-GMP distribution is independent of the genetic regulation around the glycerol metabolizing enzymes (chapter 3). It is still recommended to additionally assess Cra activity during growth on glycerol in the CSGC 6300 background, to verify the Cra activity distribution and draw conclusions on the interplay between metabolism and c-di-GMP heterogeneity.

4.4.7 Conclusions on substrate-dependent c-di-GMP distributions

We focused in this chapter on the role of Cra in setting c-di-GMP distributions. Cra activity is relatively low during growth on glucose, when the c-di-GMP distribution is unimodal high. During growth on glycerol, Cra activity is higher (Fig. 2c) and c-di-GMP distributions were found to be bimodal and dependent on *pdeL*, whose expression is positively regulated by Cra (chapter 2). We have investigated the effect of increasing Cra activity on c-di-GMP distributions, but have not looked at the effect of heterogeneity of Cra activity. This would be of relevance, as the evidence is growing for heterogeneity in metabolism in both prokaryotic and eukaryotic systems (Litsios et al., 2018; Martins and Locke, 2015; Vermeersch et al., 2019). What further requires attention is how exactly Cra activity and c-di-GMP distributions correlate. When comparing growth on glucose and glycerol, many other changes can be noted in addition to a change in Cra activity. Moreover, our readout of the cellular c-di-GMP concentration is the net result of the activity of potentially multiple diguanylate cyclases (DGCs) and c-di-GMP specific phosphodiesterase (PDEs). The *E. coli* K-12 MG1655 genome encodes 12 DGCs and 13 PDEs (Hengge et al., 2015) and for many of them the exact regulation is not fully known. The factors that we describe below are not necessarily expected to play a role in setting c-di-GMP distributions, but are rather meant to point out additional differences in cellular physiology between growth on glycerol and glucose. It would be of interest to verify if and how they affect c-di-GMP levels.

4.4.7.1 Cyclic AMP

In the absence of glucose, synthesis of the signaling molecule cyclic AMP (cAMP) is induced in *E. coli* cells (Kremling et al., 2015; Notley-McRobb et al., 1997). cAMP in complex with the CRP protein activates transcription of genes encoding metabolic enzymes for catabolism of alternative carbon sources. cAMP furthermore activates non-metabolic systems including flagellar expression in *E. coli* (Keseler et al., 2017). Although little is known so far about crosstalk between c-di-GMP and cAMP in *E. coli*, it cannot be excluded that altered cAMP levels directly or indirectly affect c-di-GMP regulation and thus the c-di-GMP concentration. Interplay between c-di-GMP and cAMP has been observed in other species. In *P. putida*, the cAMP-CRP complex represses transcription of the diguanylate cyclase *gcsA*, that is involved in *P. putida* swimming motility (Xiao et al., 2020). In pathogenic species cross regulation between the two signaling molecules was also identified. In pathogenic *Leptospira* species, cAMP functions as allosteric activator of a diguanylate cyclase activity by cAMP binding to the GAF domain adjacent to the GGDEF domain required for diguanylate cyclase catalytic activity (Vasconcelos et al., 2017). In *Vibro cholerae*, cyclic AMP negatively regulates biofilm formation through c-di-GMP (Fong and Yildiz, 2008; Heo et al., 2019). As the physiological role of cAMP across species is diverse, no direct link can be made to similar mechanisms in *E. coli*.

4.4.7.2 Growth rate

A difference between growth on glucose and glycerol one directly observes is the growth rate. The growth rate on glycerol is lower as compared to glucose. This directly affects the dilution rate of c-di-GMP regulating proteins as well as c-di-GMP molecules. Growth rate was found to impact cellular processes including the most fundamental parts of a living cell: protein expression and proteome allocation (Hui et al., 2015; Schmidt et al., 2016). These system-wide datasets on condition-dependent protein levels could at least present insight in the relative abundance of diguanylate cyclases and c-di-GMP dependent phosphodiesterases.

4.5 Materials and Methods

Strains and plasmids

The bacterial strains and plasmids used in this study are listed in Table S1 and S2 respectively. Throughout this study, *E. coli* strain K-12 MG1655 (Blattner et al., 1997) and the MG1655 stock CGSC 6300 from the Coli Genetic Stock Center (Guyer et al., 1981) were used as the strain backgrounds. The single cell Cra activity reporter plasmid was a kind gift from Matthias Heinemann. The single cell c-di-GMP biosensor plasmids were kindly provided by Andreas Kaczmarczyk.

λ-Red-mediated chromosomal engineering

Chromosomal deletions and modifications were carried out by means of λ-Red recombineering. We essentially followed the method for pKD46-mediated recombineering as described (Datsenko & Wanner 2000). In short, a recipient strain transformed with pKD46 was grown overnight in LB supplemented with ampicillin at 30°C. The next morning, LB culture was 1:100 diluted in 50 ml LB medium without NaCl supplemented with ampicillin. Culture was grown up to OD_{600} of 0.3-0.4 and expression of genes from pKD46 was induced for 50 minutes with 0.24% L-arabinose. After induction, culture was cooled on ice for at least 30 minutes prior to centrifugation. Pellet was washed 4x with ice cold 10% glycerol. Cell suspension was transformed with linear PCR product by means of electroporation. Phenotypic expression was allowed for at least two hours by growth on LB at 30°C before plating.

P1 lysate preparation and transduction

P1 phage lysate preparation and transduction were carried out as described in (Miller, 1992). In short, P1 lysates were prepared by addition of 10 μL MG1655 starter lysate to the desired strain grown in LB supplemented with 10 mM $CaCl_2$ to and OD_{600} of 0.1-0.2 at 37°C. Incubation was continued at 37°C for circa 3-4 hours until cell lysis was clearly visible. Chloroform was added to the culture before centrifugation. 3-4 ml of the supernatant was stored with chloroform in a glass vial at 4°C. For P1 transduction, recipient strain was grown in LB supplemented with 10 mM $CaCl_2$ at 37°C until an OD_{600} of 0.3-0.5. 1 mL of the culture was sampled and mixed with 100 μL of the desired P1 lysate and incubated for 15 minutes at room temperature. Directly afterwards, 100 μl 1M sodium citrate was added and the sample was incubated for at least 1h at 37°C for phenotypic expression before plating on LB Agar supplemented with 20 mM sodium citrate and the appropriate antibiotic. Single colonies appearing were restreaked three times on the same type of plate. Correct insertion of P1 transduced DNA sequence was confirmed by PCR and Sanger sequencing.

Growth Media

LB (Luria-Bertani) medium was used for growth of pre-cultures and for strain construction unless stated otherwise.

LB Agar plates consisted of LB medium supplemented with 1.5% agar (Agar Bacteriology Grade – PanReac AppliChem).

TB (Tryptone Broth) medium consisted of 10 g/l tryptone (BD Bacto Tryptone) and 5 g/l NaCl.

Minimal medium consisted of M9-based minimal medium (Gerosa et al. 2013). Minimal medium was supplemented with sterile, concentrated stock solutions of carbon sources where required adjusted to pH 7. Carbon sources were used at final concentrations of 0.5% glucose, 0.4% glycerol, 0.5% casamino acids (CAA), 0.5% galactose, 0.5% fructose, 0.2% alpha-ketoglutarate and 0.2% fumarate in minimal medium.

When needed antibiotics were supplemented at the following final concentrations: to LB, 100 μg/ml ampicillin and 30 μg/ml kanamycin; to TB overnight cultures, 100 μg/ml ampicillin and 30 μg/ml kanamycin; to TB day cultures, 40 μg/ml ampicillin and 30 μg/ml kanamycin; to glucose, 40 μg/ml ampicillin for low copy plasmids (pNDM220), 100μg/ml ampicillin for high copy plasmids (p2H12ref) and 30 μg/ml kanamycin. The inducers IPTG (in ddH_2O) and anhydrotetracyline (in DMSO) were added to final concentration as indicated in each experiment. Concentrated antibiotic stock solution and inducer solutions were added to the medium just before use.

Cultivation conditions

M9-based minimal medium was always inoculated with liquid pre-cultures grown in LB medium. Single colonies from LB Agar plates were used to inoculate 5 ml aliquots of LB medium in glass tubes. LB cultures were grown for 6-7h at 37°C. For growth in glucose minimal medium, grown LB cultures were diluted directly 1:200 in 5 ml M9-based minimal medium supplemented with 0.5% glucose. Glucose cultures were incubated overnight at 37°C and again diluted 1:500 in 5 ml fresh medium and grown for 6-7h to obtain exponentially growing cultures at an OD_{600} 0.3-0.5. Induction of gene expression (P_{lac}-*cra*, P_{lac}-*ptsG*, P_{tet}-*ptsG*, p2H12ref(-blind)) was always initiated in the overnight culture in minimal medium and continued after dilution the next morning in fresh medium at the same inducer concentration. For growth in minimal medium supplemented with CAA (Fig. 2c, 4d), a similar procedure to glucose minimal medium was followed. Day cultures required less time to reach an OD_{600} of 0.3-0.5, at which cultures were sampled.

For growth in minimal medium supplemented with other carbon sources, cultivation conditions were adapted to growth characteristics on each carbon source. Cultivation always started with an LB pre-culture followed by an overnight culture in minimal medium supplemented with the specific substrate. Depending on the turbidity of the culture the next morning, cultures were diluted or growth of the overnight culture was extended until cultures reached turbidity in exponential phase.

TB medium was inoculated with single colonies from LB Agar plates and cultures were incubated overnight at 37°C. Overnight cultures were diluted 1:500 in fresh TB medium and grown for another 4h to an OD_{600} of 0.4-0.5.

Continuous optical density measurements for growth comparison or growth rate determination were performed in 96-well plates incubated at 37°C in an Epoch2 or Synergy2 plate reader (Biotek Instruments Inc.) controlled with Gen5 software (version 3.10, Biotek Instruments Inc.). Shaking speed was set to maximum, double-orbital shaking. Plates were incubated at 37°C. OD_{600} was measured with an interval of 15 or 30 minutes.

Beta-galactosidase assays

Cultures were grown up to exponential phase in TB medium or minimal medium and optical density (OD_{600}) was measured. 500 μL of culture was mixed with 380 μL Z-buffer (75 mM Na_2HPO_4, 40 mM NaH_2PO_4, 1 mM KCl, 1 mM $MgSO_4$), 100 μL 0.1 % SDS and 20 μL chloroform. Samples were vortexed for 10 sec and left on the bench for 15 min. 200 μL sample were transferred into a clear 96-well plate. As substrate, 25 μL 4 mg/mL 2-nitrophenyl- β-D-galactopyranoside (σNPG) dissolved in Z-buffer were added. The initial velocity of the color reaction was determined at a wavelength of 420 nm.

cra expression assay

A Δ*cra* strain (AB3472) harboring the P_{lac}-*cra-3xFlag* plasmid (pAR218) or a *cra-3xFlag* strain (AB2292) were grown up to exponential phase in TB medium or glucose minimal medium supplemented with ampicillin and different IPTG concentrations if required. Cells were harvested and normalized to the same OD_{600} in SDS loading buffer (2X concentrated final composition: 100 mM Tris-HCl pH 6.8, 50% glycerol, 4% sodium dodecyl sulfate (SDS), 3% DTT, 0.2% Bromophenol blue). Samples were boiled for 15 minutes and stored at -20°C until further use. Samples were separated on a 12.5% acrylamide (SDS-PAGE) gel and transferred onto PVDF membranes (Immobilon-P, Millipore, USA). Membrane was 3x blocked with fresh 1x PMT (PBS supplemented with 5% milk powder and 0.01% Tween20) before addition of primary antibody (1:10'000 mouse anti-Flag). After incubation with the secondary antibody (1:10000 rabbit anti-mouse-HRP), membranes were incubated with LumiGLO (KPL, USA). Proteins were visualized with enhanced chemiluminescence (ECL) detection reagent and imaged in a gel imager (GE ImageQuant LAS 4000). The mean grey scale values of bands on the immunoblot images were quantified with ImageJ.

Microscopy

Cells were placed on a PBS pad solidified with 1% agarose. Images were acquired using softWoRx 6.0 (GE Healthcare) on a DeltaVision Core (Applied Precision) microscope equipped with an Olympus 100X/1.30 Oil objective and an EDGE/sCMOS CCD camera. Exposure time for microscopy frames was 0.05 sec for bright field (POL), 0.2 sec for CFP and 0.2 sec for YFP. For all settings, the ND filter was set to 100% transmission.

Fluorescent Image Processing

Image analysis was performed with custom made MATLAB (versions 2018 and 2019) scripts combined with the open-source platform SuperSegger (Stylianidou et al., 2016).

Statistics

Unpaired Two-Samples Wilcoxon Test was performed using R (version 4.1.0) to calculate P-values for *pdeL* expression measurements. An alternative, one-sided hypothesis was used to assess whether *pdeL* expression in a cra^{FBD} background is increased as compared to *cra*.

Protein sequence alignment

Full length protein sequences of the *E. coli* Cra protein (P0ACP1), *P. putida* Cra protein (Q88PQ6) and *E. coli* Lactose operon repressor LacI (P03023) were retrieved from the UniProt Knowledgebase (The Uniprot Consortium, 2021) and aligned using the multiple sequence alignment (MSA) tool MUSCLE (Madeira et al. 2019 Nucl Acids Research). The MSA output format was set to 'ClustalW (strict)'. For all other parameters, default settings were used. The visualization of the ClutalW sequence alignment including sequence logo (Fig S1b) was created in Geneious Prime (version 2019.0.4, Biomatters Ltd.).

Protein structure alignment / superimposition

Structural homology of the *E. coli* Cra protein to other known protein structures was assessed with a biosequence analysis using phmmer (Potter et al., 2018; https://www.ebi.ac.uk/Tools/hmmer/search/phmmer). Settings: the full-length *E. coli* Cra protein sequence (UniProt: P0ACP1) was used as query and compared to the PDB database. All taxa were included in the search. Significance E-values were set to 0.01 for sequence and 0.03 for hit. Report E-values were set to 1 for both sequence and hit.

Here, the top-2 results included Cra protein of *E. coli* K-12 (PDB: 2iks) and Cra protein of *P. putida* KT2440 (3o74).

PDB 2iks and the Cra protein of *P. putida* KT2440 in complex with fructose-1-phosphate (PDB: 3o75) were superimposed in the UCSF ChimeraX software (version 1.1.1) (Pettersen et al., 2021) using α-carbons of the structures. The structures are very similar as demonstrated by the low rmsd (1.6 angstroms with 247 Calpha atoms).

Final superimposition (Fig. 3b, S1a) contains residue 59-334 of *E. coli* Cra and residue 60-329 of *P. putida* Cra.

Protein purification

PdeL-strepII, Cra-strepII and Cra (R197A)-strepII were purified by StrepII-tag purification followed by a heparin purification step.

Protein overexpression: All proteins were cloned into a pET28a vector (Novagen) and overexpressed in strain BL21 (AI). BL21 (AI) strains harboring the plasmid with the StrepII-tagged protein were grown overnight at in 20 ml LB (for cultivation temperature, see Table 2). Overnight cultures were diluted 1:100 in 2 L LB in a 5 L Erlenmeyer and incubation continued in a shaking incubator (Multitron, INFORS HT) pre-warmed at the same temperature as overnight culture

(Table 2). At an OD_{600} of 0.6, the culture was induced with 0.1 % L-arabinose. Cells were harvested 4 h after starting induction by centrifugation at 6000 g for 30 min at 4°C. The pellet was resuspended in 1x PBS and centrifuged at 4000 rpm for 20 minutes. The pellet was flash-frozen in liquid nitrogen and stored at -20°C until further use.

StrepII-tag affinity purification: The cell pellet was resuspended in 7 mL Buffer A (see Table 2) including a tablet of cOmplete mini EDTA-free protease inhibitor and a spatula tip of DNaseI. Cells were lysed by three to four passages of French press and the lysate was cleared at 4°C in a table-top centrifuge set at full-speed for 40 min. 1 mL Strep-Tactin Superflow Plus (Qiagen) was loaded on a bench-scale column and equilibrated in 10 mL Buffer A. The cleared supernatant was loaded on the equilibrated column. The supernatant was reloaded another two times before washing with Buffer A followed by washing with Buffer B (see Table 2). 500 μL aliquots of proteins were eluted with Buffer B supplemented with 2.5 mM d-Desthiobiotin. Protein concentration in elution fractions was estimated by mixing 3 μL from each fraction with 100 μL 1x Bio-Rad Protein Assay Dye Reagent (Bio-Rad).

Heparin purification and dialysis: A 1 mL HiTrap Heparin HP column (GE Healthcare) was washed with 10 mL ddH_2O, followed by an equilibration with 10 mL Buffer B. The eluate from the StrepII-tag affinity purification was loaded three times. After loading, the column was washed with 10 mL Buffer A (PdeL-strepII) or 20 ml Buffer B (for Cra-strepII and Cra (R197A)-strepII) followed by a second washing step with 10 mL Buffer C (Table 2). The protein was eluted in 500 μL fractions with a total of 10-20 mL Heparin elution buffer (Table 2). Protein concentration per fraction was estimated as described in StrepII-tag purification step. Purified protein fractions were pooled and transferred to a Dialysis Membrane (MWCO 12'000-14'000 Daltons, Spectra/Por). Purified protein solution was dialyzed overnight at 4 °C under constant stirring against 1.5 L dialysis & storage buffer (Table 2). The final protein concentration was recorded at 280 nm and the content of co-purified nucleotide contaminants determined as a ration of 260/280 nm.

Electrophoretic mobility shift assay (EMSA)

5' Cy3-labeled input DNA was generated either via oligonucleotide annealing or PCR. For oligonucleotides used see Table S3. 10 nM of the input DNA and purified proteins were incubated for 10 min at room temperature in buffer consisting of 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM $MgCl_2$, 10 % Glycerol, 1 mM DTT, 0.01 % Triton X-100, 0.1 mg/mL BSA and 25 μg/mL λ-DNA. Samples were run on 8 % polyacrylamide gel. DNA-protein complexes were analyzed using Typhoon FLA 7000 (GE Healthcare).

Thermal Shift Assay (TSA)

Thermal denaturation of Cra-strepII and Cra (R197A)-strepII was monitored by extrinsic dye fluorescence in a Rotor-Gene Q thermocycler (Qiagen) combined with Rotor-Gene software (version 2.1.0.9). Samples were prepared containing 8 μM (final concentration) Cra-StrepII or

Table 2: Buffer composition for protein purification of PdeL-strepII, Cra-strepII and Cra (R197A)-strepII.

Protein construct:	**PdeL-strepII**	**Cra-strepII and Cra (R197A)-strepII**
Plasmid:	pAR1	pAR19 (wt), pAR194 (R197A)
BL21 cultivation and protein overexpression:	At 30°C	At 37°C
Buffer A	100 mM Tris-HCl pH 8.0 250 mM NaCl 5 mM $MgCl_2$ 1 mM DTT **0.5 mM EDTA**	100 mM Tris-HCl pH 8.0 250 mM NaCl 5 mM $MgCl_2$ 1 mM DTT
Washing volume Buffer A	50	-
Buffer B	100 mM Tris-HCl pH 8.0 50 mM NaCl 5 mM $MgCl_2$ 1 mM DTT **0.5 mM EDTA**	100 mM Tris-HCl pH 8.0 50 mM NaCl 5 mM $MgCl_2$ 1 mM DTT
Washing volume Buffer B	10 ml	100 ml
Buffer C	100 mM Tris-HCl pH 8.0 **350 mM NaCl** 5 mM $MgCl_2$ 1 mM DTT **0.5 mM EDTA**	100 mM Tris-HCl pH 8.0 **100 mM NaCl** 5 mM $MgCl_2$ 1 mM DTT
Heparin elution buffer	100 mM Tris-HCl pH 8.0 **2 M NaCl** 5 mM $MgCl_2$ 1 mM DTT **0.5 mM EDTA**	100 mM Tris-HCl pH 8.0 **1 M NaCl** 5 mM $MgCl_2$ 1 mM DTT
Dialysis & storage buffer	100 mM Tris-HCl pH 8.0 250 mM NaCl 5 mM $MgCl_2$ 1 mM DTT **0.5 mM EDTA**	100 mM Tris-HCl pH 8.0 250 mM NaCl 5 mM $MgCl_2$ 1 mM DTT

Cra (R197A)-StrepII, 1X SYPRO orange and 0 to 2 mM Fructose-1-phosphate in a final volume of 50 μL dialysis & storage buffer (100 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM MgCl2, 1 mM DTT). Measurements were performed in triplicates. The average denaturation temperature was taken as the reference to determine denaturation temperature shift.

Proteomics

The proteome was analyzed from five biological replicates. A volume of 1 ml was sampled from *E. coli* cultures grown in glucose minimal medium to an OD_{600} of 0.3. Sample volume was adjusted in case of a slightly higher/lower OD_{600} to sample similar number of cells. Samples were directly put on ice, pelleted and washed with 1 mL ice-cold Phosphate-buffered saline (PBS) (8 g/l NaCl, 0.2 g/l KCl, 1.44 g/l $Na_2HPO_4 \cdot 2H_2O$, 0.24 g/l KH_2PO_4, adjusted to pH 7.4). Pellets were flash-frozen in liquid nitrogen and stored at -80°C until further processing.

Protein extraction and digestion. Cells were lysed in 50 μl lysis buffer (1% sodium deoxycholate, 10mM TCEP, 100mM Tris base pH 8.5, supplemented with 15 mM chloroacetamide). Samples were sonicated for 20 minutes using a 30 seconds on / 30 seconds off program (Bioruptor Pico sonication device, Diagenode) and heated for 10 minutes at 95°C and 300 rpm (ThermoMixer C, Eppendorf). Porcine trypsin was added to a final enzyme:protein ratio of 1:50 and samples were digested overnight at 37°C and 300 rpm (ThermoMixer C, Eppendorf).

Solid phase extraction. Digested protein samples were spun down, mixed with 50 µl 5% TFA (solution in water) and 100 µl 1% TFA in 2-propanol. Samples were loaded on PR-Sulfonate Cartridges (SDB-RPS, PreOmics) and spun down for 4 minutes at 2000 rcf (Eppendorf Centrifuge). Cartridges were twice washed with 200 µl 1% TFA in 2-propanol and spun down as above. Cartridges were twice washed with 200 µl 0.2% TFA in water and spun down as above. Peptides were eluted in two centrifugation rounds with 100 µl Elution buffer (1% (v/v) ammonium hydroxide, 19% water and 80% acetonitrile (Elution) each, for 2 minutes at 2000 rcf. Peptide solutions were dry eluted in a vacuum concentrator. Dried peptides were dissolved via ultrasonification and shaking in LC buffer (0.15% Formic acid, 2% acetonitrile) to a final peptide concentration of 0.5 µg/µl. Peptide concentration was verified with a SpectroStar Nanodrop analyzer. Samples were stored at -20°C until measurement.

Relative protein concentration determination using LC-MS. Samples were analyzed by LC-MS on an Orbitrap Elite system (Thermo Scientific). Proteins were identified and quantified over all samples by MASCOT search against the *E. coli* K-12 MG1655 reference proteome UP000000625 (The Uniprot Consortium, 2021, Nucl. Acids Research). Protein abundances were quantified by MS1-based Label Free quantitation and protein fold changes between conditions were statistically analyzed by SafeQuant (version 2.3.4). The qValue represents adjusted pValue for multiple testing using Benjamini-Hochberg method.

Flow cytometry

Exponentially growing cultures were sampled and kept on ice in Eppendorf tubes. Samples were diluted into 1x PBS just before analysis. Cells were analyzed on a BD LSRFortessa flow cytometer (Fig. 6b, c, e, f, Fig. S5-6) or BD FACS AriaIII Cell Sorter (Fig. 6d, Fig. S7) at medium flow rate and a maximum event rate of 10'000/s. Per sample, 100'000 events were recorded. Parameters measured were forward scatter (FSC-H), side scatter (SSC-H and SSC-W), 'mCherry-H' for mScarlet-I. For GFP, either 'GFP-H' or specifically 'Ex488_505LP_512_25-H' (excitation at 488nm, longpass filter of 505 nm) for GFP. Data was collected using the Diva (BD Biosciences) software. FlowJo software (Tree Star, Inc.) was used for import and gating of raw data. The forward-scatter (FSC-H) and side-scatter (SSC-H and SSC-W) were used to separate cells from background particles. A third gate was applied in which we used the 'mCherry' channel to gate for cells expressing mScarlet-I. Only mScarlet-I-positive cells were included in the analysis of the GFP signal coming from the c-di-GMP sensor. Gated populations were exported from FlowJo as 'Scale Values' in csv-files and further processed using MATLAB (version 2019, MathWorks) scripts. In MATLAB, outliers were excluded by removing the top 0.01% as well as the bottom 0.01% of datapoints. Next, the function 'histcounts' was used to bin the datapoints. The bin size was adapted to the range spanned by the dataset. The results were visualized in MATLAB or using GraphPad Prism (Version 9.1.0 (216)).

4.6 Supplementary material

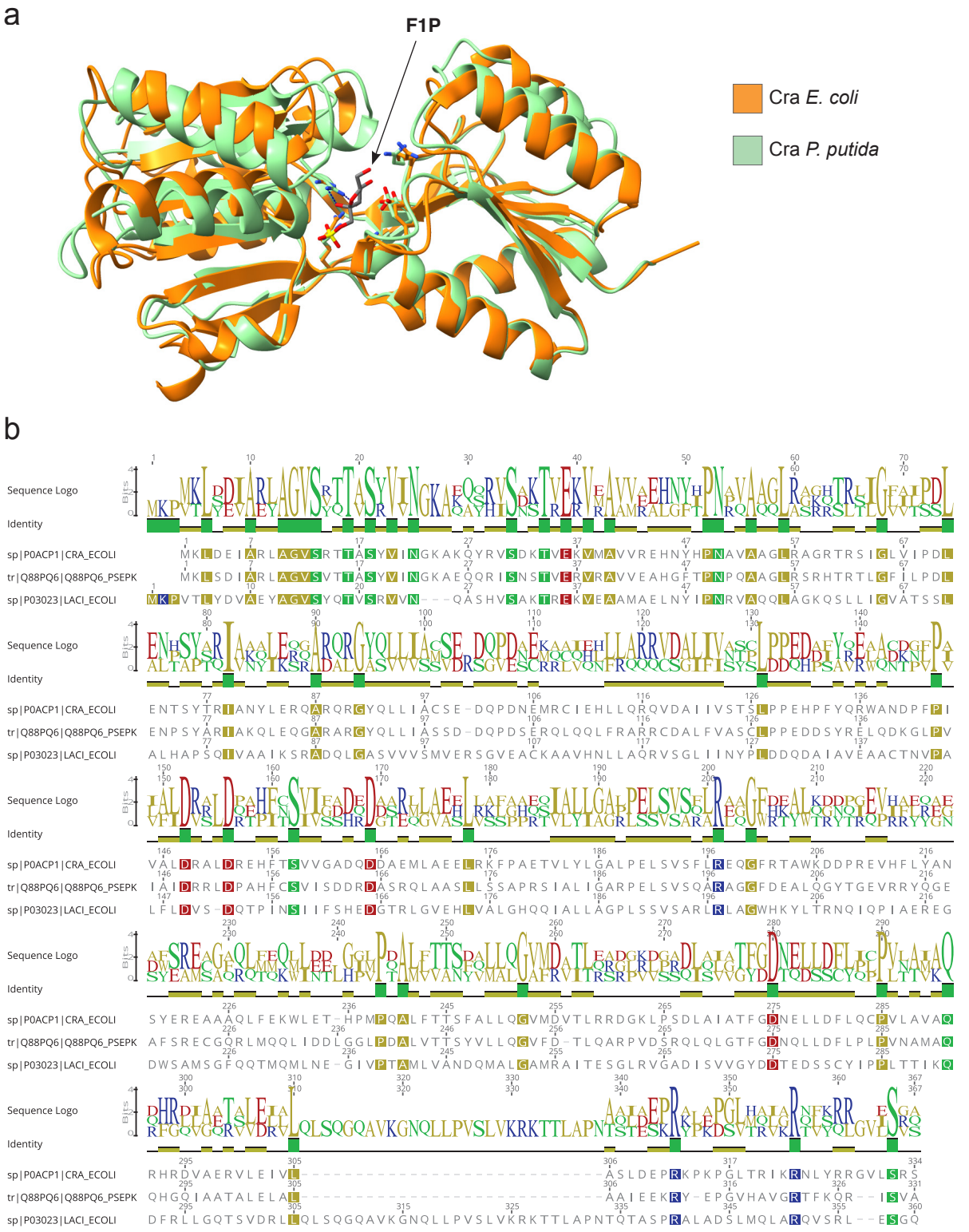


Figure S1: Cra superimposition and multiple sequence alignment. **(a)** Superimposition of the crystal structures of the sugar binding domains of Cra^{Ec} (orange) (PDB: 2iks) and Cra^{Pp} (light green) bound to F1P (grey) (PDB: 3o75). The arrow indicates the position of the ligand fructose-1-phosphate (F1P, grey). Cra^{Pp} residues shown in sticks interact with F1P. The F1P-interacting residues that are conserved between Cra^{Pp} and Cra^{Ec} are also shown in sticks in the Cra^{Ec} structure. **(b)** Multiple sequence alignment (MSA) of Cra^{Ec} (UniProt: P0ACP1), Cra^{Pp} (UniProt: Q88PQ6) and the *E. coli* LacI protein (UniProt: P03023). Sequence logo shows the consensus sequence of the three proteins. Sequence logo color coding indicates polarity and charge: brown, non-polar; green, polar; blue, positively charged; red, negatively charged. Height and coloring of identity bar indicates degree of sequence conservation: residue at position conserved among all three sequences (highest bar in green and residue colored in MSA), residue conserved among two out of three proteins (brown, lower bar), or residue not conserved (no identify bar).

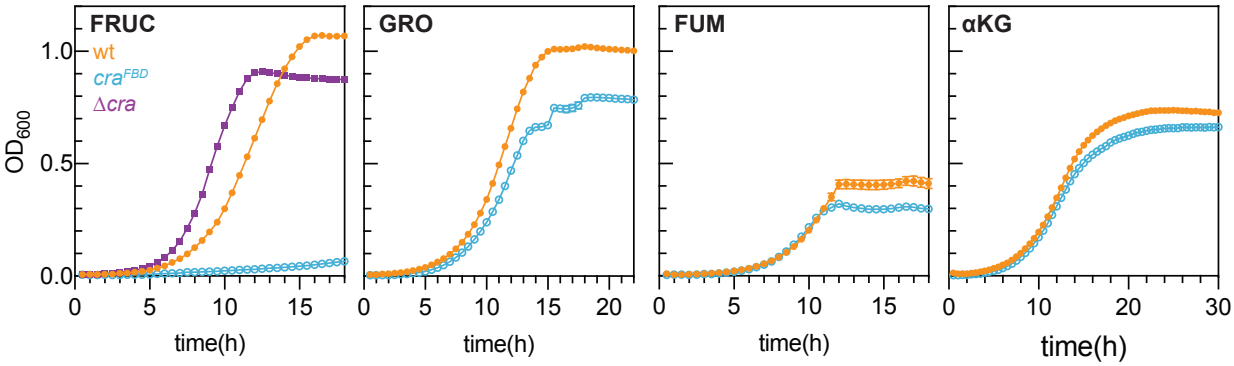


Figure S2: Growth of *cra*^{FBD} on different carbon sources. Growth curves (optical density (OD_{600}) over time) of wt *cra* (orange), *cra*^{FBD} (blue) and *Δcra* (purple) on minimal medium supplemented with fructose (FRUC), glycerol (GRO), fumarate (FUM) or alpha-ketoglutarate (αKG) as the sole carbon source.

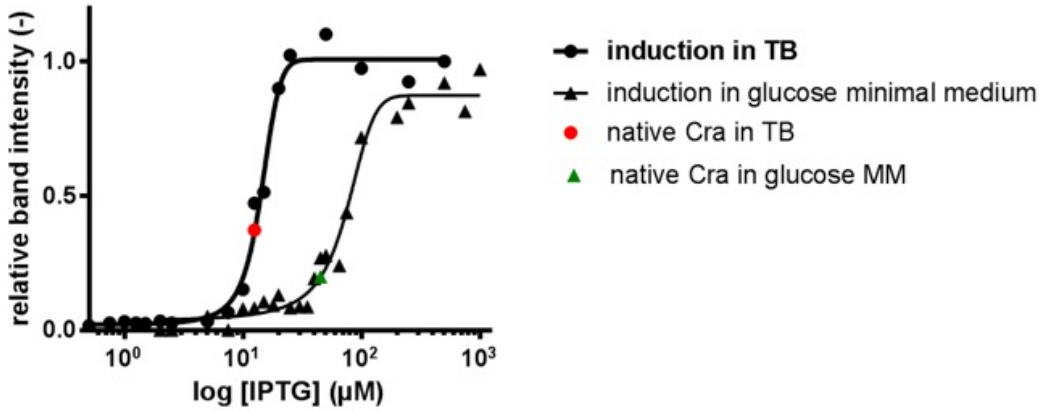


Figure S3: Immunoblot of Cra-3xFlag. Relative band intensities of antibody labeled Cra-3xflag on western blots. P_{lac}-*cra-3xflag* construct on a low copy plasmid (pAR218) was induced at increasing IPTG concentrations during growth on TB (circles) and glucose minimal medium (triangles). Samples were taken during exponential growth phase. To compare Cra levels obtained through IPTG induction with native Cra level under each growth condition, relative band intensity was compared to that of a chromosomal *cra-3xflag* construct for growth on TB (red) and glucose (green).

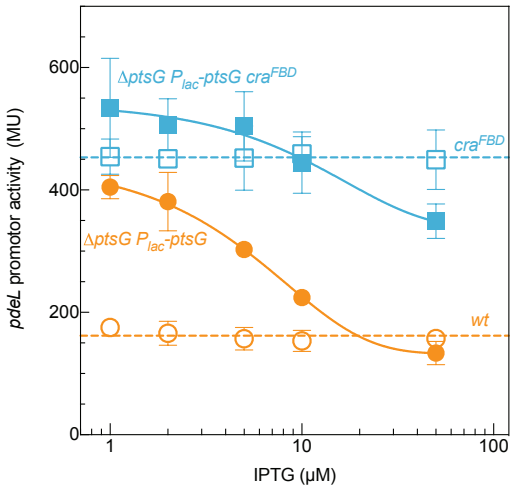


Figure S4: *pdeL* promoter activity at increasing induction of P_{lac}-*ptsG* expression. *pdeL* promoter activity during growth on glucose minimal medium at increasing IPTG concentration, as measured with the P_{pdeL}-*lacZ* reporter construct in the following strains: Δ*ptsG* transduced with a low copy plasmid carrying the P_{lac}-*ptsG* construct (pAR350; orange, closed circles), Δ*ptsG cra*^{FBD} P_{lac}-*ptsG* (blue, closed squares), wt + EV (orange, open circles) and *cra*^{FBD} + EV (blue, open squares). Broken lines indicate the average *pdeL* promoter activity taken over all IPTG concentrations for wt (orange) and *cra*^{FBD} (blue). EV: pNDM220 empty vector.

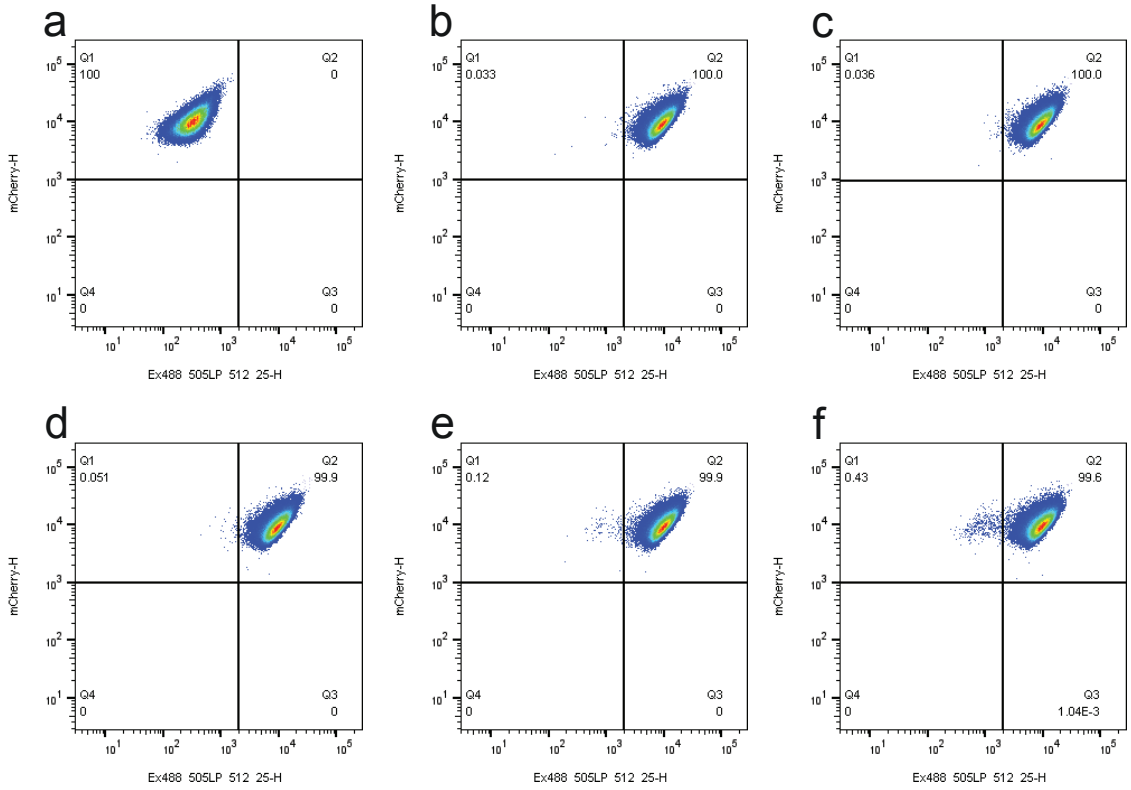


Figure S5: C-di-GMP measurements at increasing *cra* expression. Single cell mScarlet-I intensity (mCherry-H) plotted against the corresponding GFP intensity (Ex488_505LP_512_25-H) of the c-di-GMP biosensor as measured by flow cytometry. Expression of c-di-GMP biosensor and mScarlet-I was induced with 200 nM anhydrotetracycline from plasmid p2H12ref(-blind) in a *Δcra* + P_{lac}-*cra* (pAR347) background during growth on glucose under the following inducing conditions:
(a) no IPTG (blind sensor);
(b) no IPTG;
(c) 2 μM IPTG;
(d) 10 μM IPTG;
(e) 20 μM IPTG;
(f) 100 μM IPTG.

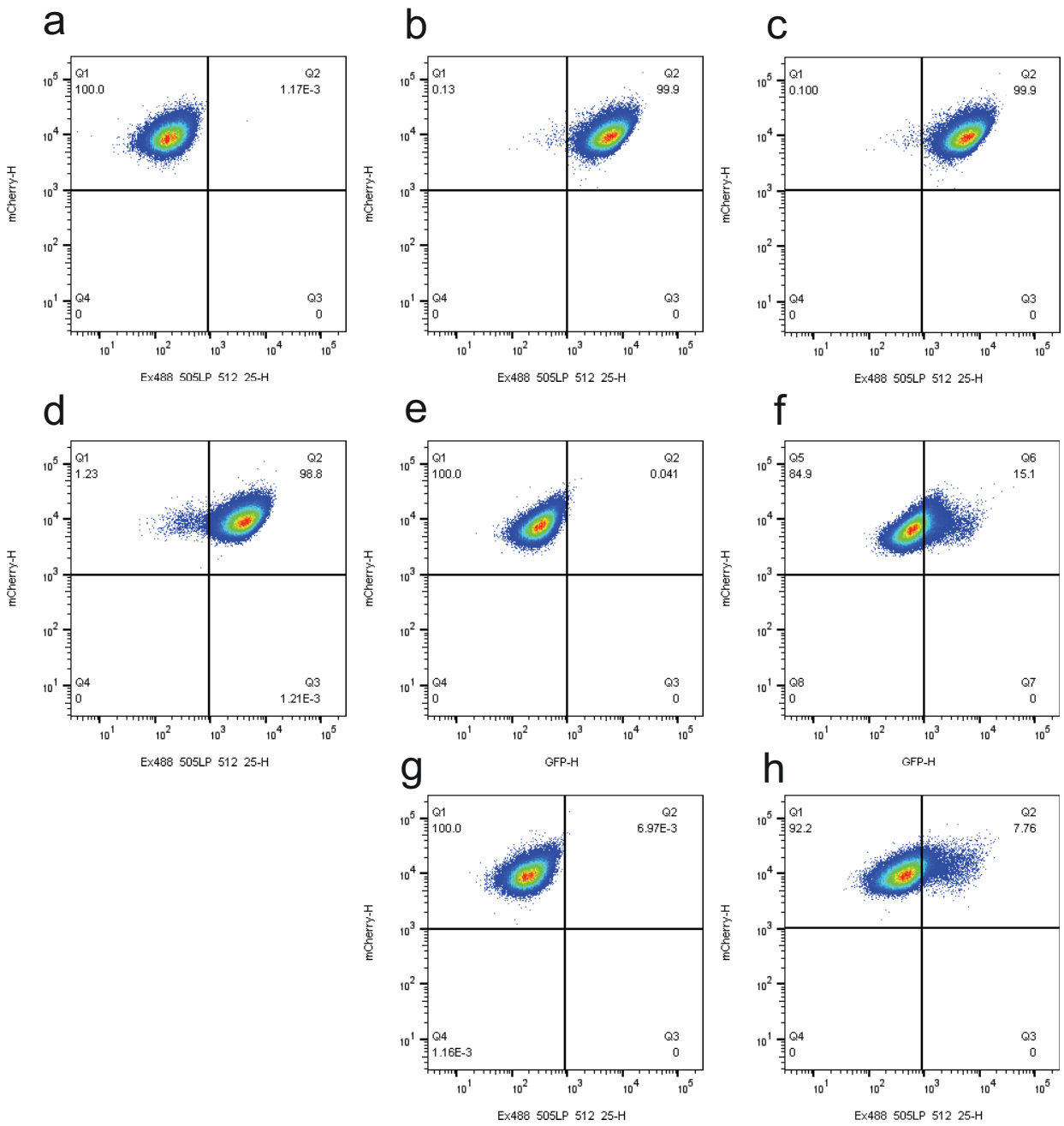


Figure S6: C-di-GMP measurements in *cra*^{FBD} and *pdeL* (D295N) mutant strains. Single cell mScarlet-I intensity (mCherry-H) plotted against the corresponding GFP intensity (Ex488_505LP_512_25-H) of the c-di-GMP biosensor as measured by flow cytometry. Expression of the c-di-GMP biosensor and mScarlet-I was induced with 200 nM anhydrotetracycline from plasmid p2H12ref(-blind) during growth on glucose in the following strains:
(a) wt (blind sensor);
(b) wt;
(c) *cra*::Frt;
(d) *cra*^{FBD}::Frt;
(e) *pdeL* (D295N) (blind sensor);
(f) *pdeL* (D295N);
(g) MG1655 CGSC 7740 (blind sensor);
(h) MG1655 CGSC 7740

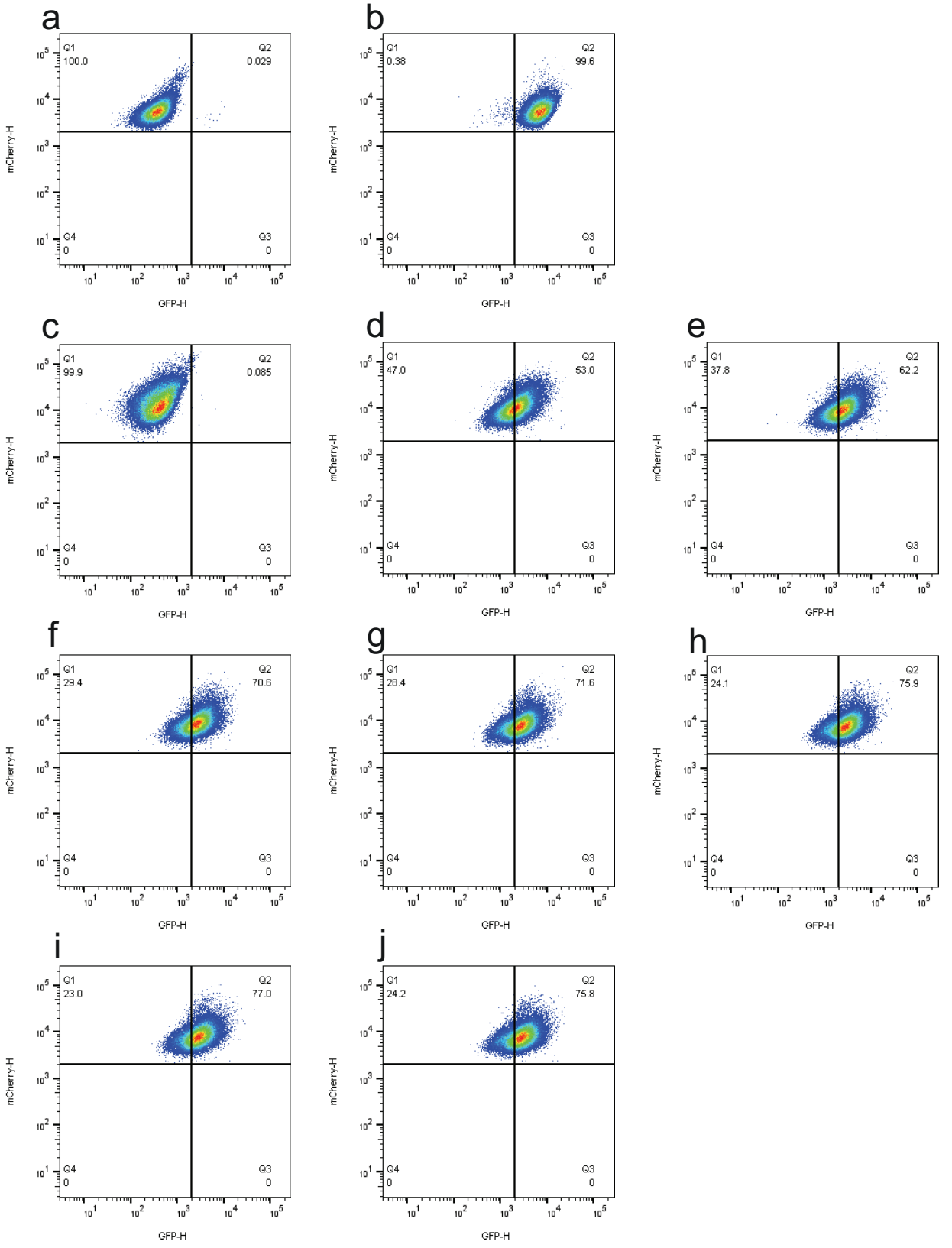


Figure S7: C-di-GMP measurements under increasing *ptsG* induction. Single cell mScarlet-I intensity (mCherry-H) plotted against the corresponding GFP intensity (GFP-H) of the c-di-GMP biosensor as measured by flow cytometry. Constitutive expression of the c-di-GMP (blind) biosensor and mScarlet-I from plasmid P_{const}-p2H12ref(-blind) during growth on glucose in the following strains and conditions:

(a) wt (blind sensor)
(b) wt
(c) P_{tet}-*ptsG* (blind sensor), 0 aTc
(d) P_{tet}-*ptsG*, 0 aTc
(e) P_{tet}-*ptsG* + 0.4 nM aTc
(f) P_{tet}-*ptsG* + 0.8 nM aTc
(g) P_{tet}-*ptsG* + 1.6 nM aTc
(h) P_{tet}-*ptsG* + 3.2 nM aTc
(i) P_{tet}-*ptsG* + 6.4 nM aTc
(j) P_{tet}-*ptsG* + 20 nM aTc

Table S1: Strains used in this study

Strain	Genotype	Source
E. coli K-12 MG1655		Blattner et al. 1997
AB3472	Δcra::Frt	This study
AB3609	cra (R197A)	This study
AB4576	Δcra::kan Prha::ccdB	This study
AB3678	ΔptsG::Frt	This study
AB2376	kan::PyahA-lacZ (merodiploid translational fusion)	A. Reinders
AB2377	ΔpdeH kan::PyahA-lacZ (merodiploid translational fusion)	This study
AB3850	Cra (R197A) pdeL (D295N) kan::PpdeL-lacZ (merodiploid translational phusion)	This study
AB3731	yahA (D295N)::Frt kan::PyahA-lacZ (merodiploid translational fusion)	This study
AB2905	yahA (K283R)-3xFlag::Frt kan::PyahA-lacZ (merodiploid translational fusion)	This study
AB3685	yahA (E141A)::Frt kan::PyahA-lacZ (merodiploid translational fusion)	This study
AB3521	ΔptsG kan::PpdeL-lacZ (merodiploid translational phusion)	This study
AB3613	Cra (R197A) ΔptsG::FRT kan::PpdeL-lacZ (merodiploid translational phusion)	This study
AB3683	Cra (R197A) kan::PpdeL-lacZ (merodiploid translational phusion)	This study
AB2292	MG1655 cra-3XFlag::frt	This study
AB2806	Δcra kan::PpdeL-lacZ (merodiploid translational phusion)	This study
AB4408	frt-Ptet1-ptsG	This study
AB4410	frt-Ptet1-ptsG in CGSC6300	This study
AB3812	CGSC 6300	This study
AB4546	CGSC 6300 Δcra::frt	This study
AB4549	cra::Frt in CGSC6300	This study
AB4550	cra (R197A)::Frt in CGSC6300	This study
AB4400	yahA(D295N)-3xFlag::Frt	This study

Table S2: Plasmids used in this study

Plasmid	Genotype (*resistence*)	Source
p_temp	Ptet-[cdG-binding-domains] (kan)	UJ11205, Kaczmarczyk and Jenal unpublished
p2H12ref	Ptet-[cdG-sensor]-[mScarlet-I] (amp)	UJ11206, Kaczmarczyk and Jenal unpublished
p2H12ref-blind	Ptet-[cdG-sensor*]-[mScarlet-I] (amp)	UJ11207, Kaczmarczyk and Jenal unpublished
p-const-p2H12ref	Pconstitutive-[cdG-sensor]-[mScarlet-I] (amp)	UJ11213, Kaczmarczyk and Jenal unpublished
p-const-p2H12ref-blind	Pconstitutive-[cdG-sensor*]-[mScarlet-I] (amp)	UJ11214, Kaczmarczyk and Jenal unpublished
pAR1	Plac-pdeL-strepII in pET28a (kan)	Reinders et al., 2016
pAR19	Cra-strepII in pET28 (kan)	A. Reinders
pAR194	cra (R197A)-strepII in pET28a (kan)	A. Reinders and S. Ozaki
pAR213	Plac-cra(R197A)-3xFlag in pNDM220 (amp)	A. Reinders
pAR218	Plac-cra-3xFlag in pNDM220 (amp)	this study
pAR282	Cra-activity reporter plasmid: PpykF-cfp P*pykF-yfp	M. Heinemann
pAR305	pAR282, rfp removed from backbone	this study
pAR347	Plac-cra in pNDM220k (kan)	this study
pAR347	Plac-cra in pNDM220k (kan)	this study
pAR350	Plac-ptsG in pNDM220 (amp)	this study
pAR351	Plac-ptsG-3xFlag in pNDM220 (amp)	this study
pAR81	P_{lac}-$RBS_{synth.}$-*pdeH-3xflag* in pNDM220 (*amp*)	A. Reinders
pCP20	FLP+ (amp)	Cherepanov and Wackernagel, 1995
pET28a	pBR332 lacI P_{T7} (*kan*) 6xHis expression vector	Novagen
pKD3	cat (used for MB013::cat)	Datsenko and Wanner, 2000
pKD45	Prha-kan-ccdB	Datsenko and Wanner, 2000
pKD46	λ RED+ (amp)	Datsenko and Wanner, 2000
pNDM220	*repA parR parM* P_{lac} (amp)	Gotfredsen et al., 1998
pNDM220k	*repA parR parM* P_{lac} (kan)	A. Reinders

Discussion and outlook

We have characterized a switch protein that generates binary patterns of the signaling molecule c-di-GMP in *E. coli* populations. The c-di-GMP specific phosphodiesterase PdeL converts gradual changes of c-di-GMP in a switch-like response. During transitions from a high to low c-di-GMP regime, PdeL activity drives a c-di-GMP switch on the single cell level. Intermediate c-di-GMP levels are avoided: instead, PdeL generates two subpopulations with distinct c-di-GMP levels, of which the ratio shifts over time. Besides, PdeL generates bimodal c-di-GMP distributions during steady state growth on alternative carbon sources in a homogeneous environment. Hence, PdeL acts both as a c-di-GMP noise quencher and amplifyer. PdeL ensures that cells robustly and unambiguously transit between different behavioral states and that populations phenotypically diversify as a potential bet-hedging strategy.

PdeL functions as an ultrasensitive c-di-GMP sensor, a catalyst and stimulator of its own expression in response to prevailing c-di-GMP levels. Using a set of *pdeL* mutants, my collaborators in the Jenal lab elucidated the molecular mechanisms that give rise to these charachteristics of PdeL. They identified two elements of feedback regulation and a high level of cooperativiy.

With the DNA-binding domain, PdeL dimers bind a high and low affinity binding site in the *pdeL* promoter region. PdeL binding activates *pdeL* transcription, creating a double-positive feedback loop.

While PdeL degrades c-di-GMP, PdeL activity is inhibited at high c-di-GMP as binding of the signaling molecule to each PdeL protomer stabilizes an inert state ('T-state'). This double negative feedback loop is interrupted upon loss of c-di-GMP binding or detabilization of the T-state.

pdeL expression shows a high degree of cooperativity. The second, low affinity PdeL binding site in the promoter region is crucial to establish cooperativity. Moreover, our collaborators from structural biology demonstrated PdeL can form tetramers in the absence of c-di-GMP. PdeL cooperativity was lost in a tetramerization-deficient mutant, showing the importance of PdeL multimerization in non-linear behavior.

When investigating the role of PdeL regulatory components *in vivo*, I found that all these aspects of PdeL, i.e. catalytic activity, promoter binding, tetramerization as well as c-di-GMP feedback inhibition are crucial to establish bimodal c-di-GMP distributions. Our results show that the simplest forms of gene regulation that could potentially demonstrate multistability, i.e. feedback regulation and non-linearity (Ferrell, 2002; Smits et al., 2006), are required for PdeL to generate binary signaling output.

During c-di-GMP transitions, I could show that at the level of *pdeL* expression and c-di-GMP, two distinct subpopulations are formed. For bimodal c-di-GMP distributions during steady-state growth, however, insight in *pdeL* expression dynamics is so far missing. A new, strong, and c-di-GMP sensor-compatible fluorescent marker for *pdeL* expression should be developed.

Moreover, data on bimodal c-di-GMP distributions so far only provide us with a snapshot in time. It would be of interest to know whether the high and low c-di-GMP subpopulations as observed in steady state growth are stable or showing dynamic switching patterns. Therefore, it is recommended to follow the c-di-GMP sensor signal in cell lineages in a set-up of for example a microfluidic device combined with time lapse microscopy. Temporal information on the c-di-GMP level across cell lineages could show whether bimodal signaling patterns are rather a form of stable, cellular differentiation as for *B. subtilis* sporulation (Chung et al., 1994; Veening et al., 2008a), or whether c-di-GMP stochastically pulses as observed for *E. coli* flagellar biosynthesis (Kim et al., 2020).

Bimodal c-di-GMP distributions did however not only depend on PdeL activity, but also on nutrient availability. Whereas bimodal distributions were observed during growth on glycerol and TCA cycle intermediates, c-di-GMP levels in *E. coli* populations were unimodal high when the medium was supplemented with amino acids or when a gluconeogenic substrate was replaced by glucose. What cellular pathways cause this phenotype, that mimics a *pdeL* deletion mutant, is so far unknown.

For the alteration in the c-di-GMP distribution upon addition of amino acids, I hypothesize this to be driven by DGC and PDE activity that affects the global c-di-GMP pool directly and additionally reduces PdeL activity through c-di-GMP. The contribution of DGCs and PDEs to the c-di-GMP pool could be systematically assessed with single deletion mutants as done previously in the context of motility (Boehm et al., 2010; Pesavento et al., 2008). However, I am highlighting here DgcQ for its potential impact on c-di-GMP in response to amino acids. In a direct screen for c-di-GMP modulators in *Salmonella enterica* serovar Thyphimurium, L-arginine sensing was found to increase cellular c-di-GMP levels through the DGC STM1987 (DgcQ) (Mills et al., 2015). Consistently, colleagues in the Jenal lab recently found that *E. coli* DgcQ, a homolog of STM1987 (García et al., 2004), requires Arginine for its activation. Activation of the DGC DgcQ in the presence of amino acids, including L-arginine, may explain the increase in cytosolic c-di-GMP concentration under these conditions. These hypotheses should be further verified by analyzing c-di-GMP levels in a *dgcQ* deletion mutant grown in the presence amino acids and L-arginine specifically.

Interestingly, in the same screen for c-di-GMP modulators (Mills et al., 2015), addition of 1 mM glucose was also found to rapidly increase c-di-GMP concentrations in *Salmonella*. Although the alteration in c-di-GMP was not as strong as for L-arginine, the glucose effect also required STM1987. For glucose however, the requirement of additional c-di-GMP modulating enzymes could not be excluded (Mills et al., 2015).

As the activity of Cra is reduced during growth on glucose (Kochanowski et al., 2013), we verified in chapter 4 for an effect of increased Cra activity on c-di-GMP distributions. C-di-GMP distributions however remained primarily unimodal and c-di-GMP levels high at increased

Cra activity. As also mentioned in the discussion of chapter 4, the role of cyclic AMP could be investigated, as this signaling molecule affects PDE and DGC expression and/or activity in multiple bacterial species including but not limited to *E. coli* (Fong and Yildiz, 2008; Heo et al., 2019; Ko and Park, 2000; Soutourina et al., 1999; Vasconcelos et al., 2017; Xiao et al., 2020).

Besides, the role of the carbon storage system (CSR) in the substrate- and nutrient-dependent c-di-GMP distribution in *E. coli* is worth further investigating. The *E. coli* CSR system consists of the RNA-binding protein CsrA, the regulatory RNA molecules *csrB* and *csrC* that sequester and thus inhibit CsrA (Liu et al., 1997; Weilbacher et al., 2003) and the regulatory protein CsrD (Jonas et al., 2006) (see also chapter 1). Apart from posttranslational modification of multiple metabolic genes (Timmermans and Melderen, 2010), CsrA negatively regulates *dgcT* and *dgcZ* (Jonas et al., 2008), two DGCs involved in the regulation of motility and biofilm formation (Boehm et al., 2009; Jonas et al., 2008). Additionally, CrsA downregulates mRNA levels of *pdeI*, *dgcI*, *dgcO*, *pdeO* and *csrD* (Jonas et al., 2008). A *csrA* deletion slightly increased c-di-GMP levels (Jonas et al., 2008). As the levels of *csrB* and *csrC* are in turn affected by substrate and amino acid availability (Chavez et al., 2010; Jonas and Melefors, 2009; Jonas et al., 2006), the CSR system shows to be a complex regulatory system that couples metabolic input to posttranslational modification of c-di-GMP modulating enzymes. How c-di-GMP levels are precisely affected by substrate availability in a CSR-dependent manner, remains to be investigated.

Overall, verification of the contribution of *dgcQ* and systematic analysis of single *dgc* and *pde* deletion mutants would be a first step to reveal which c-di-GMP modulating enzymes contribute to elevated c-di-GMP levels on glucose or in the presence of amino acids. Such condition-dependent understanding of c-di-GMP distributions and PDE/DGC activity would also be of interest for the broader c-di-GMP research community.

The identification of a metabolic regulator as essential activator of a second messenger switch raises questions about the possible benefits of such regulatory mechanism. The nutrient-dependent c-di-GMP distribution raises as many questions about the function of these distributions in corresponding environmental conditions.

Cra is a modulator of carbon flow in *E. coli* cells and pleiotropically regulates transcription of a range of central metabolic genes (Ramseier, 1996; Ramseier et al., 1995). Expression of *pdeL* is among the few non-metabolic targets of Cra. Furthermore, Cra can stimulate curli gene expression (Remhaswala and Noronha, 2011) and in Enterohemorrhagic *E. coli*, Cra is required for the activation of multiple virulence factors (Njoroge et al., 2012). The expression of the LEE pathogenicity island is activated during gluconeogenesis, but inhibited during glycolysis through co-regulation of Cra and a second transcription factor KdpE (Njoroge et al., 2012). Not only Cra, but also other global pleotropic regulators are involved in both bacterial metabolism and pathogenesis control, including the RNA binding protein CsrA of the CSR system and the carbon catabolite protein A (CcpA) in Gram positive bacteria (Abranches et al., 2008; Bhatt et al., 2009; Mey et al., 2015; Potts et al., 2019; Seidl et al., 2006). From these regulatory networks, it could be concluded that the metabolic state of the cell can largely determine activation of virulence traits.

Virulence has many links to c-di-GMP, as c-di-GMP signaling has a clear and multifactorial role in virulence and pathogenesis (Valentini and Filloux, 2019). Interestingly, a screen for GGDEF- and EAL-domain proteins in pathogenic *E. coli* revealed that in all 25 enteric strains analyzed (EAEC, ETEC, EHEC, STEC and EPEC), expression of *pdeL* is affected by a genetic rearrangement (Richter et al., 2014; Povolotsky and Hengge, 2016). The *pdeL* promoter region is in these cases replaced by an open reading frame encoding the *ehaA* autotransporter (Richter et al., 2014). Expression of *ehaA* is associated with increased adhesion and biofilm formation (Wells et al., 2008). Binding sites of the *pdeL* regulatory elements PdeL, Cra and H-NS but also the RNA polymerase binding site as described in chapter 2 have hence been replaced by the *ehaA* coding sequence. Because experimental data on *pdeL* expression in the P_{pdeL}::*ehaA* genotype seem not generated yet, the exact consequences on *pdeL* expression are unknown. Insertion of *ehaA* may silence *pdeL*, or *pdeL* may be expressed from a new bicistronic *ehaA-pdeL* operon (Richter et al., 2014). Regulation of the *ehaA* intergenic region is unclear. A putative promoter and surprisingly a Cra binding site were identified in this region, but remain to be verified (Povolotsky and Hengge, 2016). Modification of *pdeL* regulation, together with mutations in other PDEs and DGSs across commensal and pathogenic *E. coli* strains (Povolotsky and Hengge, 2016), suggest that c-di-GMP regulation is adaptive and that certain environments may select for strains with potentially highly modified c-di-GMP regulation.

A selection of nine extraintestinal pathogenic *E. coli* (ExPEc) contain the intact *pdeL* regulation as is for the commensal MG1655 (Povolotsky and Hengge, 2016). ExPEc adapt to grow as a harmless commensal in the human intestinal tract (Alteri and Mobley, 2012). Upon entering the urinary tract, however, they rapidly acquire virulence traits and cause urinary tract infections, bacteremia or septicemia. The transition from harmless strain to pathogen follows an environmental change from a relatively nutrient-abundant and carbon rich intestinal tract to nutritionally limited environment in the blatter (Alteri and Mobley, 2012). Consistently, functional gluconeogenic pathways are required for *in vivo* fitness of UPEC (Alteri et al., 2009, 2015). Not only was metabolic pathway adaptation proposed as a signal to activate pathogenic properties in the cell (Alteri and Mobley, 2012), the gluconeogenic environment may additionally generate c-di-GMP heterogeneity as observed in our study. While residing in a dynamic and uncertain host environment, successful colonization and survival require highly specific and dynamic c-di-GMP control (Valentini and Filloux, 2019). ExPEc might adopt a bet-hedging strategy during infection to increase the likeliness that part of the population induces the right program at the right time and hence survives as a species. If further investigated, conclusions may be drawn on how pathogens on one hand tightly control c-di-GMP signaling to establish precise timing (Valentini and Filloux, 2019), but on the other hand may as well induce signaling heterogeneity.

More generally, growth on alternative carbon sources may be a signal for scarcity in energy sources that induces heterogeneity in bacterial behavior. Growth on glucose, the favorite sugar of *E. coli*, may then be a signal for a distinct cellular program that includes proliferation in a high c-di-GMP regime.

The near-ubiquitous nature of c-di-GMP (Jenal et al., 2017) makes that bacterial risk-spreading through c-di-GMP is not limited to virulence, pathogenesis, motility or biofilm formation. Instead, new phenotypes as N4 phage resistance, briefly mentioned in chapter 2 and 3, show the potential of this small signaling molecule to define bacterial phenotypes.

While showing signaling heterogeneity at 37°C, we have not yet assessed the conservation of the PdeL-driven signaling heterogeneity across different temperature ranges. A larger screen across commensal and pathogenic *E. coli* strains may further confirm conservation of functional PdeL and *pdeL* regulation across strains. Together, this information could further indicate in what ecological niches binary patterns of c-di-GMP may be established.

List of Figures

List of Tables

Acknowledgements

First of all, I would like to thank my supervisor Urs Jenal, for supporting me in many aspects and throughout my PhD in his lab. Thank you, for the years filled with support, enthusiasm, questions, constructive feedback and advice in the role of mentor and First Supervisor.

Furthermore, I would like to thank my Second Supervisor Dirk Bumann and External expert, Matthias Heinemann. Besides supporting and assessing my PhD with his expertise, I highly acknowledge Matthias Heinemann for informing our lab about the *IS1* element in *dgcJ* and the MG1655 growth phenotypes. I also thank his lab for kindly sharing protocols, plasmids and strains.

For support in data acquisition and analysis, I thank Alexander Schmidt, Thomas Bock and Ulrike Lanner of the Proteomics Core Facility, Timothy Sharpe and Xiaochun Li Blatter of the Biophysics Facility, Janine Boegli and Stella Stefanova of the FACS Core Facility and Kai Schleicher and Alexia Loynton-Ferrand of the Imaging Core Facility at the Biozentrum.

I highly recognize and thank Andreas Kaczmarczyk for his efforts in succesfully developing novel c-di-GMP biosensors. Fabienne Hamburger was always there for support with molecular cloning and for lab support of all kind. A big thanks to all group members for the years in the Jenal lab. Restrictions in our way of working due to the pandemic put challenges on the way we work and interact, but I think we manged to find alternative ways to communicate science.

I also want to thank the current and former colleages on the 4th Floor: the ladies in the media kitchen, the lab members of the Dehio lab, Basler lab and Bumann lab and Marina, Patric, Claudia, Sarah, Anne-Cécile and Michaela.

I want to thank Benoit, Alex, Alex and Pablo for reviewing and commenting on sections of this thesis. I want to thank my family and friends for the support, especially Lara, Eric, my sister and my parents.

References

Abel, S., Bucher, T., Nicollier, M., Hug, I., Kaever, V., Wiesch, P.A. zur, and Jenal, U. (2013). Bi-modal Distribution of the Second Messenger c-di-GMP Controls Cell Fate and Asymmetry during the Caulobacter Cell Cycle. Plos Genet 9, e1003744.

Abel, S., Chien, P., Wassmann, P., Schirmer, T., Kaever, V., Laub, M.T., Baker, T.A., and Jenal, U. (2011). Regulatory Cohesion of Cell Cycle and Cell Differentiation through Interlinked Phosphorylation and Second Messenger Networks. Mol Cell 43, 550–560.

Abranches, J., Nascimento, M.M., Zeng, L., Browngardt, C.M., Wen, Z.T., Rivera, M.F., and Burne, R.A. (2008). CcpA Regulates Central Metabolism and Virulence Gene Expression in Streptococcus mutans. J Bacteriol 190, 2340–2349.

Ahrné, E., Glatter, T., Viganò, C., Schubert, C. von, Nigg, E.A., and Schmidt, A. (2016). Evaluation and Improvement of Quantification Accuracy in Isobaric Mass Tag-Based Protein Quantification Experiments. Journal of Proteome Research 15, 2537–2547.

Alteri, C.J., and Mobley, H.L. (2012). Escherichia coli physiology and metabolism dictates adaptation to diverse host microenvironments. Curr Opin Microbiol 15, 3–9.

Alteri, C.J., Smith, S.N., and Mobley, H.L.T. (2009). Fitness of Escherichia coli during Urinary Tract Infection Requires Gluconeogenesis and the TCA Cycle. Plos Pathog 5, e1000448.

Alteri, C.J., Himpsl, S.D., and Mobley, H.L.T. (2015). Preferential Use of Central Metabolism In Vivo Reveals a Nutritional Basis for Polymicrobial Infection. Plos Pathog 11, e1004601.

Alteriis, E. de, Cartenì, F., Parascandola, P., Serpa, J., and Mazzoleni, S. (2018). Revisiting the Crabtree/Warburg effect in a dynamic perspective: a fitness advantage against sugar-induced cell death. Cell Cycle 17, 1–43.

Altier, C., Suyemoto, M., and Lawhon, S.D. (2000). Regulation of Salmonella entericaSerovar Typhimurium Invasion Genes by csrA. Infect Immun 68, 6790–6797.

Amikam, D., and Galperin, M.Y. (2006). PilZ domain is part of the bacterial c-di-GMP binding protein. Bioinformatics 22, 3–6.

Ammar, E.M., Wang, X., and Rao, C.V. (2018). Regulation of metabolism in Escherichia coli during growth on mixtures of the non-glucose sugars: arabinose, lactose, and xylose. Sci Rep-Uk 8, 609.

Andrade, M.O., Alegria, M.C., Guzzo, C.R., Docena, C., Rosa, M.C.P., Ramos, C.H.I., and Farah, C.S. (2006). The HD-GYP domain of RpfG mediates a direct linkage between the Rpf quorum-sensing pathway and a subset of diguanylate cyclase proteins in the phytopathogen Xanthomonas axonopodis pv citri. Molecular Microbiology 62, 537–551.

Apel, D., and Surette, M.G. (2008). Bringing order to a complex molecular machine: The assembly of the bacterial flagella. Biochimica Et Biophysica Acta Bba - Biomembr 1778, 1851–1858.

Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S., and Ramphal, R. (1997). A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in Pseudomonas aeruginosa in a cascade manner. J Bacteriol 179, 5574–5581.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., Castro, E.D., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., et al. (2012). ExPASy: SIB bioinformatics resource portal. Nucleic Acids Research 40, W597--W603.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2, 2006.0008.

Babitzke, P., and Romeo, T. (2007a). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Curr Opin Microbiol 10, 156–163.

Babitzke, P., and Romeo, T. (2007b). CsrB sRNA family: sequestration of RNA-binding regulatory proteins. Curr Opin Microbiol 10, 156–163.

Bai, Y., Yang, J., Zhou, X., Ding, X., Eisele, L.E., and Bai, G. (2012). Mycobacterium tuberculosis Rv3586 (DacA) Is a Diadenylate Cyclase That Converts ATP or ADP into c-di-AMP. Plos One 7, e35206.

Balaban, N.Q., Merrin, J., Chait, R., Kowalik, L., and Leibler, S. (2004). Bacterial Persistence as a Phenotypic Switch. Science 305, 1622–1625.

Bandekar, D., Chouhan, O.P., Mohapatra, S., Hazra, M., Hazra, S., and Biswas, S. (2017). Putative protein VC0395_0300 from Vibrio cholerae is a diguanylate cyclase with a role in biofilm formation. Microbiol Res 202, 61–70.

Barker, C.S., Prüß, B.M., and Matsumura, P. (2004). Increased Motility of Escherichia coli by Insertion Sequence Element Integration into the Regulatory Region of the flhD Operon. J Bacteriol 186, 7529–7537.

Barthe, M., Tchouanti, J., Gomes, P.H., Bideaux, C., Lestrade, D., Graham, C., Steyer, J.-P., Meleard, S., Harmand, J., Gorret, N., et al. (2020). Availability of the Molecular Switch XylR Controls Phenotypic Heterogeneity and Lag Duration during Escherichia coli Adaptation from Glucose to Xylose. Mbio 11.

Bartlett, T.M., Bratton, B.P., Duvshani, A., Miguel, A., Sheng, Y., Martin, N.R., Nguyen, J.P., Persat, A., Desmarais, S.M., VanNieuwenhze, M.S., et al. (2017). A Periplasmic Polymer Curves Vibrio cholerae and Promotes Pathogenesis. Cell 168, 172-185.e15.

Basan, M., Honda, T., Christodoulou, D., Hörl, M., Chang, Y.-F., Leoncini, E., Mukherjee, A., Okano, H., Taylor, B.R., Silverman, J.M., et al. (2020). A universal trade-off between growth and lag in fluctuating environments. Nature 584, 470–474.

Basan, M., Hui, S., Okano, H., Zhang, Z., Shen, Y., Williamson, J.R., and Hwa, T. (2015). Overflow metabolism in Escherichia coli results from efficient proteome allocation. Nature 528, 99–104.

Bassler, J., Schultz, J.E., and Lupas, A.N. (2018). Adenylate cyclases: Receivers, transducers, and generators of signals. Cell Signal 46, 135–144.

Bateman, A., Martin, M.-J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., et al. (2020). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research.

Bell, C.E., and Lewis, M. (2000). A closer view of the conformation of the Lac repressor bound to operator. Nat Struct Biol 7, 209–214.

Bettenworth, V., Steinfeld, B., Duin, H., Petersen, K., Streit, W.R., Bischofs, I., and Becker, A. (2019). Phenotypic Heterogeneity in Bacterial Quorum Sensing Systems. J Mol Biol 431, 4530–4546.

Bhatt, S., Edwards, A.N., Nguyen, H.T.T., Merlin, D., Romeo, T., and Kalman, D. (2009). The RNA Binding Protein CsrA Is a Pleiotropic Regulator of the Locus of Enterocyte Effacement Pathogenicity Island of Enteropathogenic Escherichia coli. Infect Immun 77, 3552–3568.

Bi, S., and Sourjik, V. (2018). Stimulus sensing and signal processing in bacterial chemotaxis. Curr Opin Microbiol 45, 22–29.

Blattner, F.R., Plunkett, G., Bloch, C.A., and Perna, N.T. (1997). The complete genome sequence of Escherichia coli K-12. Science 277, 1453-1462

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., et al. (1997). The Complete Genome Sequence of Escherichia coli K-12. Science 277, 1453–1462.

Bledig, S.A., Ramseier, T.M., and Saier, M.H. (1996). Frur mediates catabolite activation of pyruvate kinase (pykF) gene expression in Escherichia coli. J Bacteriol 178, 280–283.

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., et al. (2020). The InterPro protein families and domains database: 20 years on. Nucleic Acids Res 49, gkaa977-.

Boehm, A., Kaiser, M., Li, H., Spangler, C., Kasper, C.A., Ackermann, M., Kaever, V., Sourjik, V., Roth, V., and Jenal, U. (2010). Second Messenger-Mediated Adjustment of Bacterial Swimming Velocity. Cell 141, 107–116.

Boehm, A., Steiner, S., Zaehringer, F., Casanova, A., Hamburger, F., Ritz, D., Keck, W., Ackermann, M., Schirmer, T., and Jenal, U. (2009). Second messenger signalling governs Escherichia coli biofilm induction upon ribosomal stress. Molecular Microbiology 72, 1500–1516.

Bordeleau, E., Purcell, E.B., Lafontaine, D.A., Fortier, L.-C., Tamayo, R., and Burrus, V. (2015). Cyclic Di-GMP Riboswitch-Regulated Type IV Pili Contribute to Aggregation of Clostridium difficile. J Bacteriol 197, 819–832.

Botsford, J.L., and Harman, J.G. (1992). Cyclic AMP in prokaryotes. Microbiol Rev 56, 100–122.

Bougdour, A., Lelong, C., and Geiselmann, J. (2004). Crl, a Low Temperature-induced Protein in Escherichia coli That Binds Directly to the Stationary Phase σ Subunit of RNA Polymerase. J Biol Chem 279, 19540–19550.

Branchu, P., Hindré, T., Fang, X., Thomas, R., Gomelsky, M., Claret, L., Harel, J., Gobert, A.P., and Martin, C. (2013). The c-di-GMP phosphodiesterase VmpA absent in Escherichia coli K12 strains affects motility and biofilm formation in the enterohemorrhagic O157:H7 serotype. Vet Immunol Immunop 152, 132–140.

Brauner, A., Fridman, O., Gefen, O., and Balaban, N.Q. (2016). Distinguishing between resistance, tolerance and persistence to antibiotic treatment. Nat Rev Microbiol 14, 320–330.

Bren, A., Park, J.O., Towbin, B.D., Dekel, E., Rabinowitz, J.D., and Alon, U. (2016). Glucose becomes one of the worst carbon sources for E.coli on poor nitrogen sources due to suboptimal levels of cAMP. Sci Rep-Uk 6, 24834.

Bruhn-Olszewska, B., Molodtsov, V., Sobala, M., Dylewski, M., Murakami, K.S., Cashel, M., and Potrykus, K. (2018). Structure-function comparisons of (p)ppApp vs (p)ppGpp for Escherichia coli RNA polymerase binding sites and for rrnB P1 promoter regulatory responses in vitro. Biochimica Et Biophysica Acta Bba - Gene Regul Mech 1861, 731–742.

Calos, M.P., Johnsrud, L., and Miller, J.H. (1978). DNA sequence at the integration sites of the insertion element IS1. Cell 13, 411--418.

Capra, E.J., and Laub, M.T. (2012). Evolution of Two-Component Signal Transduction Systems. Annual Review of Microbiology 66, 325–347.

Carraro, N., Richard, X., Sulser, S., Delavat, F., Mazza, C., and Meer, J.R. van der (2020). An analog to digital converter controls bistable transfer competence of a widespread integrative and conjugative element. eLife 9,e57915.

Cashel, M., and Gallant, J. (1969). Two Compounds implicated in the Function of the RC Gene of Escherichia coli. Nature 221, 838–841.

Chan, C., Paul, R., Samoray, D., Amiot, N.C., Giese, B., Jenal, U., and Schirmer, T. (2004). Structural basis of activity and allosteric control of diguanylate cyclase. P Natl Acad Sci Usa 101, 17084–17089.

Chavarría, M., Durante-Rodríguez, G., Krell, T., Santiago, C., Brezovsky, J., Damborsky, J., and Lorenzo, V. de (2014). Fructose 1phosphate is the one and only physiological effector of the Cra (FruR) regulator of Pseudomonas putida. Febs Open Bio 4, 377–386.

Chavarría, M., Santiago, C., Platero, R., Krell, T., Casasnovas, J.M., and Lorenzo, V. de (2011). Fructose 1-Phosphate Is the Preferred Effector of the Metabolic Regulator Cra of Pseudomonas putida *. J Biol Chem 286, 9351–9359.

Chavez, R.G., Alvarez, A.F., Romeo, T., and Georgellis, D. (2010). The Physiological Stimulus for the BarA Sensor Kinase. J Bacteriol 192, 2009–2012.

Cherepanov, P.P., and Wackernagel, W. (1995). Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene 158, 9–14.

Chiang, S.M., Dong, T., Edge, T.A., and Schellhorn, H.E. (2011). Phenotypic Diversity Caused by Differential RpoS Activity among Environmental Escherichia coli Isolates. Appl Environ Microb 77, 7915–7923.

Chilcott, G.S., and Hughes, K.T. (2000). Coupling of Flagellar Gene Expression to Flagellar Assembly in Salmonella enterica Serovar Typhimurium andEscherichia coli. Microbiol Mol Biol R 64, 694–708.

Chin, A.M., Feldheim, D.A., and Saier, M.H. (1989). Altered transcriptional patterns affecting several metabolic pathways in strains of Salmonella typhimurium which overexpress the fructose regulon. J Bacteriol 171, 2424–2434.

Chin, A.M., Feucht, B.U., and Saier, M.H. (1987). Evidence for regulation of gluconeogenesis by the fructose phosphotransferase system in Salmonella typhimurium. J Bacteriol 169, 897–899.

Choi, K.Y., Lu, F., and Zalkin, H. (1994). Mutagenesis of amino acid residues required for binding of corepressors to the purine repressor. J Biol Chem 269, 24066–24072.

Chou, S.-H., and Galperin, M.Y. (2016). Diversity of Cyclic Di-GMP-Binding Proteins and Mechanisms. J Bacteriol 198, 32–46.

Chu, D., and Barnes, D.J. (2016). The lag-phase during diauxic growth is a trade-off between fast adaptation and high growth rate. Sci Rep-Uk 6, 25191.

Chung, J.D., Stephanopoulos, G., Ireton, K., and Grossman, A.D. (1994). Gene expression in single cells of Bacillus subtilis: evidence that a threshold mechanism controls the initiation of sporulation. J Bacteriol 176, 1977–1984.

Colmer, J.A., and Hamood, A.N. (1998). Characterization of ptxS, a Pseudomonas aeruginosa gene which interferes with the effect of the exotoxin A positive regulatory gene, ptxR. Mol Gen Genetics Mgg 258, 250–259.

Colvin, K.M., Irie, Y., Tart, C.S., Urbano, R., Whitney, J.C., Ryder, C., Howell, P.L., Wozniak, D.J., and Parsek, M.R. (2012). The Pel and Psl polysaccharides provide Pseudomonas aeruginosa structural redundancy within the biofilm matrix. Environ Microbiol 14, 1913–1928.

Conner, J.G., Zamorano-Sánchez, D., Park, J.H., Sondermann, H., and Yildiz, F.H. (2017). The ins and outs of cyclic di-GMP signaling in Vibrio cholerae. Curr Opin Microbiol 36, 20–29.

Conway, T., and Cohen, P.S. (2015). Commensal and Pathogenic Escherichia coli metabolism in the gut. Microbiol Spectr 3, 343–362.

Cortay, J.-C., Negre, D., Scarabel, M., Ramseier, T.M., Vartak, N.B., Reizer, J., Jr, M.H.S., and Cozzone, A.J. (1994). In vitro asymmetric binding of the pleiotropic regulatory protein, FruR, to the ace operator controlling glyoxylate shunt enzyme synthesis. Journal of Biological Chemistry 269, 14885--14891.

Covert, M.W., Xiao, N., Chen, T.J., and Karr, J.R. (2008). Integrating metabolic, transcriptional regulatory and signal

transduction models in Escherichia coli. Bioinformatics 24, 2044–2050.

Crépin, S., Porcheron, G., Houle, S., Harel, J., and Dozois, C.M. (2017). Altered Regulation of the Diguanylate Cyclase YaiC Reduces Production of Type 1 Fimbriae in a Pst Mutant of Uropathogenic Escherichia coli CFT073. J Bacteriol 199, e00168-17.

da, P.A.D., de, A.N.M., de, A.G.G., de, S.R.F., and Rodrigues, G.C. (2020). The World of Cyclic Dinucleotides in Bacterial Behavior. Molecules 25, 2462.

Daber, R., Stayrook, S., Rosenberg, A., and Lewis, M. (2007). Structural Analysis of Lac Repressor Bound to Allosteric Effectors. J Mol Biol 370, 609–619.

Dahlstrom, K.M., and O'Toole, G.A. (2016). A Symphony of Cyclases: Specificity in Diguanylate Cyclase Signaling. Annu Rev Microbiol 71, 1–17.

Dahlstrom, K.M., Giglio, K.M., Collins, A.J., Sondermann, H., and O'Toole, G.A. (2015). Contribution of Physical Interactions to Signaling Specificity between a Diguanylate Cyclase and Its Effector. Mbio 6, e01978-15.

Dahlstrom, K.M., Giglio, K.M., Sondermann, H., and O'Toole, G.A. (2016). The Inhibitory Site of a Diguanylate Cyclase Is a Necessary Element for Interaction and Signaling with an Effector Protein. J Bacteriol 198, 1595–1603.

Daniel, P.B., Walker, W.H., and Habener, J.F. (1998). CYCLIC AMP SIGNALING AND GENE REGULATION. Annu Rev Nutr 18, 353–383.

Dasgupta, N., Wolfgang, M.C., Goodman, A.L., Arora, S.K., Jyot, J., Lory, S., and Ramphal, R. (2003). A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in Pseudomonas aeruginosa. Mol Microbiol 50, 809–824.

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

Davies, B.W., Bogard, R.W., Young, T.S., and Mekalanos, J.J. (2012). Coordinated Regulation of Accessory Genetic Elements Produces Cyclic Di-Nucleotides for V. cholerae Virulence. Cell 149, 358–370.

Diard, M., Garcia, V., Maier, L., Remus-Emsermann, M.N.P., Regoes, R.R., Ackermann, M., and Hardt, W.-D. (2013). Stabilization of cooperative virulence by the expression of an avirulent phenotype. Nature 494, 353–356.

Dong, T., and Schellhorn, H.E. (2009). Control of RpoS in global gene expression of Escherichia coli in minimal media. Mol Genet Genomics 281, 19–33.

Driever, W., Thoma, G., and Nüsslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. Nature 340, 363–367.

Dubey, B.N., Lori, C., Ozaki, S., Fucile, G., Plaza-Menacho, I., Jenal, U., and Schirmer, T. (2016). Cyclic di-GMP mediates a histidine kinase/phosphatase switch by noncovalent domain cross-linking. Sci Adv 2, e1600823.

Dubnau, D., and Losick, R. (2006). Bistability in bacteria. Molecular Microbiology 61, 564–572.

Duss, O., Michel, E., Konté, N.D. dit, Schubert, M., and Allain, F.H.-T. (2014). Molecular basis for the wide range of affinity found in Csr/Rsm protein–RNA recognition. Nucleic Acids Res 42, 5332–5346.

Düvel, J., Bense, S., Moller, S., Bertinetti, D., Schwede, F., Morr, M., Eckweiler, D., Genieser, H.-G., Jansch, L., Herberg, F.W., et al. (2016). Application of synthetic peptide arrays to uncover cyclic di-GMP binding motifs. Journal of Bacteriology 198, 138--146.

Düvel, J., Bertinetti, D., Möller, S., Schwede, F., Morr, M., Wissing, J., Radamm, L., Zimmermann, B., Genieser, H.-G., Jänsch, L., et al. (2012). A chemical proteomics approach to identify c-di-GMP binding proteins in Pseudomonas aeruginosa. J Microbiol Meth 88, 229–236.

Eisenberg, R.C., and Dobrogosz, W.J. (1967). Gluconate Metabolism in Escherichia coli1. J Bacteriol 93, 941–949.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic Gene Expression in a Single Cell. Science 297, 1183–1186.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallographica. Section D, Biological Crystallography 60, 2126–2132.

England, P., Westblade, L.F., Karimova, G., Robbe-Saule, V., Norel, F., and Kolb, A. (2008). Binding of the Unorthodox Transcription Activator, Crl, to the Components of the Transcription Machinery. J Biol Chem 283, 33455–33464.

Fahrner, K.A., and Berg, H.C. (2015). Mutations That Stimulate flhDC Expression in Escherichia coli K-12. J Bacteriol 197, 3087–3096.

Fang, X., Ahmad, I., Blanka, A., Schottkowski, M., Cimdins, A., Galperin, M.Y., Römling, U., and Gomelsky, M. (2014). GIL, a new c-di-GMP-binding protein domain involved in regulation of cellulose synthesis in enterobacteria. Mol Microbiol 93, 439–452.

Ferenci, T. (2005). Maintaining a healthy SPANC balance through regulatory and mutational adaptation. Mol

Microbiol 57, 1–8.

Fernandes, S., and São-José, C. (2018). Enzymes and Mechanisms Employed by Tailed Bacteriophages to Breach the Bacterial Cell Barriers. Viruses 10, 396.

Fernandez, N.L., Hsueh, B.Y., Nhu, N.T.Q., Franklin, J.L., Dufour, Y.S., and Waters, C.M. (2020). Vibrio cholerae adapts to sessile and motile lifestyles by cyclic di-GMP regulation of cell shape. Proc National Acad Sci 117, 29046–29054.

Ferrell, J.E. (2002). Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Curr Opin Cell Biol 14, 140–148.

Fic, E., Bonarek, P., Gorecki, A., Kedracka-Krok, S., Mikolajczak, J., Polit, A., Tworzydlo, M., Dziedzicka-Wasylewska, M., and Wasylewski, Z. (2009). cAMP Receptor Protein from Escherichia coli as a Model of Signal Transduction in Proteins – A Review. J Mol Microb Biotech 17, 1–11.

Fischer, E., and Sauer, U. (2003). A Novel Metabolic Cycle Catalyzes Glucose Oxidation and Anaplerosis in Hungry Escherichia coli *. J Biol Chem 278, 46446–46451.

Flemming, H.-C., and Wingender, J. (2010). The biofilm matrix. Nat Rev Microbiol 8, 623–633.

Folly, B.B., Ortega, A.D., Hubmann, G., Bonsing-Vedelaar, S., Wijma, H.J., Meulen, P. van der, Milias-Argeitis, A., and Heinemann, M. (2018). Assessment of the interaction between the flux-signaling metabolite fructose-1,6-bisphosphate and the bacterial transcription factors CggR and Cra. Mol Microbiol 109, 278–290.

Fong, J.C.N., and Yildiz, F.H. (2008). Interplay between Cyclic AMP-Cyclic AMP Receptor Protein and Cyclic di-GMP Signaling in Vibrio cholerae Biofilm Formation. J Bacteriol 190, 6646–6659.

Francez-Charlot, A., Laugel, B., Gemert, A.V., Dubarry, N., Wiorowski, F., Castanié-Cornet, M., Gutierrez, C., and Cam, K. (2003). RcsCDB His-Asp phosphorelay system negatively regulates the flhDC operon in Escherichia coli. Mol Microbiol 49, 823–832.

Freddolino, P.L., Amini, S., and Tavazoie, S. (2012). Newly Identified Genetic Variations in Common Escherichia coli MG1655 Stock Cultures. J Bacteriol 194, 303–306.

Galinier, A., and Deutscher, J. (2017). Sophisticated Regulation of Transcriptional Factors by the Bacterial Phosphoenolpyruvate: Sugar Phosphotransferase System. J Mol Biol 429, 773–789.

Galperin, M.Y. (2018). What bacteria want. Environ Microbiol 20, 4221–4229.

Galperin, M.Y., and Chou, S.-H. (2019). Structural Conservation and Diversity of PilZ-Related Domains. J Bacteriol 202.

Galperin, M.Y., Higdon, R., and Kolker, E. (2010). Interplay of heritage and habitat in the distribution of bacterial signal transduction systems. Mol Biosyst 6, 721–728.

Galperin, M.Y., Natale, D.A., Aravind, L., and Koonin, E.V. (1999). A specialized version of the HD hydrolase domain implicated in signal transduction. Journal of Molecular Microbiology and Biotechnology 1, 303.

Galperin, M.Y., Nikolskaya, A.N., and Koonin, E.V. (2001). Novel domains of the prokaryotic two-component signal transduction systems. Fems Microbiol Lett 203, 11–21.

García, B., Latasa, C., Solano, C., Portillo, F.G., Gamazo, C., and Lasa, I. (2004). Role of the GGDEF protein family in Salmonella cellulose biosynthesis and biofilm formation. Mol Microbiol 54, 264–277.

García-Betancur, J.C., and Lopez, D. (2019). Cell Heterogeneity in Staphylococcal Communities. J Mol Biol 431, 4699–4711.

Geerse, R., Ruig, C., Schuitema, A., and Postma, P. (1986). Relationship between pseudo-HPr and the PEP: fructose phosphotransferase system in Salmonella typhimurium and Escherichia coli. Molecular and General Genetics MGG 203, 435--444.

Geerse, R.H., Pluijm, J. van der, and Postma, P.W. (1989). The repressor of the PEP: Fructose phosphotransferase system is required for the transcription of the pps gene of Escherichia coli. Mol Gen Genetics Mgg 218, 348–352.

Gerosa, L., Kochanowski, K., Heinemann, M., and Sauer, U. (2013). Dissecting specific and global transcriptional regulation of bacterial gene expression. Mol Syst Biol 9, 658.

Giacalone, D., Smith, T.J., Collins, A.J., Sondermann, H., Koziol, L.J., and O'Toole, G.A. (2018). Ligand-Mediated Biofilm Formation via Enhanced Physical Interaction between a Diguanylate Cyclase and Its Receptor. mBio 9, e01254-18.

Göhler, A.-K., Kökpinar, Ö., Schmidt-Heck, W., Geffers, R., Guthke, R., Rinas, U., Schuster, S., Jahreis, K., and Kaleta, C. (2011). More than just a metabolic regulator - elucidation and validation of new targets of PdhR in Escherichia coli. Bmc Syst Biol 5, 197.

Gotfredsen, M., and Gerdes, K. (1998). The Escherichia coli relBE genes belong to a new toxin–antitoxin gene family. Mol Microbiol 29, 1065–1076.

Grainger, D.C., Hurd, D., Goldberg, M.D., and Busby, S.J.W. (2006). Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome. Nucleic Acids Res 34, 4642–4652.

Green, J., Stapleton, M.R., Smith, L.J., Artymiuk, P.J., Kahramanoglou, C., Hunt, D.M., and Buxton, R.S. (2014). Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches. Curr Opin Microbiol 18, 1–7.

Gregor, T., Wieschaus, E.F., McGregor, A.P., Bialek, W., and Tank, D.W. (2007). Stability and Nuclear Dynamics of the Bicoid Morphogen Gradient. Cell 130, 141–152.

Grenier, F., Matteau, D., Baby, V., and Rodrigue, S. (2014). Complete Genome Sequence of Escherichia coli BW25113. Genome Announc 2, e01038-14.

Grindley, N.D. (1978). IS1 insertion generates duplication of a nine base pair sequence at its target site. Cell 13, 419--426.

Guyer, M.S., Reed, R.R., Steitz, J.A., and Low, K.B. (1981). Identification of a Sex-factor-affinity Site in E. coli as gamma delta. Cold Spring Harb Symp 45, 135–140.

Guzzo, C.R., Salinas, R.K., Andrade, M.O., and Farah, C.S. (2009). PILZ Protein Structure and Interactions with PILB and the FIMX EAL Domain: Implications for Control of Type IV Pilus Biogenesis. J Mol Biol 393, 848–866.

Habazettl, J., Allan, M.G., Jenal, U., and Grzesiek, S. (2011). Solution Structure of the PilZ Domain Protein PA4608 Complex with Cyclic di-GMP Identifies Charge Clustering as Molecular Readout. J Biol Chem 286, 14304–14314.

Hall-Stoodley, L., Costerton, J.W., and Stoodley, P. (2004). Bacterial biofilms: from the Natural environment to infectious diseases. Nat Rev Microbiol 2, 95–108.

Hammar, M., Arnqvist, A., Bian, Z., Olsen, A., and Normark, S. (1995). Expression of two csg operons is required for production of fibronectin-and congo red-binding curli polymers in Escherichia coli K-12. Molecular Microbiology 18, 661--670.

Hamoen, L.W., Werkhoven, A.F.V., Bijlsma, J.J.E., Dubnau, D., and Venema, G. (1998). The competence transcription factor of Bacillus subtilis recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. Gene Dev 12, 1539–1550.

Harms, A., Maisonneuve, E., and Gerdes, K. (2016). Mechanisms of bacterial persistence during stress and antibiotic exposure. Science 354, aaf4268.

Hauryliuk, V., Atkinson, G.C., Murakami, K.S., Tenson, T., and Gerdes, K. (2015). Recent functional insights into the role of (p)ppGpp in bacterial physiology. Nature Reviews. Microbiology 13, 298–309.

Hayashi, K., Morooka, N., Yamamoto, Y., Fujita, K., Isono, K., Choi, S., Ohtsubo, E., Baba, T., Wanner, B.L., Mori, H., et al. (2006). Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110. Mol Syst Biol 2.

Heinrich, A.K., Glaeser, A., Tobias, N.J., Heermann, R., and Bode, H.B. (2016). Heterogeneous regulation of bacterial natural product biosynthesis via a novel transcription factor. Heliyon 2, e00197.

Hengge, R. (2020). Linking bacterial growth, survival, and multicellularity – small signaling molecules as triggers and drivers. Curr Opin Microbiol 55, 57–66.

Hengge, R. (2021). High-Specificity Local and Global c-di-GMP Signaling. Trends Microbiol.

Hengge, R., Galperin, M.Y., Ghigo, J.-M., Gomelsky, M., Green, J., Hughes, K.T., Jenal, U., and Landini, P. (2016). Systematic Nomenclature for GGDEF and EAL Domain-Containing Cyclic Di-GMP Turnover Proteins of Escherichia coli. J Bacteriol 198, 7–11.

Hengge, R., Häussler, S., Pruteanu, M., Stülke, J., Tschowri, N., and Turgay, K. (2019). Recent Advances and Current Trends in Nucleotide Second Messenger Signaling in Bacteria. J Mol Biol 431, 908–927.

Heo, K., Park, Y.-H., Lee, K.-A., Kim, J., Ham, H.-I., Kim, B.-G., Lee, W.-J., and Seok, Y.-J. (2019). Sugar-mediated regulation of a c-di-GMP phosphodiesterase in Vibrio cholerae. Nat Commun 10, 5358.

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.

Hoch, J.A. (2000). Two-component and phosphorelay signal transduction James A Hoch. Curr Opin Microbiol 3, 165–170.

Hochstrasser, R., and Hilbi, H. (2020). Legionella quorum sensing meets cyclic-di-GMP signaling. Curr Opin Microbiol 55, 9–16.

Hollinshead, W.D., Rodriguez, S., Martin, H.G., Wang, G., Baidoo, E.E.K., Sale, K.L., Keasling, J.D., Mukhopadhyay, A., and Tang, Y.J. (2016). Examining Escherichia coli glycolytic pathways, catabolite repression, and metabolite channeling using Δpfk mutants. Biotechnol Biofuels 9, 212.

Hug, I., Deshpande, S., Sprecher, K.S., Pfohl, T., and Jenal, U. (2017). Second messenger–mediated tactile response

by a bacterial rotary motor. Science 358, 531–534.

Hui, S., Silverman, J.M., Chen, S.S., Erickson, D.W., Basan, M., Wang, J., Hwa, T., and Williamson, J.R. (2015). Quantitative proteomic analysis reveals a simple strategy of global resource allocation in bacteria. Mol Syst Biol 11, 784.

Hunter, J.L., Severin, G.B., Koestler, B.J., and Waters, C.M. (2014). The Vibrio cholerae diguanylate cyclase VCA0965 has an AGDEF active site and synthesizes cyclic di-GMP. Bmc Microbiol 14, 22.

Ikebe, T., Iyoda, S., and Kutsukake, K. (1999). Structure and expression of the fliA operon of Salmonella typhimurium. Microbiology+ 145, 1389–1396.

Irving, S.E., and Corrigan, R.M. (2018). Triggering the stringent response: signals responsible for activating (p) ppGpp synthesis in bacteria. Microbiology+ 164, 268–276.

Jackson, D.W., Suzuki, K., Oakford, L., Simecka, J.W., Hart, M.E., and Romeo, T. (2002). Biofilm Formation and Dispersal under the Influence of the Global Regulator CsrA of Escherichia coli. J Bacteriol 184, 290–301.

Jaeger, J. (2011). The gap gene network. Cell Mol Life Sci 68, 243–274.

Jahreis, K., Postma, P., and Lengeler, J. (1991). Nucleotide sequence of the ilvH-fruR gene region of Escherichia coli K12 and Salmonella typhimurium LT2. Molecular and General Genetics MGG 226, 332--336.

Jenal, U., Reinders, A., and Lori, C. (2017). Cyclic di-GMP: second messenger extraordinaire. Nature Reviews. Microbiology 15, 271–284.

Jishage, M., Kvint, K., Shingler, V., and Nyström, T. (2002). Regulation of ς factor competition by the alarmone ppGpp. Gene Dev 16, 1260–1270.

Johnson, J.G., Murphy, C.N., Sippy, J., Johnson, T.J., and Clegg, S. (2011). Type 3 Fimbriae and Biofilm Formation Are Regulated by the Transcriptional Regulators MrkHI in Klebsiella pneumoniae. J Bacteriol 193, 3453–3460.

Jonas, K., and Melefors, Ö. (2009). The Escherichia coli CsrB and CsrC small RNAs are strongly induced during growth in nutrient-poor medium. Fems Microbiol Lett 297, 80–86.

Jonas, K., Tomenius, H., Römling, U., Georgellis, D., and Melefors, Ö. (2006). Identification of YhdA as a regulator of the Escherichia coli carbon storage regulation system. Fems Microbiol Lett 264, 232–237.

Jonas, K., Edwards, A.N., Simm, R., Romeo, T., Römling, U., and Melefors, Ö. (2008). The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. Mol Microbiol 70, 236–257.

Kabsch, W. (2010). XDS. Acta Crystallographica. Section D, Biological Crystallography 66, 125–132.

Kaiser, M., Jug, F., Julou, T., Deshpande, S., Pfohl, T., Silander, O.K., Myers, G., and Nimwegen, E. van (2018). Monitoring single-cell gene regulation under dynamically controllable conditions with integrated microfluidics and software. Nat Commun 9, 212.

Kalia, D., Merey, G., Nakayama, S., Zheng, Y., Zhou, J., Luo, Y., Guo, M., Roembke, B.T., and Sintim, H.O. (2013). Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. Chemical Society Reviews 42, 305–341.

Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M.G., and Alon, U. (2001). Ordering Genes in a Flagella Pathway by Analysis of Expression Kinetics from Living Bacteria. Science 292, 2080–2083.

Kanjee, U., Gutsche, I., Alexopoulos, E., Zhao, B., Bakkouri, M.E., Thibault, G., Liu, K., Ramachandran, S., Snider, J., Pai, E.F., et al. (2011). Linkage between the bacterial acid stress and stringent responses: the structure of the inducible lysine decarboxylase. Embo J 30, 931–944.

Karaolis, D.K.R., Rashid, M.H., Chythanya, R., Luo, W., Hyodo, M., and Hayakawa, Y. (2005). c-di-GMP (3--5--Cyclic Diguanylic Acid) Inhibits Staphylococcus aureus Cell-Cell Interactions and Biofilm Formation. Antimicrob Agents Ch 49, 1029–1038.

Keseler, I.M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martínez, C., Caspi, R., Fulcher, C., Gama-Castro, S., Kothari, A., Krummenacker, M., et al. (2017). The EcoCyc database: reflecting new knowledge about Escherichia coli K-12. Nucleic Acids Res 45, D543–D550.

Kiino, D.R., Licudine, R., Wilt, K., Yang, D.H., and Rothman-Denes, L.B. (1993). A cytoplasmic protein, NfrC, is required for bacteriophage N4 adsorption. Journal of Bacteriology 175, 7074–7080.

Kim, D., Seo, S.W., Gao, Y., Nam, H., Guzman, G.I., Cho, B.-K., and Palsson, B.O. (2018). Systems assessment of transcriptional regulation on central carbon metabolism by Cra and CRP. Nucleic Acids Res 46, gky069-.

Kim, J., Oliveros, J.C., Nikel, P.I., Lorenzo, V., and Silva-Rocha, R. (2013). Transcriptomic fingerprinting of Pseudomonas putida under alternative physiological regimes. Env Microbiol Rep 5, 883–891.

Kim, J.M., Garcia-Alcala, M., Balleza, E., and Cluzel, P. (2020). Stochastic transcriptional pulses orchestrate flagellar

biosynthesis in Escherichia coli. Sci Adv 6, eaax0947.

King, T., Ishihama, A., Kori, A., and Ferenci, T. (2004). A Regulatory Trade-Off as a Source of Strain Variation in the Species Escherichia coli†. J Bacteriol 186, 5614–5620.

Kirkpatrick, C.L., and Viollier, P.H. (2012). Decoding Caulobacter development. Fems Microbiol Rev 36, 193–205.

Klauck, G., Serra, D.O., Possling, A., and Hengge, R. (2018). Spatial organization of different sigma factor activities and c-di-GMP signalling within the three-dimensional landscape of a bacterial biofilm. Open Biol 8, 180066.

Kleina, L.G., and Miller, J.H. (1990). Genetic studies of the lac repressor: XIII. Extensive amino acid replacements generated by the use of natural and synthetic nonsense suppressors. Journal of Molecular Biology 212, 295--318.

Ko, M., and Park, C. (2000). Two novel flagellar components and H-NS are involved in the motor function of Escherichia coli. J. Mol. Biol. 303, 371--382.

Kochanowski, K., Gerosa, L., Brunner, S.F., Christodoulou, D., Nikolaev, Y.V., and Sauer, U. (2017). Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli. Mol Syst Biol 13, 903.

Kochanowski, K., Volkmer, B., Gerosa, L., Rijsewijk, B.R.H. van, Schmidt, A., and Heinemann, M. (2013). Functioning of a metabolic flux sensor in Escherichia coli. Proc National Acad Sci 110, 1130–1135.

Koirala, S., Mears, P., Sim, M., Golding, I., Chemla, Y.R., Aldridge, P.D., and Rao, C.V. (2014). A Nutrient-Tunable Bistable Switch Controls Motility in Salmonella enterica Serovar Typhimurium. Mbio 5, e01611-14.

Komeda, Y. (1982). Fusions of flagellar operons to lactose genes on a mu lac bacteriophage. Journal of Bacteriology 150, 16--26.

Komeda, Y. (1986). Transcriptional control of flagellar genes in Escherichia coli K-12. J Bacteriol 168, 1315–1318.

Kornberg, H.L. (2001). Routes for fructose utilization by Escherichia coli. J Mol Microb Biotech 3, 355–359.

Koshland, D., Goldbeter, A., and Stock, J. (1982). Amplification and adaptation in regulatory and sensory systems. Science 217, 220–225.

Kotte, O., Volkmer, B., Radzikowski, J.L., and Heinemann, M. (2014). Phenotypic bistability in Escherichia coli's central carbon metabolism. Mol Syst Biol 10, 736.

Kotte, O., Zaugg, J.B., and Heinemann, M. (2010). Bacterial adaptation through distributed sensing of metabolic fluxes. Mol Syst Biol 6, 355.

Krasteva, P.V., Fong, J.C.N., Shikuma, N.J., Beyhan, S., Navarro, M.V.A.S., Yildiz, F.H., and Sondermann, H. (2010). Vibrio cholerae VpsT Regulates Matrix Production and Motility by Directly Sensing Cyclic di-GMP. Science 327, 866–868.

Kremling, A., Geiselmann, J., Ropers, D., and Jong, H. de (2015). Understanding carbon catabolite repression in Escherichia coli using quantitative models. Trends Microbiol 23, 99–109.

Kristensen, H.-H., Valentin-Hansen, P., and Soøgaard-Andersen, L. (1996). CytR/cAMP-CRP Nucleoprotein Formation inE. coli: The CytR Repressor Binds its Operator as a Stable Dimer in a Ternary Complex with cAMP-CRP. J Mol Biol 260, 113–119.

Kropinski, A.M., Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P. (2009). Bacteriophages, Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions. Methods Mol Biology 501, 69–76.

Kunz, S., and Graumann, P.L. (2020). Spatial organization enhances versatility and specificity in cyclic di-GMP signaling. Biol Chem 401, 1323–1334.

Kunz, S., Tribensky, A., Steinchen, W., Oviedo-Bocanegra, L., Bedrunka, P., and Graumann, P.L. (2020). Cyclic di-GMP Signaling in Bacillus subtilis Is Governed by Direct Interactions of Diguanylate Cyclases and Cognate Receptors. Mbio 11, e03122-19.

Kutsukake, K., Ohya, Y., and Iino, T. (1990). Transcriptional analysis of the flagellar regulon of Salmonella typhimurium. J Bacteriol 172, 741–747.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nat Rev Microbiol 8, 317–327.

Lange, R., and Hengge-Aronis, R. (1991). Identification of a central regulator of stationary phase gene expression in Escherichia coli. Mol Microbiol 5, 49–59.

Latka, A., Maciejewska, B., Majkowska-Skrobek, G., Briers, Y., and Drulis-Kawa, Z. (2017). Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. Appl Microbiol Biot 101, 3103–3119.

Laventie, B.-J., and Jenal, U. (2020). Surface Sensing and Adaptation in Bacteria. Annu Rev Microbiol 74, 735–760.

Laventie, B.-J., Glatter, T., and Jenal, U. (2017). Pull-down with a c-di-GMP-specific capture compound coupled to mass spectrometry as a powerful tool to identify novel effector proteins. In C-Di-GMP Signaling, (Springer), pp.

361--376.

Laventie, B.-J., Sangermani, M., Estermann, F., Manfredi, P., Planes, R., Hug, I., Jaeger, T., Meunier, E., Broz, P., and Jenal, U. (2019). A Surface-Induced Asymmetric Program Promotes Tissue Colonization by Pseudomonas aeruginosa. Cell Host & Microbe 25, 140-152.e6.

Leclerc, G., Noel, G., and Drapeau, G.R. (1990). Molecular cloning, nucleotide sequence, and expression of shl, a new gene in the 2-minute region of the genetic map of Escherichia coli. Journal of Bacteriology 172, 4696--4700.

Lee, C.K., Anda, J. de, Baker, A.E., Bennett, R.R., Luo, Y., Lee, E.Y., Keefe, J.A., Helali, J.S., Ma, J., Zhao, K., et al. (2018). Multigenerational memory and adaptive adhesion in early bacterial biofilm communities. Proceedings of the National Academy of Sciences 104, 4471-4476.

Lehnen, D., Blumer, C., Polen, T., Wackwitz, B., Wendisch, V.F., and Unden, G. (2002). LrhA as a new transcriptional key regulator of flagella, motility and chemotaxis genes in Escherichia coli. Mol Microbiol 45, 521–532.

Lehning, C.E., Siedler, S., Ellabaan, M.M.H., and Sommer, M.O.A. (2017). Assessing glycolytic flux alterations resulting from genetic perturbations in E. coli using a biosensor. Metab Eng 42, 194–202.

Leifson, E. (1960). Atlas of Bacterial Flagellation (New York & London, Academic Press.).

Lim, H.G., Seo, S.W., and Jung, G.Y. (2013). Engineered Escherichia coli for simultaneous utilization of galactose and glucose. Bioresource Technol 135, 564–567.

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. The EMBO Journal 32, 2001–2014.

Litsios, A., Ortega, Á.D., Wit, E.C., and Heinemann, M. (2018). Metabolic-flux dependent regulation of microbial physiology. Curr Opin Microbiol 42, 71–78.

Liu, J., Lestrade, D., Arabaciyan, S., Cescut, J., François, J.-M., and Capp, J.-P. (2018). A GRX1 Promoter Variant Confers Constitutive Noisy Bimodal Expression That Increases Oxidative Stress Resistance in Yeast. Front Microbiol 9, 2158.

Liu, M.Y., Gui, G., Wei, B., Preston, J.F., Oakford, L., Yüksel, Ü., Giedroc, D.P., and Romeo, T. (1997). The RNA Molecule CsrB Binds to the Global Regulatory Protein CsrA and Antagonizes Its Activity in Escherichia coli. J Biol Chem 272, 17502–17510.

Liu, X., and Matsumura, P. (1994). The FlhD/FlhC complex, a transcriptional activator of the Escherichia coli flagellar class II operons. J Bacteriol 176, 7345–7351.

Liu, X., Beyhan, S., Lim, B., Linington, R.G., and Yildiz, F.H. (2010). Identification and Characterization of a Phosphodiesterase That Inversely Regulates Motility and Biofilm Formation in Vibrio cholerae J Bacteriol 192, 4541–4552.

Loewen, P.C., and Hengge-Aronis, R. (1994). The role of the sigma factor σS (KatF) in bacterial global regulation. Annual Review of Microbiology 48, 53--80.

Lori, C., Ozaki, S., Steiner, S., Böhm, R., Abel, S., Dubey, B.N., Schirmer, T., Hiller, S., and Jenal, U. (2015). Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. Nature 523, 236–239.

Losick, R., and Desplan, C. (2008). Stochasticity and Cell Fate. Science 320, 65–68.

Lucas, K.A., Pitari, G.M., Kazerounian, S., Ruiz-Stewart, I., Park, J., Schulz, S., Chepenik, K.P., and Waldman, S.A. (2000). Guanylyl cyclases and signaling by cyclic GMP. Pharmacol Rev 52, 375–414.

Madigan, M.T., Clark, D.P., Stahl, D., and Martinko, J.M. (2010). Brock biology of microorganisms 13th edition (Benjamin Cummings).

Mahillon, J., and Chandler, M. (1998). Insertion Sequences. Microbiol Mol Biol R 62, 725–774.

Majewski, R.A., and Domach, M.M. (1990). Simple constrained-optimization view of acetate overflow in E. coli. Biotechnol Bioeng 35, 732–738.

Markiewicz, P., Kleina, L.G., Cruz, C., Ehret, S., and Miller, J.H. (1994). Genetic studies of the lac repressor. XIV. Analysis of 4000 altered Escherichia coli lac repressors reveals essential and non-essential residues, as well as" spacers" which do not require a specific sequence. Journal of Molecular Biology 240, 421--433.

Martins, B.M., and Locke, J.C. (2015). Microbial individuality: how single-cell heterogeneity enables population level strategies. Curr Opin Microbiol 24, 104–112.

Matsuoka, Y., and Shimizu, K. (2011). Metabolic regulation in Escherichia coli in response to culture environments via global regulators. Biotechnol J 6, 1330–1341.

Mattevi, A., Valentini, G., Rizzi, M., Speranza, M.L., Bolognesi, M., and Coda, A. (1995). Crystal structure of Escherichia coli pyruvate kinase type I: molecular basis of the allosteric transition. Structure 3, 729–741.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser

crystallographic software. Journal of Applied Crystallography 40, 658–674.

McKee, A.E., Rutherford, B.J., Chivian, D.C., Baidoo, E.K., Juminaga, D., Kuo, D., Benke, P.I., Dietrich, J.A., Ma, S.M., Arkin, A.P., et al. (2012). Manipulation of the carbon storage regulator system for metabolite remodeling and biofuel production in Escherichia coli. Microb Cell Fact 11, 79.

McPartland, J., and Rothman-Denes, L.B. (2008). The Tail Sheath of Bacteriophage N4 Interacts with the Escherichia coli Receptor. Journal of Bacteriology 191, 525–532.

Meng, H., Wang, C., Yuan, Q., Ren, J., and Zeng, A.-P. (2021). An Aldolase-Based New Pathway for Bioconversion of Formaldehyde and Ethanol into 1,3-Propanediol in Escherichia coli. Acs Synth Biol 10, 799–809.

Meng, L.M., and Nygaard, P. (1990). Identification of hypoxanthine and guanine as the co-repressors for the purine regulon genes of Escherichia coli. Mol Microbiol 4, 2187–2192.

Merighi, M., Lee, V.T., Hyodo, M., Hayakawa, Y., and Lory, S. (2007). The second messenger bis-(3'-5')-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in Pseudomonas aeruginosa. Molecular Microbiology 65, 876–895.

Mey, A.R., Butz, H.A., and Payne, S.M. (2015). Vibrio cholerae CsrA Regulates ToxR Levels in Response to Amino Acids and Is Essential for Virulence. Mbio 6, e01064-15.

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 arginine-sensing pathway. Science Signaling 8, ra57.

Milo, R., Jorgensen, P., Moran, U., Weber, G., and Springer, M. (2010). BioNumbers—the database of key numbers in molecular and cell biology. Nucleic Acids Res 38, D750–D753.

Minasov, G., Padavattan, S., Shuvalova, L., Brunzelle, J.S., Miller, D.J., Basle, A., Massa, C., Collart, F.R., Schirmer, T., and Anderson, W.F. (2009). Crystal Structures of YkuI and Its Complex with Second Messenger Cyclic Di-GMP Suggest Catalytic Mechanism of Phosphodiester Bond Cleavage by EAL Domains. Journal of Biological Chemistry 284, 13174–13184.

Molenaar, D., Berlo, R. van, Ridder, D. de, and Teusink, B. (2009). Shifts in growth strategies reflect tradeoffs in cellular economics. Mol Syst Biol 5, 323.

Monod, J. (1949). The Growth of Bacterial Cultures. Annu Rev Microbiol 3, 371–394.

Morgan, J.L.W., McNamara, J.T., and Zimmer, J. (2014). Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nat Struct Mol Biol 21, 489–496.

Morin, M., Ropers, D., Letisse, F., Laguerre, S., Portais, J., Cocaign-Bousquet, M., and Enjalbert, B. (2016). The post-transcriptional regulatory system CSR controls the balance of metabolic pools in upper glycolysis of Escherichia coli. Mol Microbiol 100, 686–700.

Mouali, Y.E., Kim, H., Ahmad, I., Brauner, A., Liu, Y., Skurnik, M., Galperin, M.Y., and Römling, U. (2017). Stand-Alone EAL Domain Proteins Form a Distinct Subclass of EAL Proteins Involved in Regulation of Cell Motility and Biofilm Formation in Enterobacteria. J Bacteriol 199, e00179-17.

Müller, J.E.N., Meyer, F., Litsanov, B., Kiefer, P., Potthoff, E., Heux, S., Quax, W.J., Wendisch, V.F., Brautaset, T., Portais, J.-C., et al. (2015). Engineering Escherichia coli for methanol conversion. Metab Eng 28, 190–201.

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., and Jahn, D. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics 21, 4187–4189.

Mutalik, V.K., Adler, B.A., Rishi, H.S., Piya, D., Zhong, C., Koskella, B., Calendar, R., Novichkov, P., Price, M.N., Deutschbauer, A.M., et al. (2020). High-throughput mapping of the phage resistance landscape in E. coli. Biorxiv 2020.02.15.951020.

Naas, T., blot, M., Fitch, W.M., and Arber, W. (1994). Insertion Sequence-Related Genetic Variation in Resting Escherichia coli K-12. Genetics 136, 721–730.

Nadler, D.C., Morgan, S.-A., Flamholz, A., Kortright, K.E., and Savage, D.F. (2016). Rapid construction of metabolite biosensors using domain-insertion profiling. Nature Communications 7, 12266.

Nanchen, A., Schicker, A., Revelles, O., and Sauer, U. (2008). Cyclic AMP-Dependent Catabolite Repression Is the Dominant Control Mechanism of Metabolic Fluxes under Glucose Limitation in Escherichia coli J Bacteriol 190, 2323–2330.

Navarro, M.V.A.S., De, N., Bae, N., Wang, Q., and Sondermann, H. (2009). Structural Analysis of the GGDEF-EAL Domain-Containing c-di-GMP Receptor FimX. Structure 17, 1104–1116.

Navarro, M.V.A.S., Newell, P.D., Krasteva, P.V., Chatterjee, D., Madden, D.R., O'Toole, G.A., and Sondermann, H. (2011). Structural Basis for c-di-GMP-Mediated Inside-Out Signaling Controlling Periplasmic Proteolysis. PLoS Biology 9, e1000588.

Nègre, D., Oudot, C., Prost, J.-F., Murakami, K., Ishihama, A., Cozzone, A.J., and Cortay, J.-C. (1998). FruR-mediated

transcriptional activation at the ppsA promoter of Escherichia coli. J Mol Biol 276, 355–365.

Nelson, D.L., and Cox, M.M. (2008). Lehninger Principles of Biochemistry - fifth edition (New York: W.H. Freeman and Company).

Nesper, J., Hug, I., Kato, S., Hee, C.-S., Habazettl, J.M., Manfredi, P., Grzesiek, S., Schirmer, T., Emonet, T., and Jenal, U. (2017). Cyclic di-GMP differentially tunes a bacterial flagellar motor through a novel class of CheY-like regulators. eLife 6, e28842.

Nesper, J., Reinders, A., Glatter, T., Schmidt, A., and Jenal, U. (2012). A novel capture compound for the identification and analysis of cyclic di-GMP binding proteins. J Proteomics 75, 4874–4878.

Ni, B., Colin, R., Link, H., Endres, R.G., and Sourjik, V. (2020). Growth-rate dependent resource investment in bacterial motile behavior quantitatively follows potential benefit of chemotaxis. Proc National Acad Sci 117, 595–601.

Nikel, P.I., Romero-Campero, F.J., Zeidman, J.A., Goñi-Moreno, Á., and Lorenzo, V. de (2015). The Glycerol-Dependent Metabolic Persistence of Pseudomonas putida KT2440 Reflects the Regulatory Logic of the GlpR Repressor. Mbio 6, e00340-15.

Ninfa, A.J., and Mayo, A.E. (2004). Hysteresis vs. Graded Responses: The Connections Make All the Difference. Sci Signal 2004, pe20–pe20.

Nishino, T., Gallant, J., Shalit, P., Palmer, L., and Wehr, T. (1979). Regulatory nucleotides involved in the Rel function of Bacillus subtilis. J Bacteriol 140, 671–679.

Njoroge, J.W., Gruber, C., and Sperandio, V. (2013). The Interacting Cra and KdpE Regulators Are Involved in the Expression of Multiple Virulence Factors in Enterohemorrhagic Escherichia coli. J Bacteriol 195, 2499–2508.

Njoroge, J.W., Nguyen, Y., Curtis, M.M., Moreira, C.G., and Sperandio, V. (2012). Virulence Meets Metabolism: Cra and KdpE Gene Regulation in Enterohemorrhagic Escherichia coli. Mbio 3, e00280-12.

Nobelmann, B., and Lengeler, J.W. (1996). Molecular analysis of the gat genes from Escherichia coli and of their roles in galactitol transport and metabolism. J Bacteriol 178, 6790–6795.

Norman, T.M., Lord, N.D., Paulsson, J., and Losick, R. (2015). Stochastic Switching of Cell Fate in Microbes. Annual Review of Microbiology 69, 381–403.

Notley-McRobb, L., Death, A., and Ferenci, T. (1997). The relationship between external glucose concentration and cAMP levels inside Escherichia coli: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. Microbiology+ 143, 1909–1918.

Nystrom, T. (2004). MicroReview: Growth versus maintenance: a trade-off dictated by RNA polymerase availability and sigma factor competition? Molecular Microbiology 54, 855--862.

O'Brien, E.J., Lerman, J.A., Chang, R.L., Hyduke, D.R., and Palsson, B.Ø. (2013). Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. Mol Syst Biol 9, 693.

Ogura, M., Shindo, K., and Kanesaki, Y. (2020). Bacillus subtilis Nucleoid-Associated Protein YlxR Is Involved in Bimodal Expression of the Fructoselysine Utilization Operon (frlBONMD-yurJ) Promoter. Front Microbiol 11, 2024.

Okano, H., Hermsen, R., Kochanowski, K., and Hwa, T. (2020). Regulation underlying hierarchical and simultaneous utilization of carbon substrates by flux sensors in Escherichia coli. Nat Microbiol 5, 206–215.

Olsen, A., Arnqvist, A., Hammar, M. rten, Sukupolvi, S., and Normark, S. (1993). The RpoS sigma factor relieves H-NS-mediated transcriptional repression of csgA, the subunit gene of fibronectin-binding curli in Escherichia coli. Molecular Microbiology 7, 523--536.

Omasits, U., Ahrens, C.H., Müller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics 30, 884–886.

Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., and Jacobs-Wagner, C. (2015). Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. Molecular Microbiology. 99,767-777.

Paliwal, S., Iglesias, P.A., Campbell, K., Hilioti, Z., Groisman, A., and Levchenko, A. (2007). MAPK-mediated bimodal gene expression and adaptive gradient sensing in yeast. Nature 446, 46–51.

Park, J., Estrada, J., Johnson, G., Vincent, B.J., Ricci-Tam, C., Bragdon, M.D., Shulgina, Y., Cha, A., Wunderlich, Z., Gunawardena, J., et al. (2019). Dissecting the sharp response of a canonical developmental enhancer reveals multiple sources of cooperativity. eLife 8, e41266.

Parker, D.J., Demetci, P., and Li, G.-W. (2019). Rapid Accumulation of Motility-Activating Mutations in Resting Liquid Culture of Escherichia coli. J Bacteriol 201.

Paul, K., Nieto, V., Carlquist, W.C., Blair, D.F., and Harshey, R.M. (2010). The c-di-GMP Binding Protein YcgR

Controls Flagellar Motor Direction and Speed to Affect Chemotaxis by a "Backstop Brake" Mechanism. Mol Cell 38, 128–139.

Paul, R., Abel, S., Wassmann, P., Beck, A., Heerklotz, H., and Jenal, U. (2007). Activation of the Diguanylate Cyclase PleD by Phosphorylation-mediated Dimerization*. J Biol Chem 282, 29170–29177.

Paul, R., Jaeger, T., Abel, S., Wiederkehr, I., Folcher, M., Biondi, E.G., Laub, M.T., and Jenal, U. (2008). Allosteric Regulation of Histidine Kinases by Their Cognate Response Regulator Determines Cell Fate. Cell 133, 928.

Perraud, A.-L., Weiss, V., and Gross, R. (1999). Signalling pathways in two-component phosphorelay systems. Trends Microbiol 7, 115–120.

Perrenoud, A., and Sauer, U. (2005). Impact of Global Transcriptional Regulation by ArcA, ArcB, Cra, Crp, Cya, Fnr, and Mlc on Glucose Catabolism in Escherichia coli†. J Bacteriol 187, 3171–3179.

Persat, A., Inclán, Y.F., Engel, J.N., Stone, H.A., and Gitai, Z. (2015a). Type IV pili mechanochemically regulate virulence factors in Pseudomonas aeruginosa. Proceedings of the National Academy of Sciences 112, 7563–7568.

Persat, A., Nadell, C.D., Kim, M.K., Ingremeau, F., Siryaporn, A., Drescher, K., Wingreen, N.S., Bassler, B.L., Gitai, Z., and Stone, H.A. (2015b). The mechanical world of bacteria. Cell 161, 988–997.

Pesavento, C., and Hengge, R. (2009). Bacterial nucleotide-based second messengers. Curr Opin Microbiol 12, 170–176.

Pesavento, C., and Hengge, R. (2012). The global repressor FliZ antagonizes gene expression by σ S -containing RNA polymerase due to overlapping DNA binding specificity. Nucleic Acids Res 40, 4783–4793.

Pesavento, C., Becker, G., Sommerfeldt, N., Possling, A., Tschowri, N., Mehlis, A., and Hengge, R. (2008). Inverse regulatory coordination of motility and curli-mediated adhesion in Escherichia coli. Gene Dev 22, 2434–2446.

Peterson, A.C., Russell, J.D., Bailey, D.J., Westphall, M.S., and Coon, J.J. (2012). Parallel Reaction Monitoring for High Resolution and High Mass Accuracy Quantitative, Targeted Proteomics. Mol Cell Proteomics 11, 1475–1488.

Pires, D.P., Oliveira, H., Melo, L.D.R., Sillankorva, S., and Azeredo, J. (2016). Bacteriophage-encoded depolymerases: their diversity and biotechnological applications. Appl Microbiol Biot 100, 2141–2151.

Potrykus, K., Murphy, H., Philippe, N., and Cashel, M. (2011). ppGpp is the major source of growth rate control in E. coli. Environ Microbiol 13, 563–575.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., Stevenson, K., et al. (2018). CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallographica. Section D, Structural Biology 74, 68–84.

Potts, A.H., Guo, Y., Ahmer, B.M.M., and Romeo, T. (2019). Role of CsrA in stress responses and metabolism important for Salmonella virulence revealed by integrated transcriptomics. Plos One 14, e0211430.

Povolotsky, T.L., and Hengge, R. (2016). Genome-Based Comparison of Cyclic Di-GMP Signaling in Pathogenic and Commensal Escherichia coli Strains. J Bacteriol 198, 111–126.

Pratt, L.A., and Silhavy, T.J. (1998). Crl stimulates RpoS activity during stationary phase. Mol Microbiol 29, 1225–1236.

Prüss, B.M., and Matsumura, P. (1996). A regulator of the flagellar regulon of Escherichia coli, flhD, also affects cell division. J Bacteriol 178, 668–674.

Prüss, B.M., Campbell, J.W., Dyk, T.K.V., Zhu, C., Kogan, Y., and Matsumura, P. (2003). FlhD/FlhC Is a Regulator of Anaerobic Respiration and the Entner-Doudoroff Pathway through Induction of the Methyl-Accepting Chemotaxis Protein Aer. J Bacteriol 185, 534–543.

Prüss, B.M., Liu, X., Hendrickson, W., and Matsumura, P. (2001). FlhD/FlhC-regulated promoters analyzed by gene array and lacZ gene fusions. Fems Microbiol Lett 197, 91–97.

Prüss, B.M., Markovic, D., and Matsumura, P. (1997). The Escherichia coli flagellar transcriptional activator flhD regulates cell division through induction of the acid response gene cadA. J Bacteriol 179, 3818–3821.

Pultz, I.S., Christen, M., Kulasekara, H.D., Kennard, A., Kulasekara, B., and Miller, S.I. (2012). The response threshold of Salmonella PilZ domain proteins is determined by their binding affinities for c-di-GMP. Mol Microbiol 86, 1424–1440.

Pursley, B.R., Maiden, M.M., Hsieh, M.-L., Fernandez, N.L., Severin, G.B., and Waters, C.M. (2018). Cyclic di-GMP Regulates TfoY in Vibrio cholerae To Control Motility by both Transcriptional and Posttranscriptional Mechanisms. J Bacteriol 200, e00578-17.

Ramseier, T.M. (1996). Cra and the control of carbon flux via metabolic pathways. Res Microbiol 147, 489–493.

Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R., and Saier, M.H. (1995). The global regulatory protein FruR modulates the direction of carbon flow in Escherichia coli. Mol Microbiol 16, 1157–1169.

Ramseier, T.M., Nègre, D., Cortay, J.-C., Scarabel, M., Cozzone, A.J., and Saier, M.H. (1993). In Vitro Binding of the Pleiotropic Transcriptional Regulatory Protein, FruR, to the fru, pps, ace, pts and icd Operons of Escherichia coli and Salmonella typhimurium. J Mol Biol 234, 28–44.

Rangarajan, A.A., and Schnetz, K. (2018). Interference of transcription across H-NS binding sites and repression by H-NS. Mol Microbiol 108, 226–239.

Rao, F., Qi, Y., Chong, H.S., Kotaka, M., Li, B., Li, J., Lescar, J., Tang, K., and Liang, Z.X. (2009). The Functional Role of a Conserved Loop in EAL Domain-Based Cyclic di-GMP-Specific Phosphodiesterase. Journal of Bacteriology 191, 4722–4731.

Reinders, A., Hee, C.-S., Ozaki, S., Mazur, A., Boehm, A., Schirmer, T., and Jenal, U. (2016). Expression and Genetic Activation of Cyclic Di-GMP-Specific Phosphodiesterases in Escherichia coli. J Bacteriol 198, 448–462.

Renilla, S., Bernal, V., Fuhrer, T., Castaño-Cerezo, S., Pastor, J.M., Iborra, J.L., Sauer, U., and Cánovas, M. (2012). Acetate scavenging activity in Escherichia coli: interplay of acetyl–CoA synthetase and the PEP–glyoxylate cycle in chemostat cultures. Appl Microbiol Biot 93, 2109–2124.

Reshamwala, S.M.S., and Noronha, S.B. (2011). Biofilm formation in Escherichia colicra mutants is impaired due to down-regulation of curli biosynthesis. Arch Microbiol 193, 711.

Rhee, H.S., and Pugh, B.F. (2012). ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy. Current Protocols in Molecular Biology 100, 21--24.

Richter, A.M., Possling, A., Malysheva, N., Yousef, K.P., Herbst, S., Kleist, M. von, and Hengge, R. (2020). Local c-di-GMP Signaling in the Control of Synthesis of the E. coli Biofilm Exopolysaccharide pEtN-Cellulose. J Mol Biol 432, 4576–4595.

Richter, A.M., Povolotsky, T.L., Wieler, L.H., and Hengge, R. (2014). Cyclic-di-GMP signalling and biofilm-related properties of the Shiga toxin-producing 2011 German outbreak Escherichia coli O104:H4. Embo Mol Med 6, 1622–1637.

Rijsewijk, B.R.B.H. van, Nanchen, A., Nallet, S., Kleijn, R.J., and Sauer, U. (2011). Large-scale 13C-flux analysis reveals distinct transcriptional control of respiratory and fermentative metabolism in Escherichia coli. Mol Syst Biol 7, 477.

Riley, M., Abe, T., Arnaud, M.B., Berlyn, M.K.B., Blattner, F.R., Chaudhuri, R.R., Glasner, J.D., Horiuchi, T., Keseler, I.M., Kosuge, T., et al. (2006). Escherichia coli K-12: a cooperatively developed annotation snapshot—2005. Nucleic Acids Res 34, 1–9.

Roelofs, K.G., Jones, C.J., Helman, S.R., Shang, X., Orr, M.W., Goodson, J.R., Galperin, M.Y., Yildiz, F.H., and Lee, V.T. (2015). Systematic Identification of Cyclic-di-GMP Binding Proteins in Vibrio cholerae Reveals a Novel Class of Cyclic-di-GMP-Binding ATPases Associated with Type II Secretion Systems. Plos Pathog 11, e1005232.

Romeo, T., Gong, M., Liu, M.Y., and Brun-Zinkernagel, A.M. (1993). Identification and molecular characterization of csrA, a pleiotropic gene from Escherichia coli that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. J Bacteriol 175, 4744–4755.

Römling, U. (2005). Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. Cell Mol Life Sci Cmls 62, 1234–1246.

Romling, U., Galperin, M.Y., and Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiology and Molecular Biology Reviews 77, 1--52.

Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., Vroom, E. de, Marel, G.A. van der, Boom, J.H. van, et al. (1987). Regulation of cellulose synthesis in Acetobacter xylinum by cyclic diguanylic acid. Nature 325, 279–281.

Rumbaugh, K.P., and Sauer, K. (2020). Biofilm dispersion. Nat Rev Microbiol 18, 571–586.

Ryan, R.P., Lucey, J., O'Donovan, K., McCarthy, Y., Yang, L., Tolker-Nielsen, T., and Dow, J.M. (2009). HD-GYP domain proteins regulate biofilm formation and virulence in Pseudomonas aeruginosa. Environ Microbiol 11, 1126–1136.

Ryder, C., Byrd, M., and Wozniak, D.J. (2007). Role of polysaccharides in Pseudomonas aeruginosa biofilm development. Curr Opin Microbiol 10, 644–648.

Ryjenkov, D.A., Simm, R., Römling, U., and Gomelsky, M. (2006). The PilZ Domain Is a Receptor for the Second Messenger c-di-GMP THE PilZ DOMAIN PROTEIN YcgR CONTROLS MOTILITY IN ENTEROBACTERIA *. J Biol Chem 281, 30310–30314.

Sabnis, N.A., Yang, H., and Romeo, T. (1995). Pleiotropic Regulation of Central Carbohydrate Metabolism in Escherichia coli via the Gene csrA (*). J Biol Chem 270, 29096–29104.

Saini, S., Koirala, S., Floess, E., Mears, P.J., Chemla, Y.R., Golding, I., Aldridge, C., Aldridge, P.D., and Rao, C.V. (2010). FliZ Induces a Kinetic Switch in Flagellar Gene Expression. J Bacteriol 192, 6477–6481.

Santiveri, M., Roa-Eguiara, A., Kühne, C., Wadhwa, N., Hu, H., Berg, H.C., Erhardt, M., and Taylor, N.M.I. (2020). Structure and Function of Stator Units of the Bacterial Flagellar Motor. Cell 183, 244-257.e16.

Santos-Zavaleta, A., Salgado, H., Gama-Castro, S., Sánchez-Pérez, M., Gómez-Romero, L., Ledezma-Tejeida, D., García-Sotelo, J.S., Alquicira-Hernández, K., Muñiz-Rascado, L.J., Peña-Loredo, P., et al. (2018). RegulonDB v 10.5: tackling challenges to unify classic and high throughput knowledge of gene regulation in E. coli K-12. Nucleic Acids Res 47, gky1077-.

Sarenko, O., Klauck, G., Wilke, F.M., Pfiffer, V., Richter, A.M., Herbst, S., Kaever, V., and Hengge, R. (2017). More than Enzymes That Make or Break Cyclic Di-GMP—Local Signaling in the Interactome of GGDEF/EAL Domain Proteins of Escherichia coli. Mbio 8, e01639-17.

Schaller, G.E., Kieber, J.J., and Shiu, S.-H. (2008). Two-Component Signaling Elements and Histidyl-Aspartyl Phosphorelays. Arabidopsis Book 6, e0112.

Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent Escherichia coli proteome. Nat Biotechnol 34, 104–110.

Schmidt, A.J., Ryjenkov, D.A., and Gomelsky, M. (2005). The Ubiquitous Protein Domain EAL Is a Cyclic Diguanylate-Specific Phosphodiesterase: Enzymatically Active and Inactive EAL Domains†. J Bacteriol 187, 4774–4781.

Schniederberend, M., Williams, J.F., Shine, E., Shen, C., Jain, R., Emonet, T., and Kazmierczak, B.I. (2019). Modulation of flagellar rotation in surface-attached bacteria: A pathway for rapid surface-sensing after flagellar attachment. PLoS Pathogens 15, e1008149.

Schröter, L., and Dersch, P. (2019). Phenotypic diversification of microbial pathogens - cooperating and preparing for the future. J Mol Biol 431, 4645–4655.

Schuetz, R., Zamboni, N., Zampieri, M., Heinemann, M., and Sauer, U. (2012). Multidimensional Optimality of Microbial Metabolism. Science 336, 601–604.

Schuhmacher, J.S., Thormann, K.M., and Bange, G. (2015). How bacteria maintain location and number of flagella? Fems Microbiol Rev 39, 812–822.

Schumacher, M.A., and Zeng, W. (2016). Structures of the activator of K. pneumonia biofilm formation, MrkH, indicates PilZ domains involved in c-di-GMP and DNA binding. Proc National Acad Sci 113, 10067–10072.

Schumacher, M.A., Seidel, G., Hillen, W., and Brennan, R.G. (2007). Structural Mechanism for the Fine-tuning of CcpA Function by The Small Molecule Effectors Glucose 6-Phosphate and Fructose 1,6-Bisphosphate. J Mol Biol 368, 1042–1050.

Seidl, K., Stucki, M., Ruegg, M., Goerke, C., Wolz, C., Harris, L., Berger-Bächi, B., and Bischoff, M. (2006). Staphylococcus aureus CcpA Affects Virulence Determinant Production and Antibiotic Resistance. Antimicrob Agents Ch 50, 1183–1194.

Serra, D.O., and Hengge, R. (2019). A c-di-GMP-Based Switch Controls Local Heterogeneity of Extracellular Matrix Synthesis which Is Crucial for Integrity and Morphogenesis of Escherichia coli Macrocolony Biofilms. J Mol Biol 431, 4775–4793.

Serra, D.O., Richter, A.M., and Hengge, R. (2013). Cellulose as an Architectural Element in Spatially Structured Escherichia coli Biofilms. J Bacteriol 195, 5540–5554.

Seshasayee, A.S.N., Fraser, G.M., and Luscombe, N.M. (2010). Comparative genomics of cyclic-di-GMP signalling in bacteria: post-translational regulation and catalytic activity. Nucleic Acids Res 38, 5970–5981.

Seyll, E., and Melderen, L.V. (2013). The Ribonucleoprotein Csr Network. Int J Mol Sci 14, 22117–22131.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498, 236–240.

Shimada, T., Fujita, N., Maeda, M., and Ishihama, A. (2005). Systematic search for the Cra-binding promoters using genomic SELEX system. Genes Cells 10, 907–918.

Shimada, T., Yamamoto, K., and Ishihama, A. (2011). Novel Members of the Cra Regulon Involved in Carbon Metabolism in Escherichia coli. J Bacteriol 193, 649–659.

Shimizu, K. (2009). Toward systematic metabolic engineering based on the analysis of metabolic regulation by the integration of different levels of information. Biochem Eng J 46, 235–251.

Shin, S., and Park, C. (1995). Modulation of flagellar expression in Escherichia coli by acetyl phosphate and the osmoregulator OmpR. J Bacteriol 177, 4696–4702.

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.

Simm, R., Remminghorst, U., Ahmad, I., Zakikhany, K., and Römling, U. (2009). A Role for the EAL-Like Protein STM1344 in Regulation of CsgD Expression and Motility in Salmonella enterica Serovar Typhimurium. J Bacteriol 191, 3928–3937.

Simmons, E.L., Bond, M.C., Koskella, B., Drescher, K., Bucci, V., and Nadell, C.D. (2020). Biofilm Structure Promotes Coexistence of Phage-Resistant and Phage-Susceptible Bacteria. Msystems 5, e00877-19.

Sinderen, D. van, and Venema, G. (1994). comK acts as an autoregulatory control switch in the signal transduction route to competence in Bacillus subtilis. J Bacteriol 176, 5762–5770.

Singh, D., Fairlamb, M.S., Harrison, K.S., Weeramange, C., Meinhardt, S., Tungtur, S., Rau, B.F., Hefty, P.S., Fenton, A.W., and Swint-Kruse, L. (2017). Protein-protein interactions with fructose-1-kinase alter function of the central Escherichia coli transcription regulator, Cra. Biorxiv 201277.

Sjöström, A.E., Sondén, B., Müller, C., Rydström, A., Dobrindt, U., Wai, S.N., and Uhlin, B.E. (2009). Analysis of the sfaXII locus in the Escherichia coli meningitis isolate IHE3034 reveals two novel regulatory genes within the promoter-distal region of the main S fimbrial operon. Microb Pathogenesis 46, 150–158.

Smits, W.K., Kuipers, O.P., and Veening, J.-W. (2006). Phenotypic variation in bacteria: the role of feedback regulation. Nat Rev Microbiol 4, 259–271.

Sobala, M., Bruhn-Olszewska, B., Cashel, M., and Potrykus, K. (2019). Methylobacterium extorquens RSH Enzyme Synthesizes (p)ppGpp and pppApp in vitro and in vivo, and Leads to Discovery of pppApp Synthesis in Escherichia coli. Front Microbiol 10, 859.

Sommerfeldt, N., Possling, A., Becker, G., Pesavento, C., Tschowri, N., and Hengge, R. (2009). Gene expression patterns and differential input into curli fimbriae regulation of all GGDEF/EAL domain proteins in Escherichia coli. Microbiology+ 155, 1318–1331.

Soutourina, O., Kolb, A., Krin, E., Laurent-Winter, C., Rimsky, S., Danchin, A., and Bertin, P. (1999). Multiple Control of Flagellum Biosynthesis in Escherichia coli: Role of H-NS Protein and the Cyclic AMP-Catabolite Activator Protein Complex in Transcription of the flhDC Master Operon. J Bacteriol 181, 7500–7508.

Spangler, C., Böhm, A., Jenal, U., Seifert, R., and Kaever, V. (2010). A liquid chromatography-coupled tandem mass spectrometry method for quantitation of cyclic di-guanosine monophosphate. Journal of Microbiological Methods 81, 226–231.

Sprecher, K.S., Hug, I., Nesper, J., Potthoff, E., Mahi, M.-A., Sangermani, M., Kaever, V., Schwede, T., Vorholt, J., and Jenal, U. (2017). Cohesive Properties of the Caulobacter crescentus Holdfast Adhesin Are Regulated by a Novel c-di-GMP Effector Protein. MBio 8, e00294-17.

Srivastava, D., Hsieh, M., Khataokar, A., Neiditch, M.B., and Waters, C.M. (2013). Cyclic di GMP inhibits Vibrio cholerae motility by repressing induction of transcription and inducing extracellular polysaccharide production. Mol Microbiol 90, 1262–1276.

Steinchen, W., Zegarra, V., and Bange, G. (2020). (p)ppGpp: Magic Modulators of Bacterial Physiology and Metabolism. Front Microbiol 11, 2072.

Steiner, S., Lori, C., Boehm, A., and Jenal, U. (2013). Allosteric activation of exopolysaccharide synthesis through cyclic di-GMP-stimulated protein-protein interaction. The EMBO Journal 32, 354–368.

Stulke, J., and Hillen, W. (1999). Carbon catabolite repression in bacteria. Current Opinion in Microbiology 2, 195--201.

Stülke, J., and Krüger, L. (2020). Cyclic di-AMP Signaling in Bacteria. Annu Rev Microbiol 74, 1–21.

Stylianidou, S., Brennan, C., Nissen, S.B., Kuwada, N.J., and Wiggins, P.A. (2016). SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells. Mol Microbiol 102, 690–700.

Subramanian, S., Gao, X., Dann, C.E., and Kearns, D.B. (2017). MotI (DgrA) acts as a molecular clutch on the flagellar stator protein MotA in Bacillus subtilis. Proc National Acad Sci 114, 13537–13542.

Sundriyal, A., Massa, C., Samoray, D., Zehender, F., Sharpe, T., Jenal, U., and Schirmer, T. (2014). Inherent regulation of EAL domain-catalyzed hydrolysis of second messenger cyclic di-GMP. Journal of Biological Chemistry 289, 6978–6990.

Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc National Acad Sci 99, 12795–12800.

Swint-Kruse, L., and Matthews, K.S. (2009). Allostery in the LacI/GalR family: variations on a theme. Curr Opin Microbiol 12, 129–137.

Tamar, E., Koler, M., and Vaknin, A. (2016). The role of motility and chemotaxis in the bacterial colonization of protected surfaces. Sci Rep-Uk 6, 19616.

Tamman, H., Nerom, K.V., Takada, H., Vandenberk, N., Scholl, D., Polikanov, Y., Hofkens, J., Talavera, A., Hauryliuk, V., Hendrix, J., et al. (2020). A nucleotide-switch mechanism mediates opposing catalytic activities of

Rel enzymes. Nature Chem Biol 16, 834–840.

Tanaka, K., Takayanagi, Y., Fujita, N., Ishihama, A., and Takahashi, H. (1993). Heterogeneity of the principal σ factor in Escherichia coli The rpoS gene product, σ38, is a second principal σ f.pdf. Proc. Natl. Acad. Sci. USA 90, 3511–3515.

Taymaz-Nikerel, H., Mey, M. de, Ras, C., Pierick, A. ten, Seifar, R.M., Dam, J.C. van, Heijnen, J.J., and Gulik, W.M. van (2009). Development and application of a differential method for reliable metabolome analysis in Escherichia coli. Anal Biochem 386, 9–19.

Thattai, M., and Oudenaarden, A. van (2004). Stochastic Gene Expression in Fluctuating Environments. Genetics 167, 523–530.

Thompson, C.M., and Malone, J.G. (2020). Nucleotide second messengers in bacterial decision making. Curr Opin Microbiol 55, 34–39.

Thongsomboon, W., Serra, D.O., Possling, A., Hadjineophytou, C., Hengge, R., and Cegelski, L. (2018). Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. Science 359, 334–338.

Timmermans, J., and Melderen, L.V. (2009). Conditional Essentiality of the csrA Gene in Escherichia coli J Bacteriol 191, 1722–1724.

Timmermans, J., and Melderen, L.V. (2010). Post-transcriptional global regulation by CsrA in bacteria. Cell Mol Life Sci 67, 2897–2908.

Tobias, N.J., and Bode, H.B. (2019). Heterogeneity in Bacterial Specialized Metabolism. J Mol Biol 431, 4589–4598.

Tschowri, N., Schumacher, M.A., Schlimpert, S., Chinnam, N.B., Findlay, K.C., Brennan, R.G., and Buttner, M.J. (2014). Tetrameric c-di-GMP Mediates Effective Transcription Factor Dimerization to Control Streptomyces Development. Cell 158, 1136–1147.

Tuckerman, J.R., Gonzalez, G., Sousa, E.H.S., Wan, X., Saito, J.A., Alam, M., and Gilles-Gonzalez, M.-A. (2009). An Oxygen-Sensing Diguanylate Cyclase and Phosphodiesterase Couple for c-di-GMP Control. Biochemistry-Us 48, 9764–9774.

Typas, A., Barembruch, C., Possling, A., and Hengge, R. (2007). Stationary phase reorganisation of the Escherichia coli transcription machinery by Crl protein, a fine-tuner of σs activity and levels. Embo J 26, 1569–1578.

Umenhoffer, K., Fehér, T., Balikó, G., Ayaydin, F., Pósfai, J., Blattner, F.R., and Pósfai, G. (2010). Reduced evolvability of Escherichia coli MDS42, an IS-less cellular chassis for molecular and synthetic biology applications. Microb Cell Fact 9, 38.

Uyar, E., Kurokawa, K., Yoshimura, M., Ishikawa, S., Ogasawara, N., and Oshima, T. (2009). Differential Binding Profiles of StpA in Wild-Type and hns Mutant Cells: a Comparative Analysis of Cooperative Partners by Chromatin Immunoprecipitation-Microarray Analysis. J Bacteriol 191, 2388–2391.

Uzzau, S., Figueroa-Bossi, N., Rubino, S., and Bossi, L. (2001). Epitope tagging of chromosomal genes in Salmonella. Proc National Acad Sci 98, 15264–15269.

Valentini, G., Chiarelli, L., Fortin, R., Speranza, M.L., Galizzi, A., and Mattevi, A. (2000). The Allosteric Regulation of Pyruvate Kinase A SITE-DIRECTED MUTAGENESIS STUDY*. J Biol Chem 275, 18145–18152.

Valentini, M., and Filloux, A. (2019). Multiple Roles of c-di-GMP Signaling in Bacterial Pathogenesis. Annu Rev Microbiol 73, 387–406.

Valgepea, K., Adamberg, K., Nahku, R., Lahtvee, P.-J., Arike, L., and Vilu, R. (2010). Systems biology approach reveals that overflow metabolism of acetate in Escherichia coli is triggered by carbon catabolite repression of acetyl-CoA synthetase. Bmc Syst Biol 4, 166.

Vartak, N.B., Reizer, J., Reizer, A., Gripp, J.T., Groisman, E.A., Wu, L.-F., Tomich, J.M., and Saier, M.H. (1991). Sequence and evolution of the FruR protein of Salmonella typhimurium: a pleiotropic transcriptional regulatory protein possessing both activator and repressor functions which is homologous to the periplasmic ribose-binding protein. Res Microbiol 142, 951–963.

Vasconcelos, F.N. da C., Maciel, N.K., Favaro, D.C., Oliveira, L.C. de, Barbosa, A.S., Salinas, R.K., Souza, R.F. de, Farah, C.S., and Guzzo, C.R. (2017). Structural and Enzymatic Characterization of a cAMP-Dependent Diguanylate Cyclase from Pathogenic Leptospira Species. J Mol Biol 429, 2337–2352.

Veening, J., Hamoen, L.W., and Kuipers, O.P. (2005). Phosphatases modulate the bistable sporulation gene expression pattern in Bacillus subtilis. Mol Microbiol 56, 1481–1494.

Veening, J., Igoshin, O.A., Eijlander, R.T., Nijland, R., Hamoen, L.W., and Kuipers, O.P. (2008a). Transient heterogeneity in extracellular protease production by Bacillus subtilis. Mol Syst Biol 4, 184.

Veening, J.-W., Smits, W.K., and Kuipers, O.P. (2008b). Bistability, Epigenetics, and Bet-Hedging in Bacteria. Annu Rev Microbiol 62, 193–210.

Vermeersch, L., Perez-Samper, G., Cerulus, B., Jariani, A., Gallone, B., Voordeckers, K., Steensels, J., and Verstrepen, K.J. (2019). On the duration of the microbial lag phase. Curr Genet 65, 721–727.

Vidakovic, L., Singh, P.K., Hartmann, R., Nadell, C.D., and Drescher, K. (2017). Dynamic biofilm architecture confers individual and collective mechanisms of viral protection. Nature Microbiology 3, 26-31.

Volkmer, B., and Heinemann, M. (2011). Condition-Dependent Cell Volume and Concentration of Escherichia coli to Facilitate Data Conversion for Systems Biology Modeling. Plos One 6, e23126.

Wada, T., Hatamoto, Y., and Kutsukake, K. (2012). Functional and expressional analyses of the anti-FlhD4C2 factor gene ydiV in Escherichia coli. Microbiology+ 158, 1533–1542.

Wada, T., Morizane, T., Abo, T., Tominaga, A., Inoue-Tanaka, K., and Kutsukake, K. (2011). EAL Domain Protein YdiV Acts as an Anti-FlhD4C2 Factor Responsible for Nutritional Control of the Flagellar Regulon in Salmonella enterica Serovar Typhimurium. J Bacteriol 193, 1600–1611.

Wang, S., Fleming, R.T., Westbrook, E.M., Matsumura, P., and McKay, D.B. (2006). Structure of the Escherichia coli FlhDC Complex, a Prokaryotic Heteromeric Regulator of Transcription. J Mol Biol 355, 798–808.

Wang, X., and Wood, T.K. (2011). IS5 inserts upstream of the master motility operon flhDC in a quasi-Lamarckian way. Isme J 5, 1517–1525.

Wang, X., Dubey, A.K., Suzuki, K., Baker, C.S., Babitzke, P., and Romeo, T. (2005). CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of Escherichia coli. Mol Microbiol 56, 1648–1663.

Wang, X., Koirala, S., Aldridge, P.D., and Rao, C.V. (2020). Two Tandem Mechanisms Control Bimodal Expression of the Flagellar Genes in Salmonella enterica. J Bacteriol 202.

Wang, Y.-C., Chin, K.-H., Tu, Z.-L., He, J., Jones, C.J., Sanchez, D.Z., Yildiz, F.H., Galperin, M.Y., and Chou, S.-H. (2016). Nucleotide binding by the widespread high-affinity cyclic di-GMP receptor MshEN domain. Nat Commun 7, 12481.

Weber, H., Pesavento, C., Possling, A., Tischendorf, G., and Hengge, R. (2006). Cyclic-di-GMP-mediated signalling within the σS network of Escherichia coli. Mol Microbiol 62, 1014–1034.

Weber, H., Polen, T., Heuveling, J., Wendisch, V.F., and Hengge, R. (2005). Genome-Wide Analysis of the General Stress Response Network in Escherichia coli: σS-Dependent Genes, Promoters, and Sigma Factor Selectivity. J Bacteriol 187, 1591–1603.

Wei, B., Shin, S., LaPorte, D., Wolfe, A.J., and Romeo, T. (2000). Global Regulatory Mutations in csrA and rpoS Cause Severe Central Carbon Stress in Escherichia coli in the Presence of Acetate. J Bacteriol 182, 1632–1640.

Wei, B.L., Brun-Zinkernagel, A., Simecka, J.W., Prüß, B.M., Babitzke, P., and Romeo, T. (2001). Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli. Mol Microbiol 40, 245–256.

Wei, L.-N., Zhu, L.-W., and Tang, Y.-J. (2016). Succinate production positively correlates with the affinity of the global transcription factor Cra for its effector FBP in Escherichia coli. Biotechnol Biofuels 9, 264.

Weickert, M.J., and Adhya, S. (1992). A family of bacterial regulators homologous to Gal and Lac repressors. J Biol Chem 267, 15869–15874.

Wells, T.J., Sherlock, O., Rivas, L., Mahajan, A., Beatson, S.A., Torpdahl, M., Webb, R.I., Allsopp, L.P., Gobius, K.S., Gally, D.L., et al. (2008). EhaA is a novel autotransporter protein of enterohemorrhagic Escherichia coli O157:H7 that contributes to adhesion and biofilm formation. Environ Microbiol 10, 589–604.

Westermayer, S.A., Fritz, G., Gutiérrez, J., Megerle, J.A., Weißl, M.P.S., Schnetz, K., Gerland, U., and Rädler, J.O. (2016). Single-cell characterization of metabolic switching in the sugar phosphotransferase system of Escherichia coli. Mol Microbiol 100, 472–485.

Whiteley, A.T., Eaglesham, J.B., Mann, C.C. de O., Morehouse, B.R., Lowey, B., Nieminen, E.A., Danilchanka, O., King, D.S., Lee, A.S.Y., Mekalanos, J.J., et al. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature 567, 194–199.

Whitney, J.C., and Howell, P.L. (2013). Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. Trends Microbiol 21, 63–72.

Whitney, J.C., Colvin, K.M., Marmont, L.S., Robinson, H., Parsek, M.R., and Howell, P.L. (2012). Structure of the Cytoplasmic Region of PelD, a Degenerate Diguanylate Cyclase Receptor That Regulates Exopolysaccharide Production in Pseudomonas aeruginosa. J Biol Chem 287, 23582–23593.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Science: A Publication of the Protein Society 27, 293–315.

Winkler, A., Udvarhelyi, A., Hartmann, E., Reinstein, J., Menzel, A., Shoeman, R.L., and Schlichting, I. (2014).

Characterization of elements involved in allosteric light regulation of phosphodiesterase activity by comparison of different functional BlrP1 states. Journal of Molecular Biology 426, 853–868.

Witte, G., Hartung, S., Büttner, K., and Hopfner, K.-P. (2008). Structural Biochemistry of a Bacterial Checkpoint Protein Reveals Diadenylate Cyclase Activity Regulated by DNA Recombination Intermediates. Mol Cell 30, 167–178.

Wolfe, A.J. (2005). The Acetate Switch. Microbiol Mol Biol R 69, 12–50.

Wu, J., Sun, L., Chen, X., Du, F., Shi, H., Chen, C., and Chen, Z.J. (2013). Cyclic GMP-AMP Is an Endogenous Second Messenger in Innate Immune Signaling by Cytosolic DNA. Science 339, 826–830.

Xiao, Y., Liu, H., He, M., Nie, L., Nie, H., Chen, W., and Huang, Q. (2020). A crosstalk between c-di-GMP and cAMP in regulating transcription of GcsA, a diguanylate cyclase involved in swimming motility in Pseudomonas putida. Environ Microbiol 22, 142–157.

Xu, Z., Zhang, H., Zhang, X., Jiang, H., Liu, C., Wu, F., Qian, L., Hao, B., Czajkowsky, D.M., Guo, S., et al. (2019). Interplay between the bacterial protein deacetylase CobB and the second messenger c-di-GMP. Embo J 38, e100948.

Yakhnin, A.V., Baker, C.S., Vakulskas, C.A., Yakhnin, H., Berezin, I., Romeo, T., and Babitzke, P. (2013). CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage. Mol Microbiol 87, 851–866.

Yamada, Y., Hoshino, K., and Ishikawa, T. (1997). The Phylogeny of Acetic Acid Bacteria Based on the Partial Sequences of 16S Ribosomal RNA : The Elevation of the Subgenus Gluconoacetobacter to the Generic Level. Biosci Biotechnology Biochem 61, 1244–1251.

Yang, H., Liu, M.Y., and Romeo, T. (1996). Coordinate genetic regulation of glycogen catabolism and biosynthesis in Escherichia coli via the CsrA gene product. J Bacteriol 178, 1012–1017.

Yilmaz, C., Rangarajan, A.A., and Schnetz, K. (2020). The Transcription Regulator and c-di-GMP Phosphodiesterase PdeL Represses Motility in Escherichia coli. J Bacteriol 203.

Yoon, S. hun, and Waters, C.M. (2021). The ever-expanding world of bacterial cyclic oligonucleotide second messengers. Curr Opin Microbiol 60, 96–103.

You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C.W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of bacterial proteome with metabolism by cyclic AMP signalling. Nature 500, 301–306.

Zähringer, F., Lacanna, E., Jenal, U., Schirmer, T., and Boehm, A. (2013). Structure and Signaling Mechanism of a Zinc-Sensory Diguanylate Cyclase. Structure 21, 1149–1157.

Zaver, S.A., and Woodward, J.J. (2020). Cyclic dinucleotides at the forefront of innate immunity. Curr Opin Cell Biol 63, 49–56.

Zhang, Z., Kukita, C., Humayun, M.Z., and Jr, M.H.S. (2017). Environment-directed activation of the Escherichia coliflhDC operon by transposons. Microbiology+ 163, 554–569.

Zhao, K., Liu, M., and Burgess, R.R. (2007). Adaptation in bacterial flagellar and motility systems: from regulon members to ‘foraging’-like behavior in E. coli. Nucleic Acids Res 35, 4441–4452.

Zhu, M., Pan, Y., and Dai, X. (2019). (p)ppGpp: the magic governor of bacterial growth economy. Curr Genet 65, 1121–1125.