Elucidating the regulatory function of LysR-type transcription regulators in *Escherichia coli*

Doctoral thesis

for

the award of the doctoral degree

of the Faculty of Mathematics and Natural Sciences

of the University of Cologne

by

Ria Basu Roy

Accepted in the year 2025

This research work has been done in the group of Prof. Dr. Karin Schnetz, Institute for Genetics, Department of Biology, Faculty of Mathematics and Natural Sciences, University of Cologne.

Table of content

Abstract 3
Main Introduction 4
Stress responses in *E. coli* 4
LysR-type transcription regulators 7
H-NS and LTTRs as prone H-NS antagonist 10
YdcI 12
Objective of the thesis 13
Results Chapter 1: 14
Characterisation of LysR-type transcription regulators (LTTRs) in *E. coli* 14
1.1 Expression analyses of LTTRs using 3xFLAG tagged alleles 17
1.2 Autoregulation of LTTRs using promoter lacZ reporters 19
1.3 YafC, YeiE, YhaJ and YnfL bind to their own promoter fragments 23
1.4 Transcriptome analysis (RNA-seq) of ΔlrhA, ΔydcI, ΔyafC and ΔyhaJ strains 24
Conclusion 36
Materials and Methods 38
Bacterial strains, plasmids and media 38
Protein purification 38
Electrophoretic mobility shift assays (EMSA) 39
β-galactosidase expression analysis 39
Western blotting 39
RNA-seq and data analysis 40
Data availability 40
Tables 41
Supplementary Information 45
Results Chapter 2: 56
YdcI, a LysR-type transcription regulator related to stress response and catabolism 56
Abstract 57
Introduction 58
Results 61
Identification of YdcI target loci 61
In vitro DNA-binding of presumptive target loci by YdcI and identification of an YdcI DNA-binding motif 65
YdcI DNA-binding motif 69
Role of YdcI in the phosphate stress control of *iraP* 70
Regulation of *ydcI* transcription 74
Growth phase dependent YdcI synthesis 76

Regulation of YdcI targets 80
Discussion 82
Material and Methods 86
Bacterial strains and growth conditions 86
Plasmid construction 86
β-galactosidase expression analysis 87
Growth curve analysis 87
Western blotting 87
Electrophoretic mobility shift assays for DNA-binding analyses 87
$YdcI_{His6}$ protein purification 88
RNA-seq and data analysis 88
Data availability 89
Identification of YdcI binding motif 89
Supplementary Information 90
Tables 93
Main Discussion 101
Approaches to characterize LTTRs 101
The regulatory and functional role of YdcI 103
Alternative approaches to characterise the LTTRs 105
References 106

Abstract

LysR-type transcriptional regulators (LTTRs) constitute the most abundant group of transcription regulators in bacteria. In *E. coli* K-12, out of the 46 LTTRs, 29 are still partially characterised or completely uncharacterised with respect to their regulatory and functional role. For the majority of these poorly and uncharacterised LTTRs, information regarding the control of the LTTR's expression and activity, and their target genes remains limited. In this work, I characterized a subset of these partially characterized or uncharacterized LTTRs with a focus on systematic characterization of the LTTR YdcI. Out of the 17 LTTRs selected for this study, 8 LTTRs were shown to be autoregulators and for 4 of them binding to their own promoter regions was shown. A comparative RNA-seq analysis revealed the presumptive target genes of 4 LTTRs (LrhA, YafC, YdcI and YhaJ). Additionally, validation of published ChIP-exo data of YdcI by EMSA identified 10 loci including *iraP* (responsible for stabilization of stress sigma factor), genes related to transport across the membrane and the TCA cycle as specific targets of this LTTR. A consensus DNA-binding motif for YdcI has also been identified in this work. Another finding of this study elucidates the role of alarmone ppGpp in induction of *ydcI* gene expression and the role of YdcI possibly acting as an H-NS antagonist for its target, *iraP*. Hence, this work highights the potential role of YdcI in *E. coli* in regulating the carbon flux in the TCA cycle in response to stress. Taken together, the findings of this work are of relevance to understand and decipher the complex regulatory network and potential functional implications of the partially characterized LTTRs in *E. coli*.

Main Introduction

A complex gene regulatory network involving many different transcription regulators exist in bacteria to enable quick responses to environmental stresses like changes in temperature, pH or availability of nutrients. Bacteria like *E. coli* respond to the various stress signals by eliciting the expression of certain stress responsive genes. This adaption to various type of stresses is tightly controlled from the level of transcription to protein synthesis. The LysR-type transcription regulators (LTTRs) constitute the largest family of transcription regulators in bacteria and are involved in various type of stress responses. H-NS, which is a global gene repressor in *Enterobacteriaceae* with a preference for AT-rich DNA, plays an important role in repressing horizontally acquired genes including some pathogenicity islands and virulence factors as well as stress response genes. Repression by H-NS can be relieved gene-specifically by transcription regulators including LTTRs which are prone H-NS antagonists due to their AT-rich DNA-binding motifs. However, of the 46 LTTRs in *E. coli* K-12 many are still uncharacterized. Hence different approaches were undertaken in this work to characterize the functional and regulatory role of some of the partially characterized or completely uncharacterized LTTRs in *E. coli* with emphasis on the LTTR YdcI. In the following sections, current knowledge of stress responses of *E. coli* with relevance to this study, features of LTTRs, repression by H-NS and de-repression, and the LTTR YdcI are introduced.

Stress responses in *E. coli*

Bacteria encounter various types of changing environmental conditions, depriving nutrient availability and stresses. Hence, bacteria like *E. coli* have evolved different mechanisms to survive and adapt to the various types of changing environment and stresses encountered. Many of these adaptive responses are specific depending upon the induced stress, like chaperons and proteases are induced in response to heat shock and misfolded proteins, DNA damage induces the SOS response invoking the proteins and polymerases which deals with the DNA damage (Gottesman, 2019). Also, bacteria can cope with the existing stress at the level of transcriptional regulation, by activation of some alternate sigma factors which bind to RNA polymerase, thereby expressing a subset of stress responsive genes. This type of stress response is called the general stress response (Battesti et al., 2011, Gottesman, 2019).

In *E. coli*, σS (RpoS) is the sigma factor for general stress responses which drives the transcription of genes required to respond to various types of stresses like osmotic, heat shock, pH and starvation of various nutrients in the stationery phase (Cho et al., 2014, Peano et al., 2015, Schellhorn, 2020, Wong et al., 2017). The regulation of σS in itself is complex and goes from the level of transcription to translation and RpoS proteolysis to maintain low cellular levels of this sigma factor when the cells are

not under stress. The degradation of σS by the protease ClpXP is facilitated by adaptor protein RssB (Figure 1). The ATP dependent protease ClpXP (Gottesman, 2019) is composed of ClpP (proteolytic core) and ClpX (ATPase subunit). Substrates like σ^S require special adaptors, like RssB, in order to mediate their recognition by ClpXP. RssB activity is controlled by several small proteins identified as anti-adaptors which compete with RpoS for interacting with RssB thereby preventing ClpXP protease from degrading RpoS (Figure 1). In *E. coli* three such anti-adaptor proteins have been identified, which stabilize RpoS. IraP, is one such anti-adaptor molecule, which is specifically active when the cells are exposed to phosphate starvation conditions. The close proximity of *iraP* to *phoA*, an important gene of phosphate regulon is an indicator for the crucial role of this anti-adaptor molecule in response to phosphate stresses in bacteria. IraP has been shown to inhibit the proteolysis of RpoS through a direct interaction with RssB and this process is independent of phosphorylation of RssB (Bougdour and Gottesman, 2007). In response to phosphate starvation, levels of the alarmone ppGpp increase, which activates *iraP* expression. As a consequence, σS is stabilized (Bougdour and Gottesman, 2007). Other anti-adaptors are IraD, which responds to DNA damage or UV stress, and IraM, which stabilizes RpoS when the cells are under magnesium starvation and low pH stress conditions (Bougdour et al., 2008, Bougdour et al., 2006, Bouillet et al., 2024, Eguchi et al., 2011, Xu et al., 2019).

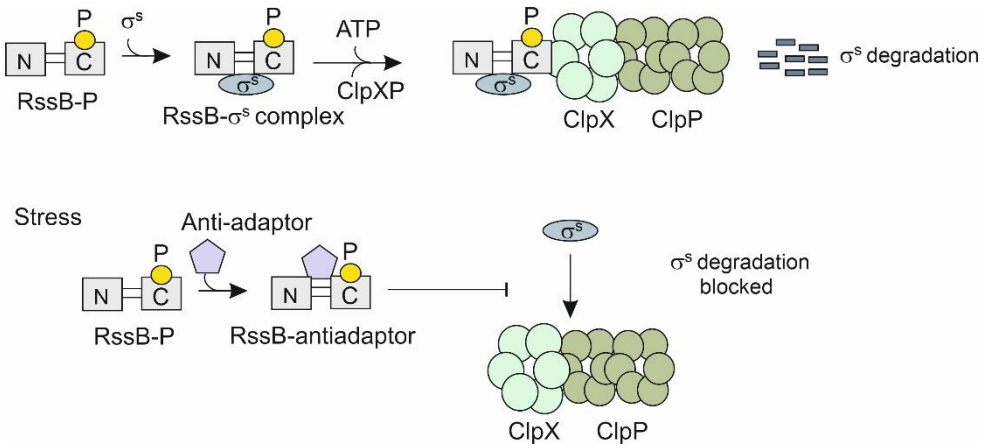


Figure 1: RssB dependent degradation of σ^s by ClpXP. RssB binds to σ^S, this complex then binds to ClpX, subjecting σ^S to proteolytic degradation. In contrast for the cells under stress conditions, anti-adaptors like IraP bind to RssB blocking access and degradation of σ^S by ClpXP. Figure adapted from Brugger et al. (2023).

Another signaling pathway which helps bacteria to survive and adapt to nutrient and abiotic stresses is the stringent response, which is mediated by the second messengers guanosine 5'-diphosphate 3'-diphosphate (ppGpp) and guanosine 5'-triphosphate 3'-diphosphate (pppGpp), collectively referred to as alarmone (p)ppGpp (Steinchen et al., 2020). The alarmone (p)ppGpp reprograms bacterial physiology to adapt to sudden nutrient downshift and is involved in regulation of bacterial virulence,

persistence and antibiotic resistance (Geiger et al., 2012, Huttener et al., 2018, Patacq et al., 2020, Verstraeten et al., 2015, Wu et al., 2015). The intracellular levels of ppGpp are in turn regulated by proteins of the RelA/SpoT homology (RSH) protein family. RSH proteins typically consist of a hydrolase and a synthetase domain comprising the N-terminal part and additional regulatory domain in the C-terminal part (Bange et al., 2021). The bifunctional enzyme, RelA has an active synthesis (SYNTH) domain (the synthesis of ppGpp takes place) but inactive (HD) hydrolysis (responsible for ppGpp hydrolysis) domain (Aravind and Koonin, 1998, Atkinson et al., 2011, Heinemeyer and Richter, 1978). The other bifunctional RSH protein, SpoT, has a weak ppGpp synthetic activity in its SYNTH domain and a strong ppGpp degrading activity in the HD domain (Atkinson et al., 2011, Xiao et al., 1991). The (p)ppGpp synthetases transfer the 5'-βγ pyrophosphate group of ATP onto the 3’ hydroxyl group of the acceptor substrate of GTP or GDP generating ppGpp or pppGpp respectively, while (p)ppGpp hydrolases removes the pyrophosphate and regenerate GDP and GTP (Cashel and Kalbacher, 1970, Fernandez-Coll and Cashel, 2020, Steinchen et al., 2020, Sy and Lipmann, 1973). The stringent response mediated by RelA operates in response to amino acid starvation in association with the ribosome. In absence of amino acids, the uncharged tRNA bound in the ribosomal A-site is sensed by RelA and this stalled ribosome then induces the ppGpp production by a direct interaction of RelA with ribosome (Fernandez-Coll and Cashel, 2020). The stringent response by SpoT is elicited in response to fatty acid starvation, iron, phosphate and carbon limitation conditions (Battesti and Bouveret, 2006, Bougdour and Gottesman, 2007, Murray and Bremer, 1996, Spira et al., 1995, Vinella et al., 2005). At this stress conditions interaction of SpoT with YtfK (and possibly other proteins) shifts the SpoT activity from hydrolase to synthetase activity (Germain et al., 2019, Meyer et al., 2021). SpoT is essential in *relA*$^{+}$ strains, since only SpoT can hydrolyse ppGpp, while *relA spoT* double mutants are viable.

The alarmone ppGpp, produced after starvation, alters gene expression by binding directly to RNA polymerase at two sites. The ppGpp binding “site 1” lies at the interface of the ω and the β’ subunit and “site 2” is at the interface of the RNA Polymerase binding small protein DksA and the β’ subunit (Sanchez-Vazquez et al., 2019). Transcription of rRNA encoding genes is downregulated upon ppGpp binding to RNA polymerase, while transcription of genes required for amino acid synthesis is up-regulated for ppGpp-bound RNA polymerase, due to specific differences in the promoter discriminator’ sequences located between the -10 and the transcription start site. Apart from regulating transcription, ppGpp can also regulate different other cellular processes like DNA replication, nucleotide metabolism, translation, ribosome biogenesis (Bennison et al., 2019, Corrigan et al., 2016, Liu et al., 2015, Srivatsan et al., 2008, Steinchen et al., 2020). Similar to directly interacting with RNA polymerase, ppGpp can bind to DNA primase, GTP biosynthetic enzymes and ribosome

assembly factors (Diez et al., 2020). A recent study has shown that depending upon the state the cell is in, ppGpp can act as an alarmone or as a second messenger (Figure 2). During starvation or stress, the cells tend to go into survival mode then (p)ppGpp acts as an alarmone while as the cellular levels of (p)ppGpp is lower in absence of starvation like during exponential phase, ppGpp adjusts the external growth conditions with bacterial growth and function primarily via regulating cellular mechanisms like amino acid and ribosome biosynthesis (Figure 2) therefore acting as a second messenger (Fernandez-Coll and Cashel, 2020, Zhu et al., 2024). Additionally, during the transition from exponential to stationery growth, the elevated ppGpp production is associated with mechanisms like recruitment of alternate sigma factors and prevention of σS proteolysis (Zhu et al., 2024).

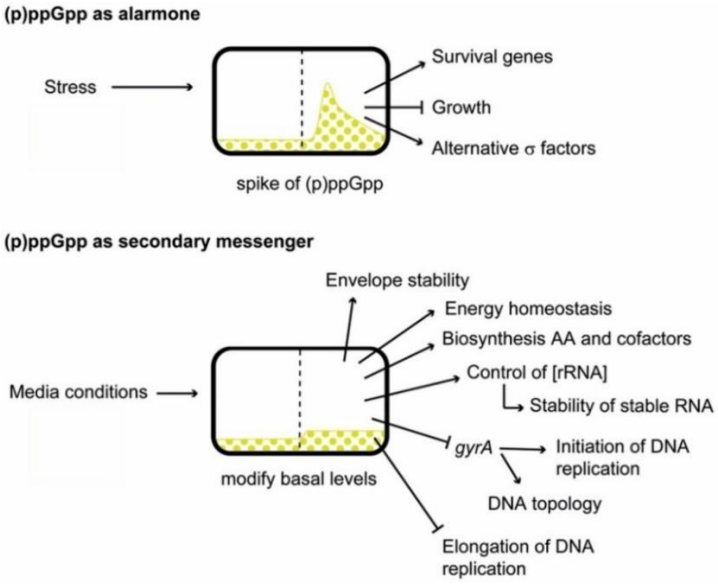


Figure 2: Function of (p)ppGpp. ppGpp acts as an alarmone during starvation conditions or as a second messenger to modulate the changes in basal level (adapted from Fernandez-Coll and Cashel, 2020).

LysR-type transcription regulators

LysR-type transcription regulators represents the largest family of transcription regulators in bacteria and in *E. coli* K-12, 46 of approximately 200 transcription regulators are LTTRs (Karp et al., 2014). LTTRs are conserved, ubiquitous among bacteria and different orthologues have been identified in eukaryotic organisms and Archae families as well (Maddocks and Oyston, 2008). LTTRs are involved in a broad spectrum of important processes including regulation of virulence, biofilm formation, stress responses, utilization of various carbon sources and metabolites, horizontal gene transfer and others (Henikoff et al., 1988, Karp et al., 2014, Keseler et al., 2017). LTTRs are homotetramers with each monomer consisting of a smaller N-terminal winged helix-turn-helix DNA-binding domain (wHTH-DBD) that connects to the larger C-terminal regulatory effector binding domain (EBD) via a helical linker helix

and a flexible hinge (Maddocks and Oyston, 2008, Momany and Neidle, 2012). The tetramer is a dimer of dimers as the wHTH-DBD domains dimerize as well as the EBDs, i.e. the tetramer contains two wHTH-DBD dimers as well as two EBD dimers, and can bind to two adjacent sites on the DNA. The wHTH-DBD binds to the specific DNA target with the 3rd helix known as the recognition helix inserting itself in the major groove of DNA and forming the DNA-protein interface (Aravind et al., 2005). LTTRs can act as activators or repressors and an effector which can be a coactivator or a co-repressor can bind to the inducer binding cavity (IBC) within the EBD between the EBD's regulatory subdomain 1 (RD1) and regulatory subdomain 2 (RD2). These effectors can be small ions, amino acids or metabolites. In addition, modification of an amino acid can also affect the LTTRs activity, as in case of the oxidative stress regulator OxyR (Christman et al., 1989). The two effector binding subdomains namely RD-1 and RD-2 are connected with antiparallel cross-over β-strands. These β-strands form the dimerization interface in an antiparallel head-to-tail orientation (Figure 3).

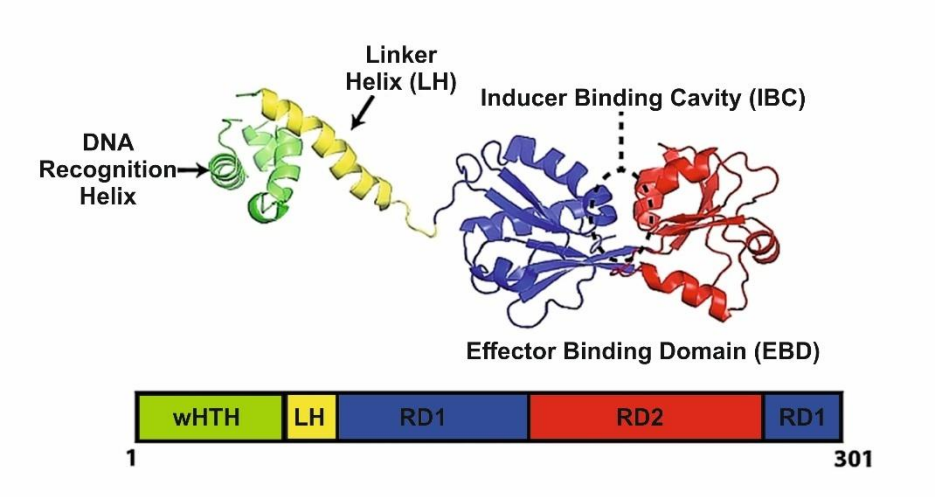


Figure 3: Structure of LTTRs. The monomers of LTTR consisting of an N-terminal winged HTH DNA-binding (wHTH) domain and a C-terminal effector binding domain (EBD), which are connected by a flexible linker helix (LH). A possible co-effector can bind in the inducer binding cavity (IBC), which is located between regulatory domain 1 (RD1) and regulatory domain 2 (RD2). Figure adapted from Lerche et al. (2016).

A general DNA-binding motif for LTTRs has the consensus sequence T-N11-A (Maddocks and Oyston, 2008, Parsek et al., 1994). Co-crystallization of the wHTH-DBD with the specific DNA for LTTRs like BenM (Alanazi et al., 2013) and CbnR (Koentjoro et al., 2018) has helped to clarify the DNA-binding mode. Each wHTH-DBD dimer binds a palindromic DNA that consists of two half sites, and thus tetrameric LTTR binds two sites, namely the regulatory binding site (RBS) and the activator binding site (ABS) (Baugh et al., 2023). LTTRs, as other transcription regulators, can act as both activators and repressors by interacting with RNA polymerase or preventing access of RNA polymerase to the

promoter depending on the positioning of the binding sites relative to the promoter (Busby, 2019). A peculiarity of LTTRs is that they always bind specifically, independent of an effector, but that they 'slide' on the DNA depending on effector binding. As an example, BenM, in the absence of the effector, bind to the *benA* promoter and represses transcription of *benA*, whereas in presence of an effector, one of the two dimeric wHTH-DBDs is repositioned and this causes the tetramer to shift and binding to the activator binding site takes place thereby promoting interaction with RNA polymerase and activating the transcription of *benA* (Baugh et al., 2023). This mode of regulation of different target genes by sliding along the sequence is called 'sliding dimer mechanism' and has been studied for DntR, BenM, ArgP and OxyR (Bundy et al., 2002, Kullik et al., 1995, Laishram and Gowrishankar, 2007, Lerche et al., 2016, Toledano et al., 1994). Unlike most of the other families of transcription regulators, the variable DNA-binding mode and structural property of the LTTRs has limited the techniques which can be used to characterize them. However, techniques which utilizes the property of direct DNA-binding of these regulators like Chromatin Immunoprecipitation (ChIP) coupled to high-throughput sequencing (ChIP-seq) in combination with exonuclease treatment (ChIP-exo) has led to successful identification of DNA-binding sites of some of the uncharacterized LTTRs (Gao et al., 2018).

For some LTTRs, amino acid exchanges have been identified which presumably mimick an effector-bound activated form of these LTTRs. This includes OxyR-C199D in *Pseudomonas aeruginosa* and LeuO-S120D in *E. coli* (Fragel et al., 2019, Jo et al., 2015). Additionally, the exchange of amino acid residue changes the angle between two DBDs as elucidated for OxyR (Jo et al., 2015) thereby regulating the expression of relevant antioxidant genes in *Pseudomonas aeruginosa*. Many effectors have been identified for some of the LTTRs in *E. coli* and additionally some of these LTTRs can act as autoregulators or regulate the expression of their divergent genes (Table S1, Chapter 1). For the LTTRs of MetR, CysB, CatR, GcvA, and OxyR interactions are identified between the C-terminal domains of the α subunit of RNA polymerase (Baugh et al., 2023). Some LTTRs are shown to have pleiotropic effect and can regulate many genes. For example, LeuO is one of the LTTRs which regulates a number of functions from pathogenicity, CRISPR/Cas immunity and regulating multi-drug efflux system in *E. coli* and *S. enterica* (Espinosa and Casadesus, 2014, Fragel et al., 2019, Medina-Aparicio et al., 2011, Shimada et al., 2009). However, many of the LTTRs in *E. coli* are still uncharacterized with respect to their regulatory and functional role and also in context of different conditions which can regulate their own expression.

H-NS and LTTRs as prone H-NS antagonist

The nucleoid-associated protein of H-NS is a pleiotropic regulator present in *Enterobacteriaceae* which plays a dual role in genome organization and global gene repression (Dorman, 2004, 2014, Luijsterburg et al., 2006). The mechanism of the transcriptional repression by H-NS involves binding of H-NS at specific AT-rich nucleation sites, which are dispersed throughout the genome, followed by polymerizing along the DNA into nearby low affinity sites, thereby forming nucleoprotein complexes (Bouffartigues et al., 2007, Kahramanoglou et al., 2011, Rimsky et al., 2001). The mechanism of repression by H-NS takes place either by forming linear extended nucleoprotein complexes along the DNA (filaments) or by building DNA-H-NS-DNA bridges (Dame et al., 2005, Maurer et al., 2009). H-NS functions as a xenogenic silencer by repressing approximately 5% of all genes in the *E. coli* chromosome, many of these include horizontally acquired genes including pathogenicity islands and several virulence factors (Dorman, 2014, Hommais et al., 2001, Lucchini et al., 2006, Navarre et al., 2007).

Chromatin Immunoprecipitation-on-chip (ChIP-on-chip) studies have revealed 350 loci for H-NS binding which are spread across the *E. coli* genome (Grainger et al., 2006, Kahramanoglou et al., 2011, Lucchini et al., 2006, Uyar et al., 2009). H-NS exerts its effect by inhibiting transcription either by occluding or trapping the RNA polymerase or by modulating RNA polymerase activity at the promoter regions of the repressed genes (Grainger et al., 2006). Transcription factors can relieve repression by H-NS by disrupting the oligomerization state of H-NS or competing with H-NS for the binding site, this can occur by displacing H-NS or induced topological changes in the DNA which relieves this repression (Navarre et al., 2007, Stoebel et al., 2008, Will et al., 2015, Winardhi et al., 2015). This mechanism of relieving the repression by H-NS using DNA-binding activators, which binds near the promoters by displacing or rearranging H-NS is called counter-silencing (Figure 4). Moreover, the bridged H-NS complexes can also inhibit transcription at the level of elongation by promoting pausing and termination of RNA polymerase (Figure 4). Many species of *Enterobacteriaceae* contain *StpA* which is a paralogue of H-NS sharing 59% sequence identity on amino acid level with H-NS (Zhang and Belfort, 1992). StpA also binds AT-rich DNA sequences via bridged and linear filaments thereby silencing the expression of horizontally acquired genes (Lim et al., 2012, Muller et al., 2006, Sonnenfield et al., 2001, Srinivasan et al., 2013). Deletion of only StpA does not have any phenotype in *E. coli*. However, altered genome wide transcription and crippled cell growth is observed in strains having both *hns* and *stpA* deletion (Hustmyer et al., 2022). In the exponential phase, *stpA*, being repressed by H-NS, directs expression of low protein levels of StpA. Further, StpA is degraded by Lon protease (Zhang *et al.*, 1996,

Johansson *et al*., 2001). In *E. coli*, a subset of genes are bound and regulated by both H-NS and StpA (Uyar et al., 2009).

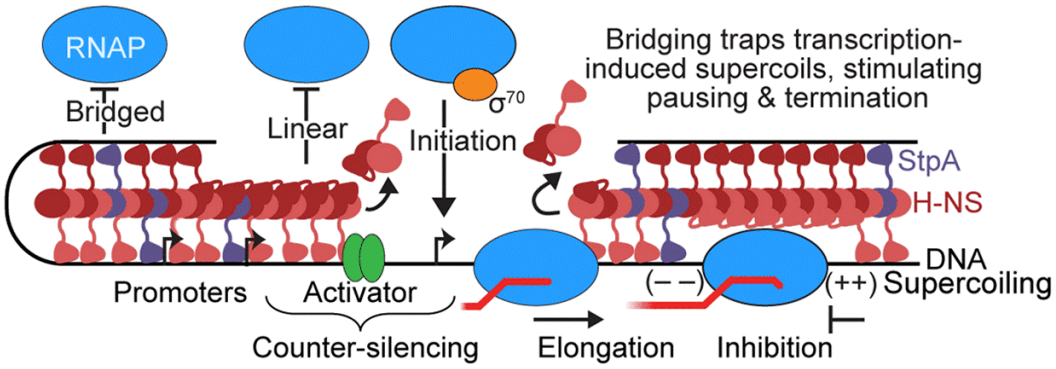


Figure 4: H-NS exerts an effect on RNA polymerase (RNAP) initiation and elongation. Both H-NS (dark red and light red monomers) and StpA (purple) prevent RNAP–σ^{70} (blue–orange) binding by forming either bridged or linear filaments. The DNA binding activators (in green) also known as counter silencers can displace H-NS and initiate transcription by RNAP. The bridged H-NS filaments can also inhibit elongation of transcription by RNA polymerase pausing and termination. Figure adapted from Hustmyer et al. (2022).

The LTTRs have a preferential binding specificity for AT-rich DNA and hence can act as the counter-silencers, which relieve the repression by H-NS. Like, LeuO is one of the LTTRs which activates expression of the H-NS repressed *yjjQ-bglJ* operon in *E. coli* (Stratmann et al., 2008). Further LeuO activates the genes coding for CRISPR/Cas immunity system in *E. coli* and *S. enterica*, these genes are otherwise repressed by H-NS (Medina-Aparicio et al., 2011, Pul et al., 2010, Westra et al., 2010). The mechanism of relieving the repression of H-NS by LeuO at the *cas* operon is still unknown. From previous studies, LeuO has been shown to act as an H-NS antagonist by presumably competing with H-NS for DNA-binding by preventing the formation of oligomeric H-NS complexes (Chen and Wu, 2005, De la Cruz et al., 2007, Westra et al., 2010). Additionally, environmental factors like osmolarity can also cause derepression by H-NS by modulating the DNA-binding by H-NS (Qin et al., 2020). Some pathogenic genes are shown to be repressed by H-NS at low temperature and this repression is relieved at higher temperature (Ono et al., 2005, Trachman and Yasmin, 2004, Yang et al., 2005). LTTRs are presumably prone to compete with H-NS for the AT-rich binding site and to relieve repression by H-NS thereby acting as H-NS antagonist.

YdcI

The *ydcI* gene encodes for a DNA-binding protein belonging to the LTTR family. YdcI is a highly conserved gene which is found in a wide range of gram-negative bacteria and is associated with a number of phenotypes in *Salmonella enterica* serovar Typhimurium (Jennings et al., 2011, Solomon et al., 2014). YdcI from *E. coli* K-12 MG1655 shares 80% of its identity with YdcI from *S. enterica* (Gao et al., 2018). Hence, it was thought to regulate similar cellular functions in both the related species. YdcI has been found to be autorepressed, involved in biofilm formation and acid stress resistance in *S.* Typhimurium (Jennings et al., 2011, Romiyo and Wilson, 2020). The ectopic expression of YdcI in species of gram-negative bacteria like *E. coli* and *Enterobacter cloacae* makes the cells more susceptible to antibiotic stress (Solomon et al., 2014). YdcI is needed for thermal stress resistance in *E. coli*, but not in *S.* Typhimurium (Solomon et al., 2014). Thus, YdcI confers completely different phenotype in *Salmonella* and *E. coli* (Table 1, Chapter 2). Surprisingly, RNA-seq analysis of *ydcI* deletion and wild-type strain revealed none of the gene hits could be related to YdcI's previously identified function in *E. coli* and *S.* Typhimurium (Romiyo and Wilson, 2020). Additionally, the expression level of YdcI protein in *S.* Typhimurium has been found to be five times higher than in *E. coli* suggesting that *E. coli* in contrast to other species, does not tolerate increase of cellular YdcI above a certain level (Solomon et al., 2014). Recently, Ribo-seq analysis has shown that translation of YdcI is enhanced in *E. coli* in response to acute acid stress (Schumacher et al., 2023). Intriguingly, in repeated long-term starvation a combination of *rho-R109H* (transcription termination factor) and loss of function *ydcI* mutants in *E. coli* have been shown to lead to intracellular alkalization (Worthan et al., 2024). A genome wide binding profile of YdcI in *E. coli* K-12 MG1655 using ChIP-exo has revealed 16 binding sites for YdcI (Gao et al., 2018). The functions of the different targets of YdcI are either linked to regulation of metabolic pathways like the TCA cycle, transporting metabolites across the cell or to different stress responses in *E. coli* (Table 2, Chapter 2). Since the link between the target genes and physiological role of YdcI in *E. coli* is not very clear, this work aims to address the discrepency existing in literature regarding the regulatory and functional role of YdcI.

Objective of the thesis

The aim of the thesis revolves around systematic characterization of LysR-type transcription regulators in *E. coli*. Since many LTTRs are still uncharacterized and not much information is available in the literature regarding the genes regulated by them or the mechanism of regulation of the target loci by these LTTRs, this makes it important to study and elucidate the regulatory and functional role of the LTTRs in more detail.

In order to characterize the LTTRs, at the first step, the approach was to test the expression of uncharacterized LTTRs. This was followed by trying to identify the downstream targets of the LTTRs by RNA-seq analysis to get an idea about the regulatory role of the LTTRs. In parallel, the autoregulatory function of the LTTRs was studied using promoter *lacZ* reporters expression analysis. Additionally, *in vitro* DNA binding analyses was done to detect the binding of the LTTRs to their own promoters (Results Chapter 1).

Another aspect of this study involves detailed characterization of the LTTR YdcI. From the previously available ChIP-exo data in the literature and the set of specific differentially expressed genes identified in the comparative RNA-seq data, the targets for YdcI were identified and validated by *in vitro* DNA binding analyses. In addition, the expression analysis for different targets of YdcI and the growth conditions required for expression of YdcI were studied. Additionally, regulation of transcription of *ydcI* and YdcI synthesis were addressed, transcription and regulation of *iraP* by YdcI was also characterized under different growth conditions (Results Chapter 2).

Results Chapter 1:

Characterisation of LysR-type transcription regulators (LTTRs) in *E. coli*

E. coli K-12 encodes 46 LysR-type transcriptional regulators (LTTRs) as listed in supplementary table S1. For 17 of these LTTRs, target genes and signals that modulate their activity have been described, for 10 of these LTTRs some aspects of their function are known, but regulation of the LTTR's activity by possible effectors are unidentified, and the remaining 19 LTTRs are of unknown function (Table S1). In order to gain more insights about the regulatory and functional role of the LTTRs, genome wide techniques like microarray transcriptome analysis, genomic SELEX (*in vitro* technique to enrich DNA fragments bound by a transcription regulator) (Ishihama et al., 2016) and chromatin immunoprecipitation coupled with microarray analysis (ChIP-chip) have been used. More recently, chromatin immunoprecipitation coupled with high-throughput sequencing (ChIP-seq) and a combination of the ChIP-seq protocol coupled with exonuclease treatment (ChIP-exo) have been used to identify DNA-binding sites of transcription regulators including some of the LTTRs (as for example Gao et al., 2018, 2021) (Table S1). Further, the combination of ChIP-exo and RNA-seq has been used for some transcription regulators including some LTTRs to further decipher the target genes of these transcription regulators and their transcription regulatory network in *E. coli* (Gao et al., 2018). However, experimental validation for the targets of poorly and uncharacterized LTTRs still remains elusive. Moreover, several LTTRs have been shown to act as autoregulators and in some cases the LTTRs regulate genes that map close to the *lttr* gene or their divergent genes (Table S1). Thus, analyzing the gene context of a specific *lttr* gene may be helpful for elucidating their biological function. From all previous studies 19 of the 46 LTTRs (~40%) in *E. coli* K-12 still remain poorly characterized (Table S1), hence a subset of the partially characterized or uncharacterized LTTRs has been selected to study their regulatory function in this work, as specified in Table 1.

In this study a combinational approach has been used to characterize the regulatory role of LTTRs (Figure 1). From the subset of 17 partially characterized or uncharacterized LTTRs (Table 1), 4 LTTRs (LrhA, YafC, YdcI and YhaJ) were found to be synthesized at significant levels (LTTR-3xFLAG expression analysis and transcriptome data) and hence were chosen for RNA-seq analysis. Comparative analysis of RNA-seq data revealed a high number of non-LTTR specific changes in the transcriptome levels. Further comparison with published ChIP-exo data for 2 LTTRs (YdcI and YafC) revealed consistency for one LTTR (YdcI) but not for the other one (YafC). In addition, 8 LTTRs (out of total 16 LTTRs analyzed for autoregulation) were found to autoregulate their expression (using promoter *lacZ* reporter fusions), out of which for 4 LTTRs, the binding to their promoter region was validated by EMSA (Table

1). For the LTTR YdcI, presumable target genes have been identified with ChIP-exo, and further data indicate that YdcI is relevant in the response to acid and alkaline stress and the regulation of *gltA* encoding the enzyme citrate synthase of the TCA cycle (Gao et al., 2018). Nonetheless, the regulatory role for this LTTR still remains an open question which is addressed in Chapter 2. Taken together, the reverse genetic approach for the chracterization of LTTRs in this chapter revealed how challenging it is to characterize transcription regulatory factors of unknown function.

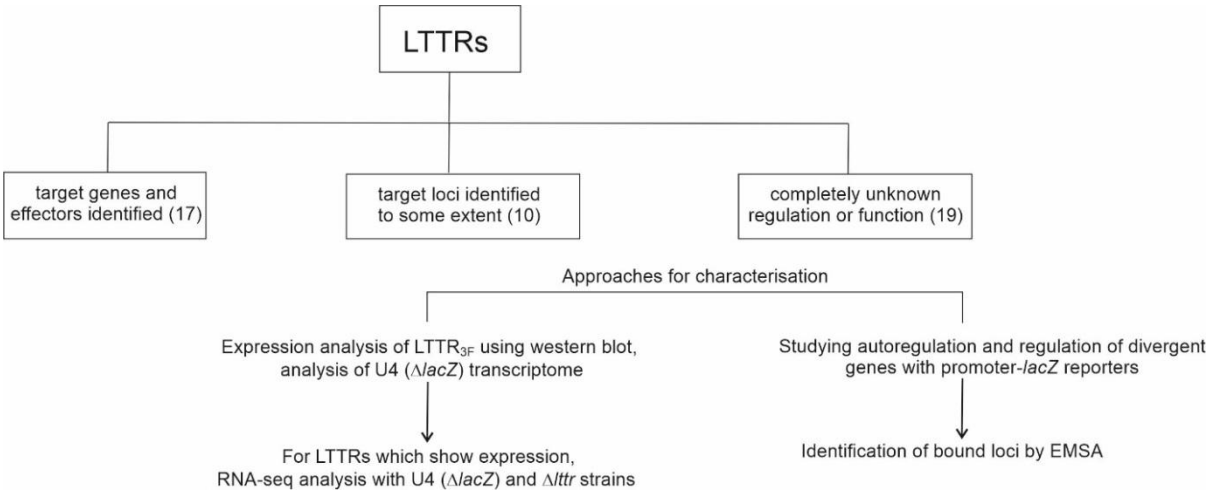


Figure 1: Approaches to characterize LTTRs. In order to characterize the LTTRs of unknown regulatory or functional role, two parallel approaches were taken. The LTTR-3xFLAG ($LTTR_{3F}$) expression analysis was followed by RNA-seq analysis of LTTRs with significant expression. The autoregulatory role and DNA-binding to the LTTR's promoter fragment with EMSA was studied.

Table 1: Subset of poorly and uncharacterized LTTRs of *E. coli* K-12

LTTRs	Expression with Western Blot[1]	Expression with Transcriptome data[2]	DNA binding using EMSA[3]	Regulatory role studied using promoter *lacZ* fusions[4]	RNA-seq analysis
LrhA	high	high	N.A	N.A	yes
YafC	low	low	yes	Autorepressor	yes
YdcI	low	high	no	no	yes
YeiE	no	intermediate	yes	Autorepressor	N.A
YhaJ	low	high	yes	Autorepressor	yes
YnfL	no	very low	yes[5]	autorepressor, upregulates divergent *ynfM*	N.A
YahB	no	low	N.A	no	N.A
YbdO/CitR	no	very low[6]	N.A	no	N.A
YbeF	no	very low[6]	N.A	no	N.A
YbhD	no	very low[6]	N.A	autoactivator	N.A
YcaN	no	very low	N.A	no	N.A
YeeY	low	very low	N.A	no	N.A
YfeR	no	very low[6]	N.A	autorepressor	N.A
YfiE	no	low	N.A	autorepressor	N.A
YgfI/SrsR	no	very low	N.A	no	N.A
YhjC/RcdB	no	low[6]	N.A	autorepressor	N.A
YiaU	N.A	very low[6]	N.A	no	N.A

[1]Western Blot using $LTTR_{3F}$ (LTTR-3xFLAG) and strains are grown in tryptone upto stationery phase, see figure 2.
[2]Based on data visualized in IGV (Integrative Genomics Viewer) using RNA-seq data from *E. coli* K-12 strain U4 (Δ*lacZ*) grown in tryptone upto exponential phase, see figure 8.
[3]EMSA performed using promoter fragments of LTTRs, see figure 5.
[4]Autoregulation was analyzed using plasmidic P_{lttr} *lacZ* reporters, see figure 3 and 4.
[5]YnfL-DBD_{His6} could be purified and used for EMSA.
[6] H-NS bound, presumably silenced
N.A (Not available)

1.1 Expression analyses of LTTRs using 3xFLAG tagged alleles

To analyze the expression of LTTRs, genomic alleles were constructed that encode C-terminal 3xFLAG (3F) tagged LTTR proteins (Table 3). Then the expression of these *lttr-3xFLAG* alleles were analyzed by Western blotting for *lrhA*$_{3F}$, *ydcI*$_{3F}$, *yafC*$_{3F}$, *yeiE*$_{3F}$, *yhaJ*$_{3F}$ and *ynfL*$_{3F}$ in dependence of the growth phase. The samples for Western blot were harvested after growing the strains in tryptone medium upto exponential (2h, 3h and 4h) and early stationary (6h and 8h) growth phase as well as over-night (o/n). The results revealed that there is high expression of LrhA$_{3F}$ independent of the growth phase (Figure 2A). The expression levels of YafC$_{3F}$ and YhaJ$_{3F}$ are lower than of LrhA$_{3F}$ and likewise not dependent on the growth phase (Figure 2A). The expression of YdcI$_{3F}$ is absent in the over-night cultures but is detected during the exponential growth phase (from two hours onwards). In contrast the expression of YeiE$_{3F}$ and YnfL$_{3F}$ could not be detected in any of the growth phases (Figure 2A). Additionally, a second set of LTTRs were chosen and their expression was compared with *lrhA*$_{3F}$ which was used as reference in the subsequent Western blot analysis (since its expression was high, Figure 2A). For this second subset of uncharacterized LTTRs including YahB, YbdO, YbeF, YbhD, YcaN, YeeY, YfeR, YfiE, YgfI, YhjC and YiaU, a similar approach was taken to analyze their expression using 3xFLAG tagged chromosomal alleles. The samples for Western blot were harvested after growing the *lttr*$_{3F}$ strains in tryptone medium to exponential (3h) and early stationary (7h) growth phase as well as over-night (o/n). The results show that only a low level of protein could be detected for YeeY$_{3F}$ (o/n and 7h), but no protein could be detected for all other LTTRs (Figure 2D).

From the results of the LTTR-3xFLAG expression analysis, constant cellular levels were detected for the LTTRs LrhA, YafC and YhaJ independent of the growth phase. The expression of YdcI is dependent on the growth phase and detected from the early to late exponential growth phase (Figure 2A). For the second subset of LTTRs, low expression could only be detected for YeeY, which was dependent on the growth phase and detectable only for the early and late stationary growth phases (7h and o/n) (Figure 2D). Based on these results, YafC, YhaJ and YdcI were considered to be good candidates for further analysis in respect to their biological role by RNA-seq, using LrhA as a reference (described in section 1.4).

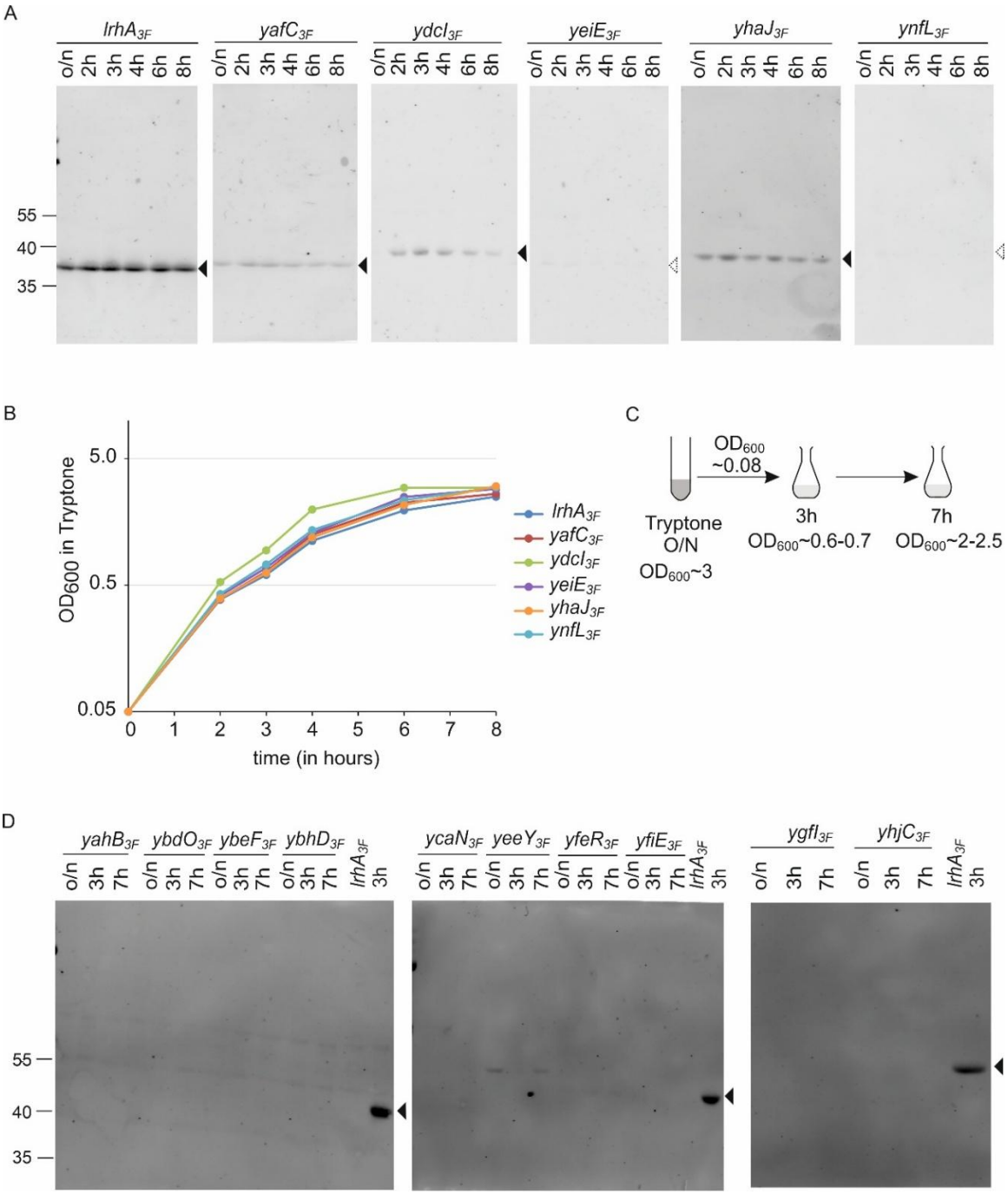


Figure 2: Expression of strains carrying 3xFLAG-tagged (3F) alleles of *lttr* genes. A) *lrhA*$_{3F}$ (U253), *ydcI*$_{3F}$ (U255), *yafC*$_{3F}$ (U254), *yeiE*$_{3F}$ (U256), *yhaJ*$_{3F}$ (U263), and *ynfL*$_{3F}$ (U258) were grown in tryptone medium and aliquots were harvested from over-night cultures (o/n) and from the exponential (2h, 3h, 4h) and early stationary growth phase (6h, 8h). B) Growth curve of *lrhA*$_{3F}$ (U253), *yafC*$_{3F}$ (U254), *ydcI*$_{3F}$ (U255), *yeiE*$_{3F}$ (U256), *yhaJ*$_{3F}$ (U263), and *ynfL*$_{3F}$ (U258) grown in tryptone upto 8 hours. C) Schematics showing the growth of *lttr*$_{3F}$ strains in tryptone medium upto 7 hours. D) *yahB*$_{3F}$ (U511), *ybdO*$_{3F}$ (U512), *ybeF*$_{3F}$ (U513), *ybhD*$_{3F}$ (U514), *ycaN*$_{3F}$ (U515), *yeeY*$_{3F}$ (U516), *yfeR*$_{3F}$ (U517), *yfiE*$_{3F}$ (U518), *ygfI*$_{3F}$ (U519) and *yhjC*$_{3F}$ (U520) were grown in tryptone medium and aliquots were harvested from over-night cultures (o/n) and from the exponential (3h) and early stationary growth phase (7h). Of each sample the equivalent of OD_{600} 0.08 was separated by SDS-PAGE (12% with TCE). Blots incubated with the primary monoclonal ANTI-FLAG® M2 antibody (1:4000) and secondary Goat Anti-Mouse IgG H&L (Alexa Fluor® 680) antibody (1:10000).

1.2 Autoregulation of LTTRs using promoter lacZ reporters

In a second approach, conducted in parallel to determining the LTTR's protein levels (described in section 1.1), I analyzed whether the same set of LTTRs autoregulate their expression. The rational of this approach was that transcription regulators may autoregulate their expression to balance their cellular level, as in the classical example of phage Lambda repressor cI (Ptashne et al., 1976). Further, if autoregulation was detected, a DNA fragment encompassing the promoter and regulatory region would be available to characterize the *in vitro* DNA-binding of the respective LTTR. In order to address the activity of LTTRs in autoregulation of their own promoters as well the regulation on the promoter of the divergent genes (in case present) plasmidic promoter *lacZ* reporter fusions were constructed (Figures 3 and 4). Then autoregulation was analyzed in Δ*lacZ* strains carrying a deletion of the respective *lttr* gene, while the LTTR was provided using plasmids expressing it under the control of *tac* promoter. Thus, autoregulation could be studied independent of the expression and control of the *lttr* gene itself. The strains were then co-transformed with the respective promoter *lacZ* reporter plasmids and the plasmids expressing the LTTR under the control of the *tac* promoter. The strains were grown in LB medium without induction of the *tac* promoter by IPTG to allow only the basal level synthesis of the LTTR proteins, which is sufficient for studying their regulatory function. The results shown in Figure 3 suggest that YafC represses its own promoter, P_{yafC}, by almost 3-fold (compare expression in presence of YafC, P_{tac} *yafC*, and the control, Fig. 3A). YafC does not regulate the promoter of the divergent gene, *yafD*. Likewise, YhaJ does exhibit autorepression, but does not regulate the expression of the divergent *yhaK* (Figure 3B). YeiE also represses ten-fold its own promoter (P_{yeiE} *lacZ* reporter, Figure 3C), while YeiE does not appear to regulate the expression of the divergent P_{yeiH} promoter (P_{yeiH} *lacZ* reporter, Figure 3C). YnfL likewise autorepresses its own promoter and additionally YnfL upregulates the divergent gene, *ynfM* (Figure 3D). Taken together, the LTTRs YafC, YeiE, YhaJ and YnfL negatively autoregulate their promoters and only in case of YnfL a weak positive regulation of the divergent gene was detected under the tested growth conditions (late exponential growth phase in LB medium).

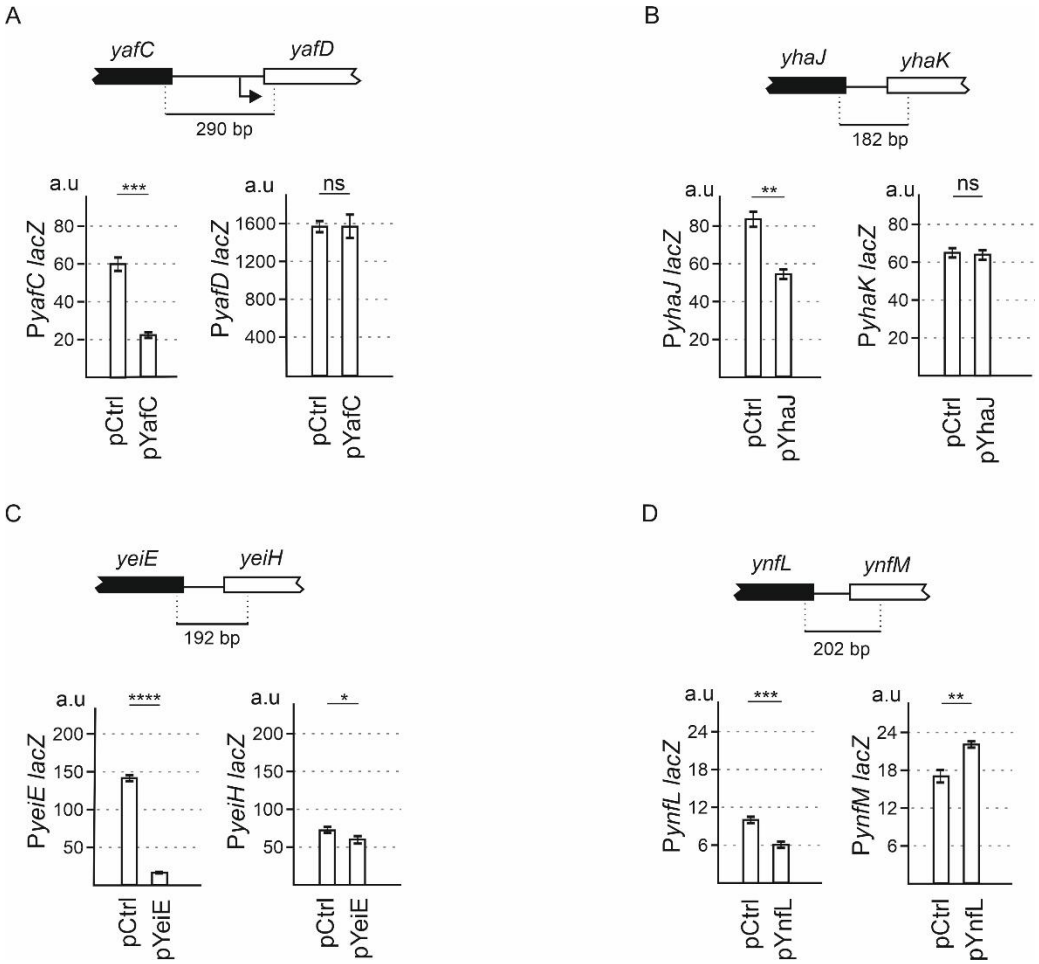


Figure 3: Autoregulation of *yafC*, *yeiE*, *yhaJ* and *ynfL* promoters. Expression levels directed from P_{lttr} *lacZ* promoter fusion were determined in Δ*lttr* strains. Empty vector pKESK22 served as vector control (pCtrl) and the respective LTTRs were provided in trans from the plasmids (pLTTR) expressing the LTTRs under the control of *tac* promoter. Transformants were grown in LB kanamycin and chloramphenicol medium to an OD_{600} of 0.8. Mean values of at least three independent biological replicates are shown and error bars indicate standard deviations. Statistically significant differences are indicated with asterisks (****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05, n.s., not significant). The P_{lttr} *lacZ* fusions include A) P_{yafC} (pKERBR25), P_{yafD} (pKERBR24), B) P_{yhaJ} (pKERBR29), P_{yhaK} (pKERBR28), C) P_{yeiE} (pKERBR27), P_{yeiH} (pKERBR26), and D) P_{ynfL} (pKERBR32), P_{ynfM} (pKERBR31) as listed in Table 4. These reporters were analyzed in combination with pCtrl (pKESK22) as vector control or plasmids providing respective LTTRs as pYafC (pKERBR19), pYeiE (pKERBR20), pYhaJ (pKERBR21), pYnfL (pKERBR23) (Table 4) in the Δ*lacZ* mutant strain background having deletions of *yafC* (U456), *yeiE* (U457), *yhaJ* (U458) and *ynfL* (U460) as listed in Table 3.

In a similar line of experiments, the second set of uncharacterized LTTRs were tested if they could also possibly regulate their own promoters (Figure 4). To address the question, deletion mutants of *yahB, ybdO, ybeF, ybhD, ycaN, yeeY, yfeR, yfiE, ygfI, yhjC* and *yiaU*, respectively were constructed in the Δ*lacZ* strain background. The respective mutant strains were then co-transformed with reporter plasmids having the respective promoters fused to *lacZ* and plasmids expressing YahB, YbdO, YbeF, YbhD, YcaN, YeeY, YfeR, YfiE, YgfI, YhjC and YiaU, respectively, under the control of *tac* promoter. Again, the strains were grown in LB medium without induction of the *tac* promoter by IPTG to allow

only the basal level synthesis of the respective LTTRs which is sufficient to study their regulatory function. The results of β-galactosidase assay show that YbhD activates its own promoter (Figure 4D). YfeR, YfiE and YhjC function as autorepressors of their promoters (Figure 4 G, H, and J). However, none of the other LTTRs exhibits any autoregulatory function. Taken together, these results suggest that YbhD activates its own promoter and that YfeR, YfiE and YhjC show autorepression of their own promoters. However, these LTTRs were not further studied since they are not expressed as shown by Western analyses (section 1.1) and transcriptome data of the wild-type strain (U4) (section 1.4).

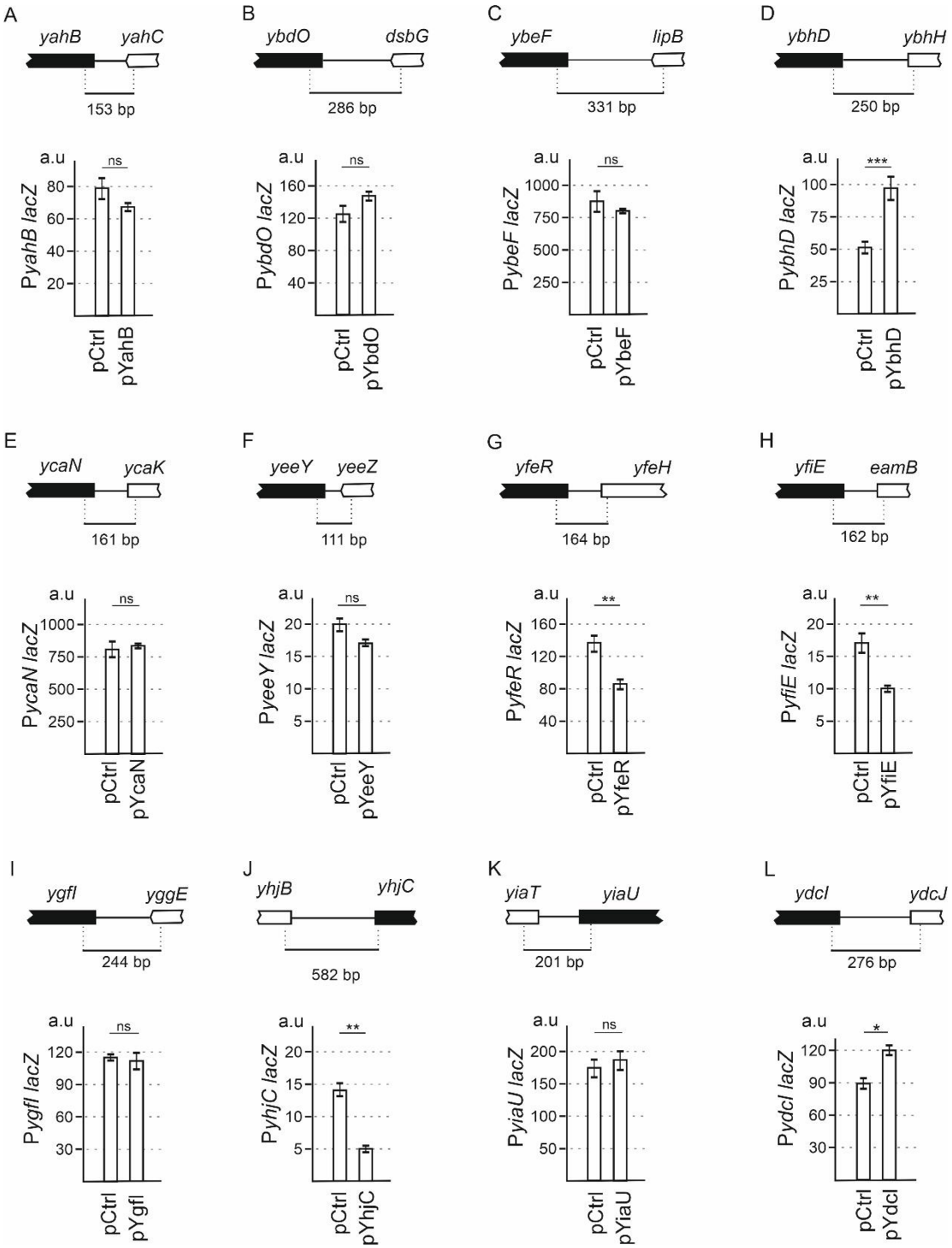


Figure 4: Autoregulation of *lttr* promoters. Expression levels directed from P_{lttr} *lacZ* reporter fusions were determined in the respective Δ*lttr* strains. Empty vector pKESK22 served as vector control (pCtrl) and the respective LTTRs were provided in trans from the plasmids (pLTTR) expressing the LTTRs under the control of *tac* promoter. Transformants were grown in LB kanamycin and chloramphenicol medium to an OD_{600} of 0.8. Mean values of at least three independent biological replicates are shown and error bars indicate standard deviations. Statistically significant differences are indicated with asterisks (****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05, n.s., not significant). The P_{lttr} *lacZ* fusions include A) P_{yahB} (pKERBR80), B) P_{ybdO} (pKERBR75), C) P_{ybeF} (pKERBR81), D) P_{ybhD} (pKERBR82), E) P_{ycaN} (pKERBR83), F) P_{yeeY} (pKERBR78), G) P_{yfeR} (pKERBR79), H) P_{yfiE} (pKERBR76), I) P_{ygfI} (pKERBR84), J) P_{yhjC} (pKERBR77), K) P_{yiaU} (pKERBR85) L) P_{ydcI} (pKERBR93) as listed in Table 4.

These reporters were analyzed in combination with pCtrl (pKESK22) as vector control or plasmids providing respective LTTRs as pYahB (pKERBR69), pYbdO (pKERBR64), pYbeF (pKERBR70), pYbhD (pKERBR71), pYcaN (pKERBR72), pYeeY (pKERBR67), pYfeR (pKERBR68), pYfiE (pKERBR65), pYgfI (pKERBR73), pYhjC (pKERBR66), pYiaU (pKERBR74), pYdcI (pKEHB6) (Table 4) in the Δ*lacZ* mutant strain background having deletions of Δ*yahB* (U533), Δ*ybdO* (U534), Δ*ybeF* (U535), Δ*ybhD* (U541), Δ*ycaN* (U536), Δ*yeeY* (U540), Δ*yfeR* (U537), Δ*yfiE* (U542), Δ*ygfI* (U539), Δ*yhjC* (U543), Δ*yiaU* (U538), Δ*ydcI* (U505) as listed in Table 3.

1.3 YafC, YeiE, YhaJ and YnfL bind to their own promoter fragments

To address the question of whether the LTTRs, which show autoregulation, can also bind to their own promoter DNA fragments, plasmids encoding C-terminally His6-tagged versions of YafC, YeiE, YhaJ and YnfL were constructed and purified by Ni-NTA batch purification. For the LTTRs YbhD, YfeR, YfiE and YhjC, the plasmids encoding C-terminal His6-tagged proteins were cloned, and protein expression was tested. However, the purification for these four LTTRs and their DNA-binding studies were not conducted. The Ni-NTA batch purification was successful for three proteins (YafC, YeiE and YhaJ) except for His6-tagged full length protein of YnfL. For YnfL the His6-tagged DNA-binding domain, YnfL-DBD was purified. The DNA binding domain includes a wHTH motif, the dimerization helix, and the hinge region. The DNA binding domain of LTTRs can dimerise alone and in case of the LTTR LeuO specific DNA-binding of the DBD has been shown (Fragel et al., 2019). Purified $YafC_{His6}$, $YeiE_{His6}$, $YhaJ_{His6}$ and YnfL-DBD_{His6} were then used for *in vitro* DNA-binding assays with EMSA using different set of DNA fragments to include specific and non-specific fragments, as shown in Figure 5. The different sets of DNA fragments were incubated with increasing concentrations of $YafC_{His6}$, $YeiE_{His6}$, $YhaJ_{His6}$ and YnfL-DBD_{His6}, respectively, separated by gel electrophoresis on 8% polyacrylamide gel, and visualized by ethidium bromide staining. $YafC_{His6}$, $YeiE_{His6}$ and $YhaJ_{His6}$ and YnfL-DBD_{His6} bind to the DNA fragments containing their own promoter regions of *yafC* (Figure 5A), *yeiE* (Figure 5B), *yhaJ* (Figure 5C), *ynfL* (Figure 5D) and *ydcI* (Figure 5E), respectively, which is clear from the visible shift of the *yafC* promoter fragment and the disappearance of DNA fragments encompassing promoter regions of *yeiE*, *yhaJ* and *ynfL* (Figure 5). $YdcI_{His6}$ does not bind to the DNA fragment containing *ydcI* promoter region but binds to a fragment encompassing *leuO* promoter region (Figure 5E), which was shown previously (Breddermann and Schnetz, 2017). Moreover, at higher protein concentration of the respective LTTRs, a shift or disappearance of non-specific fragments can be detected, which indicates non-specific binding. Taken together, these results identify four LTTRs that function as autoregulators by binding to their own promoter fragments.

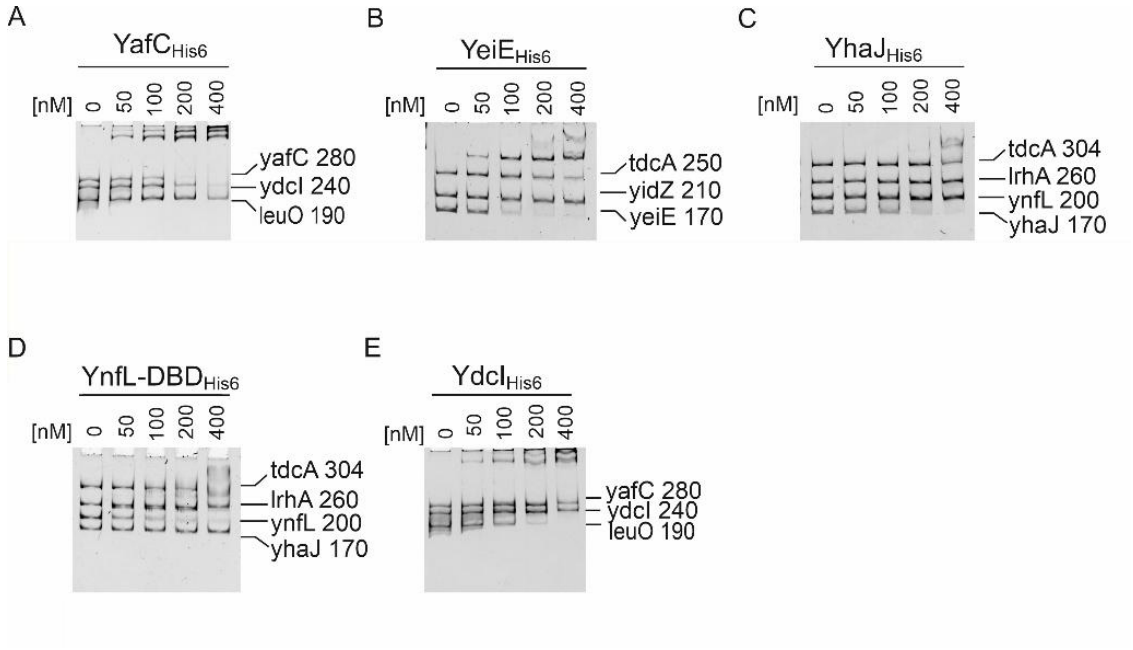


Figure 5: YafC, YeiE, YhaJ and YnfL-DBD bind to their own promoter regions *in vitro*. Electrophoretic mobility shift assays of A) YafC$_{His6}$, B) Yei$_{His6}$, C) YhaJ$_{His6}$, D) YnfL-DBD$_{His6}$ and E) YdcI$_{His6}$ with mixture of DNA fragments of A) *yafC, ydcI, leuO*, B) *tdcA, yidZ, yeiE* C) *tdcA, lrhA, ynfL, yhaJ* D) *tdcA, lrhA, ynfL, yhaJ* and E) *yafC, ydcI, leuO*. The mixture of DNA fragments were incubated with increasing concentrations of purified YafC$_{His6}$, YeiE$_{His6}$, YhaJ$_{His6}$, YnfL-DBD$_{His6}$ and YdcI$_{His6}$. Oligonucleotides used to amplify DNA fragments (length in bps shown next to the fragments) are listed in Table S2. The samples were separated on 8% native polyacrylamide gels, and the gels were stained with ethidium bromide. Gel pictures are shown with inverted colors.

1.4 Transcriptome analysis (RNA-seq) of ΔlrhA, ΔydcI, ΔyafC and ΔyhaJ strains

For LrhA, YdcI, YafC and YhaJ the LTTR-3xFLAG Western analysis showed that these proteins are synthesized under the tested conditions (Figure 2A). Hence, to study the regulatory function of these LTTRs, RNA-seq was performed to identify the potential target genes regulated by them. At first, the *E. coli* K-12 strain U4 (Δ*lacZ*) was used as a control to identify the differentially expressed genes in Δ*lrhA* (U504), Δ*ydcI* (U505), Δ*yafC* (U456) and Δ*yhaJ* (U458) strains under the tested growth conditions (tryptone, OD_{600}~0.6-0.8). The analyzed transcriptome data shows a large number of differentially expressed genes with log2 fold change greater than (+/-) 1.5 for LrhA, YafC, YdcI and YhaJ in comparison to the control strain U4. However, most of these genes either are common or unspecific targets for all four LTTRs (Table 2). Therefore, the data were screened for target genes that were specific to each individual LTTR by performing a comparative RNA-seq data analysis evaluating the data for one LTTR at a time, and comparing the log2 fold change with the other three LTTRs. This method revealed five genes/loci specific for LrhA, five genes/loci specific for YafC, three loci specific for YdcI and one gene, *yqjF*, which was specific for YhaJ (Table 2). The target locus, *dkgB*, which is the adjacent gene to *yafC*, serves as an example which shows that LTTRs can potentially regulate the

expression of genes which maps close to the *lttr* genes (Figure 6). However, this change in expression could also result from the deletion of the *yafC* gene.

The RNA-seq data for YafC and YdcI were further evaluated in respect to ChIP-exo analyses (Gao et al., 2018) that identified five loci for YafC (Table 2 and Figure 7) and sixteen loci for YdcI (Table 2 and further data shown in Chapter 2). The ChIP-exo data were re-analyzed and visualized using Integrative Genomics Viewer (IGV). For YafC, out of the five loci identified by Gao et al. (2018), *yoaC*, *cysJ-queD* and *ygjZ-mdaB* had very high peaks from IGV analysis, while the number of reads for *ybiB and the yjiST* loci was lower (Figure 7). However, for none of these five loci significant changes in the RNA-seq data were detected (marked in blue in Table 2). For YafC, studies show a reduced survival of *yafC* deletion strains on exposure to IR radiation (Byrne et al., 2014) and correlation of YafC expression to increased ethanol production (Shi et al., 2016). However, no correlation has been found between the target genes for YafC identified from comparative RNA-seq analysis in this study and the ChIP-exo targets identified by Gao et al. (2018). The comparative RNA-seq analysis identified the adjacent gene, *dkgB* as one putative target for YafC. Four other specific targets for YafC include the *gladH-lhgD-gabDTP* locus, the *gudDX* locus, and the *glcGFED* locus.

For YdcI, out of the 16 loci identified by ChIP-exo (Gao et al., 2018) 11 have strong peaks in IGV analysis (data shown in Chapter 2). Further, out of the 16 ChIP-exo targets for YdcI, four targets, *iraP*, *yhdV-yhdWXYZ*, *yobA-yebZ-yebY* and *mqo* were identified as specific targets from the comparative RNA-seq anaylsis in this study (Table 2).

For YhaJ, a study using RNA-seq analysis reported a role of YhaJ in locus of enterocyte effacement (LEE) associated pathogenesis (Connolly et al., 2016). In another study, YhaJ has been shown to regulate different subset of genes in enterohemorrhagic *E. coli* (EHEC) and extraintestinal uropathogenic *E. coli* (UPEC) (Connolly et al., 2019). Moreover, a recent study identifies YhaJ as being involved in the activation of *yqjF* (Palevsky et al., 2016), which correlates with the identification of *yqjF* as the target gene of YhaJ in this study.

For LrhA, which was used as a reference, a previous study using microarray analysis has shown that it represses the *fimAICDFGH* operon encoding proteins associated with fimbriae synthesis (Lehnen et al., 2002). This correlates with the identification of the *fimAICDFGH* operon genes from the comparative RNA-seq analysis in this study. Similarly, the identification of a linkage between the GAD system and flagellar synthesis via LrhA and HdeD (Yamanaka et al., 2022) could explain the detection of the transcriptional activator for the acid response system, *gadE*, as a LrhA specific target in this

study. LrhA was also identified as repressor of the *flhDC* locus, which encodes the master regulator of flagella synthesis (Lehnen et al., 2002). However, this was not evident in the RNA-seq data.

In addition, the RNA-seq data from *E. coli* K-12 strain U4 (Δ*lacZ*) was used to obtain an idea about the amounts of transcript of various *lttr*s using IGV genome browser. The data shows that varying amounts of transcript were detected for the partially and uncharacterized LTTRs (Figure 8). High amounts of transcripts were detected for LrhA, YdcI and YhaJ, while transcript levels for other LTTRs were very low to intermediate (Figure 8, Table 1).

Taken together, the approach of comparative RNA-seq analysis was helpful for the identification of few presumptive target genes of YafC, YdcI and YhaJ. However, a discrepancy to ChIP-exo data and functional data is apparent questioning the vailidity of the RNA-seq data set. Nonetheless, the transcriptome analysis of the reference strain U4 gives an idea about the variable levels of transcript under the tested growth conditons for all the LTTRs in this study.

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
nhaA	17489	18655	Na+:H+ antiporter NhaA	-0.38	0.120	-0.32	0.128	-0.19	0.614	-0.35	0.243
carA	29651	30799	carbamoyl phosphate synthetase subunit α	1.72	0.000	2.34	0.000	1.35	0.006	1.09	0.034
rrlH	225759	228662	23S ribosomal RNA	1.00	0.000	1.63	0.000	0.86	0.053	0.72	0.177
dkgB	229167	229970	methylglyoxal reductase DkgB	0.33	0.380	2.44	0.000	0.05	0.937	0.44	0.258
gpt	255977	256435	xanthine-guanine phsophoribosyltransferase	1.43	0.000	1.68	0.000	0.77	0.174	1.45	0.000
codB	354922	356181	cytosine transporter	1.54	0.000	2.00	0.000	1.08	0.030	1.16	0.014
ddlA	399829	400923	D-alanine-D-alanine ligase A	-0.09	0.776	0.02	0.951	0.11	0.777	-0.06	0.924
iraP	401386	401646	anti-adaptor protein IraP	-0.27	0.509	-0.36	0.266	2.58	0.000	-0.06	0.929
phoA	401747	403162	alkaline phosphatase	-0.48	0.013	-0.55	0.001	2.07	0.000	-0.42	0.100
tomB	480334	480708	Hha toxicity modulator TomB	-0.91	0.000	-0.64	0.000	-0.35	0.468	-0.58	0.088
purK	551527	552594	5-(carboxyamino)imidazole ribonucleotide synthase	1.42	0.002	2.12	0.000	0.20	0.791	0.99	0.105
purE	552591	553100	N5-carboxyaminoimidazole ribonucleotide mutase	1.84	0.000	2.33	0.000	0.18	NA	1.77	0.000
borD	578600	578893	DLP12 prophage; prophage lipoprotein BorD	4.92	0.000	1.47	0.000	0.32	0.632	0.91	0.059
dcuC	654583	655968	anaerobic C4-dicarboxylate transporter DcuC	-3.29	0.000	-3.48	0.000	-1.21	0.031	-2.44	0.000
glnU	696865	696939	tRNA-Gln (UUG)	1.10	0.003	1.72	0.000	0.96	0.032	0.68	0.199
leuW	696963	697047	tRNA-Leu (UAG)	1.31	0.001	1.75	0.000	1.19	0.009	0.96	0.065
abrB	746723	747769	putative regulator	-1.62	0.000	-1.20	0.000	-0.99	0.042	-1.23	0.001
gltA	753185	754468	citrate synthase	-0.06	0.914	-0.44	0.028	-0.65	0.052	-0.09	0.859
sdhC	755177	755566	succinate:quinone oxidoreductase, membrane protein SdhC	0.82	0.003	0.44	0.020	0.10	0.827	0.43	0.309
sdhD	755560	755907	succinate:quinone oxidoreductase, membrane protein SdhD	0.58	0.077	0.27	0.363	-0.09	0.842	0.29	0.600
sdhA	755907	757673	succinate:quinone oxidoreductase, FAD binding protein	0.33	0.313	0.13	0.721	-0.21	0.523	0.20	0.703
sdhB	757689	758405	succinate:quinone oxidoreductase, iron-sulfur cluster binding protein	0.19	0.578	0.02	0.956	-0.31	0.293	0.08	0.893
valT	780765	780840	tRNA-Val(Gao et al.)	1.16	0.000	1.74	0.000	1.21	0.000	0.88	0.039
valZ	781068	781143	tRNA-Val(Gao et al.)	1.03	0.000	1.52	0.000	1.26	0.000	0.88	0.025
lysY	781147	781222	tRNA-Lys (UUU)	1.22	0.000	1.70	0.000	1.48	0.000	0.85	0.062
lysZ	781369	781444	tRNA-Lys (UUU)	1.15	0.000	1.72	0.000	1.37	0.000	0.91	0.032
lysQ	781577	781652	tRNA-Lys (UUU)	0.97	0.005	1.70	0.000	1.43	0.000	0.86	0.021
ybhI	801887	803320	putative tricarboxylate transporter	0.54	0.348	0.35	0.599	0.08	0.920	0.56	0.372
bioF	810381	811535	8-amino-7-oxononanoate synthase	-1.63	0.000	-1.55	0.000	-1.08	0.039	-1.46	0.001
bioC	811522	812277	malonyl-acyl carrier protein methyltransferase	-1.61	0.000	-1.59	0.000	-1.06	0.045	-1.44	0.000
bioD	812270	812947	dethiobiotin synthetase	-1.62	0.000	-1.73	0.000	-1.04	0.050	-1.50	0.000
ybiB	835248	836210	nonspecific DNA-binding protein YbiB	-0.36	0.148	-0.57	0.004	-0.44	0.201	-0.21	0.642
bssR	878248	878631	regulator of biofilm formation	-3.14	0.000	-3.35	0.000	-1.20	0.032	-2.50	0.000
hcp	912162	913814	protein S-nitrosylase	-1.41	0.000	-1.86	0.000	-1.16	0.023	-1.10	0.018
dmsA	940959	943403	dimethyl sulfoxide reductase subunit A	-2.27	0.000	-2.92	0.000	-1.19	0.039	-1.78	0.000
dmsB	943414	944031	dimethyl sulfoxide reductase subunit B	-1.56	0.000	-2.16	0.000	-1.13	0.036	-1.03	0.051
ycbJ	971752	972645	putative phosphotransferase YcbJ	-2.26	0.000	-2.49	0.000	-1.55	0.002	-1.76	0.000
wrbA	1067112	1067708	NAD(P)H:quinone oxidoreductase	-0.19	0.693	-0.03	0.948	0.48	0.343	0.53	0.311
ymdF	1068081	1068254	stress-induced bacterial acidophilic repeat motifs-containing protein YmdF	-0.46	0.150	0.53	0.055	0.87	0.003	0.82	0.003

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
ymdG	1079120	1079146	protein YmdG	0.65	0.338	1.91	0.001	1.21	NA	0.99	NA
putP	1079305	1080813	proline:Na+ symporter	0.77	0.227	2.42	0.000	2.13	0.000	0.91	0.102
pgaA	1089866	1092289	partially deacetylated poly-beta-1,6-N-acetyl-D-glucosamine outer membrane porin	-1.55	0.000	-0.95	0.000	-0.66	0.086	-1.09	0.000
trhO	1116807	1117859	tRNA U34 hydroxylase TrhO	1.54	0.000	1.71	0.000	0.85	0.083	0.93	0.036
ychH	1258791	1259069	stress-induced protein	-1.73	0.000	-1.57	0.000	-1.17	0.015	-1.06	0.003
narK	1277957	1279348	nitrate:nitrite antiporter NarK	-1.67	0.002	-1.97	0.001	-0.81	0.159	-1.09	0.067
narG	1279864	1283607	nitrate reductase A subunit α	-1.77	0.001	-2.43	0.000	-0.95	0.100	-1.05	0.080
narH	1283604	1285142	nitrate reductase A subunit beta	-1.18	0.043	-1.70	0.002	-1.13	0.039	-0.84	0.136
tyrV	1287244	1287328	tRNA-Tyr (GUA)	1.30	0.001	1.56	0.000	1.23	0.004	0.87	0.088
ompW	1314020	1314658	outer membrane protein W	-1.51	0.000	-1.96	0.000	-1.45	0.001	-1.20	0.001
pyrF	1341921	1342658	orotidine-5'-phosphate decarboxylase	1.21	0.000	1.53	0.000	0.84	0.068	0.88	0.019
yciH	1342658	1342984	putative translation factor	1.28	0.000	1.57	0.000	0.87	0.084	0.72	0.126
trkG	1423782	1425239	Rac prophage; K+ transporter TrkG	0.21	0.396	0.42	0.033	0.23	0.501	0.20	0.574
fdnG	1547401	1550448	formate dehydrogenase N subunit α	-1.62	0.000	-1.94	0.000	-1.20	0.022	-1.21	0.001
pqqL	1572407	1575202	periplasmic metalloprotease	2.00	0.000	-0.61	0.002	-0.98	0.000	-0.52	0.126
yddB	1575247	1577619	putative TonB-dependent receptor YddB	2.93	0.000	-0.43	0.032	-0.92	0.000	-0.41	0.245
yddA	1577657	1579342	ABC transporter family protein YddA	3.65	0.000	0.59	0.037	-0.18	0.766	0.01	0.989
ydeT	1588853	1590001	fimbrial usher domain-containing protein YdeT	1.54	0.000	1.40	0.000	0.98	0.072	0.84	0.089
ynfO	1636756	1636989	Qin prophage; DUF3950 domain-containing protein YnfO	-0.97	NA	-1.52	0.005	-0.62	NA	-0.89	NA
ydfO	1637047	1637457	Qin prophage; DUF1398 domain-containing protein YdfO	-1.22	0.013	-1.51	0.001	-1.31	0.008	-0.90	NA
ynfE	1658069	1660495	putative selenate reductase YnfE	-1.47	0.000	-1.96	0.000	-1.13	0.034	-1.08	0.034
ynfK	1666524	1667219	putative dethiobiotin synthetase	-1.85	0.000	-1.82	0.000	-1.26	0.007	-1.55	0.000
Nth	1711523	1712158	endonuclease III	-0.23	0.603	-0.11	0.780	-0.12	0.799	-0.51	0.200
dtpA	1712769	1714271	dipeptide/tripeptide:H+ symporter DtpA	-0.17	0.460	0.26	0.340	-0.02	0.955	-0.14	0.843
purR	1737844	1738869	DNA-binding transcriptional repressor PurR	1.00	0.001	1.54	0.000	0.41	0.535	0.88	0.063
ydhY	1753851	1754477	putative 4Fe-4S ferredoxin-like protein YdhY	-1.84	0.000	-2.03	0.000	-1.22	0.021	-1.34	0.004
yoaC	1894133	1894432	DUF1889 domain-containing protein YoaC	0.06	0.899	0.30	0.421	0.79	0.027	0.62	0.088
yobF	1907448	1907591	DUF2527 domain-containing protein YobF	-0.55	0.051	-0.31	0.060	0.08	0.803	-0.53	0.087
yebY	1923365	1923706	DUF2511 domain-containing protein YebY	-0.19	0.383	0.02	0.962	-0.63	0.007	-0.07	0.883
yebZ	1923719	1924591	putative inner membrane protein	-0.20	0.582	-0.04	0.910	-0.84	0.004	-0.14	0.799
yobA	1924595	1924969	CopC domain-containing protein	-0.39	0.159	-0.30	0.286	-1.04	0.000	-0.25	0.556
purT	1930881	1932059	phosphoribosylglycinamide formyltransferase 2	1.02	0.093	1.89	0.001	0.04	0.958	1.28	0.024
flhC	1977266	1977844	DNA-binding transcriptional dual regulator FlhC	0.33	0.098	0.16	0.486	0.08	0.828	-0.08	0.892
flhD	1977847	1978197	DNA-binding transcriptional dual regulator FlhD	0.40	0.067	0.17	0.503	0.10	0.792	-0.03	0.970
isrC	2071317	2071511	CP4-44 prophage; small RNA IsrC	4.47	0.000	4.21	0.000	3.76	0.000	0.22	0.832
Flu	2071539	2074658	CP4-44 prophage; self recognizing antigen 43 (Ag43) autotransporter	4.56	0.000	4.33	0.000	3.83	0.000	0.10	0.929
yeeR	2074779	2076311	CP4-44 prophage; inner membrane protein YeeR	4.31	0.000	4.07	0.000	3.89	0.000	0.11	0.924
yeeS	2076308	2076754	CP4-44 prophage; RadC-like JAB domain-containing protein YeeS	3.60	0.000	3.51	0.000	3.31	0.000	0.08	0.947
yeeT	2076817	2077038	CP4-44 prophage; DUF987 domain-containing protein YeeT	1.62	NA	1.37	0.027	1.23	NA	-0.10	NA

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
yehC	2190926	2191645	putative fimbrial chaperone YehC	-1.28	NA	-1.82	0.001	-0.91	NA	-0.98	NA
yehD	2191680	2192222	putative fimbrial protein YehD	-1.68	0.000	-2.33	0.000	-1.20	0.030	-1.09	0.037
napG	2299565	2300260	ferredoxin-type protein NapG	-1.26	0.001	-1.69	0.000	-1.10	0.012	-0.89	0.047
napA	2300267	2302753	periplasmic nitrate reductase subunit NapA	-1.43	0.000	-2.01	0.000	-1.26	0.006	-1.09	0.013
napD	2302750	2303013	NapA signal peptide-binding chaperone NapD	-1.42	0.000	-1.56	0.000	-1.16	0.022	-1.27	0.001
napF	2303003	2303497	ferredoxin-type protein	-1.77	0.000	-1.87	0.000	-1.32	0.005	-1.66	0.000
mqo	2305108	2306754	malate:quinone oxidoreductase	0.34	0.236	0.55	0.038	-1.24	0.000	0.08	0.913
purF	2428721	2430238	Amidophosphoribosyltransferase	1.30	0.003	2.13	0.000	0.22	0.784	1.40	0.003
cvpA	2430275	2430763	colicin V production protein	1.34	0.001	2.06	0.000	0.33	0.665	1.41	0.002
ypdK	2496921	2496992	uncharacterized membrane protein YpdK	1.33	0.000	1.70	0.000	1.06	0.012	0.97	0.035
valU	2520931	2521006	tRNA-Val(Gao et al.)	1.57	0.000	2.12	0.000	1.10	0.015	1.31	0.000
valX	2521051	2521126	tRNA-Val(Gao et al.)	1.62	0.000	1.90	0.000	1.11	0.007	1.30	0.000
valY	2521173	2521248	tRNA-Val(Gao et al.)	1.45	0.000	2.17	0.000	1.06	0.029	1.18	0.012
ypfM	2590807	2590866	uncharacterized protein YpfM	-1.31	0.000	-1.84	0.000	-1.31	0.002	-1.35	0.001
purC	2596905	2597618	phosphoribosylaminoimidazole-succinocarboxamide synthase	1.14	0.046	1.76	0.001	0.16	0.828	1.55	0.001
purM	2621197	2622234	phosphoribosylformylglycinamide cyclo-ligase	1.19	0.036	2.25	0.000	0.20	NA	1.15	0.049
purN	2622234	2622872	phosphoribosylglycinamide formyltransferase 1	1.35	0.001	1.94	0.000	0.19	0.804	1.37	0.002
guaB	2632604	2634070	inosine 5'-monophosphate dehydrogenase	1.41	0.010	1.75	0.003	0.13	0.840	1.20	0.027
ndk	2644433	2644864	nucleoside diphosphate kinase	1.85	0.000	1.69	0.000	0.56	0.361	0.64	0.306
csiE	2665435	2666715	stationary phase-inducible protein CsiE	-1.85	0.000	-0.11	0.665	0.11	0.789	0.18	0.687
purL	2691656	2695543	phosphoribosylformylglycinamide synthetase	1.22	0.008	1.72	0.001	0.21	0.787	1.34	0.006
grcA	2716066	2716449	stress-induced alternate pyruvate formate-lyase subunit	-1.48	0.000	-1.71	0.000	-1.02	0.040	-1.04	0.025
raiA	2737154	2737495	ribosome-associated inhibitor A	-1.84	0.000	-1.98	0.000	-1.60	0.000	-1.42	0.000
ypjF	2777453	2777782	CP4-57 prophage; toxin of the YpjF-YfjZ toxin-antitoxin system	0.57	0.280	1.34	0.002	1.10	0.024	1.51	0.001
glaH	2788985	2789962	glutarate dioxygenase GlaH	-0.74	0.249	-1.63	0.000	-0.70	0.212	-0.62	0.200
lhgD	2789982	2791250	L-2-hydroxyglutarate dehydrogenase	-1.12	0.012	-1.80	0.000	-0.80	0.136	-0.82	0.037
gabD	2791273	2792721	succinate-semialdehyde dehydrogenase (NADP+) GabD	-1.26	0.001	-2.02	0.000	-0.88	0.105	-0.98	0.012
gabT	2792735	2794015	4-aminobutyrate aminotransferase GabT	-1.20	0.001	-2.00	0.000	-0.82	0.144	-0.92	0.024
gabP	2794253	2795653	4-aminobutanoate:H+ symporter	-1.18	0.003	-2.05	0.000	-0.80	0.159	-0.88	0.054
stpA	2798091	2798495	DNA-binding transcriptional repressor StpA with RNA chaperone activity	2.18	0.000	2.45	0.000	0.92	0.106	1.62	0.000
argV	2818473	2818549	tRNA-Arg (ACG)	1.35	0.000	1.57	0.000	1.01	0.010	1.03	0.000
serV	2818553	2818645	tRNA-Ser (GCU)	1.20	0.000	1.52	0.000	1.14	0.001	1.15	0.000
gutM	2828621	2828980	DNA-binding transcriptional activator GutM	-0.10	0.848	-0.04	0.945	-0.23	0.669	-0.23	0.702
hypA	2850647	2850997	hydrogenase 3 nickel incorporation protein HypA	-1.28	0.000	-1.51	0.000	-1.29	0.003	-1.32	0.001
hypB	2851001	2851873	hydrogenase isoenzymes nickel incorporation protein HypB	-1.37	0.000	-1.82	0.000	-1.35	0.001	-1.44	0.000
cysJ	2890099	2891898	sulfite reductase, flavoprotein subunit	1.03	0.016	1.31	0.000	0.76	0.174	1.25	0.015
queD	2892214	2892579	6-carboxy-5,6,7,8-tetrahydropterin synthase	0.73	0.037	0.90	0.000	0.72	0.064	0.38	0.503
gudD	2918045	2919385	D-glucarate dehydratase	-1.11	0.013	-1.90	0.000	-1.05	0.039	-0.87	0.125
gudX	2919406	2920746	glucarate dehydratase-related protein	-1.48	0.007	-3.03	0.000	-1.08	0.064	-0.93	0.113
gudP	2920748	2922100	galactarate/D-glucarate transporter GudP	-4.44	0.000	-4.80	0.000	-1.23	0.024	-3.39	0.000

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
ygeW	3006262	3007452	putative carbamoyltransferase YgeW	-3.26	0.000	-3.42	0.000	-1.13	0.046	-1.30	0.021
ygeX	3007510	3008706	2,3-diaminopropionate ammonia-lyase	-2.30	0.000	-2.62	0.000	-1.48	0.005	-1.48	0.004
ygeY	3008764	3009975	putative peptidase YgeY	-1.39	0.000	-1.65	0.000	-1.07	0.015	-0.91	0.053
yqeB	3012614	3014239	XdhC-CoxI family protein YqeB	-1.21	0.002	-1.57	0.000	-0.82	0.156	-0.91	0.082
yqeC	3014287	3015057	uncharacterized protein YqeC	-1.37	0.001	-1.79	0.000	-0.96	0.095	-1.19	0.014
mocA	3015160	3015738	molybdenum cofactor cytidylyltransferase	-1.60	0.000	-1.57	0.000	-0.79	0.136	-1.23	0.001
ygfK	3016060	3019158	putative oxidoreductase, Fe-S subunit	-2.96	0.000	-3.29	0.000	-1.19	0.038	-1.16	0.045
ssnA	3019161	3020489	putative aminohydrolase SsnA	-1.59	0.000	-1.98	0.000	-0.96	0.088	-1.07	0.035
uacT	3031367	3032815	urate:H+ symporter	-1.97	0.000	-1.97	0.000	-1.07	0.023	-1.11	0.019
yqfA	3042489	3043148	hemolysin-III family protein YqfA	-1.28	0.000	-1.60	0.000	-1.16	0.010	-1.33	0.002
ansB	3099682	3100728	L-asparaginase 2	-2.54	0.000	-3.33	0.000	-1.30	0.022	-2.17	0.000
glcG	3123827	3124231	putative heme-binding protein GlcG	-0.67	0.024	-1.58	0.000	-0.81	0.084	-0.72	0.064
glcF	3124236	3125459	glycolate dehydrogenase, putative iron-sulfur subunit	-0.95	0.011	-1.78	0.000	-0.81	0.136	-1.11	0.002
glcE	3125470	3126522	glycolate dehydrogenase, putative FAD-binding subunit	-1.48	0.000	-2.43	0.000	-1.23	0.022	-1.62	0.000
glcD	3126522	3128021	glycolate dehydrogenase, putative FAD-linked subunit	-1.55	0.000	-2.43	0.000	-1.22	0.024	-1.67	0.000
ygiZ	3171879	3172211	DUF2645 domain-containing inner membrane protein YgiZ	0.37	NA	0.53	0.484	0.28	NA	0.11	NA
mdaB	3172530	3173111	NADPH:quinone oxidoreductase MdaB	0.49	0.065	0.89	0.000	0.14	0.798	0.52	0.088
ttdR	3205324	3206256	Dan transcriptional activator // DNA-binding transcriptional activator Dan	-2.14	0.000	-2.04	0.000	-1.35	0.006	-1.63	0.000
rpsU	3210781	3210996	30S ribosomal subunit protein S21	1.50	0.000	1.69	0.000	0.94	0.047	1.04	0.004
yqjF	3250562	3250954	DoxX family protein	-0.01	0.985	0.04	0.955	0.21	0.746	-1.83	0.000
tdcB	3265039	3266028	catabolic threonine dehydratase	-1.13	0.053	-1.84	0.000	-1.54	0.000	-1.29	0.008
tdcA	3266127	3267065	DNA-binding transcriptional activator TdcA	-2.05	0.000	-2.11	0.000	-1.50	0.003	-1.87	0.000
garK	3270625	3271770	glycerate 2-kinase 1	-1.17	0.012	-1.85	0.000	-0.99	0.068	-0.69	0.177
garR	3271867	3272751	tartronate semialdehyde reductase	-1.30	0.004	-2.13	0.000	-1.21	0.025	-0.91	0.096
garL	3272787	3273557	alpha-dehydro-beta-deoxy-D-glucarate aldolase	-1.65	0.003	-3.69	0.000	-1.20	0.032	-1.24	0.031
garP	3273573	3274907	galactarate/D-glucarate transporter GarP	-4.12	0.000	-5.22	0.000	-1.16	0.029	-2.21	0.000
garD	3275282	3276853	galactarate dehydratase GarD	-1.70	0.000	-2.01	0.000	-1.21	0.014	-1.32	0.001
ubiU	3301485	3302480	ubiquinone biosynthesis protein UbiU	-1.98	0.000	-2.07	0.000	-1.53	0.004	-1.76	0.000
ubiV	3302489	3303367	ubiquinone biosynthesis protein UbiV	-1.74	0.000	-2.25	0.000	-1.30	0.012	-1.26	0.006
yhcC	3353121	3354050	radical SAM family oxidoreductase YhcC	-2.00	0.000	-2.04	0.000	-1.28	0.014	-1.64	0.000
yhdV	3418390	3418611	lipoprotein YhdV	-0.63	0.044	-0.13	0.757	0.01	0.981	-0.25	0.634
yhdX	3420134	3421315	putative ABC transporter membrane subunit YhdX	-0.33	0.336	0.25	0.504	0.90	0.024	0.10	0.898
yhdY	3421325	3422428	putative ABC transporter membrane subunit YhdY	-0.30	0.487	0.16	0.721	0.88	0.032	0.55	0.217
yhdZ	3422436	3423194	putative ABC transporter membrane subunit YhdZ	-0.19	0.615	-0.23	0.482	0.13	0.784	0.22	0.672
nirB	3494011	3496554	nitrite reductase catalytic subunit NirB	-2.57	0.000	-3.12	0.000	-1.10	0.053	-1.35	0.016
nirD	3496551	3496877	nitrite reductase subunit NirD	-1.66	0.001	-2.37	0.000	-1.31	0.019	-1.17	0.036
nikA	3613667	3615241	Ni (2+) ABC transporter periplasmic binding protein	-1.56	0.004	-2.14	0.000	-1.10	0.058	-1.85	0.000
nikB	3615241	3616185	Ni (2+) ABC transporter membrane subunit NikB	-2.14	0.000	-2.78	0.000	-1.57	0.003	-1.73	0.000
nikC	3616182	3617015	Ni (2+) ABC transporter membrane subunit NikC	-2.16	0.000	-2.82	0.000	-1.44	0.006	-1.75	0.000
nikD	3617015	3617779	Ni (2+) ABC transporter ATP binding subunit NikD	-2.09	0.000	-2.61	0.000	-1.51	0.003	-1.54	0.001
nikE	3617776	3618582	Ni (2+) ABC transporter ATP binding subunit NikE	-1.80	0.000	-2.31	0.000	-1.31	0.014	-1.34	0.006

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
gadE	3658366	3658893	DNA-binding transcriptional activator GadE	-1.58	0.000	-0.08	0.879	0.02	0.979	-0.09	0.917
xylA	3729443	3730765	xylose isomerase	1.65	0.000	1.54	0.000	0.56	0.380	1.56	0.000
ysaA	3741109	3741582	putative electron transport protein YsaA	-1.72	0.000	-1.90	0.000	-1.25	0.016	-1.53	0.000
lldP	3777399	3779054	lactate/glycolate:H+ symporter LldP	-0.67	0.144	-0.56	0.071	-0.84	0.010	-0.40	0.296
lldR	3779054	3779830	DNA-binding transcriptional dual regulator LldR	-0.87	0.019	-0.69	0.009	-0.90	0.024	-0.62	0.014
lldD	3779827	3781017	L-lactate dehydrogenase	-1.03	0.001	-1.06	0.000	-1.07	0.010	-0.72	0.043
pyrE	3815127	3815768	orotate phosphoribosyltransferase	1.66	0.000	2.04	0.000	0.68	0.248	1.17	0.007
yicG	3818874	3819491	PF03458 family inner membrane protein YicG	0.64	0.024	0.66	0.000	-0.06	0.908	0.65	0.016
xanP	3828945	3830336	xanthine:H+ symporter XanP	0.90	0.134	2.12	0.000	0.15	NA	1.13	0.044
ilvN	3850802	3851092	acetohydroxy acid synthase I subunit IlvN	-1.24	0.001	-1.62	0.000	-0.94	0.099	-1.13	0.004
asnA	3927155	3928147	asparagine synthetase A	-1.30	0.000	-1.65	0.000	-1.40	0.000	-1.09	0.003
leuT	3982606	3982692	tRNA-Leu (CAG)	1.23	0.003	1.72	0.000	1.24	0.002	0.73	0.182
alaT	4037260	4037335	L-alanyl-tRNAalaT // tRNA-Ala (UGC)	1.14	NA	1.54	0.011	1.03	NA	0.22	NA
ptsA	4139720	4142221	putative PTS multiphosphoryl transfer protein PtsA	-1.15	0.000	-1.61	0.000	-1.38	0.000	-1.17	0.003
tyrU	4175472	4175556	tRNA-Tyr (GUA)	1.54	0.000	1.72	0.000	0.87	0.059	1.15	0.000
glyT	4175673	4175747	tRNA-Gly (UCC)	1.51	0.000	1.70	0.000	0.95	0.025	1.07	0.001
thrT	4175754	4175829	tRNA-Thr (GGU)	1.62	0.000	1.76	0.000	0.81	0.102	1.10	0.004
purH	4205943	4207532	bifunctional AICAR transformylase/IMP cyclohydrolase	0.89	0.147	1.64	0.003	0.08	NA	0.92	0.144
ghxP	4278479	4279828	guanine/hypoxanthine transporter GhxP	1.61	0.000	2.48	0.000	0.52	0.439	1.36	0.007
nrfA	4287764	4289200	cytochrome c552 nitrite reductase	-1.96	0.000	-2.45	0.000	-1.53	0.002	-1.53	0.000
nrfB	4289245	4289811	periplasmic nitrite reductase penta-heme c-type cytochrome	-1.61	0.000	-2.27	0.000	-1.39	0.007	-1.32	0.010
nrfC	4289808	4290479	putative menaquinol-cytochrome c reductase 4Fe-4S subunit	-1.33	0.007	-2.18	0.000	-1.25	0.020	-0.96	0.087
nrfD	4290476	4291432	putative menaquinol-cytochrome c reductase subunit NrfD	-1.17	0.003	-1.57	0.000	-0.84	0.124	-0.79	0.109
gltP	4294481	4295794	glutamate/aspartate: H+ symporter GltP	-0.78	0.000	-0.97	0.000	-1.02	0.002	-0.72	0.024
nrdD	4460522	4462660	anaerobic ribonucleoside-triphosphate reductase	-1.13	0.009	-1.57	0.000	-1.00	0.051	-0.73	0.206
pyrI	4470986	4471447	aspartate carbamoyltransferase, PyrI subunit	1.33	0.018	1.65	0.005	0.61	0.349	1.55	0.001
pyrB	4471460	4472395	aspartate carbamoyltransferase catalytic subunit	1.39	0.014	1.68	0.004	0.59	0.365	1.15	0.050
fimA	4543115	4543663	type 1 fimbriae major subunit	-2.04	0.000	-0.51	0.191	-0.52	0.317	-0.31	0.612
fimI	4543728	4544267	putative fimbrial protein FimI	-1.89	0.000	-0.16	0.599	-0.12	0.804	-0.10	0.883
fimC	4544304	4545029	type 1 fimbriae periplasmic chaperone	-1.76	0.000	-0.18	0.600	-0.14	0.750	-0.11	0.883
fimD	4545096	4547732	type I fimbriae usher protein	-1.79	0.000	-0.02	0.968	-0.06	0.918	-0.03	0.974
fimF	4547742	4548272	type 1 fimbriae minor subunit FimF	-1.92	0.000	-0.28	0.480	-0.37	0.472	-0.28	0.595
fimG	4548285	4548788	type 1 fimbriae minor subunit FimG	-1.72	0.000	-0.13	0.791	-0.14	0.811	0.07	0.933
yjiS	4571751	4571915	DUF1127 domain-containing protein YjiS	0.13	NA	0.28	NA	0.00	NA	0.23	NA
yjiT	4572414	4573931	putative uncharacterized protein YjiT	-0.51	0.001	-0.37	0.013	-0.32	0.149	-0.17	0.566
leuV	4606079	4606165	tRNA-Leu (CAG)	1.15	0.008	1.59	0.000	1.09	0.010	0.82	0.078
leuP	4606200	4606286	tRNA-Leu (CAG)	1.58	0.000	1.58	0.000	1.23	0.001	0.82	0.100
leuQ	4606315	4606401	tRNA-Leu (CAG)	1.19	0.001	1.59	0.000	1.10	0.010	1.04	0.018
yjjW	4614680	4615543	putative glycyl-radical enzyme activating enzyme YjjW	-1.01	0.015	-1.51	0.000	-0.91	0.060	-0.88	0.074

Table 2: RNA-seq data of LTTRs

Gene	pos 1	pos 2	Function	WT/ΔlrhA log2 (FC)	P-adj	WT/ΔyafC log2 (FC)	P-adj	WT/ΔydcI log2 (FC)	P-adj	WT/ΔyhaJ log2 (FC)	P-adj
yjjI	4615515	4617065	DUF3029 domain-containing protein YjjI	-2.11	0.000	-2.63	0.000	-1.74	0.000	-1.71	0.000

The log2 fold change (FC) values marked in green are specific for the respective LTTR as compared to other LTTRs. The cut off for log2 fold change (FC) for all LTTRs is taken as > (+/-)1.5.

The log2 fold change (FC) and P-adj values marked in grey have P-adj > 0.05 or NA (Not available).

The log2 fold change (FC) values in blue highlight the targets which are identified from ChIP-exo analysis (Gao et al., 2018). The target genes corresponding to the log2 fold change (FC) values enclosed in strong black bold frames have strong peaks in IGV analysis whereas the log2 fold change (FC) values in black frames correspond to targets having weak peaks in IGV analysis.

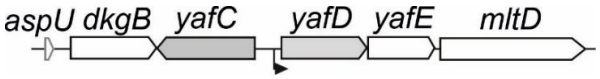


Figure 6: Schematics showing the genome organization for *yafC*. The gene *dkgB* is the adjacent gene to *yafC,* with *yafD* being the divergent gene for *yafC*.

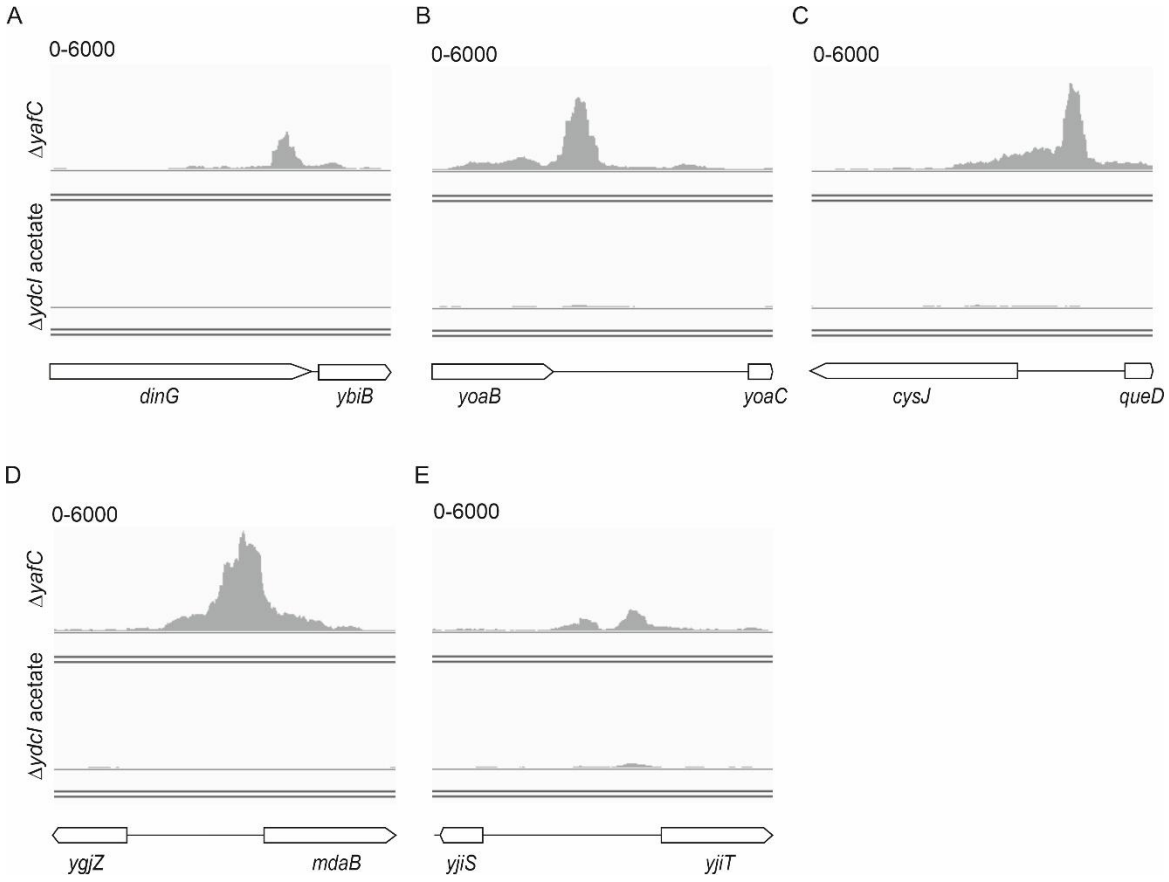


Figure 7: Visualisation of ChIP-exo data for YafC targets in IGV browser. The whole genome BAM file from Gao et al. (2018) for *E. coli* K-12 Δ*yafC and* Δ*ydcI* (grown in acetate medium and taken as control) are used for the visualization. The reads are shown as the grey graphs and are depicted for A) *ybiB*, B) *yoaC*, C) *cysJ-queD*, D) *ygjZ-mdaB* and E) *yjiS-yjiT*. The window size of each graph is around 900 bp and the data range is selected from 0-6000.

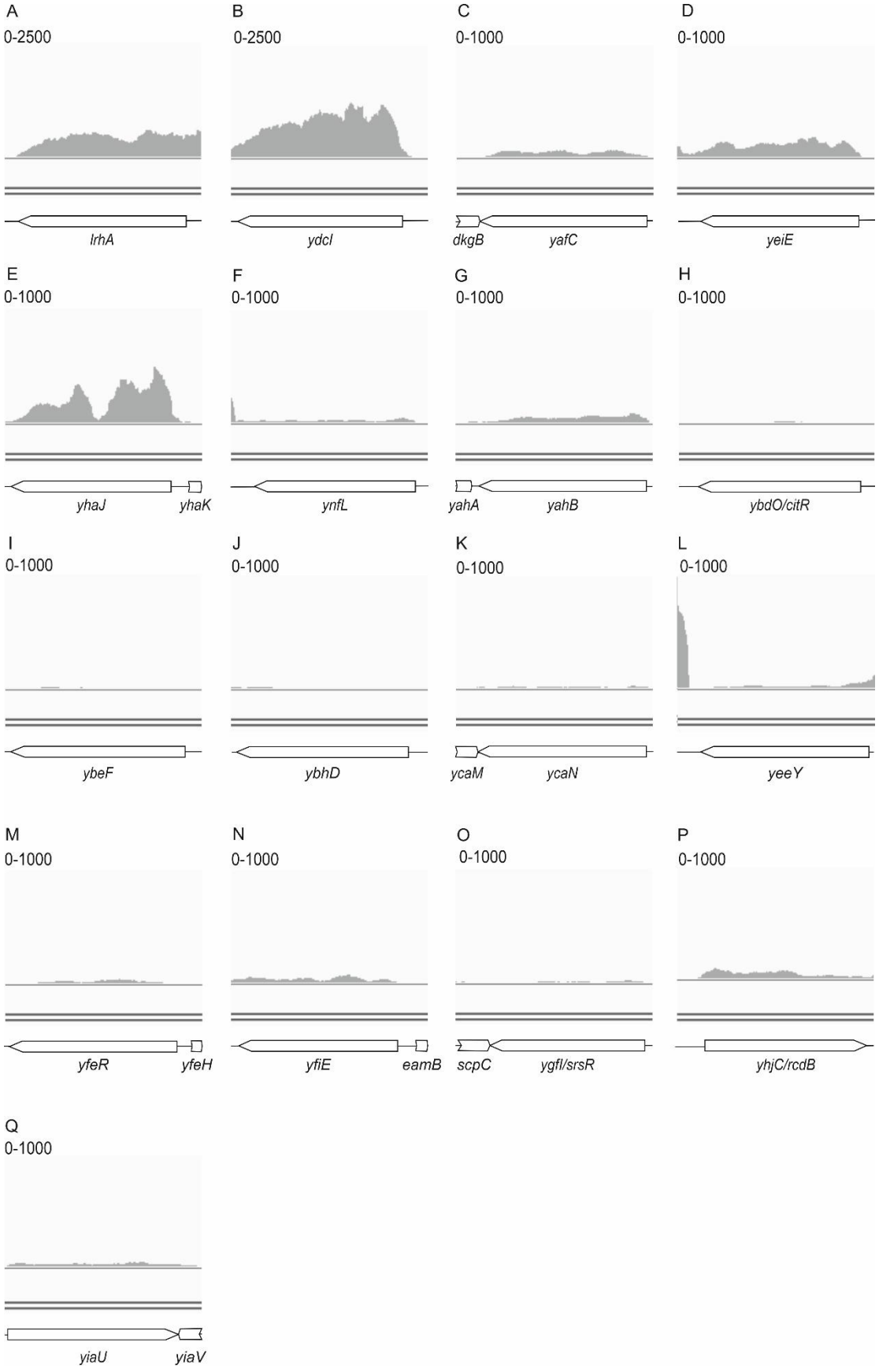

A
0-2500
lrhA
B
0-2500
ydcI
C
0-1000
dkgB
yafC
D
0-1000
yeiE
E
0-1000
yhaJ
yhaK
F
0-1000
ynfL
G
0-1000
yahA
yahB
H
0-1000
ybdO/citR
I
0-1000
ybeF
J
0-1000
ybhD
K
0-1000
ycaM
ycaN
L
0-1000
yeeY
M
0-1000
yfeR
yfeH
N
0-1000
yfiE
eamB
O
0-1000
scpC
ygfI/srsR
P
0-1000
yhjC/rcdB
Q
0-1000
yiaU
yiaV

Figure 8 (previous page): Visualisation of RNA-seq data for the LTTRs in IGV browser. The BAM file for the plus or minus strands for *E. coli* K-12 strain U4 (Δ*lacZ*) are used for the anaylsis. The reads are shown as the grey graphs and are depicted for A) *lrhA*, B) *ydcI,* C) *yafC*, D) *yeiE*, E) *yhaJ*, F) *ynfL*, G) *yahB*, H) *ybdO/citR*, I) *ybeF*, J) *ybhD*, K) *ycaN*, L) *yeeY*, M) *yfeR*, N) *yfiE*, O) *ygfI/srsR*, P) *yhjC/rcdB* and Q) *yiaU*. The window size of each graph is around 1080 bp and the data range is selected from 0-2500 for *lrhA*, *ydcI* and from 0-1000 for all other LTTRs.

Conclusion

Out of 46 LTTRs known for *E. coli* K-12, 19 remain partially characterized or completely uncharacterised (Tables 1 and S1). This underlines the need of further work to understand the regulatory and functional role of these LTTRs. In this study a combinational approach of characterizing the LTTRs expression (protein and transcript level), identification of putative target genes (RNA-seq transcriptome analysis), and their potential autoregulation (using promoter *lacZ* reporter fusions and *in vitro* DNA binding analyses) was used for characterization of the LTTRs. The LTTR-3xFLAG expression analysis shows detectable expression for LrhA, YafC and YhaJ independent of the growth phase and expression of YdcI and YeeY being dependent on the growth phase (Figure 2). In addition, as evident no expression was detected for the majority of the uncharacterized LTTRs under laboratory growth conditions (Figure 2D), which may be attributable to the absence of an inducer or their repression by the pleiotropic nucleoid-associated H-NS protein under laboratory growth conditions. Based on the LTTR-3xFLAG expression analysis, LrhA, YafC, YhaJ and YdcI were selected for RNA-seq analyses to identify the putative target genes for these LTTRs and further decipher their regulatory role. However, the analyzed RNA-seq data shows a large number of non-specific targets and only very few specific targets for each of the LTTRs (section 1.4). In parallel, promoter *lacZ* reporter expression analysis identified YafC, YeiE, YhaJ, YnfL, YfeR, YfiE and YhjC as autorepressors and the LTTR, YbhD as an activator of its own promoter (section 1.2). Also, a weak positive regulation of the divergent gene is detected for YnfL (section 1.2). Additionally, out of the LTTRs which function as autoregulators, for YafC, YeiE, Yhaj and YnfL binding to their promoter fragments was confirmed by EMSA (Figure 5).

Taken together, this study allowed characterising the LTTRs with repect to their autoregulatory function and to some extent identifying putative target genes of LTTRs. However, a major challenge remains in identifying the environmental conditions for expression of the LTTRs. In addition, lack of knowledge about the small molecule ligands/effectors which presumably are required to modulate LTTRs activity makes it difficult to identify specific LTTRs target genes. In this study, the wide spectrum of the non-specific target genes of the LTTRs identified redundantly by RNA-seq analysis renders this transcriptome analysis as a not very effective method for the identification of the target genes of the LTTRs. The comparative RNA-seq analysis instead helped to identify the candidate target genes for the selected LTTRs. Alternately, high throughput techniques such as ChIP-seq or ChIP-exo, which takes into consideration the property of direct DNA-binding of the LTTRs would be more advantageous for identifying the specific target genes of the

LTTRs and should be used for future studies to elucidate their regulatory and functional roles. Nonetheless, it is quite evident that a reverse genetic approach to characterise LTTRs or other transcriptional regulators without prior idea about their biological function is not the best approach for their characterisation, and has been successful only in very few cases such as identification of the biological function of only 3 out of 16 and 10 out of 34 transcription regulators (Gao et al., 2021, Gao et al., 2018).

Materials and Methods

Bacterial strains, plasmids and media

Escherichia coli K-12 strains and strain constructions are described in Table 3. Strains were constructed by P1*vir* transduction (Miller, 1992) and Lambda Red mediated chromosomal recombineering (Datsenko and Wanner, 2000). The plasmid constructions are described in Table 4 and oligonucleotides are given in supplementary Table S2. Bacterial cultures were grown in LB medium (5 g/l Bacto Yeast extract, 10 g/l Bacto Tryptone, 5 g/l NaCl), tryptone medium (10 g/L Bacto-Tryptone, 5 g/L NaCl) and for plates 15 g/l Bacto Agar was added. Antibiotics were added in final concentrations of 50 µg/ml ampicillin, 15 µg/ml chloramphenicol and 25 µg/ml kanamycin. IPTG (isopropyl-β-D-thiogalactopyranoside) was added to liquid cultures at a concentration of 200 µM.

Protein purification

For purification of histidine-tagged $LTTR_{His6}$ and LTTR-DBD_{His6}, expression strain C41(DE3) was transformed with plasmids carrying the respective genes under the control of the *T7* promoter (Table 4). For protein purification, over-night cultures of the transformants were used to inoculate 400 ml culture of LB medium with ampicillin to an OD_{600} of 0.08 and grown at 37°C to OD_{600} of 0.6. At this point, protein expression was induced by adding 200 µM IPTG and the cultures were grown further for 2 hours at 37°C. The cultures were harvested and the bacteria were pelleted by centrifugation. The cells were resuspended in resuspension buffer (20 mM Tris-HCl pH 8.0, 300 mM KCl), and again pelleted for storage at −80°C. Lysate was prepared at 4°C; the cell pellets were resuspended in lysis buffer [4 ml/g of cells; 20mM Tris-HCl, pH 8.0, 300 mM KCl, 10 mM Imidazole, DNaseI (10000 U/µl; Roche)] and lysed by sonication (40% amplitude, 4 min with a 2-s pulse/2-s pause; Sonics Vibracell VCX750 High-Volume Ultrasonic Cell Disrupter). The lysate was cleared by centrifugation (2X; 40 000 rpm, 30 min, Sorvall Centrifuge with SS34). The supernatant (cleared lysate) is bound to Ni-NTA slurry for an hour and was then loaded onto polypropylene column. The column was washed with wash buffer (20 mM Tris-HCl, pH 8.0, 300 mM KCl) with increasing imidazole concentrations (20 and 60 mM) with 10 times the resin volume at each step. The proteins were eluted with 20 mM Tris-HCl, pH 8.0, 300 mM KCl, 250 mM imidazole and 600 µl fractions were collected and analyzed by SDS-PAGE. Proteins that were purified were stored in storage buffer (20 mM Tris-HCl, pH 8.0, 100 mM KCl and 50 mM NDSB-256). Protein concentrations were measured by Pierce bicinchoninic acid (BCA) protein assay (Thermo Scientific, Germany). Aliquots of the proteins were stored at -80°C.

Electrophoretic mobility shift assays (EMSA)

For DNA-binding analyses by electrophoretic mobility shift assays (EMSAs), DNA fragments were amplified by PCR using the primers indicated in Table S2. After clean-up of the fragments (NucleoSpin Gel and PCR Clean-up Mini kit, Macherey-Nagel, Germany), the DNA samples were stored at -20°C. DNA-binding reactions of purified $LTTR_{His6}$ and LTTR-DBD_{His6} were set up on ice in 1X binding buffer (20mM Tris-HCl [pH 7.5], 100mM KCl, 2mM dithiothreitol [DTT] and 10% glycerol) with 10 nM DNA fragment for each DNA fragment in a reaction volume of 10 µl. The samples were incubated at 30°C for 20 min and then loaded onto an 8% polyacrylamide gel (acrylamide-bisacrylamide [29:1] in 0.5x TBE [45 mM Tris base, 45 mM boric acid, 1 mM EDTA]). The gel was run for 10 min at 50 V, followed by 45 min at 200 V. The gel was stained with ethidium bromide for 20 min followed by imaging with GelDoc from Biorad company.

β-galactosidase expression analysis

For expression analyses of promoter *lacZ* fusions, β-galactosidase assays were performed (Miller, 1992). Briefly, exponential cultures were inoculated from a fresh over-night culture to an OD_{600} 0.02 in LB medium that was supplemented with the specific antibiotic in case of transformants. Cultures were grown at 37°C to OD_{600} ~0.8 and then harvested on ice. The β-galactosidase assays were repeated at least three times from independent biological replicates. Mean values of the three independent biological replicates and the corresponding standard deviations are calculated. Student *t*-Test is performed for pair of samples as indicated with the asterisks in the bar graphs.

Western blotting

For analysis of the growth phase dependent expression of chromosomal *lttr-3xFLAG* under the control of its native promoter, cultures were grown over-night and used to inoculate 50 ml of tryptone medium to an OD_{600} of 0.05 and grown for 8 hours at 37°C. The samples were collected from over-night cultures and after 2, 3, 4, 6, 7 and 8 hours of growth and the harvested cells were pelleted for 2 min at 13,000 rpm. SDS-PAGE and Western blot procedures were performed. For each lane, cells equivalent to OD_{600} of 0.08 were loaded and equal loading was verified by 2,2,2-trichloroethane staining of total protein (Ladner et al., 2004). Epitope-tagged LTTR-3xFLAG was detected using primary antibody anti-FLAG M2 from mouse (diluted 1:4,000) (catalogue number F3165; Sigma) and Alexa Fluor 680 fluorescent dye-labeled secondary anti-mouse antibody from goat (diluted 1:10,000) (catalogue number A21057; Thermo Scientific). The blots were visualized with Odyssey V3.0 software.

RNA-seq and data analysis

For RNA-seq, total RNA was isolated from *E. coli* K-12 strain U4 (Δ*lacZ*) taken as a control and isogenic strains with deletions of Δ*yafC* (U456), Δ*yhaJ* (U458), Δ*ydcI* (U505) and Δ*lrhA* (U504) (Table 3). These strains were inoculated to an OD_{600} of 0.05 in 10 ml tryptone medium and grown for 2.5 h to an OD_{600} of approximately 0.8 at 37°C with shaking. Then 1.5 ml samples of the cultures were harvested by adding RNAprotect bacterial reagent (Qiagen, Germany) followed by total RNA isolation using the RNeasy minikit (Qiagen, Germany) with on-column DNA digestion by DNase I treatment, according to the manufacturer's instructions (Qiagen, Germany). The integrity of the RNA was validated by separating 1 µg of RNA on 8% denaturing acrylamide gels (7 M urea, 45 mM Tris base, 45 mM boric acid, 1 mM EDTA, 8% acrylamide-bisacrylamide [19:1, wt/wt]) and visual inspection of the 23S rRNA and 16S rRNA bands.

The preparation of RNA-seq libraries was performed according to methods described previously (Borries et al., 2011), with minor adaptations. The synthesis of cDNA libraries was performed by the Cologne Center for Genomics (CCG) (http://portal.ccg.uni-koeln.de/ccg/) using an Illumina TruSeq RNA library kit according to the manufacturer's instructions. Three biological replicates of each strain were sequenced, with 20 million reads each, using an Illumina HiSeq 4000 instrument. For data analysis, the sequencing raw data files were uploaded to the public Galaxy Web platform (https://usegalaxy.org/) (Afgan et al., 2016). First, raw sequencing reads were uploaded in FASTQ file format and aligned to the *E. coli* K-12 MG1655 genome (GenBank assembly accession number GCA_000005845.2) using Bowtie2 (Langmead and Salzberg, 2012) with default settings for paired-end libraries. These generated BAM files containing the aligned reads were used to calculate individual gene abundances. For this, HTSeq (Anders et al., 2015) was used in union mode for stranded data sets with a minimum alignment quality of 10. Tabular lists of gene abundances for each sample were used for differential expression analysis using DEseq2 (Love et al., 2014) with default settings. In addition, the aligned reads compiled in the BAM files of the individual samples were inspected using IGV (Integrative Genomics Viewer) (Robinson et al., 2011).

Data availability

Raw sequencing files (FASTQ file format) and gene abundance tables (TXT file format) have been deposited and are available in NCBI's Gene Expression Omnibus database under accession number GSE304471.

Tables

Table 3: *E. coli* K-12 strains		
Strains	**Genotypes[a]**	**Reference/ Details[b]**
T1241	BW30270 *ilvG*$^{+}$, *fhlDC*(IS1) highly motile, *dgcJ*::IS1	(Pannen et al., 2016)
C41(DE3)	BL21(DE3) derivate for toxic protein expression	(Miroux and Walker, 1996)
Donor strains for transduction		
JW0198-1	BW25113 Δ*yafC727*$_{kan}$ Lab number T2791	(Baba et al., 2006)
JW2144-1	BW25113 Δ*yeiE784* $_{kan}$ Lab number T2792	(Baba et al., 2006)
JW3076-1	BW25113 Δ*yhaJ724* $_{kan}$ Lab number T2793	(Baba et al., 2006)
JW1587-1	BW25113 Δ*ynfL743* $_{kan}$ Lab number T2795	(Baba et al., 2006)
JW0308-3	BW25113 Δ*yahB758* $_{kan}$ Lab number T2911	(Baba et al., 2006)
JW0596-1	BW25113 Δ*ybdO742* $_{kan}$ Lab number T2912	(Baba et al., 2006)
JW0624-1	BW25113 Δ*ybeF768* $_{kan}$ Lab number T2914	(Baba et al., 2006)
JW0883-1	BW25113 Δ*ycaN724* $_{kan}$ Lab number T2916	(Baba et al., 2006)
JW2400-1	BW25113 Δ*yfeR739* $_{kan}$ Lab number T2919	(Baba et al., 2006)
JW2561-1	BW25113 Δ*yfiE745* $_{kan}$ Lab number T2922	(Baba et al., 2006)
JW3557-5	BW25113 Δ*yiaU769* $_{kan}$ Lab number T2923	(Baba et al., 2006)
JW5476-2	BW25113 Δ*ygfI772* $_{kan}$ Lab number T2924	(Baba et al., 2006)
JW5834-1	BW25113 Δ*yeeY784* $_{kan}$ Lab number T2925	(Baba et al., 2006)
JW5896-1	BW25113 Δ*ybhD740* $_{kan}$ Lab number T2926	(Baba et al., 2006)
JW5226-1	BW25113 Δ*ydcI783* $_{kan}$ Lab number T2034	(Baba et al., 2006)
K-12 T1241 derivatives		
T2935	T1241 Δ*yhjC*$_{KD13\text{-}kan}$	T1241/pKD46 x OB284/OB285 (pKD13)
U3	T1241 Δ*ara, Δlac,* motile	(Breddermann and Schnetz, 2016)
U4	T1241 Δ*lacZ*	(Yilmaz et al., 2020)
U78	T1241 Δ*ara* Δ*lac* Δ*lrhA*$_{KD3\text{-}cm}$	(Breddermann and Schnetz, 2017)
U253	T1241 *lrhA*$_{3F\ kan}$	T1241/pKD46 x OA816/OA817 (pSUB11)
U254	T1241 *yafC*$_{3F\ kan}$	T1241/pKD46 x OA826/OA827 (pSUB11)

Table 3: *E. coli* K-12 strains		
Strains	**Genotypes[a]**	**Reference/ Details[b]**
U255	T1241 *ydcI*$_{3F\ kan}$	T1241/pKD46 x OA824/OA825 (pSUB11)
U256	T1241 *yeiE*$_{3F\ kan}$	T1241/pKD46 x OA828/OA829 (pSUB11)
U258	T1241 *ynfL*$_{3F\ kan}$	T1241/pKD46 x OA832/OA833 (pSUB11)
U263	T1241 *yhaJ*$_{3F\ kan}$	T1241/pKD46 x OA822/OA823 (pSUB11)
U511	T1241 *yahB*$_{3F\ kan}$	T1241/pKD46 x OB165/OB166 (pSUB11)
U512	T1241 *ybdO*$_{3F\ kan}$	T1241/pKD46 x OB167/OB168 (pSUB11)
U513	T1241 *ybeF*$_{3F\ kan}$	T1241/pKD46 x OB169/OB170 (pSUB11)
U514	T1241 *ybhD*$_{3F\ kan}$	T1241/pKD46 x OB171/OB172 (pSUB11)
U515	T1241 *ycaN*$_{3F\ kan}$	T1241/pKD46 x OB173/OB174 (pSUB11)
U516	T1241 *yeeY*$_{3F\ kan}$	T1241/pKD46 x OB175/OB176 (pSUB11)
U517	T1241 *yfeR*$_{3F\ kan}$	T1241/pKD46 x OB177/OB178 (pSUB11)
U518	T1241 *yfiE*$_{3F\ kan}$	T1241/pKD46 x OB179/OB180 (pSUB11)
U519	T1241 *ygfI*$_{3F\ kan}$	T1241/pKD46 x OB181/OB182 (pSUB11)
U520	T1241 *yhjC*$_{3F\ kan}$	T1241/pKD46 x OB183/OB184 (pSUB11)
U451	U4 Δ*yafC*$_{kan}$	U4 x P1vir [JW0198-1 Δ*yafC*$_{kan}$]
U452	U4 Δ*yeiE*$_{kan}$	U4 x P1vir [JW2144-1 Δ*yeiE*$_{kan}$]
U453	U4 Δ*yhaJ*$_{kan}$	U4 x P1vir [JW3076-1 Δ*yhaJ*$_{kan}$]
U455	U4 Δ*ynfL*$_{kan}$	U4 x P1vir [JW1587-1 Δ*ynfL*$_{kan}$]
U522	U4 Δ*yahB*$_{kan}$	U4 x P1vir [JW0308-3 Δ*yahB*$_{kan}$]
U523	U4 Δ*ybdO*$_{kan}$	U4 x P1vir [JW0596-1 Δ*ybdO*$_{kan}$]
U524	U4 Δ*ybeF*$_{kan}$	U4 x P1vir [JW0624-1 Δ*ybeF*$_{kan}$]
U525	U4 Δ*ycaN*$_{kan}$	U4 x P1vir [JW0883-1 Δ*ycaN*$_{kan}$]
U526	U4 Δ*yfeR*$_{kan}$	U4 x P1vir [JW2400-1 Δ*yfeR*$_{kan}$]
U527	U4 Δ*yfiE*$_{kan}$	U4 x P1vir [JW2561-1 Δ*yfiE*$_{kan}$]
U528	U4 Δ*yiaU*$_{kan}$	U4 x P1vir [JW3557-5 Δ*yiaU*$_{kan}$]
U529	U4 Δ*ygfI*$_{kan}$	U4 x P1vir [JW5476-2 Δ*ygfI*$_{kan}$]
U530	U4 Δ*yeeY*$_{kan}$	U4 x P1vir [JW5834-1 Δ*yeeY*$_{kan}$]
U531	U4 ΔybhD$_{kan}$	U4 x P1vir [JW5896-1 Δ*ybhD*$_{kan}$]
U532	U4 ΔyhjC$_{kan}$	U4 x P1vir [T2935 Δ*yhjC*$_{KD13\text{-}kan}$]
U456	U4 Δ*yafC*$_{FRT}$	U451 x pCP20
U457	U4 Δ*yeiE*$_{FRT}$	U452 x pCP20
U458	U4 Δ*yhaJ*$_{FRT}$	U453 x pCP20
U460	U4 Δ*ynfL*$_{FRT}$	U455 x pCP20
U502	U4 Δ*lrhA*$_{cm}$	U4 x P1vir [U78 Δ*lrhA*$_{KD3\text{-}cm}$]
U503	U4 Δ*ydcI*$_{kan}$	U4 x P1vir [JW5226-1 Δ*ydcI*$_{kan}$]
U504	U4 Δ*lrhA*$_{FRT}$	U502 x pCP20
U505	U4 Δ*ydcI*$_{FRT}$	U503 x pCP20
U533	U4 Δ*yahB*$_{FRT}$	U522 x pCP20
U534	U4 Δ*ybdO*$_{FRT}$	U523 x pCP20
U535	U4 Δ*ybeF*$_{FRT}$	U524 x pCP20
U536	U4 Δ*ycaN*$_{FRT}$	U525 x pCP20
U537	U4 Δ*yfeR*$_{FRT}$	U526 x pCP20
U538	U4 Δ*yiaU*$_{FRT}$	U528 x pCP20
U539	U4 Δ*ygfI*$_{FRT}$	U529 x pCP20
U540	U4 Δ*yeeY*$_{FRT}$	U530 x pCP20
U541	U4 Δ*ybhD*$_{FRT}$	U531 x pCP20
U542	U4 Δ*yfiE*$_{FRT}$	U527 x pCP20

Table 3: *E. coli* K-12 strains		
Strains	**Genotypes[a]**	**Reference/ Details[b]**
U543	U4 Δ*yhjC*$_{FRT}$	U532 x pCP20

[a]Abbreviations and genetic features are as follow: kanamycin resistance (kan); chloramphenicol resistance (cm); FLP recombinase target site (FRT).

[b] Strains which were constructed by transduction is stated as "x phage [donor strain]"; λ Red recombineering, stated as "x PCR primer pair (template)" followed by Flp recombinase-catalyzed deletion of the resistance marker, "x pCP20".

Table 4: Plasmids		
Plasmids	**Description***	**Reference or construction**
pCP20	*cI*$_{857}$ *P*$_R$ *flp* in pSC101-ori rep^{ts} Amp	(Cherepanov and Wackernagel, 1995)
pKD3	FRT-*cmR*-FRT oriRγ Amp	(Datsenko and Wanner, 2000)
pKD4	FRT-*kanR*-FRT oriRγ Amp	(Datsenko and Wanner, 2000)
pET22b	*P*$_{T7}$ MCS His6 in pMB1-ori Amp	Novagen
pKESK22	*lacI*^{q} *P*$_{tac}$ MCS in p15A-ori Kan	(Stratmann et al., 2008)
pKD13	FRT-*kanR*-FRT *tL3* oriRγ Amp	(Datsenko and Wanner, 2000)
pKD46	*araC Para γ-β-exo* in pSC101-ori rep^{ts} Amp	(Datsenko and Wanner, 2000)
pSUB11	*3xFLAG* in pKD4	(Shimada et al., 2005)
pKESL315	*yafC*$_{His6}$ (full-length) in pET22b(+) Amp	PCR OA862/OA863 of T1241 in pET22b(+) (NdeI, XhoI)
pKESL319	*yeiE*$_{His6}$ (full-length) in pET-22b(+) Amp	PCR OA868/OA869 of T1241 in pET22b(+) (NdeI, XhoI)
pKESL317	*ydcI*$_{His6}$ (full-length) in pET-22b(+) Amp	PCR OA865/OA866 of T1241 in pET22b(+) (NdeI, XhoI)
pKESL321	*yhaJ*$_{His6}$ (full-length) in pET-22b(+) Amp	PCR OA877/OA878 of T1241 in pET22b(+) (NdeI, XhoI)
pKESL326	*ynfL-DBD*$_{His6}$ in pET-22b(+) Amp (DBD: Met1 - Gly93)	PCR OA871/OA873 of T1241 in pET22b(+) (NdeI, XhoI)
pKEHB6	*lacI*^{q} *P*$_{tac}$ *ydcI* in pKESK22	(Breddermann and Schnetz, 2017)
pKERBR19	*lacI*^{q} *P*$_{tac}$ *yafC* in pKESK22	*yafC* (PCR OB031/OB032) in pKESK22
pKERBR20	*lacI*^{q} *P*$_{tac}$ *yeiE* in pKESK22	*yeiE* (PCR OB037/OB038) in pKESK22
pKERBR21	*lacI*^{q} *P*$_{tac}$ *yhaJ* in pKESK22	*yhaJ* (PCR OB043/OB044) in pKESK22
pKERBR23	*lacI*^{q} *P*$_{tac}$ *ynfL* in pKESK22	*ynfL* (PCR OB053/OB054) in pKESK22
pKERBR64	*lacI*^{q} *P*$_{tac}$ *ybdO* in pKESK22	*ybdO* (PCR OB219/OB220) in pKESK22
pKERBR69	*lacI*^{q} *P*$_{tac}$ *yahB* in pKESK22	*yahB* (PCR OB239/OB240) in pKESK22

Table 4: Plasmids		
Plasmids	**Description***	**Reference or construction**
pKERBR70	*lacI*^{q} P_{tac} *ybeF* in pKESK22	*ybeF* (PCR OB243/OB244) in pKESK22
pKERBR71	*lacI*^{q} P_{tac} *ybhD* in pKESK22	*ybhD* (PCR OB247/OB248) in pKESK22
pKERBR72	*lacI*^{q} P_{tac} *ycaN* in pKESK22	*ycaN* (PCR OB251/OB252) in pKESK22
pKERBR67	*lacI*^{q} P_{tac} *yeeY* in pKESK22	*yeeY* (PCR OB231/OB232) in pKESK22
pKERBR68	*lacI*^{q} P_{tac} *yfeR* in pKESK22	*yfeR* (PCR OB235/OB236) in pKESK22
pKERBR65	*lacI*^{q} P_{tac} *yfiE* in pKESK22	*yfiE* (PCR OB223/OB224) in pKESK22
pKERBR73	*lacI*^{q} P_{tac} *ygfI* in pKESK22	*ygfI* (PCR OB255/OB256) in pKESK22
pKERBR66	*lacI*^{q} P_{tac} *yhjC* in pKESK22	*yhjC* (PCR OB227/OB228) in pKESK22
pKERBR74	*lacI*^{q} P_{tac} *yiaU* in pKESK22	*yiaU* (PCR OB259/OB260) in pKESK22
pKERBR24	P_{yafD} *lacZ* in pSC-ori Cm	P_{yafD} (PCR OB033/OB034) in pKECY49
pKERBR25	P_{yafC} *lacZ* in pSC-ori Cm	P_{yafC} (PCR OB035/OB036) in pKECY49
pKERBR26	P_{yeiH} *lacZ* in pSC-ori Cm	P_{yeiH} (PCR OB039/OB040) in pKECY49
pKERBR27	P_{yeiE} *lacZ* n pSC-ori Cm	P_{yeiE} (PCR OB041/OB042) in pKECY49
pKERBR28	P_{yhaK} *lacZ* in pSC-ori Cm	P_{yhaK} (PCR OB045/OB046) in pKECY49
pKERBR29	P_{yhaJ} *lacZ* in pSC-ori Cm	P_{yhaJ} (PCR OB047/OB048) in pKECY49
pKERBR31	P_{ynfM} *lacZ* in pSC-ori Cm	P_{ynfM} (PCR OB055/OB056) in pKECY49
pKERBR32	P_{ynfL} *lacZ* in pSC-ori Cm	P_{ynfL} (PCR OB057/OB058) in pKECY49
pKERBR75	P_{ybdO} *lacZ* in pSC-ori Cm	P_{ybdO} (PCR OB261/OB262) in pKECY49
pKERBR80	P_{yahB} *lacZ* in pSC-ori Cm	P_{yahB} (PCR OB271/OB272) in pKECY49
pKERBR81	P_{ybeF} *lacZ* in pSC-ori Cm	P_{ybeF} (PCR OB273/OB274) in pKECY49
pKERBR82	P_{ybhD} *lacZ* in pSC-ori Cm	P_{ybhD} (PCR OB275/OB276) in pKECY49
pKERBR83	P_{ycaN} *lacZ* in pSC-ori Cm	P_{ycaN} (PCR OB277/OB278) in pKECY49
pKERBR78	P_{yeeY} *lacZ* in pSC-ori Cm	P_{yeeY} (PCR OB267/OB268) in pKECY49
pKERBR79	P_{yfeR} *lacZ* in pSC-ori Cm	P_{yfeR} (PCR OB269/OB270) in pKECY49
pKERBR76	P_{yfiE} *lacZ* in pSC-ori Cm	P_{yfiE} (PCR OB263/OB264) in pKECY49
pKERBR84	P_{ygfI} *lacZ* in pSC-ori Cm	P_{ygfI} (PCR OB279/OB280) in pKECY49
pKERBR77	P_{yhjC} *lacZ* in pSC-ori Cm	P_{yhjC} (PCR OB265/OB266) in pKECY49
pKERBR85	P_{yiaU} *lacZ* in pSC-ori Cm	P_{yiaU} (PCR OB281/OB282) in pKECY49
pKERBR93	P_{ydcI} *lacZ* in pSC-ori Cm	P_{ydcI} (PCR OB314/ OB315) in pKECY49

Abbreviations and features are as follow: chloramphenicol resistance (Cm); kanamycin resistance (Kan); ampicillin resistance (Amp); MCS for multiple cloning site; rep^{ts} for temperature sensitive plasmid replication; His6 for 6x histidine tag; 3x FLAG for DYKDHDGDYKDHDIDYKDDDDK (Sigma-Aldrich, Germany).

Supplementary Information

The supplementary information includes

table S1, an overview of all LTTRs in *E. coli* K-12, and table S2, oligonucleotides used in this study.

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
AaeR	3,389,520 -> 3,390,449	309	activates divergent *aaeXAB* encoding hydroxylated aromatic carboxylic acids efflux system	Unknown	Yes	No	203	(Van Dyk et al., 2004)
AbgR	1,404,741 -> 1,405,649	302	putatively regulates *abgABT* encoding p-aminobenzoyl-glutamate hydrolase subunits and symporter	Unknown	Yes	No	27	(Hussein et al., 1998)
AllS	531,295 <- 532,221	308	activates *allDCE* encoding enzymes involved in anaerobic utilization of allantoin as nitrogen source; *allS* maps divergent to *allA* encoding ureidoglycolate lyase	Unknown	No	No	1	(Rintoul et al., 2002)
ArgP	3,059,753 -> 3,060,646	297	regulates *argO* encoding arginine exporter ArgP-arginine induces *argO* (by recruitment of RNA polymerase and stimulation of open complex formation) ArgP-lysine is a repressor of *argO,* and other genes negative dominant ArgP mutants include P274S, S94L, P108S, V144M, P217L, L294F, R295C, A68V; ArgP-S94L confers constitutive *argO* expression	arginine lysine	No	No	789	(Laishram and Gowrishankar, 2007, Marbaniang and Gowrishankar, 2011, Nandineni and Gowrishankar, 2004, Nandineni et al., 2004, Ruiz et al., 2011)
Cbl	2,059,964 <- 2,060,914	316	activates *ssu* and *tau* operon coding for transport and enzymatic processing of alkane sulphonates and taurine, respecitvely co-effector adenosine 5-phosphosulphate inhibits activation by Cbl 'constitutive' Cbl mutants T102W, E150A, W166A and T202A Cbl binding site mapped by footprinting	adenosine phospho-sulphate	No	No	6	(Bykowski et al., 2002, Iwanicka-Nowicka and Hryniewicz, 1995, Stec et al., 2006, van der Ploeg et al., 2001, van Der Ploeg et al., 1999, van der Ploeg et al., 1997)
CynR	357,791 <- 358,690	299	activates divergent *cyn* operon in response to cyanate promoters of *cyn* operon and *cynR* overlap indicative of autoregulation CynR binding site mapped by footprinting	Cyanate	Yes	No	152	(Knapp and Hu, 2009, Lamblin and Fuchs, 1993, 1994, Sung and Fuchs, 1992)
CysB	1,333,855 -> 1,334,829	324	inducer N-acetyl-L-serine; anti-inducer sulfide and thiosulfate Inducer binding changes CysB-DNA-binding pattern	N-acetyl-L-serine	No	No	1291	(Hryniewicz and Kredich, 1991, Jovanovic et al., 2003, Lochowska et al., 2001, Tyrrell et al., 1997, van der Ploeg et al., 2001, van der Ploeg et al., 1997)
DmlR	1,880,886 <- 1,881,809	307	regulates aerobic and anaerobic D-malate metabolism. 49% identical with TtdR divergent to *dmlA*, encoding D-malate dehydrogenase	D-malate (or L- or meso-tartate)	?	No	53	(Lukas et al., 2010)

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
DsdC	2,476,694 <- 2,477,629	311	activates divergent *dsdXA* operon coding for D-serine transporter and D-serine deaminase (D-serine is toxic in minimal medium due to inhibition of L-serine synthesis)	D-serine	Yes	No	153	(Norregaard-Madsen et al., 1995)
GcvA	2,941,650 <- 2,942,567	305	activates divergent *gcvB* encoding sRNA and *gcvTHP* operon coding for glycine cleavage enzyme system GcvA activity is regulated by interaction with GcvR (not DNA binding), glycine binds to GcvR and inhibits interaction of GcvR with GcvA activation of transcription by GcvA involves RNA polymerase α-subunit	GcvR/glycine	Yes	No	287	(Heil et al., 2002, Jourdan and Stauffer, 1998, Stauffer and Stauffer, 2012, Wilson et al., 1995)
HcaR	2,668,006 <- 2,668,896	296	regulates genes involved in 3-phenylpropionate (hydrocinnamic acid) catabolism	phenyl-propionate	Yes	No	33	(Diaz et al., 1998, Turlin et al., 2005)
HdfR	3,947,128 <- 3,947,967	279	repressor of *flhDC* *hdfR* is H-NS repressed	Unknown	No	No	297	(Ko and Park, 2000)
HypT	4,557,378 <- 4,558,289	303	thiol based redox regulator putative dodecameric structure	hypochlorite-responsive	No	No	58	(Drazic et al., 2014, Drazic et al., 2013, Gebendorfer et al., 2012)
IlvY	3,956,927 <- 3,957,820	297	activates (inducer-dependent) divergent *ilvC* coding for acetohydroxy acid isomeroreductase (leucine, isoleucine, valine pathway) autorepresses (independent of inducer) *ilvY* effector-binding to IlvY affects DNA binding pattern	α-aceto-lactate or α-acetohydroxy butyrate	Yes	No	100	(Rhee et al., 1998, Wek and Hatfield, 1988)
LeuO	84,368 -> 85,312	314	*leuO* is H-NS and StpA regulated, activated by LrhA, *leuO* is antagonistically regulated by the heterodimeric transcription regulator BglJ-RcsB and LeuO itself, Ser120 to Asp substitution triggers a structural change that is related to effector-induced structural changes of LTTRs	Unknown	No	Yes	11	(Breddermann and Schnetz, 2016, 2017, Fragel et al., 2019, Shimada et al., 2011)
LrhA	2,405,703 <- 2,406,641	312	represses $FlhD_2C_2$, LrhA activates leuO, inhibits translation of rpoS, RcsCDB phosphorelay system and DNA motor protein FtsK represses LrhA, divergent to *yfbQ*	Unknown	?	Yes	962	(Blumer et al., 2005, Breddermann and Schnetz, 2017, Honda et al., 2009, Lehnen et al., 2002, Peterson et al., 2006)
LysR	2,979,021 -> 2,979,956	311	regulates *lysA* encoding diaminopimelate decarboxylase catalyzing decarboxylation of diaminopimelate into lysine	diamino-pimelic acid (lysine neg)	Yes	No	82	(Stragier et al., 1983, Stragier and Patte, 1983, Stragier et al., 1983)

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
MetR	4,011,863 <- 4,012,816	317	regulator of methionine pathway (genetically well studied) activates the divergent *metE* antagonistically regulates *glyA* and *hmp*, dependent on homocysteine interacts with α-subunit of RNA polymerase	Homocysteine	Yes	No	292	(Fritsch et al., 2000, Lorenz and Stauffer, 1996, Punekar et al., 2016, Weissbach and Brot, 1991)
ModE	793,856 <- 794,644		regulates operons encoding transporter of molybdenum and synthesis of molybdoenzymes and molybdate-related functions crystal structure of dimer solved, questions whether ModE is a LTTR or belongs to a different family	Molybdate	No	No	653	(Anderson et al., 1997, Hall et al., 1999, McNicholas and Gunsalus, 2002)
Nac	2,061,016 <- 2,061,933	305	regulates, without a co-effector, genes involved in nitrogen metabolism under nitrogen-limiting conditions N-terminal 100 aa sufficient for regulation possibly glutamine regulated	Unknown	No	No	0 (130 in minimal medium)	(Camarena et al., 1998, Muse and Bender, 1998, 1999)
NhaR	18,715 -> 19,620	301	activates *nhaA* coding for Na+/H+ antiporter, *pga* operon required for production of the biofilm adhesin poly-1,6-Nacetyl-D-glucosamine (PGA), and other genes *nhaA* target is H-NS repressed NhaR DNA footprint pattern changes in presence of Na^+ multicopy NhaR can act as a repressor in an *hns* strain in the absence of Na^+ but as activator in presence of Na^+	Na^+	No	No, regulates with H-NS	141	(Carmel et al., 1997, Dover et al., 1996, Goller et al., 2006, Krol, 2018, Rahav-Manor et al., 1992, Toesca et al., 2001, Wu et al., 2013)
OxyR	4,158,490 -> 4,159,407	305	activates divergent *oxyS* (sRNA) and regulates several other loci relevant for oxidative stress response, autorepesses thiol based redox regulator, reduction of cystein to SOH and possible Cys-S-S-Cys bridge alteration of DNA binding mode (shift by one helical turn in reduced state) mapped by DNA footprinting	oxidation of cysteine in effector-binding pocket	Yes	No	1867	(Choi et al., 2001, Dubbs and Mongkolsuk, 2012, Guo et al., 2019, Hou et al., 2019, Jo et al., 2015, Kim et al., 2002, Knapp et al., 2009, Kullik et al., 1995, Toledano et al., 1994, Zheng et al., 1998)
PerR	269,289 <- 270,182	297	unknown function, annotation as peroxide regulator unclear maps divergent to insertion element IS30	Unknown	?	No	10	(Dubbs and Mongkolsuk, 2012, Keseler et al., 2017, Norregaard-Madsen et al., 1995)
PgrR	1,391,991 -> 1,392,890	299	PgrR regulates the divergent *ycjY-ymjD-ymjC-mpaA* operon, which is considered to encode the enzymes for peptide glycan degradation	Unknown	Yes	No	107	(Shimada et al., 2013)

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
TdcA	3,266,127 <- 3,267,065	312	controls *tdcBC* operon encoding threonine dehydratase and a threonine-serine permease for metabolism of threonine and serine during anaerobiosis co-regulates with TdcR, which is encoded divergently *Salmonella* mutant attenuated virulence	Unknown	Yes	Yes	6	(Ganduri et al., 1993, Hagewood et al., 1994)
TtdR (Dan)	3,205,324 <- 3,206,256	310	TtdR (named Dan) stimulates expression of the divergently transcribed *ttdABT* operon in presence of L- and meso-tartrate TtdR (named Dan) implicated in nucleoid organization at anoxic conditions	L-tartrate meso-tartrate	Yes	No	2	(Lim et al., 2013, Oshima and Biville, 2006)
XapR	2,521,593 <- 2,522,477	294	activates transcription of *xapAB* operon involved in xanthosine metabolism hyperactive mutations V104, E132, P203, T205, D207, F210; V104E	Xanthosine	No	Yes	23	(Jorgensen and Dandanell, 1999, Seeger et al., 1995)
YafC	229,967 <- 230,881	304	unknown function, divergent to *yafDE* which is of unknown function D275G mutation increases resistance to a vacuum-treated sugarcane bagasse hemicellulose hydrolysate, regulates genes important in IR survival	Unknown	No*	No	316	(Byrne et al., 2014, Gao et al., 2018, Shi et al., 2016)
YahB	333,501 <- 334,433	310	unknown function, no divergent gene, possibly 2nd gene of operon	Unknown	No	No	27	(Gao et al., 2021, Keseler et al., 2017, Norregaard-Madsen et al., 1995)
YbdO (CitR)	636,716 <- 637,618	300	unknown function, no divergent gene *ybdO* is H-NS repressed	Unknown	No	Yes	5	(Gao et al., 2021, Higashi et al., 2016, Keseler et al., 2017)
YbeF	660,425 <- 661,378	317	unknown function, no divergent gene	Unknown	No	Yes	21	(Gao et al., 2021, Keseler et al., 2017)
YbhD	799,622 <- 800,575	317	unknown function, divergent to *ybhH*, which is of unknown function	Unknown	?	Yes	1	(Gao et al., 2021, Keseler et al., 2017)
YcaN	948,660 <- 949,568	302	unknown function, divergent to *ycaK* which is predicted to encode an NAD(P)H dehydrogenase and electron transfer flavoprotein-NAD/FAD/quinone oxidoreductase	Unknown	?	No	50	(Gao et al., 2021, Keseler et al., 2017)
YdcI	1,494,148 <- 1,495,071	307	unknown function, divergent to *ydcJ* coding for protein of unknown function S. Typhimurium *ΔydcI* mutants are more sensitive to organic and inorganic acid stress, details in Chapter 2	Unknown	?	No	148	(Gao et al., 2018, Hingley-Wilson et al., 2020, Jennings et al., 2011, Ogasawara et al., 2020, Solomon et al., 2014)

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
YdhB	1,738,866 <- 1,739,798	310	unknown function, divergent to *ydhC* coding for uncharacterized member of the Major Facilitator Superfamily (MFS) of transporters that may function as a proton-driven drug efflux system	Unknown	?	No	114	(Keseler et al., 2017)
YeeY	2,087,329 <- 2,088,258	309	unknown function, no divergent gene	Unknown	No	No	4	(Keseler et al., 2017)
YeiE	2,248,737 <- 2,249,618	293	presumably activates *lysP* encoding a lysine H^+ symporter; *lysP* maps downstream of *yeiE* divergent to *yeiH* (encoding conserved inner membrane protein), involved in iron uptake under iron limitation conditions	Unknown	No*	No	237	(Fujii et al., 2002, Gao et al., 2018)
YfeR	2,525,930 <- 2,526,856	308	unknown function *yfeR* gene of *S. enterica* serovar Typhimurium is induced by osmotic down-shift. Presumptive repressor of divergent *yfeH* encoding putative Na^+-dependent transporter, but in *E. coli yfeH* is annotated as putative cytochrome oxidase which is induced by Autoinducer-2	Unknown	?	No	95	(Banos et al., 2011, Keseler et al., 2017)
YfiE	2,714,439 <- 2,715,320	293	unknown function divergent to *eamB* encoding exporter of L-cysteine (O-Acetylserine/ cysteine export protein)	Unknown	?	No	111	(Gao et al., 2021, Keseler et al., 2017)
YgfI (SrsR)	3,066,277 <- 3,067,173	298	unknown function divergent to *och5* encoding sRNA possibly involved in *csgD* regulation, regulation of swimming and swarming, negatively regulated by RpoS	Unknown	?	No	15	(Gao et al., 2021, Keseler et al., 2017, Onufryk et al., 2005, Yamada et al., 1999)
YhaJ	3,253,318 <- 3,254,214	298	regulates divergent *yhaKL*, required for degradation of quinol and DNT In EHEC YhaJ is a regulator of *yhaO* and the LEE locus, *yhaO* encodes a D-serine transporter (signalling for extraintestinal environment), both *yhaJ* and *yhaO* are required for full pathogenicity of EHEC	Unknown	No*	No	534	(Connolly et al., 2016, Connolly et al., 2020, Connolly et al., 2019, Palevsky et al., 2016)
YhjC (RcdB)	3,672,414 -> 3,673,313	299	unknown function, activator of csgD promoter, large >500 bp intergenic region to divergent yhjB	Unknown	No	Yes	64	(Gao et al., 2021, Keseler et al., 2017, Ogasawara et al., 2020)
YiaU	3,751,992 -> 3,752,966	324	unknown function divergent to *yiaT* encoding a putative outer membrane protein	Unknown	?	Yes	31	(Gao et al., 2021, Keseler et al., 2017)

Table S1: LysR-type transcription regulators of *E. coli* K-12

For all LTTRs of *E. coli* K-12, as annotated in Ecocyc (Keseler et al., 2017), the genome position, the size in amino acids, and a short description of their function is given. In addition, validated or putative co-effectors that act positively or negatively are listed. H-NS bound indicates whether the regulatory region of the LTTR-encoding gene is H-NS bound according to ChIP-chip data (Uyar et al., 2009), while protein synthesis data are taken from a ribosome profiling study (Li et al., 2014). Further, links to structural data of the LTTRs or of homologues and relevant references are given.

Name	Genome position	Length Aa	Description	co-effector(s)	regulates divergent gene	H-NS bound (Uyar et al., 2009)	protein synthesis rate (Li et al., 2014)	references
YidZ	3,892,765 -> 3,893,724	319	unknown function, maps downstream of *mtdL* (multi drug efflux) encoding a member of the major facilitator superfamily (MFS)	Unknown	No	No	84	(Keseler et al., 2017) (Gao et al., 2021, Justino et al., 2005)
YneJ	1,614,804 -> 1,615,685	293	unknown function, divergent to *sad* encoding succinate semialdehyde dehydrogenases	Unknown	?	No	107	(Keseler et al., 2017)
YnfL	1,668,699 <- 1,669,592	297	unknown function, divergent to *ynfM* encoding an uncharacterized member of the major facilitator superfamily (MFS) transporters that may function as a proton-driven drug efflux system, confers hypersensitivity to acriflavine	Unknown	Yes*	No	60	(Keseler et al., 2017) (Gao et al., 2021, Nishino and Yamaguchi, 2001)

*shown using promoter *lacZ* reporter fusions in this study (see figure 3).

LTTRs shown in bold were analyzed in this study.

(adapted from a table compiled by Karin Schnetz)

Table S2: Oligos		
Oligos	**Sequence***	**Purpose**
OA267	TTTGTATTGAATTATATGTTTTGCGATTTTT	EMSA *leuO* promoter region upstream
OA268	TAATGTTTTTGTTGCGAAAACAATCTAA	EMSA *leuO* promoter region downstream
OA890	TATATTATCACTTACTGGCGGCTCATAC	EMSA *lrhA* promoter region upstream
OA889	TACAAACACGCGCATGTCATAAATAGA	EMSA *lrhA* promoter region downstream
OA892	CAGCAACAAAGTAACGCAGATGACG	EMSA *ynfL* promoter region downstream
OA846	TTGCTTGTCAGTGTCGCTTGC	EMSA *ynfL* promoter region upstream
OA891	TTTCTTTGGCCATTTCGTTCTCATTC	EMSA *yhaJ* promoter region downstream
OA850	AACCATCCGTAGTCTGCTTGTCCA	EMSA *yhaJ* promoter region upstream
OB004	ATGTGACATGCCGCATTATTTATCACT	EMSA *tdcA* promoter region (250 bp)
OB005	ACCAGGTGCTGCGTTTTCG	EMSA *tdcA* promoter region (304 bp)
OA838	ATATGGTCTCAGCCCCTTTTTTGTA	EMSA *tdcA* promoter region (304 bp)
OB006	ATATCAACAAGTTAAAGTATAAAAATCAGCATAAA	EMSA *tdcA* promoter region (250bp)
OA995	CCAGCTCGACCTTAGGTATTGCG	EMSA *yidZ* promoter region
OA996	CAGCAGATTGAGATCAAGCGTGGTG	EMSA *yidZ* promoter region
OB002	GCTAACCCCGGAATAAAATGCC	EMSA *yeiE* promoter region
OB003	TGATGTGCATAGTCGCTTACCACTTA	EMSA *yeiE* promoter region
OA103	ATAGTGTCTGCCTCCAGTGGTCAG	EMSA *ydcI* promoter region
OA999	TAAACAGACTATTTTTTTCCATAAGCGAT	EMSA *ydcI* promoter region
OA840	GCATGGCATAGGTGTTTTTTCG	EMSA *yafC* promoter region
OB001	CGACGTGGCTTTCATTTTTGCT	EMSA *yafC* promoter region
OA862	tcaccatatgAAAGCCACGTCGGAAGAACTC	Cloning $yafC_{His6}$ in pET22b(+)
OA863	tcacctcgagAGCCTCTCTGACAGCTCCTCCG	Cloning $yafC_{His6}$ in pET22b(+)
OA868	tcaccatatgCACATCACCCTCCGGCAGTT	Cloning $yeiE_{His6}$ in pET22b(+)
OA869	tcacctcgagACGCGGCACATTTGCGGGA	Cloning $yeiE_{His6}$ in pET22b(+)
OA877	tcatcatATGGCCAAAGAAAGGGCATTAACG	Cloning $yhaJ_{His6}$ in pET22b(+)
OA878	tcatctcgagTTTTCCGTTAAAAAGTTTGGGAATT	Cloning $yhaJ_{His6}$ in pET22b(+)
OA871	tcaccatatgAATATTGAACTTCGTCATCTGCG	Cloning $ynfL\text{-}DBD_{His6}$ in pET22b(+)
OA873	tcacctcgagCCCCGCTTCACCCTGATGCAGCCTT	Cloning $ynfL\text{-}DBD_{His6}$ in pET22b(+)
OA865	tcaccatATGGAAAAAAATAGTCTGTTTAGTCAGCG	Cloning $ydcI_{His6}$ in pET22b(+)
OA866	tcacctcgagGAACGGCATTGATTTTCGAATAGCG	Cloning $ydcI_{His6}$ in pET22b(+)
OB031	gcaggaattcTATCTCTTCTATTTTGCAACAGGAGCAAAAATGAAAGCC	Cloning *yafC* under P_{tac}
OB032	gcagtctagaTTAAGCCTCTCTGACAGCTCCTCC	Cloning *yafC* under P_{tac}
OB037	gcaggaattcATAATTAATCTTTATAAGTGGTAAGCGACTATGC	Cloning *yeiE* under P_{tac}
OB038	gcagtctagaTTAACGCGGCACATTTGCGGGAT	Cloning *yeiE* under P_{tac}
OB043	gcaggaattcGTATGTTCAAATTTCCTGAATGAGAACGAAATGGCCAAAG	Cloning *yhaJ* under P_{tac}
OB044	gcagtctagaTTATTTTCCGTTAAAAAGTTTGGGAATTTCCCGCAGAC	Cloning *yhaJ* under P_{tac}

Table S2: Oligos		
Oligos	**Sequence***	**Purpose**
OB053	gcaggaattcATTAATAATCTCAAATAAGATGTTTTAAATATGAATATTGAACTTCGTCATCTGCG	Cloning *ynfL* under P_{tac}
OB054	gcagtctagaTTACCTGAGAGCATTCAGCAGATGAATACGAAAG	Cloning *ynfL* under P_{tac}
OB219	gcaggaattcTATCATGGAAGAAAAATATAACCGGAGTAGT	Cloning *ybdO* under P_{tac}
OB220	gcagtctagattaAAAACTCTCATTTAGACGGTCAATAAATC	Cloning *ybdO* under P_{tac}
OB239	gcaggaattcTTCATAGTAGATAATGGCGAGAATGGT	Cloning *yahB* under P_{tac}
OB240	gcagtctagattaATTCTCCCGATGCAGTAAATCCT	Cloning *yahB* under P_{tac}
OB243	gcaggaattcTATTTTTCCATGAATAAAATTGGAGTAAGT	Cloning *ybeF* under P_{tac}
OB244	gcagtctagattaAAATGCATTACGGATAACATCTATTACTCC	Cloning *ybeF* under P_{tac}
OB247	gcaggaattcCCCAAAATCGTCGCTTCCATT	Cloning *ybhD* under P_{tac}
OB248	gcagtctagattaACCTATCTCCTGTAACGCGTGTCT	Cloning *ybhD* under P_{tac}
OB251	gcaggaattcATTCAGGATGAGTTAATGAAGGATTTTC	Cloning *ycaN* under P_{tac}
OB252	gcagtctagattaTTGTGGTCGGTGACGCTGACT	Cloning *ycaN* under P_{tac}
OB231	gcaggaattcCAGGACGCCGCATGCACTG	Cloning *yeeY* under P_{tac}
OB232	gcagtctagattaAATGGCTGATTTTCCGTCATCAG	Cloning *yeeY* under P_{tac}
OB235	gcaggaattcGAGAAATAAAATTGATTTATTGCATCCATT	Cloning *yfeR* under P_{tac}
OB236	gcagtctagattaTATCTGATACAACGGATCCCCTTCC	Cloning *yfeR* under P_{tac}
OB223	gcaggaattcAGATTCGATATTTTCGATAGGTTTGGG	Cloning *yfiE* under P_{tac}
OB224	gcagtctagattaTCCAGCCACAAAACTCTCTTCAA	Cloning *yfiE* under P_{tac}
OB255	gcaggaattcGGTTAAGCGGACAAGGATGCG	Cloning *ygfI* under P_{tac}
OB256	gcagtctagattaATGACTCGTCCCGTCACGATCT	Cloning *ygfI* under P_{tac}
OB227	gcaggaattcTTTCGCTGTGAACAACAATCGGTC	Cloning *yhjC* under P_{tac}
OB228	gcagtctagattaGTCCAGGTACTCTTTCATTACGCC	Cloning *yhjC* under P_{tac}
OB259	gcaggaattcTCTCCGGTGCAAGATTCATATTTATAT	Cloning *yiaU* under P_{tac}
OB260	gcagtctagattaATCTTCGCTTTCCGCCGCT	Cloning *yiaU* under P_{tac}
OB035	gcagtctagaTTCCGACGTGGCTTTCATTTTTGCT	*yafC* promoter *lacZ* fusion
OB036	gcaggtcgacATAGCGCATGGCATAGGTGTTTTT	*yafC* promoter *lacZ* fusion
OB033	gcaggtcgacTTCCGACGTGGCTTTCATTTTTGCT	*yafD* promoter lacZ fusion
OB034	gcagtctagaATAGCGCATGGCATAGGTGTTTTT	*yafD* promoter *lacZ* fusion
OB039	gcaggtcgacAACTTCCAACTGCCGGAGGGT	*yeiH* promoter *lacZ* fusion
OB040	gcagtctagaGGCTAACCCCGGAATAAAATGC	*yeiH* promoter *lacZ* fusion
OB041	gcagtctagaAACTTCCAACTGCCGGAGGGT	*yeiE* promoter *lacZ* fusion
OB042	gcaggtcgacGGCTAACCCCGGAATAAAATGC	*yeiE* promoter *lacZ* fusion
OB045	gcaggtcgacTTCCAGCGTTAATGCCCTTTCTT	*yhaK* promoter *lacZ* fusion
OB046	gcagtctagaTCCGTAGTCTGCTTGTCCACACTG	*yhaK* promoter *lacZ* fusion
OB047	gcagtctagaTTCCAGCGTTAATGCCCTTTCTT	*yhaJ* promoter *lacZ* fusion
OB048	gcaggtcgacTCCGTAGTCTGCTTGTCCACACTG	*yhaJ* promoter *lacZ* fusion
OB055	gcaggtcgacAACAGCAACAAAGTAACGCAGATGA	*ynfM* promoter *lacZ* fusion
OB056	gcagtctagaTTGCTTGTCAGTGTCGCTTGCC	*ynfM* promoter *lacZ* fusion
OB057	gcagtctagaAACAGCAACAAAGTAACGCAGATGA	*ynfL* promoter *lacZ* fusion
OB058	gcaggtcgacTTGCTTGTCAGTGTCGCTTGCC	*ynfL* promoter *lacZ* fusion
OB261	gcagtctagaTAAGTCGAACTTTTTCAAGTCGTAGAGA	*ybdO* promoter *lacZ* fusion
OB262	gcaggtcgacCCCGATCAGAAAACGCTTAATATCATTA	*ybdO* promoter *lacZ* fusion
OB263	gcaggctagcTTTAAGCGTAATAAAACGGCGCA	*yfiE* promoter *lacZ* fusion
OB264	gcaggtcgacCCAAAAAGCACTTAAAAGGGTCGG	*yfiE* promoter *lacZ* fusion
OB265	gcaggtcgacTATTGACTGCCTGTCAAACATGACTAT	*yhjC* promoter *lacZ* fusion

Table S2: Oligos		
Oligos	**Sequence***	**Purpose**
OB266	gcagtctagaGAACAACTGCATTGCGTGAATTTTATC	*yhjC* promoter *lacZ* fusion
OB267	gcagtctagaGAGGATCATCAACACATCCAGCAG	*yeeY* promoter *lacZ* fusion
OB268	gcaggtcgacTATCCCGATCCGCTGGTAATG	*yeeY* promoter *lacZ* fusion
OB269	gcagtctagaCGCTACTGTGACGAAAACCTTTAATTGT	*yfeR* promoter *lacZ* fusion
OB270	gcaggtcgacGAAAGGATCGAGGATACGAAAAAGTT	*yfeR* promoter *lacZ* fusion
OB271	gcagtctagaCAGATTCTCTTCGGTAAAAATTGAGTTC	*yahB* promoter *lacZ* fusion
OB272	gcaggtcgacGAAGCGACGACTAAACGCGAC	*yahB* promoter *lacZ* fusion
OB273	gcagtctagaACAGGGTTCAATTTGATTATTACTATCCAC	*ybeF* promoter *lacZ* fusion
OB274	gcaggtcgacGCGCTACTAAACAATCCGGACTT	*ybeF* promoter *lacZ* fusion
OB275	gcagtctagaGACAAATGCCTTCATACTTGATAATTCAT	*ybhD* promoter *lacZ* fusion
OB276	gcaggtcgacTCGCATCATCACGCAGGGTA	*ybhD* promoter *lacZ* fusion
OB277	gcagtctagaAGTGGCAAAGTCAGACATATTCATCC	*ycaN* promoter *lacZ* fusion
OB278	gcaggtcgacCCATACCAAATAAATACGTTCAGACTGC	*ycaN* promoter *lacZ* fusion
OB279	gcagtctagaTAATAAAATGAAATTGCGCATCTTTTTC	*ygfI* promoter *lacZ* fusion
OB280	gcaggtcgacCAACAACCCGCTAAAACACCTG	*ygfI* promoter *lacZ* fusion
OB281	gcaggtcgacCAACGCAAATAACGCCACAAT	*yiaU* promoter *lacZ* fusion
OB282	gcagtctagaAATTTTTAACTCCCGGTACTTAAGTTGTAA	*yiaU* promoter *lacZ* fusion
OB314	gcagtctagaGCGCTGACTAAACAGACTATTTTTTTCC	*ydcI* promoter *lacZ* fusion
OB315	gcaggtcgacAATCTCATCCGCCGTGATGCT	*ydcI* promoter *lacZ* fusion
OB284	AGAAATCGCCACATTCGGCATGACAACATTGTGAAACCCGGCA TTAGATGgtgtaggctggagctgcttcg	*yhjC* cloning
OB285	TCGTGGTCATTCTTAAAAAGTATAGTCGGTCAGTCCAGGTACTC TTTCATaattccggggatccgtcgacc	*yhjC* cloning
OA816	GCTCCGGAAGGGGATGATTCATTGTTGATTGAAAGGGATATCG AGgactacaaagaccatgacggtgattata	Cloning of $lrhA_{3F\ kan}$
OA817	TTTTTTACACTATTGTCTCAGGAATTATCTATCGTCCGTCGAcata tgaatatcctccttagttcctattcc	Cloning of $lrhA_{3F\ kan}$
OA822	CTGGTGTCTGCGGGAAATTCCCAAACTTTTTAACGGAAAAgacta caaagaccatgacggtgattata	Cloning of $yhaJ_{3F\ kan}$
OA823	GCGATATTTCGCCCCTTTTTATTAACAATAATAAATAcatatgaata tcctccttagttcctattcc	Cloning of $yhaJ_{3F\ kan}$
OA824	CCAGTTGATGATTAACGCTATTCGAAAATCAATGCCGTTC gactacaaagaccatgacggtgattata	Cloning of $ydcI_{3F\ kan}$
OA825	AAATGTTCAAGGAGGGATCGACAGATCCCTTCACCTTcatatgaat atcctccttagttcctattcc	Cloning of $ydcI_{3F\ kan}$
OA826	AGCATGTAAAAACAGCTCCCGGAGGAGCTGTCAGAGAGGCTga ctacaaagaccatgacggtgattata	Cloning of $yafC_{3F\ kan}$
OA827	CCTGGTTAGCCCGGAAGGTCTGGCTCCTGAATGGGAcatatgaat atcctccttagttcctattcc	Cloning of $yafC_{3F\ kan}$
OA828	GGCGCTTTCTGGACTATTGCGATCCCGCAAATGTGCCGCGTgac tacaaagaccatgacggtgattata	Cloning of $yeiE_{3F\ kan}$
OA829	AATCGACACAGCACCAGCATGTTCTTGTACAGCAACcatatgaata tcctccttagttcctattcc	*Cloning of* $yeiE_{3F\ kan}$
OA832	AAGTCCGGCTGCGCGTAACTTTCGTATTCATCTGCTGAATGCTC TCAGGgactacaaagaccatgacggtgattata	Cloning of $ynfL_{3F\ kan}$
OA833	GCATGTGATTAACAGCACATTTTTCGGGCTTTTTCGCTGAAATT TCCCcatatgaatatcctccttagttcctattcc	Cloning of $ynfL_{3F\ kan}$
OB165	GGCATATTAACGGGGATCTATCAGGATTTACTGCATCGGGAGA ATgactacaaagaccatgacggtgattata	Cloning of $yahB_{3F\ kan}$
OB166	GATGTAAACCGCTATTCACAGCGCATCGGGAGGTTGGCAGCGA catatgaatatcctccttagttcctattcc	*Cloning of* $yahB_{3F\ kan}$

Table S2: Oligos		
Oligos	**Sequence***	**Purpose**
OB167	AATTTGGTGTTCTCCCGATTTATTGACCGTCTAAATGAGAGTTT Tgactacaaagaccatgacggtgattata	Cloning of *ybdO*$_{3F\ kan}$
OB168	GTGAATAAATAGATGAAAGTTATAGTAGCTGTCTGTAGTTATAT Acatatgaatatcctccttagttcctattcc	Cloning of *ybdO*$_{3F\ kan}$
OB169	TGATCCCATCCTGCACGGAGTAATAGATGTTATCCGTAATGCAT TTgactacaaagaccatgacggtgattata	Cloning of *ybeF*$_{3F\ kan}$
OB170	CATTTTATCGTTCTGCGCTGTGATCTGGCTGTAACAATACTTTTc atatgaatatcctccttagttcctattcc	Cloning of *ybeF*$_{3F\ kan}$
OB171	CCAATATCGTCGAACGCCAGAGACACGCGTTACAGGAGATAG GTgactacaaagaccatgacggtgattata	Cloning of *ybhD*$_{3F\ kan}$
OB172	ACTAACCGCTGATTTACCCGGCGCAGTCTCTCCTGCGCCGGTGT Acatatgaatatcctccttagttcctattcc	Cloning of *ybhD*$_{3F\ kan}$
OB173	TGTAAAAACGGGTCAGATCTGCCAGAGTCAGCGTCACCGACCA CAAgactacaaagaccatgacggtgattata	*Cloning of ycaN*$_{3F\ kan}$
OB174	TAGGACTGATATTCCCGCTGCTGGCGCGTAAAGCGAATAGTAA ATAAcatatgaatatcctccttagttcctattcc	Cloning of *ycaN*$_{3F\ kan}$
OB175	CGATCGCCGGTGTTTATGCAAAATCTGATGACGGAAAATCAGC CATTgactacaaagaccatgacggtgattata	Cloning of *yeeY*$_{3F\ kan}$
OB176	GGGGGGGAAATGACAAGAGAATGAGAGGCTTGTCAGAATAAT TTTTCTcatatgaatatcctccttagttcctattcc	Cloning of *yeeY*$_{3F\ kan}$
OB177	AATGCGCTAATGGCTGGGCGGGAAGGGGATCCGTTGTATCAG ATAgactacaaagaccatgacggtgattata	Cloning of *yfeR*$_{3F\ kan}$
OB178	TTAACTTCTGATTGATTACTGCCGGGGCTCCCGGCAGTATGAAT TCACTcatatgaatatcctccttagttcctattcc	Cloning of *yfeR*$_{3F\ kan}$
OB179	CGATGCACACTTTTATTCAGTGTGTTGAAGAGAGTTTTGTGGCT GGAgactacaaagaccatgacggtgattata	Cloning of *yfiE*$_{3F\ kan}$
OB180	CGCACTTTGTCGACAAGTTAAATGCCGGGTTTAATCCCGACATT ATTTcatatgaatatcctccttagttcctattcc	Cloning of *yfiE*$_{3F\ kan}$
OB181	ACGATGCTTCAGCATATTGCTGAAGATCGTGACGGGACGAGTC ATgactacaaagaccatgacggtgattata	Cloning of *ygfI*$_{3F\ kan}$
OB182	TCTTCGACTTACACCGTAATTTAATTGCAACCGGCTCGATGCTG GGcatatgaatatcctccttagttcctattcc	Cloning of *ygfI*$_{3F\ kan}$
OB183	TAAACCTGTTTATGCAGTGGCTGGCTGGCGTAATGAAAGAGTA CCTGGACgactacaaagaccatgacggtgattata	Cloning of *yhjC*$_{3F\ kan}$
OB184	CGCTCTTTTGTCCTTCTGTCGTGGTCATTCTTAAAAAGTATAGTC GGcatatgaatatcctccttagttcctattcc	Cloning of *yhjC*$_{3F\ kan}$

*Oligonucleotide sequences are given in 5' to 3' direction. Homologous sequences to the aligned and indicated regions cloned are shown in capital letters and the sites for restriction endonucleases are underlined.

Results Chapter 2:

YdcI, a LysR-type transcription regulator related to stress response and catabolism

Ria Basu ROY and Karin SCHNETZ

Institute for Genetics, Department of Biology, Faculty of Mathematics and Natural Sciences, University of Cologne, Zülpicher Str, 47a, 50674 Cologne, Germany;

Keywords: LysR-type transcription regulator, LTTR, iraP, phosphate stress response

to be submitted

(Ria Basu Roy conducted the investigation and prepared the initial draft, Karin Schnetz conceptualized the study and the manuscript was jointly edited by Karin Schnetz and Ria Basu Roy.)

Abstract

The *ydcI* gene encodes for a protein belonging to family of LysR-type transcription regulators and is found in a wide range of gram-negative bacteria. YdcI has been partially characterized in *S*. Typhimurium and *E. coli* for its regulatory and functional role. Hence, a detailed characterization of YdcI in *E. coli* still remains elusive. In this work, EMSA confirmed the binding of YdcI to 10 loci known from previous ChIP-exo analysis and the DNA-binding motif of YdcI was identified in this work. In addition to the DNA-binding analyses, expression analysis using promoter *lacZ* reporters confirms the direct transcriptional regulation of *iraP* and *mqo* promoters by YdcI. In addition, to regulating the anti-adaptor protein, IraP which is involved in stabilization of RpoS, the other target genes of YdcI are shown to be encoding enzymes for the TCA cycle and for membrane transport. Here we could also show that the target of YdcI, *iraP* is repressed by H-NS and YdcI presumably acts as an H-NS antagonist. Furthermore, the transcription of *ydcI* gene itself was induced by the alarmone, ppGpp under phosphate starvation and there is a dramatic increase in protein levels of YdcI upon entry of the cells from early to mid-exponential phase. This work therefore unravels the different aspects of regulation of *ydcI* gene expression and its potential involvement in regulating metabolic pathways and response to various abiotic stresses encountered by *E. coli*.

Introduction

LysR-type transcription regulators (LTTRs) constitute the largest family of transcription regulators in *E. coli* K-12 (Henikoff et al., 1988, Maddocks and Oyston, 2008). The LTTRs are involved in regulation of many important biological processes like virulence, biofilm formation, stress responses, utilization of various carbon sources and other metabolites, horizontal gene transfer, and others (Henikoff et al., 1988, Karp et al., 2014, Keseler et al., 2017). LTTRs are tetramers. They consist of N-terminal DNA-binding domain (DBD) and a C-terminal effector domain (EBD) as regulatory domain. The two domains are connected via a helical linker helix and a flexible hinge (Maddocks and Oyston, 2008, Momany and Neidle, 2012). LTTRs bind DNA specifically with their two dimeric DBDs even without effector binding to the regulatory EBD. Effector binding rather modulates the DNA-binding pattern at one specific site then being a prerequisite for specific DNA-binding (Baugh et al., 2023). This offers the advantage that specific DNA-binding sites of LTTRs can be identified even if no effector and/or inducing condition is known.

YdcI is one of the LysR-type transcription regulators which is highly conserved and found in a wide range of gram-negative bacteria (Jennings et al., 2011, Solomon et al., 2014). YdcI is not yet well characterized and no effector or condition is known that modulates its protein activity and its gene expression. Recently, chromatin immunoprecipitation coupled with exonuclease treatment and high-throughput sequencing (ChIP-exo) have been used to screen target genes of several transcription regulators including YdcI (Gao et al., 2018). YdcI from *E. coli* K-12 MG1655 and *Salmonella enterica share* 80% identity (Gao et al., 2018). Hence, YdcI was thought to regulate similar cellular functions in both the related species. YdcI in *S.* Typhimurium autorepresses its own expression, enhances biofilm formation and confers acid stress resistance (Jennings et al., 2011, Romiyo and Wilson, 2020). In contrast, in *E. coli*, YdcI has been found to be responsible for thermal stress resistance, reduction in cell viability, conferring antibiotic sensitivity and oxidative stress reponse (Solomon et al., 2014). In addition, YdcI deletion has been reported to cause an increase in frequency of persister cell population in *E. coli* (Hingley-Wilson et al., 2020) and Ribo-seq analysis of *E. coli* showed that YdcI translation at pH 4.4 is enhanced, but the mRNA levels of YdcI remains constant in response to acid stress (Schumacher et al., 2023). Further, in repeated long-term starvation experiments with *E. coli* a combination of *rho-R109H* (transcription termination factor) and loss of function *ydcI* mutants have been identified and it was shown that the combination of these mutations leads to intracellular alkalization (Worthan et al., 2024). As summarized in table 1, YdcI in *S. Typhimurium* and *E. coli* K-12 are associated with different cellular functions. Surprisingly, differential gene expression by RNA-seq analysis in *ydcI* deletion strain compared to WT strain showed none of the gene hits could be related

to the previously identified function of YdcI in *E. coli* and *S.* Typhimurium (Romiyo and Wilson, 2020). However, this study identified *iraP* as one of the targets of YdcI in *E. coli* (Romiyo and Wilson, 2020).

IraP, a small protein anti-adapter, has been shown to inhibit the proteolysis of the stress Sigma factor RpoS through a direct interaction with RssB under phosphate starvation conditions (Bougdour et al., 2006). Additionally, in response to phosphate starvation conditions, direct transcriptional activation of *iraP* is mediated by increased levels of alarmone ppGpp (Bougdour and Gottesman, 2007). The alarmone ppGpp is produced by two known stringent stress response synthetases namely RelA and SpoT. The stringent response mediated by RelA is elicited in response to amino acid starvation and the stringent response by SpoT is elicited in response to fatty acid starvation, iron, phosphate and carbon limitation conditions (Battesti and Bouveret, 2006, Bougdour and Gottesman, 2007, Murray and Bremer, 1996, Spira et al., 1995, Vinella et al., 2005). Only SpoT functions also as ppGpp hydrolase keeping ppGpp levels low at non-stress conditions. Recently, Northern blot anaylsis showed that induction of *iraP* expression on phosphate starvation is presumably mediated by inhibition of SpoT's hydrolase activity (Bougdour and Gottesman, 2007).

The inconsistent findings on YdcI in the literature highlight the need for further studies to resolve its functional role in *E. coli*. In this work, ChIP-exo identified targets of YdcI (Gao et al., 2018) were validated by DNA-binding analyses (EMSA), which confirmed YdcI binding to 10 loci including *iraP,* genes related to the TCA cycle and to membrane transport. In addition, the role of YdcI in ppGpp mediated induction of *iraP* was addressed. The data suggest that YdcI activates *iraP* independent of stress conditions and possibly acting as antagonist of H-NS which represses *iraP*. Further, *ydcI* transcription was found to be induced by ppGpp and YdcI protein levels increased upon entry of the cells from early to mid-exponential phase.

Table 1: Overview of the studies done on YdcI

Phenotype	***Salmonella* spp.**	***E. coli***	**Reference**
Biofilm formation	Detected with overexpression of YdcI	Not studied	(Jennings et al., 2011, Romiyo and Wilson, 2020)
Thermal stress sensitivity	Not detected	Detected in Δ*ydcI* strain	(Solomon et al., 2014)
Autoregulation	Detected with overexpression of YdcI	Not studied	(Jennings et al., 2011)
Acid stress resistance	Detected from reduced cell survival of Δ*ydcI* strain	Growth defect in low and high pH in Δ*ydcI* strain	*Salmonella spp.* (Jennings et al., 2011, Romiyo and Wilson, 2020) *E. coli* (Gao et al., 2018)
Reduced cell viability	Not detected	Detected with overexpression of YdcI	(Solomon et al., 2014)

Table 1: Overview of the studies done on YdcI			
Phenotype	***Salmonella* spp.**	***E. coli***	**Reference**
Antibiotic sensitivity	Not detected	Detected with overexpression of YdcI	(Solomon et al., 2014)
Sensitive to oxidative stress	Not detected	Detected with overexpression of YdcI	*Salmonella spp.* (Jennings et al., 2011, Romiyo and Wilson, 2020) *E. coli* (Solomon et al., 2014)
Increase in persister cells	Not studied	Detected in Δ*ydcI* strain	(Hingley-Wilson et al., 2020)
Long term starvation and alkalization of cytoplasm	not studied	Detected in *rho-R109H* and *ydcI* mutants	(Worthan et al., 2024)
Enhanced translation upon acid stress pH 4.4	not studied	Detected in *E. coli* cells	(Schumacher et al., 2023)

Results

Identification of YdcI target loci

To characterize the functional role of YdcI we performed an RNA-seq analysis comparing the transcriptomes of the *E. coli* K-12 strain U4 (Δ*lacZ*) with its isogenic Δ*ydcI* mutant U505 (Table S2) grown to mid-exponential phase in tryptone medium. The RNA-seq data were then compared with ChIP-exo data (Gao et al., 2018), after re-evaluation of the latter and visualization using the Integrative Genomics Viewer (IGV) (Fig. 1 and Table 2). In total 4 loci were identified, for which significant changes both in transcript levels and DNA-binding by YdcI were detected (Table 2). These include 2 loci presumably activated by YdcI, which are *iraP* and *yhdV-yhdWXYZ*. IraP is a phosphate stress induced anti-adaptor of RssB and prevents RssB-mediated targeting of sigma factor RpoS for protein degradation (Bougdour et al., 2006). The *iraP* gene has been identified before by RNA-seq analysis of *ydcI* mutant compared to wild-type *E. coli* K-12 (derivate DH10B) (Romiyo and Wilson, 2020). The *yhdV-yhdWXYZ* locus putatively codes for a lipoprotein (YhdV) and presumptive charged/polar amino acid ABC-transporter (YhdWXYZ); however, *yhdW* carries a frameshift in codon 23 and is thus a pseudogene in *E. coli* K-12. In addition, 2 further loci are presumably repressed by YdcI. These are *mqo* and the *yobA-yebZ-yebY* operon, encoding the citrate cycle enzymes malate:quinone oxidoreductase (Mqo) and ABC transporter YobA-YebZ-YebY implicated in delivering copper to the copper-dependent NADH:quinone oxidoreductase II encoded by *ndh* (Hadley et al., 2022). Additional loci identified by ChIP-exo (Gao et al., 2018) (see Fig. 1) are *gltA-sdhCDAB* which encodes citrate synthase (GltA) and succinate:quinone oxidoreductase (SdhCDAB), i.e. key enzymes of the tricarboxylic acid cycle (EcoCyc), as well as *nhaA*, *yobF*, *lldPRD*, *yicG*, *gltP* and *dtpA* (Fig. 1). For the remaining YdcI bound loci identified by Gao et al. (2018) including *tomB*, *ybhI*, *wrbA*, *trkG* and *gutM*/*srlM* only a low enrichment was apparent in the IGV analysis (Fig. 1). Taken together, the results of comparative analysis of RNA-seq and ChIP-exo data suggest that 11 out of the 16 predicted target loci of YdcI (Gao et al., 2018) are YdcI-bound of which only 4 loci are presumably regulated by YdcI.

Table 2: ChIP-exo and RNA-seq data for YdcI

				RNA-seq log2 fold change[3]			
Gene(s)	**Description**	**IGV peaks[1]**	**EMSA[2]**	**WT/ ΔydcI**	**P-adj**	**WT/ ΔyafC**	**P-adj**
nhaA	Na^+:H^+ antiporter	High	++	-0.19	0.614	-0.32	0.128
iraP	anti-RssB, σS stabilization, phosphate and other stress	High	++	2.58	0.000	-0.36	0.266
ddlA	D-alanine-D-alanine ligase A			0.11	0.777	0.02	0.951
gltA	TCA cycle, citrate synthase (GltA)	High	++	-0.65	0.052	-0.44	0.028
sdhC	succinate:quinone oxidoreductase			0.10	0.827	0.44	0.020
sdhD	(SdhCDAB)			-0.09	0.842	0.27	0.363
sdhA				-0.21	0.523	0.13	0.721
sdhB				-0.31	0.293	0.02	0.956
yobF	role in envelope and acid stress	High	+	0.08	0.803	-0.31	0.060
yebY	YobA, YebZY: delivering copper to	High	++	-0.63	0.007	0.02	0.962
yebZ	copper-dependent NADH:quinone			-0.84	0.004	-0.04	0.910
yobA	oxidoreductase II (NDH-2)			-1.04	0.000	-0.30	0.286
hole	HolE: DNA polymerase III subunit Θ			0.35	0.453	0.91	0.000
Mqo	TCA cycle, malate:quinone oxidoreductase	High	++	-1.24	0.000	0.55	0.038
yhdV	YhdV: lipoprotein	High	++	0.01	0.981	-0.13	0.757
yhdW	YhdWXYZ: putative polar/charged amino						
yhdX	acid ABC transporter (frameshift in 23rd			0.90	0.024	0.25	0.504
yhdY	codon of *yhdW*)			0.88	0.032	0.16	0.721
yhdZ				0.13	0.784	-0.23	0.482
lldP	LldPRD: lactate/glycolate metabolism	High	++	-0.84	0.010	-0.56	0.071
lldR				-0.90	0.024	-0.69	0.009
lldD				-1.07	0.010	-1.06	0.000
yicG	Gly transporter, inner membrane protein	High	+	-0.06	0.908	0.66	0.000
gltP	glutamate/aspartate: H+ symporter	High	+	-1.02	0.002	-0.97	0.000
dtpA (tppB)	dipeptide/tripeptide: H+ symporter	High	-	-0.02	0.955	0.26	0.340
Tomb	protein that modulates Hha toxicity	Low	N.D	-0.35	0.468	-0.64	0.000
ybhI	putative tricarboxylate transporter	Low	N.D	0.08	0.920	0.35	0.599
wrbA	NAD(P)H:quinone oxidoreductase	Low	-	0.48	0.343	-0.03	0.948
trkG	Na^+-dependent K^+ transporter	Low	N.D	0.23	0.501	0.42	0.033
gutM (srlM)	transport and utilization of glucitol	Low	N.D	-0.23	0.669	-0.04	0.945

[1]Genes identified by ChIP-exo of YdcI (Gao et al., 2018).

[2]in EMSA ++ (strong binding), + (weak binding), - (no binding), N.D. (not determined).

[3]The log2 fold change values shown in grey are statistically non-significant (P-adj > 0.05), WT (wild-type).

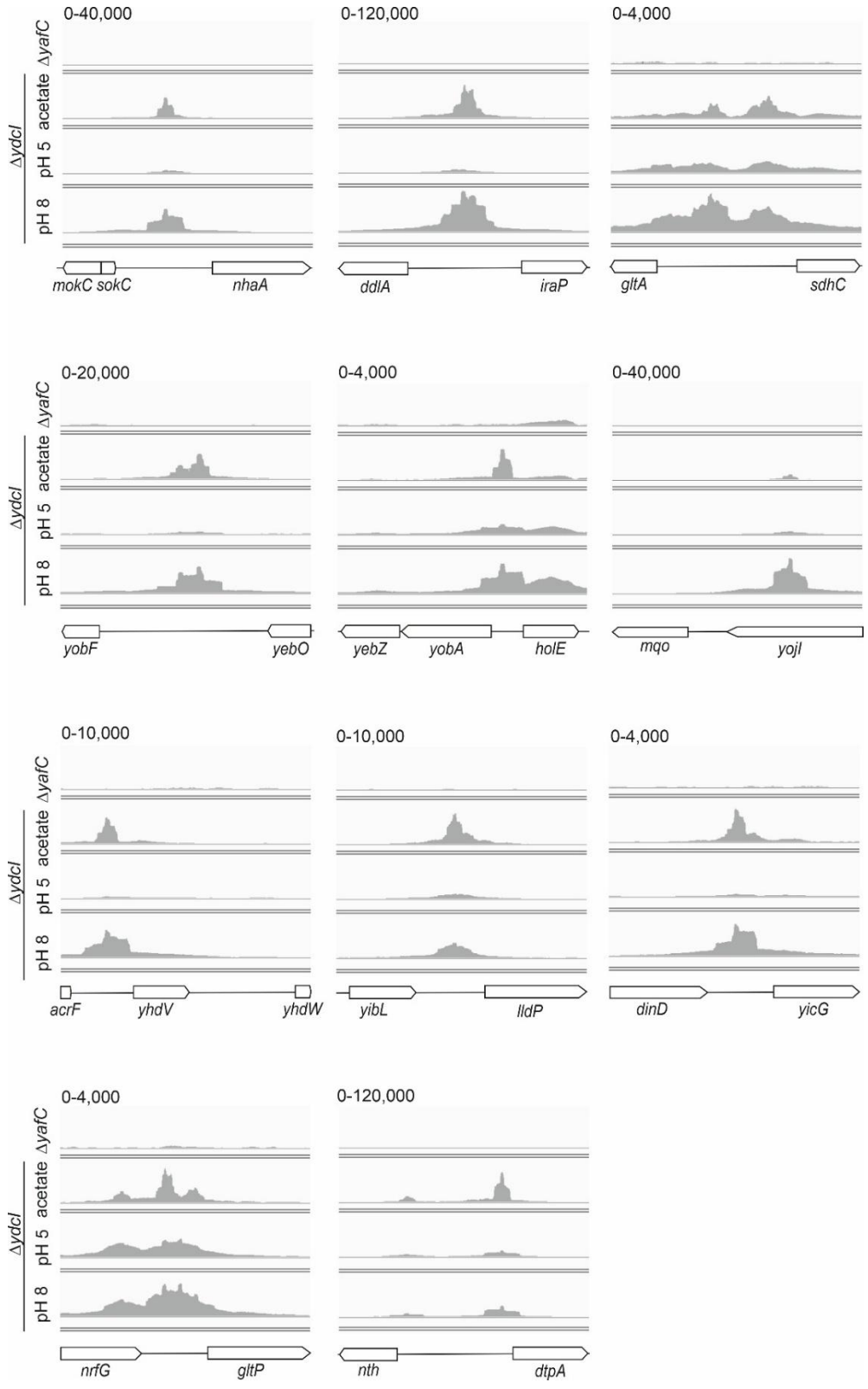
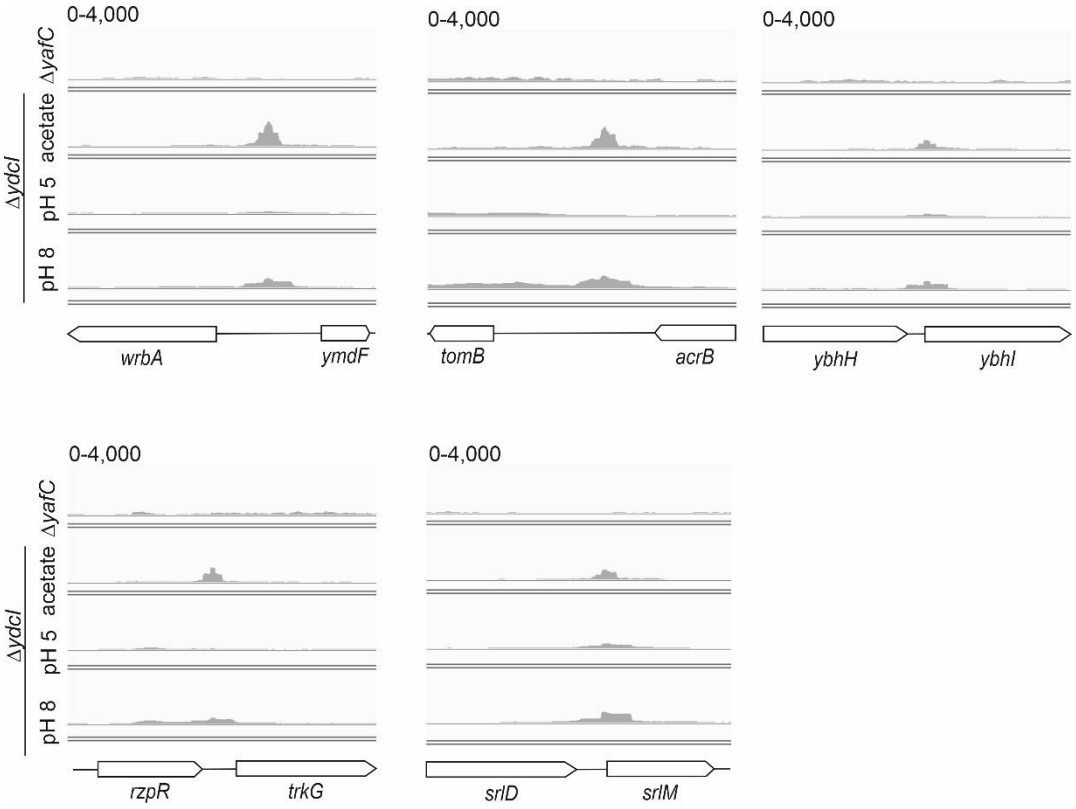


Figure 1 Visualisation of ChIP-exo data for YdcI targets in IGV browser. To visualize ChIP-exo data (Gao et al., 2018) whole genome BAM files from for *E. coli* K-12 *yafC-3xFLAG* (grown in acetate medium, and taken as control) and *ydcI-3xFLAG* (grown in acetate medium, and M9 minimal media, with pH 5 and pH 8) were used. The reads are shown as the grey graphs and are depicted for all the ChIP-exo identified target of YdcI (Gao et al., 2018). The data range of the reads is variable and depicted at the top of each graph, the window size of each graph is 1000 bp. Figure continued on next page.



Continuation of Figure 1

In vitro DNA-binding of presumptive target loci by YdcI and identification of an YdcI DNA-binding motif

To validate DNA-binding of YdcI to presumptive target loci, *in vitro* electrophoretic mobility shift assays (EMSA) were performed using FAM-labelled fragments and purified $YdcI_{His6}$ protein (Fig. 2 and Fig. 3). To assess the specificity of DNA-binding to individual FAM-labelled fragments, competitive EMSA assays with non-labelled, specific as well as non-labelled non-specific competitor fragments were also performed. EMSA were conducted using 60 bp overlapping FAM-fragments mapping in the intergenic region between *iraP* and *ddlA*, which were incubated with increasing concentration of $YdcI_{His6.}$ The results show specific binding of $YdcI_{His6}$ at the region encompassing *iraP* fragment B and non-specific binding of $YdcI_{His6}$ to fragment A and fragment C (Fig. 2B). Hence, for all subsequent assays, *iraP* fragment A and C were selected as the non-specific competitor fragments and fragment B was taken as a specific competitor. In order to further narrow down on the region upstream of *iraP* where the YdcI binding site potentially maps, several sub-fragments from the region covered by the *iraP* fragment B namely fragments B1, B2 and B3 were tested in further EMSA. The results show more specific binding of $YdcI_{His6}$ to fragment B1 and B2 as compared to fragment B3 (Fig. 2B). This suggests the presence of YdcI binding site in the region overlapping between *iraP* fragment B1 and B2.

Similarly, competitive EMSA was performed with 60 bp overlapping fragments mapping in the ChIP-exo predicted binding site for other potential target genes of YdcI. The results of the competitive EMSA shows binding of $YdcI_{His6}$ to region upstream of *nhaA*, *yobF*, *mqo*, *yhdV*, *lldP*, *yicG*, *gltP* and in the intergenic region between *gltA-sdhC* and *yobA-holE* (Fig. 3). Interestingly, the YdcI-DNA binding in correlation with ChIP-exo predicted binding site (Gao et al., 2018) for *mqo* maps within the gene *yojI* (Fig. 3E) which is located upstream of *mqo* indicating that the *mqo* promoter maps within *yojI* in this region (Fig. 3).

In addition, DNA-binding by $YdcI_{His6}$ was analyzed by EMSA for all other loci with a statistically significant log2 fold change of transcript levels greater than (+/-) 1.5 (Chapter 1, section 1.4). Further, the binding of YdcI to two target loci (*wrbA*, *dtpA*) that were identified by ChIP-exo (Gao et al., 2018) was also tested with EMSA. The results show that no binding was detected for the loci which only exhibit a change of transcript levels (Fig. S1). For the target gene, *wrbA* which has a weak peak in the IGV-analysis, EMSA shows no binding with YdcI (Fig. S1) suggesting only the targets with significant reads in ChIP-exo bind to YdcI. However, for the target locus *dtpA*, which has strong reads in IGV analysis no binding with YdcI could be detected (Fig. S1). These results could be due to the choice of the overlapping fragments in the *dtpA* region upstream with each fragment covering maximal one IGV

peak which might not contain the complete YdcI binding site. It is also possible that *in vivo* additional DNA-binding proteins support or are required for YdcI's binding to the *dtpA* regulatory region.

Taken together, the results of the competitive EMSA reveal that only those loci (except *dtpA*) which have significant peaks in the IGV analysis are bound *in vitro* by YdcI.

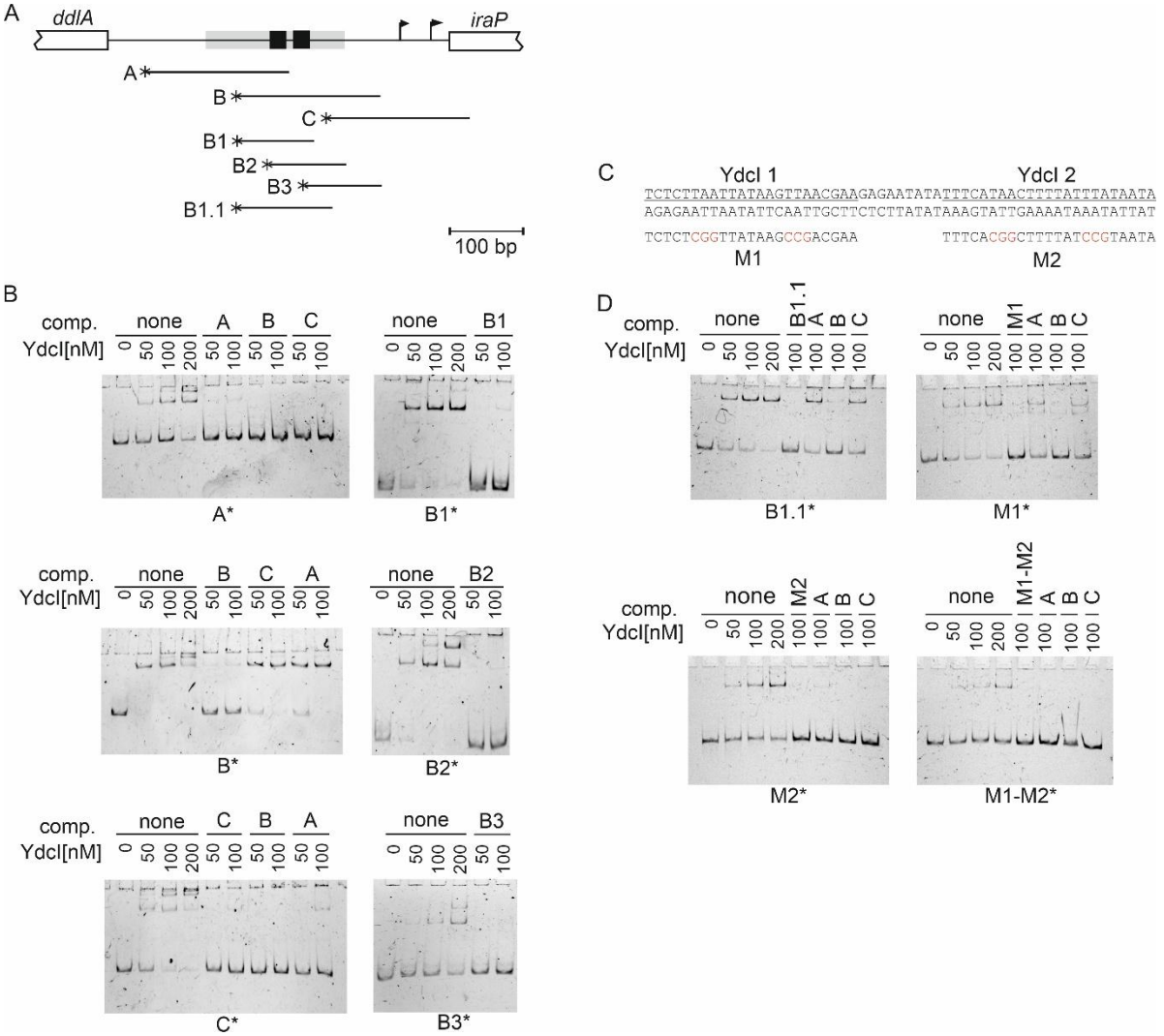


Figure 2 YdcI specifically binds *iraP* DNA fragments. For DNA-binding analyses of $YdcI_{His6}$, electrophoretic mobility shift assays (EMSAs) were performed using *iraP* upstream DNA fragments. (A) Schematic of the *iraP* upstream region. Fragments "A" to "B3" are depicted as solid lines and were amplified by PCR using oligonucleotides listed in Table S4. 5' 6-FAM labeled fragments are marked with an asterisk. Predicted binding site of YdcI (Fig. 1) is depicted as grey boxes and the motifs identified from MEME-suite (Fig. 4) are indicated with the black bars. (C) Sequence of the YdcI binding motifs as identified from MEME-suite, upstream of *iraP* with YdcI 1 as "YdcI binding site 1" and YdcI 2 as the "YdcI binding site 2". Electrophoretic mobility shift assays with increasing concentrations of purified $YdcI_{His6}$ using the (B) *iraP* fragments A, B, C, B1, B2 and B3 and (D) *iraP* fragment B1.1, mutants M1, M2 and M1-M2. Fragment A and C are used as non-specific competitors and fragment B is used as a specific competitor for the EMSA. Scale bar is applicable to the intergenic region only.

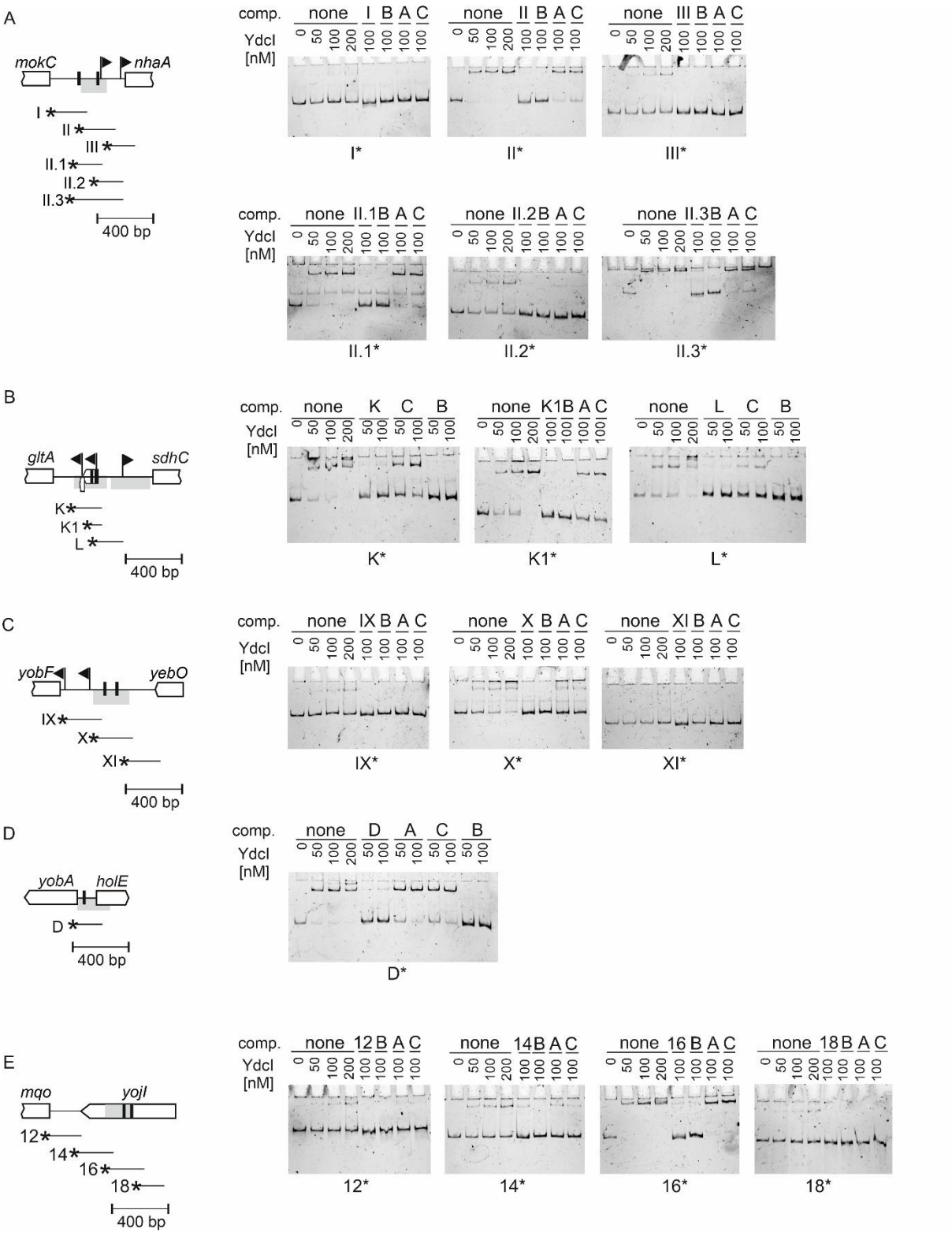


Figure 3 continued next page

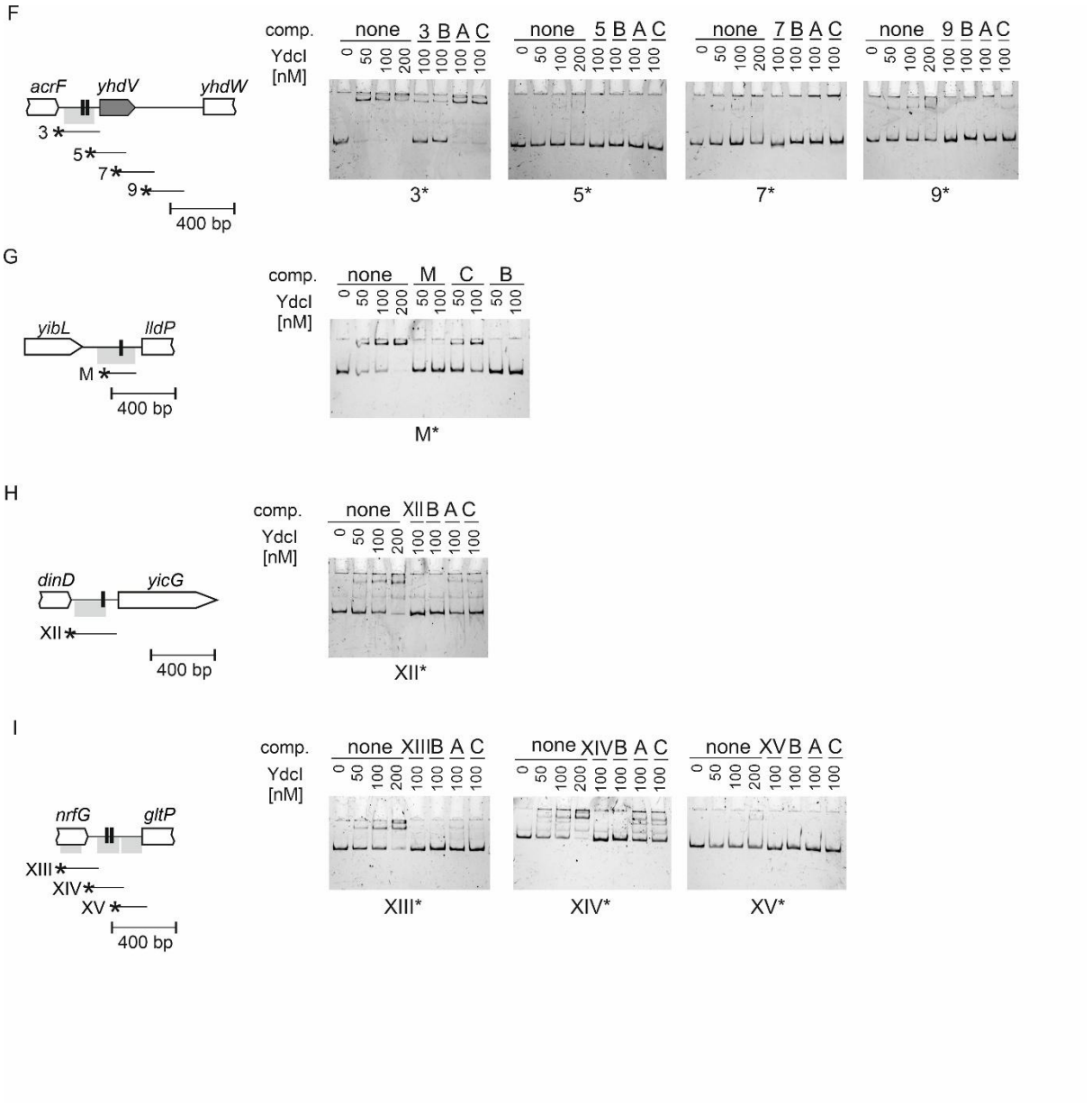


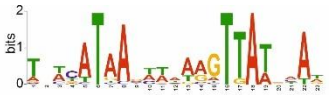
Figure 3 YdcI specifically binds to ChIP-exo identified targets of YdcI. For DNA-binding analyses of $YdcI_{His6}$, electrophoretic mobility shift assays (EMSAs) were performed using (A) *nhaA*, (B) *gltA-sdhC,* (C) *yobF*, (D) *yobA-holE*, (E) *mqo*, (F) *yhdV,* (G) *lldP,* (H) *yicG* and (I) *gltP* promoter DNA fragments. Schematic of the regions upstream of (A) *nhaA,* (C) *yobF,* (E) *mqo*, the intergenic region of (B) *gltA-sdhC* and (D) *yobA-holE,* (F) *yhdV,* (G) *lldP,* (H) *yicG* and (I) *gltP*. All the fragments are depicted as solid lines and were amplified by PCR using oligonucleotides listed in Table S4. 5' 6-FAM labeled fragments are marked with an asterisk. Binding sites of YdcI identified by ChIP-exo (Fig. 1) are depicted as grey boxes and the motifs identified from MEME-suite (Fig. 4) are indicated with black bars. Electrophoretic mobility shift assays with increasing concentrations of purified $YdcI_{His6}$ using the (A) *nhaA* fragments I, II and III*,* (B) *gltA-sdhC* fragments K, K1 and L, (C) *yobF* fragments IX, X and XI, (D) *yobA* fragment D, (E) *mqo* fragments 12, 14, 16 and 18, (F) *yhdV* fragments 3, 5, 7 and 9, (G) *lldP* fragment M, (H) *yicG* fragment XII and (I) *gltP* fragments XIII, XIV and XV. Fragment A and C are used as non-specific competitors and fragment B is used as a specific competitor for the EMSA. Scale bar is applicable to the intergenic region only.

YdcI DNA-binding motif

Based on the EMSA data, we analyzed the bound DNA-sequences for a common DNA-binding motif of YdcI using MEME-suite (Bailey and Elkan, 1994). In this analysis we considered that LysR-type transcription regulators are usually tetramers with two distinct dimeric DNA-binding domains that bind to two adjacent DNA-sites and that the distance between these two DNA-binding contact sites of each tetramer is variable. Hence, the DNA sequences of the ten loci (Table S1) showing strong binding to YdcI were used to define a sequence logo using MEME-suite (Fig. 4). The identified binding motif is AT-rich and includes the T-N11-A pattern which is conserved for LTTRs binding sites. Additionally, the YdcI motif found from this work is quite similar to the identified YdcI motif (Gao et al., 2018).

Next, in order to validate the YdcI DNA-binding motif, mutations were introduced into the two YdcI DNA binding motifs upstream of *iraP*. In fragments M1, M2, YdcI motif 1 and motif 2, respectively were mutated, and in fragment M1-M2 both motifs were mutated (Fig. 2C). Then these fragments were used for competitive EMSA. Mutation of either motif reduced specific binding by YdcI with *iraP* fragment M2 showing a more pronounced effect (Fig. 2D). When both the YdcI binding motifs were mutated (fragment M1-M2), YdcI binding was completely abolished (Fig. 2D).

Taken together, the YdcI DNA-binding motifs identified by MEME-suite were validated by competitive EMSA.



nhaA	TTATT**TAA**CTTT**AA**T**TTAT**CTCT	17151 -	17173
	CACCT**TAA**GCTATGA**TTAT**CT**A**G	17288 -	17310
iraP	**T**CTCT**TAA**TTAT**AAGTTA**ACG**AA**	401142 -	401164
	TTTC**ATAA**CTTTT**A**T**TTAT**AATA	401174 -	401196
gltA	**T**CACC**TAA**GATTAAC**TTAT**GT**A**A	754754 -	754732
	TATA**AT**GCGTAA**AAGTTAT**GA**AG**	754785 -	754763
yobF	**T**TAT**AT**C**A**CTTTTTC**T**C**AT**AA**AA**	1907920 -	1907898
	TAAA**ATAA**GAAA**AA**T**T**G**AT**GA**AT**	1907983 -	1908005
yobA	AGAC**AT**T**A**GACA**AGGTTAT**GCTC	1925026 -	1925004
mqo	**T**TTG**ATAA**CTTC**A**G**GTTA**ACA**AT**	2307261 -	2307283
	TTGCT**TAA**CTCA**A**GC**TTAT**GAGC	2307313 -	2307291
yhdV	**T**CAC**ATAA**TTTGTCA**TTAT**TC**AT**	3418269 -	3418291
	TTTT**AT**C**A**CCGG**AGGTTAT**AA**AT**	3418322 -	3418300
lldP	ATAC**ATAA**CATTT**AGTTA**ACC**AT**	3777261 -	3777283
yicG	**T**CAC**ATAA**CTAC**AAGTTAT**CA**AT**	3818770 -	3817792
gltP	**T**TCC**ATAA**CAAATG**GTTAT**TC**AT**	4294251 -	4294273
	TGCC**ATAA**ACTCGC**GTTAT**TATG	4294280 -	4294302

Figure 4 YdcI DNA-binding motif. Top: Sequence logo for the YdcI binding site. The logo was generated by MEME-SUITE (Grant et al., 2011) using the EMSA DNA sequences encompassing the mapped YdcI binding sites for all the listed loci. Middle: Sequences of the YdcI binding sites and their genome position in the MG1655 reference genome sequence NC_000913.3. Upper case letters in bold used for highly conserved bases, and the upper-case letters not in bold for ambiguous conserved bases.

Role of YdcI in the phosphate stress control of *iraP*

Previous work showed that IraP stabilizes σ^S under phosphate starvation conditions by mediating a direct interaction with RssB and preventing σ^S proteolysis (Bougdour et al., 2006). Additionally, a direct transcriptional activation of *iraP* is observed under phosphate starvation conditions by increased levels of alarmone ppGpp (Bougdour and Gottesman, 2007). Previous studies with RNA-seq transcriptome analysis in *E. coli* identified the small molecule anti-adaptor, *iraP* as one of the targets of YdcI (Romiyo and Wilson, 2020). In addition, the results of previous ChIP-exo data (Gao et al., 2018) and comparative RNA-seq analysis in this work, suggest a possible activation of *iraP* by YdcI (Table 1) which needs to be experimentally evaluated. Moreover, the binding profile of H-NS to the *E. coli* genome shows that H-NS binds *iraP* regulatory region (Uyar et al., 2009). This suggests a possible role of YdcI in phosphate stress response of *iraP* and additionally for YdcI acting as an H-NS antagonist for *iraP*.

In order to study the regulation of *iraP* promoter (P_{iraP}) by YdcI and nucleoid associated global repressor H-NS, a plasmid carrying a P_{iraP} *lacZ* transcriptional reporter fusion was used. The expression of P_{iraP} *lacZ* is at least 2.5-fold decreased in the *ydcI* mutant as compared to the wild-type strain in tryptone medium, showing the activation of *iraP* by YdcI (Fig. 5B). The expression level of the P_{iraP} *lacZ* fusion was low in the wild-type strain and at least 3-fold higher in the isogenic *hns* mutant. The deletion of *stpA* encoding the H-NS homologue, StpA did not have any statistically significant difference with the wild-type strain, but in the *hns stpA* double mutant the expression of P_{iraP} *lacZ* increased further in both the *hns, stpA* double mutant and *hns, stpA, ydcI* triple mutant (Fig. 5B). Moreover, there was no statistically significant difference in expression of P_{iraP} *lacZ* between the *hns* and *hns ydcI* mutants and also between the *hns, stpA* and *hns*, *stpA, ydcI* mutants (Fig. 5B). This suggests that the activation of *iraP* promoter by YdcI is mediated by YdcI acting as an H-NS antagonist.

To study the role of YdcI in phosphate stress regulation of *iraP*, phosphate starvation growth condition was used. The alkaline phosphatase, *phoA* has been shown to be upregulated upon phosphate stress by transcriptome sequencing (Gardner and McCleary, 2019) and is also regulated by the PhoBR two-component system (the two-component regulatory system of the Pho regulon) (Makino et al., 1986). Hence, P_{phoA} *lacZ* fusion was used at first to test the phosphate stress growth conditions. The significant activation of the P_{phoA} *lacZ* under phosphate stress and absence of this activation in the *phoBR* deletion strain (Fig. 5D) verified the use of this growth condition for all subsequent assays. Further, P_{iraP} *lacZ* expression in the wild-type and *ydcI* mutant showed activation of P_{iraP} in both low and high phosphate (Pi) MOPS glucose medium (0.1 mM and 1.32 mM Pi) similar to the activation by YdcI in tryptone medium (Fig. 5C). To further study if the activation of *iraP* promoter in response to phosphate

starvation was dependent on PhoBR, the expression analysis of P_{iraP} *lacZ* transcriptional reporter fusion in the PhoBR mutant strain was studied as well in low and high Pi MOPS glucose medium. The results clearly demonstrate that phosphate stress induced the *iraP* promoter in both wild-type and *ydcI* mutant strain and is independent of PhoBR, while the activation of P_{iraP} by YdcI is not affected by phosphate starvation (Fig. 5C).

In adition, a P_{ddlA} *lacZ* transcriptional reporter fusion was analyzed, since *ddlA* maps divergent to *iraP*, *ddlA* was also analyzed. The data show phosphate stress induction of *ddlA*, a 2-fold repression of *ddlA* by PhoBR at low phophate conditions and no regulation by YdcI (Fig. 5E). Taken together, the results show *iraP* promoter is repressed by H-NS and H-NS/StpA (in tryptone medium) and is activated by YdcI independent of PhoBR (under phosphate starvation conditions).

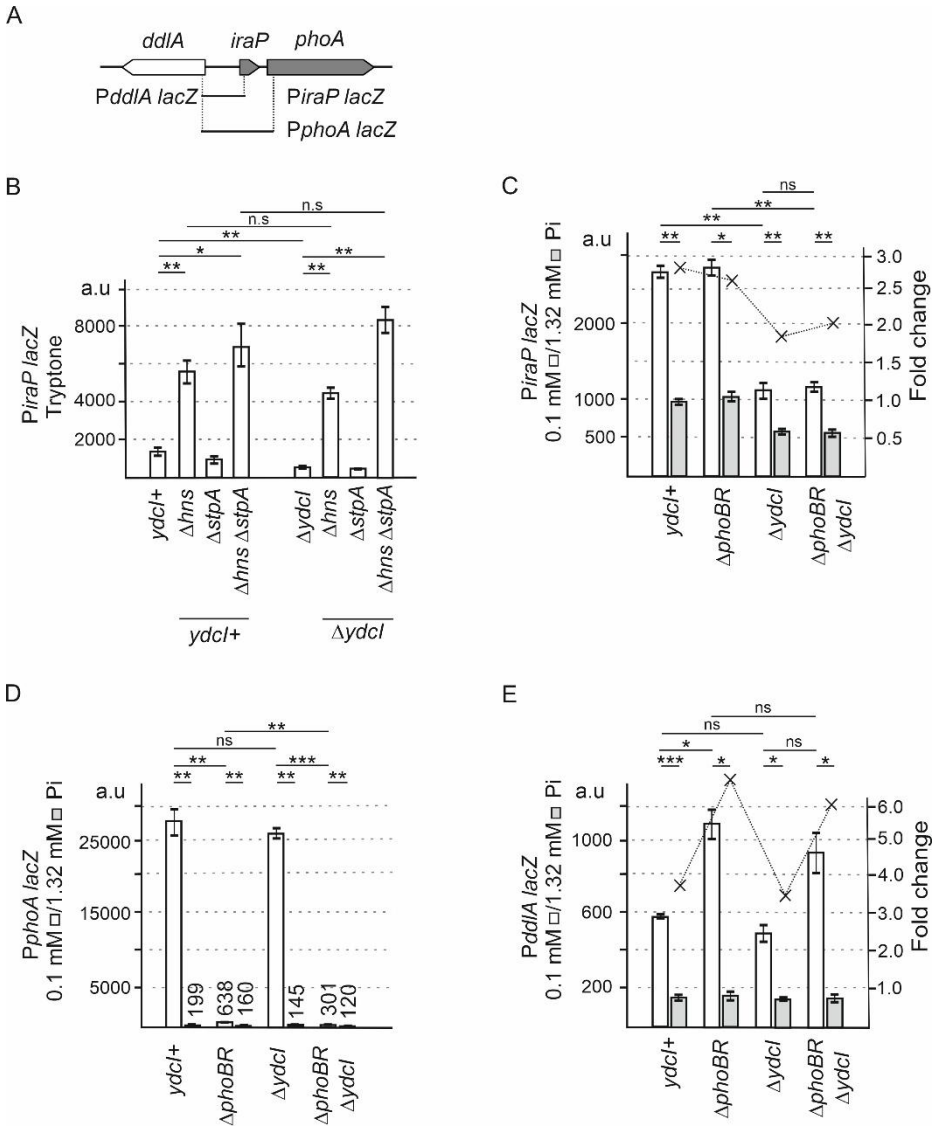


Figure 5 (previous page) YdcI and H-NS/StpA regulates *iraP* promoter. A) Schematic showing *ddlA*, *iraP* and *phoA* gene and the region used for constructing promoter *lacZ* reporter fusion is depicted. The regulation of P_{iraP} in response to deletion of B) H-NS/StpA and C) PhoBR was determined. Expression level from P_{iraP} lacZ (pKERBR091) fusion was determined in wild-type (U4), Δ*hns* (U261), Δ*stpA* (U559), Δ*hns* Δ*stpA* (U563), Δ*phoBR* (U624), Δ*ydcI* (U505), Δ*ydcI* Δ*hns* (U558), Δ*ydcI* Δ*stpA* (U560), Δ*ydcI* Δ*hns* Δ*stpA* (U564) and Δ*ydcI* Δ*phoBR* (U625) strains. Regulation by PhoBR is determined for D) P_{phoA} *lacZ* and E) P_{ddlA} *lacZ* reporter fusions with plasmid expressing P_{phoA} lacZ (pKERBR104) and P_{ddlA} lacZ (pKERBR117). Cultures for β-galactosidase assays were inoculated from overnight cultures and grown to OD_{600}~0.8 in tryptone (B), MOPS glu low (0.1 mM Pi) and high (1.32 mM Pi) medium (C, D, E). Washing steps were performed (D) or not performed (C, E) (see materials and methods). Mean values from at least three independent replicates are shown as bars, and error bars indicate standard deviations. T-test was performed and the values are shown as ****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05, ns=non-significant.

A study has shown that the alarmone, ppGpp accumulates in *E. coli* when exposed to phosphate starvation (Spira et al., 1995). And recently, Northern blot anaylsis showed that ppGpp is responsible for *iraP* induction and out of the two, phosphate stress response synthetases (RelA and SpoT), sensing of phosphate stress may occur by inhibition of hydrolase activity of SpoT (Bougdour and Gottesman, 2007).

In order to study whether RelA or SpoT plays a role in the activation of P_{iraP}, the expression analysis of P_{iraP} *lacZ* transcriptional reporter fusion in the *relA*, *relA spoT* and $spoT_{E319Q}$, *relA* $spoT_{E319Q}$ mutants were studied in low and high Pi MOPS glucose medium. The mutant $SpoT_{E319Q}$ used here retains the hydrolytic activity but is defective in ppGpp synthesis. There was no significant difference in fold change for expression of P_{iraP} *lacZ* between the wild-type and *relA* mutant and also between the *ydcI* and *relA, ydcI* mutants (Fig. 6A). Similarly, the fold induction of P_{iraP} *lacZ* expression between the wild-type and $spoT_{E319Q}$ mutant and between the *ydcI* and $spoT_{E319Q}$ *ydcI* mutants remains the same (Fig. 6B). As expected, P_{iraP} induction upon phosphate stress was completely abolished in all the *relA spoT* and *relA* $spoT_{E319Q}$ mutants (Fig. 6). The difference between the fold induction of *iraP* promoter expression in *relA* mutants (same strain) for P_{iraP} *lacZ*-cm and P_{iraP} *lacZ*-tet reporter system (Fig. 6A and 6B) could be attributed to the two different antibiotic resistance reporter plasmids for the expression analysis but this requires further studies for clarification. Taken together, the result confirms ppGpp induced activation of *iraP* expression as shown before (Bougdour and Gottesman, 2007) being dependent on either RelA or SpoT.

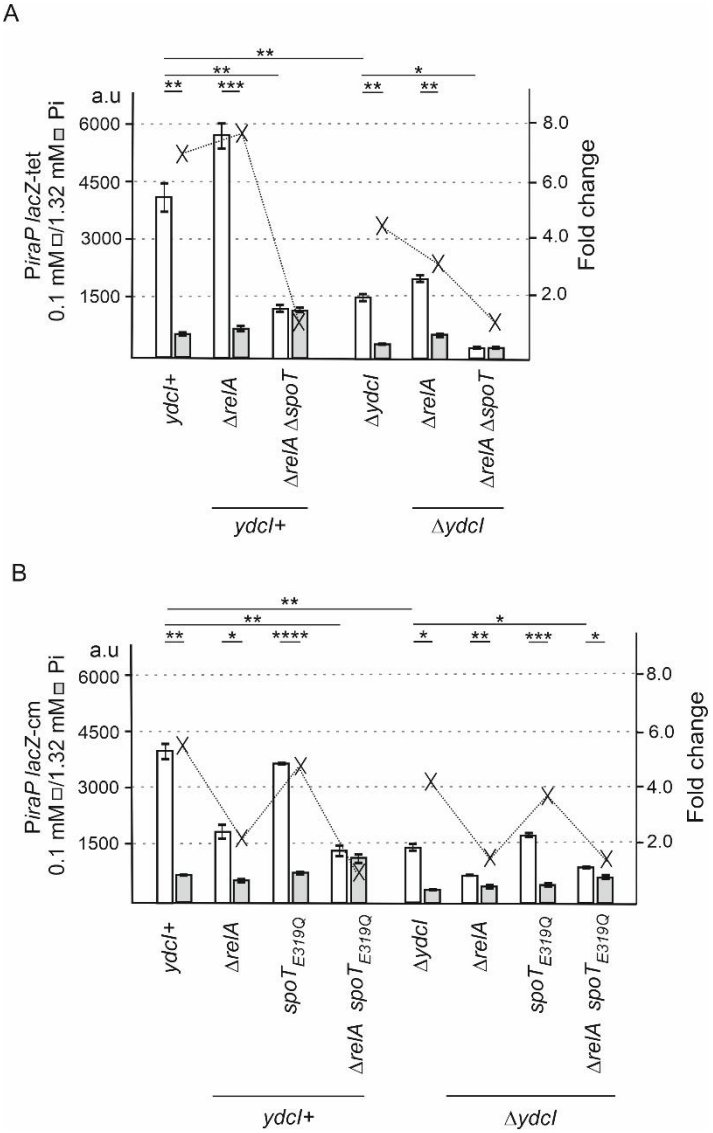


Figure 6 ppGpp regulates *iraP* promoter. The regulation of P_{iraP} by ppGpp were determined using wild-type (U4), Δ*relA* (U571), Δ*relA* Δ*spoT* (U578), Δ*ydcI* (U505), Δ*ydcI* Δ*relA* (U572), Δ*ydcI* Δ*relA* Δ*spoT* (U579) in trans with plasmid expressing A) P_{iraP} *lacZ*-tet (pKERBR115), with wild-type (U4), Δ*relA* (U571), $spoT_{E319Q}$ (U613), Δ*relA* $spoT_{E319Q}$ (U615), Δ*ydcI* (U505), Δ*ydcI* Δ*relA* (U572), Δ*ydcI* $spoT_{E319Q}$ (U614), Δ*ydcI* Δ*relA* $spoT_{E319Q}$ (U616) strains in trans with plasmid expressing B) P_{iraP} *lacZ*-cm (pKERBR091). Cultures for β-galactosidase assays were inoculated from MOPS glu high (1.32 mM Pi) overnight cultures and grown in MOPS glu low (0.1 mM Pi) and high (1.32 mM Pi) medium to an OD_{600} of 0.5-0.6. Mean values from at least three independent replicates are shown as bars, and error bars indicate standard deviations. T-test was performed and the statistically significant values are shown as ****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05.

Regulation of *ydcI* transcription

In a similar line of experiments as to studying the regulation of P_{iraP}, to study the regulation of *ydcI* promoter by H-NS/StpA and PhoBR, a plasmid carrying P_{ydcI} *lacZ* transcriptional reporter fusion was used. At first autoregulation by YdcI and regulation by H-NS and the H-NS homologue StpA was analyzed. The expression of P_{ydcI} *lacZ* did not show any statistically significant change in all mutant strains as compared to the wild-type (Fig. 7A). This suggests that no autoregulation exists for YdcI and also H-NS or H-NS/StpA does not repress the *ydcI* promoter. P_{ydcI} *lacZ* expression analysis is increased in low Pi MOPS glucose medium (0.1 mM) as compared to high Pi MOPS glucose medium (1.32 mM) in wild-type and also in *ydcI*, *phoBR*, *phoBR ydcI* mutant strains. This demonstrates the activation of *ydcI* promoter in response to phosphate stress. In addition, there is no significant change in the fold activation of *ydcI* promoter between the wild-type and *phoBR* mutant and between the *phoBR* and *phoBR ydcI* mutant strains (Fig. 7A and 7B). This shows there is no regulation of *ydcI* promoter by PhoBR and suggests that the *ydcI* promoter is induced upon phosphate stress independent of PhoBR.

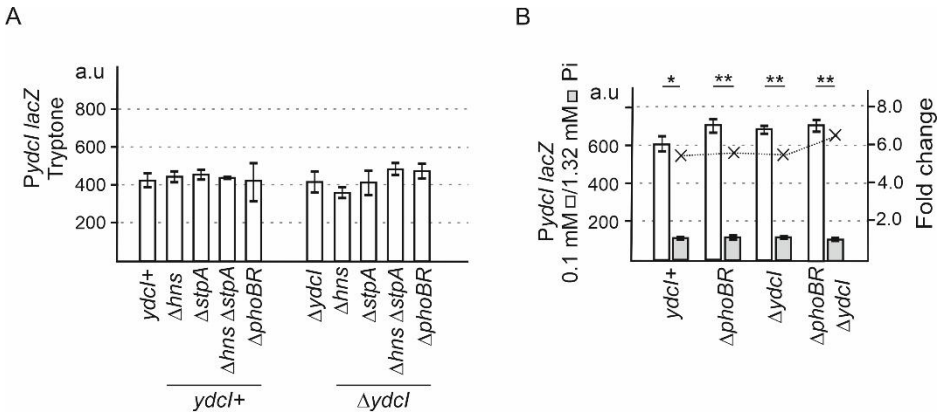


Figure 7 H-NS/StpA and PhoBR does not regulate *ydcI* promoter. The regulation of P_{ydcI} activity in response to deletion of A) H-NS/StpA and B) PhoBR was determined. Expression level was determined from P_{ydcI} *lacZ* (pKERBR093) fusion in wild-type (U4), Δ*hns* (U261), Δ*stpA* (U559), Δ*hns* Δ*stpA* (U563), Δ*phoBR* (U624), Δ*ydcI* (U505), Δ*ydcI* Δ*hns* (U558), Δ*ydcI* Δ*stpA* (U560), Δ*ydcI* Δ*hns* Δ*stpA* (U564) and Δ*ydcI* Δ*phoBR* (U625) strains. Cultures for β-galactosidase assays were inoculated from overnight cultures and grown in A) tryptone, B) MOPS glu low (0.1 mM Pi) and high (1.32 mM Pi) medium to an OD_{600}~0.8. Mean values from at least three independent replicates are shown as bars, and error bars indicate standard deviations. T-test was performed and the statistically significant values are shown as ****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05.

In order to study the role of ppGpp synthesized by RelA and SpoT in the phosphate stress induced activation of promoter of *ydcI*, the deletion strains for *relA*, *relA spoT* and $spoT_{E319Q}$, *relA* $spoT_{E319Q}$ mutants were used. A reduced induction of the P_{ydcI} *lacZ* reporter expression is observed between the wild-type and *relA* mutant and between *ydcI* and *ydcI relA* mutants (Fig. 8A). The induction of P_{ydcI} *lacZ* expression is completely abolished in *relA spoT* (Fig. 8A) and *relA* $spoT_{E319Q}$ (Fig. 8B) mutants suggesting

that the phosphate stress induced activation of the *ydcI* promoter is dependent on ppGpp synthesized by either RelA or SpoT (Fig. 8A). Similar to P_{iraP}, the difference between the activation of *ydcI* promoter in P_{ydcI} *lacZ*-cm and P_{ydcI} *lacZ*-tet reporter system could be attributed to the two different antibiotic resistance reporter plasmids used for the expression analysis but would need further studies for clarification. Taken together, the results show ppGpp induced activation of *ydcI* promoter, which is dependent on ppGpp provided by either RelA or SpoT.

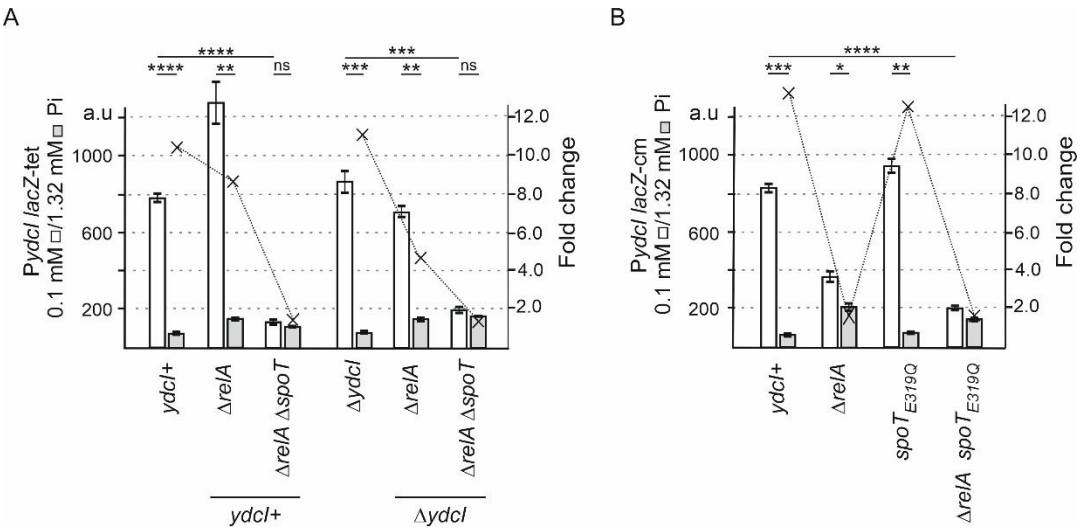
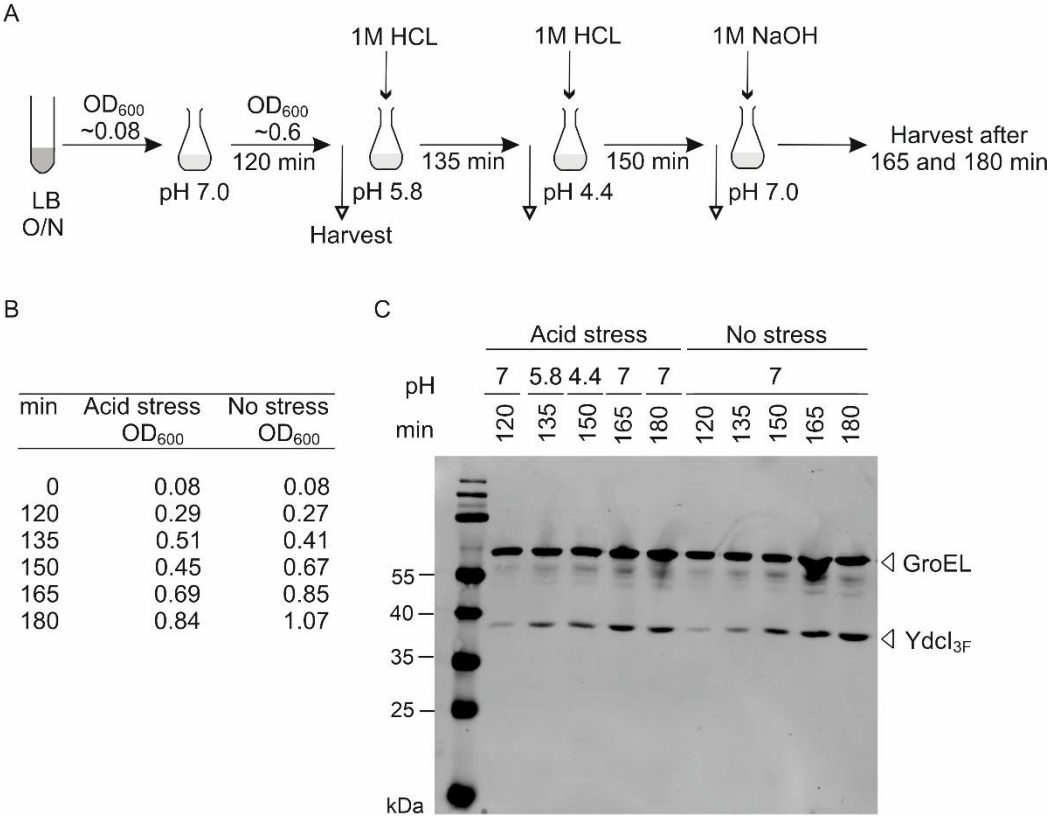


Figure 8 ppGpp regulates *ydcI* promoter. The regulation of P_{ydcI} by ppGpp were determined using wild-type (U4), Δ*relA* (U571), Δ*relA* Δ*spoT* (U578), Δ*ydcI* (U505), Δ*ydcI* Δ*relA* (U572), Δ*ydcI* Δ*relA* Δ*spoT* (U579) in trans with plasmid expressing A) P_{ydcI} *lacZ*-tet (pKERBR114), with wild-type (U4), Δ*relA* (U571), $spoT_{E319Q}$ (U613), Δ*relA* $spoT_{E319Q}$ (U615) strains in trans with plasmid expressing B) P_{ydcI} *lacZ*-cm (pKERBR093). Cultures for β-galactosidase assays were inoculated from MOPS glu high (1.32 mM Pi) overnight cultures and grown in MOPS glu low (0.1 mM Pi) and high (1.32 mM Pi) medium to an OD_{600} of 0.5-0.6. Mean values from at least three independent replicates are shown as bars, and error bars indicate standard deviations. T-test was performed and the values are shown as ****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05, ns=non-significant.

Growth phase dependent YdcI synthesis

A genome wide Ribo-seq and RNA-seq analysis of *E. coli* at pH stress conditions revealed that YdcI was differentially translated at pH 4.4 but the mRNA levels of YdcI were constant in response to acid stress (Schumacher et al., 2023). To analyse the protein levels of YdcI exposed to acid stress, a strain encoding chromosomal *ydcI-3xFLAG* (*ydcI$_{3F}$*) allele was used for Western blot analysis. Western blotting showed the level of YdcI-3xFLAG (YdcI$_{3F}$) increased when pH of the growth medium drops to 5.8 and further to 4.4 (Fig. 9C), as suggested by Schumacher et al. (2023). Interestingly, a control experiment with bacteria grown for similar time at constant pH 7 showed that YdcI$_{3F}$ levels similarly increased (Fig. 9C). Hence, the observed increase in YdcI protein level is rather dependent on the growth phase and independent of the induced acid stress as evident from the expression level of YdcI$_{3F}$ in the strain grown without an induced stress (Fig. 9C).

Further, to analyse the growth phase dependent expression of YdcI$_{3F}$, the strain carrying the chromosomal *ydcI$_{3F}$* was grown in different media (LB, tryptone and tryptone 0.2% glucose) and YdcI$_{3F}$ levels were determined over the growth phase (Fig. 10A). Western blotting showed the expression of YdcI$_{3F}$ in LB and tryptone starts from early-mid exponential phase and stays the same throughout in LB, while the protein level declines after 24 hours in tryptone (Fig. 10B and 10C). Interestingly, the protein level of YdcI$_{3F}$ is detected only from early-mid exponential to early stationary phase in tryptone medium supplemented with 0.2% glucose and declines after 5 hours of growth (Fig. 10D). Taken together the data shows that the protein level of YdcI is dependent on the growth phase and independent of acid stress.



min	Acid stress OD_{600}	No stress OD_{600}
0	0.08	0.08
120	0.29	0.27
135	0.51	0.41
150	0.45	0.67
165	0.69	0.85
180	0.84	1.07

Figure 9 Chromosomal YdcI-3xFLAG ($YdcI_{3F}$) levels in response to acid stress anaylsed by Western blotting.

A) Schematic diagram showing the growth of *ydcI-3xFLAG* (*$ydcI_{3F}$*) strain (U584). The strain was inoculated from overnight cultures of *ydcI-3xFLAG* (*$ydcI_{3F}$*) strain (U584) in LB medium and grown for 180 min. Samples were harvested as depicted in (A), following a stepwise pH reduction from 7 to 5.8 and subsequently to 4.4, after which the pH of the medium was increased back to 7. B) OD_{600} measurements of the cells in response to acid stress or no stress condition. C) Western blot analysis of *$ydcI_{3F}$* (U584) grown in LB pH 7, 5.8 and 4.4. As a control, strains were grown in parallel in LB without any induced stress and harvested over the same duration from 120 to 180 min. Lysates corresponding to bacteria at an OD_{600} of 0.08 were loaded per lane. Blots incubated with the primary monoclonal ANTI-FLAG® M2 antibody (1:2000), for uniform loading control GroEL (1:5000) and secondary Goat Anti-Mouse IgG H&L (Alexa Fluor® 680) antibody (1:10000) were used.

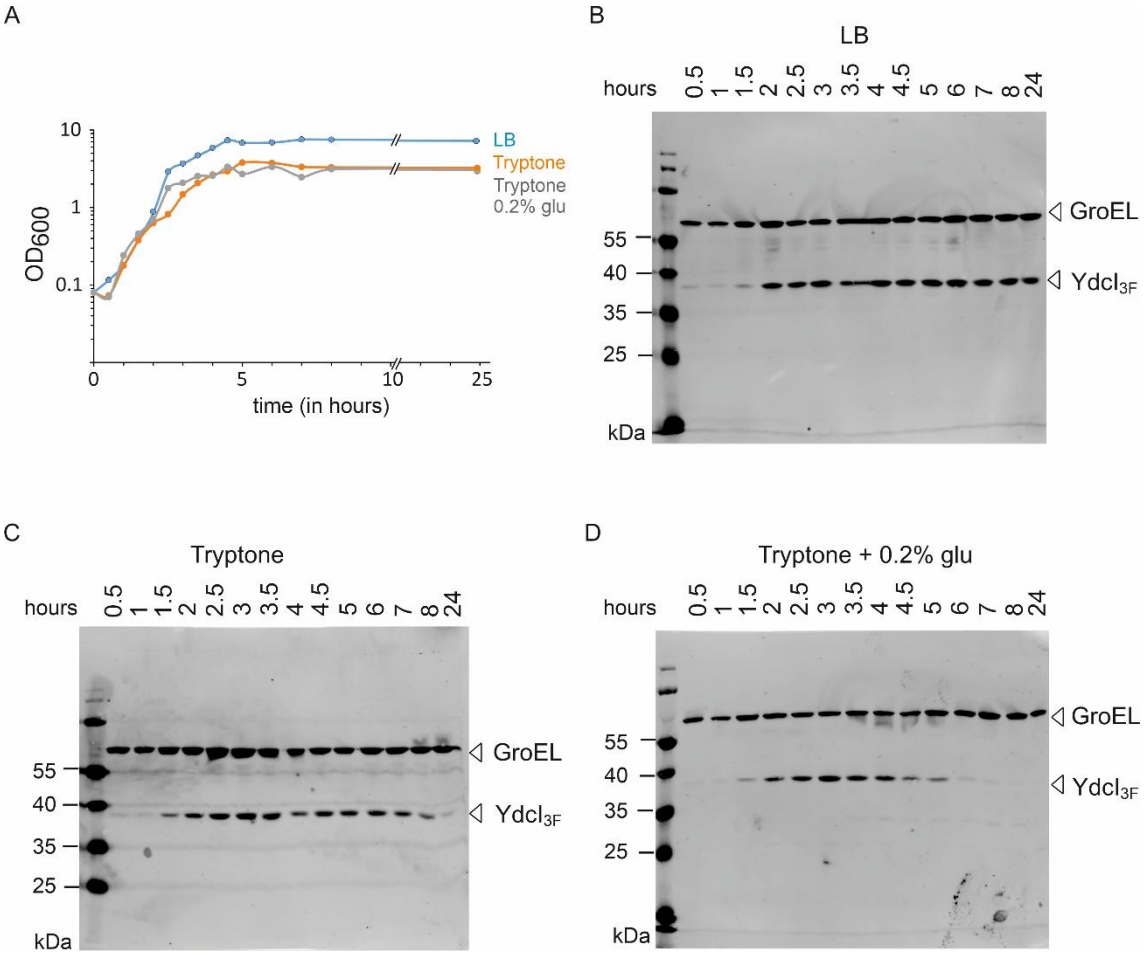


Figure 10 Chromosomal YdcI-3xFLAG ($YdcI_{3F}$) levels in different growth conditions as detected by Western blotting. Cultures of *ydcI-3xFLAG* (*$ydcI_{3F}$*) (U584) were grown in B) LB, C) tryptone, D) tryptone + 0.2% glucose and the corresponding growth curve in different medium depicted in (A). Samples were harvested after every 30 min as indicated upto 5 hours and then after 1 hour each upto 8 hours followed by harvesting of overnight culture (24 hours). Lysates corresponding to bacteria at an OD_{600} of 0.08 were loaded per lane. Blots incubated with the primary monoclonal ANTI-FLAG® M2 antibody (1:2000), for uniform loading control GroEL (1:5000) and secondary Goat Anti-Mouse IgG H&L (Alexa Fluor® 680) antibody (1:10000) were used.

Since, YdcI levels increased during the early-mid exponential phase (Fig. 10), it was investigated if the growth dependent expression of YdcI conferred any possible phenotype for YdcI. Hence, for both wild-type and *ydcI* mutant strain a growth curve analysis was conducted in tryptone medium upto 4 hours. The growth curve analysis does not show any distinct difference between the growth pattern of wild-type and *ydcI* but revealed a slightly delayed entry of the *ydcI* mutant from the early exponential to the mid-exponential growth (Fig. 11).

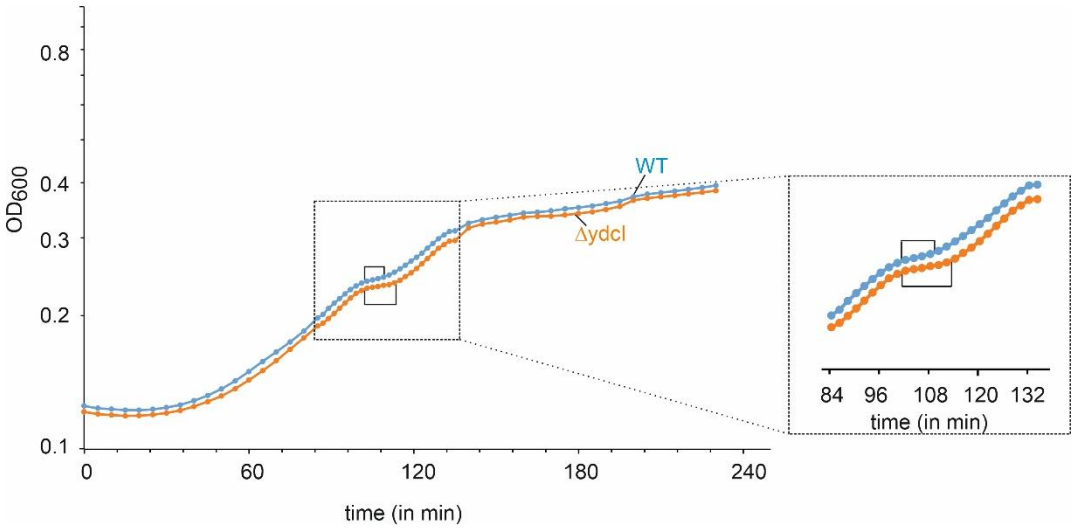


Figure 11 Growth curve analysis of wild-type and Δ*ydcI*. Growth curve measurement of strains T1241 Δ*lacZ* (U4), T1241 Δ*lacZ* Δ*ydcI* (U505) in tryptone media. The strains U4 and U505 were inoculated at an OD_{600} of 0.05 and grown in tryptone medium, the growth was monitored upto 4 hours. Each strain was grown twice as part of two different independent experiments and the average of the OD_{600} measurements were taken for plotting the growth curves of U4 and U505.

Regulation of YdcI targets

DNA-binding analysis by EMSA identified 10 YdcI bound loci. To analyse their regulation by YdcI, trancriptional promoter *lacZ* fusions of the identified targets were constructed on low copy number pSC-derived plasmids. The expression level for P_{mqo} *lacZ* reporter was repressed in the wild-type strain as compared to the *ydcI* mutant in compliance with the RNA-seq data when grown in tryptone medium for 2.5 hours (Fig. 12D). However, the data show that none of the other targets exhibited any regulation by YdcI in tryptone medium grown for 2 to 3 hours. This could be attributed to the selection of the incorrect growth conditions considering that YdcI synthesis is induced as of the transition from the early to mid-exponential phase in tryptone (starting from 2 hours) as shown (Fig. 10). In addition, a small molecule ligands/effector might be required to modulate the activity of YdcI as typical of LTTRs. In accordance with the increase of YdcI levels as of 2 hours of growth, the promoter *lacZ* reporter assays were conducted over a period of 2 to 4 hours with an interval of 30 minute between each sample collection. However, there was no statistically significant change detected in the expression of the promoter *lacZ* levels of target loci (Fig. 13). Interestingly, the promoter *lacZ* reporter fusions show that there exists tendency of downregulation for promoters of *nhaA*, *gltA*, *yobA*, *mqo*, *lldP* and upregulation for promoter of *iraP*, *sdhC*, *yhdV*, *gltP* by YdcI (Fig. 13).

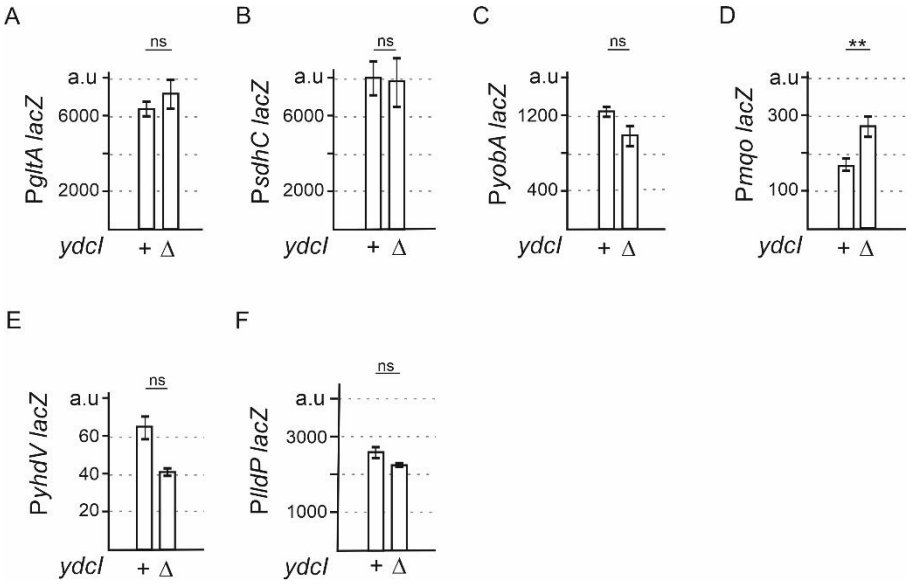


Figure 12 Regulation of YdcI targets. The regulation of the targets of YdcI were determined using wild-type (U4) and Δ*ydcI* (U505) and using promoter *lacZ* reporters expression analysis for A) P_{gltA} *lacZ* (pKERBR111), B) P_{sdhC} *lacZ* (pKERBR105), C) P_{yobA} *lacZ* (pKERBR089), D) P_{mqo} *lacZ* (pKERBR113), E) P_{yhdV} *lacZ* (pKERBR112) and F) P_{lldP} *lacZ* (pKERBR090). Cultures for β-galactosidase assays were inoculated to an OD_{600}~0.08 from overnight cultures in tryptone and grown to an OD_{600} of 0.8 in tryptone pH 7.5 for (10 mM Tricine pH 7.5 used to adjust the pH of the

medium) for all targets except F) P_{lldP} *lacZ* (pKERBR090) which was grown in tryptone. Mean values from at least three independent replicates for the β-galactosidase assays are shown as bars, and error bars indicate standard deviations. T-test was performed and the values are shown as ****P=<0.0001, ***P=<0.001, **P=<0.01, *P=<0.05, ns=non-significant.

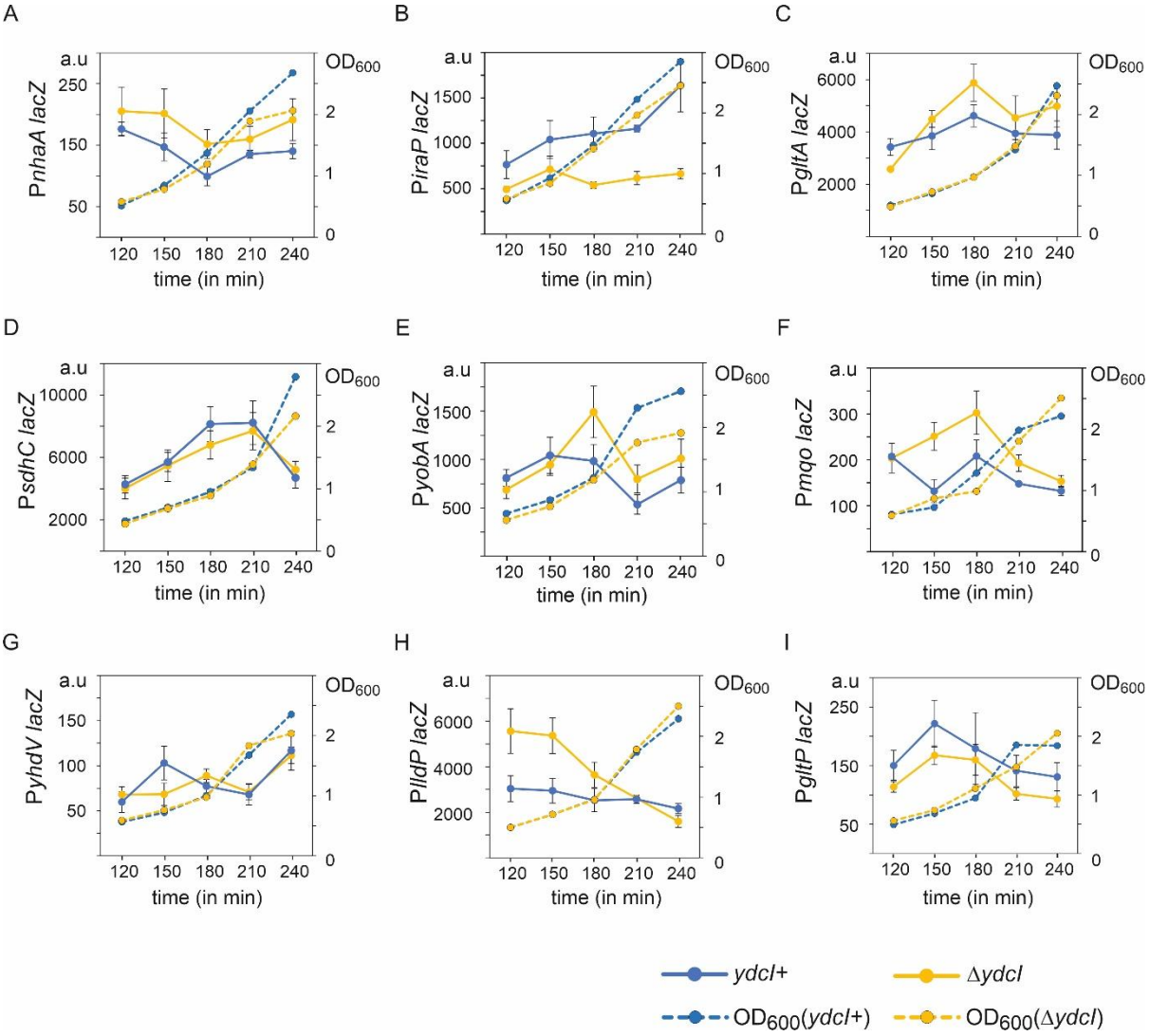


Figure 13 Time dependent regulation of YdcI targets. The regulation of the targets of YdcI were determined using wild-type (U4) and Δ*ydcI* (U505) and using promoter *lacZ* reporters expression analysis for A) P_{nhaA} *lacZ* (pKERBR122), B) P_{iraP} *lacZ* (pKERBR091), C) P_{gltA} *lacZ* (pKERBR111), D) P_{sdhC} *lacZ* (pKERBR105), E) P_{yobA} *lacZ* (pKERBR089), F) P_{mqo} *lacZ* (pKERBR113), G) P_{yhdV} *lacZ* (pKERBR112), H) P_{lldP} *lacZ* (pKERBR090) and I) P_{gltP} *lacZ* (pKERBR123). Cultures for β-galactosidase assays were inoculated to an OD_{600}~0.08 from overnight cultures and grown to an OD_{600} of 0.5-2.5 in tryptone starting from 120 min to 240 min. Mean values from at least three independent replicates for the β-galactosidase are shown as bars, and error bars indicate standard deviations. An average of the OD_{600} measurements for the respective assays is shown for wild-type (U4) and Δ*ydcI* (U505).

Discussion

This study has been undertaken to address the existing gap in literature regarding the regulatory and functional role of the LysR-type transcription regulator, YdcI. The characterisation of YdcI started from validating ChIP-exo identified target genes of YdcI (Gao et al., 2018) using *in vitro* DNA binding analyses. These revealed binding of YdcI to 10 out of 16 ChIP-exo predicted loci and a consensus DNA-binding motif that was confirmed by mutagenesis using YdcI-sites at *iraP*. Of the 10 YdcI-bound loci, *iraP* and *yhdV-yhdWXYZ* are activated by YdcI, *mqo* and *yobA-yebZ-yebY* are repressed by YdcI, while no regulation of the remaining six loci by YdcI was detected with *lacZ* reporter fusions, at least at the physiological conditions used. YdcI activates *iraP*, encoding a protein of the phosphate stress signaling response, independent of ppGpp-dependent phosphate stress induction of *iraP*, but presumably by antagonizing *iraP* repression by H-NS. Furthermore, here it was found that the *ydcI* promoter is induced by phosphate stress in a ppGpp-dependent manner like *iraP*, and that YdcI synthesis increases during the transition from the early exponential to the slowly declining growth phase (named here mid-exponential). Interestingly, two of the YdcI bound loci (*mqo*, *gltA-sdhCDAB*) encode enzymes of the TCA cycle, YobA-YebZ-YebY is important for NAD synthesis by NDH-2, five YdcI target loci encode transport proteins, and one a presumptive small stress signaling protein (YobF). Taken together the data indicate that YdcI is related to regulation of metabolic functions and stress responses.

The function of YdcI as transcription regulator seems to relate to metabolism and/or nutritional shift. First, YdcI protein levels, which are very low in the early exponential growth phase, increase significantly after approximately 2 h of growth in LB and tryptone media, at an OD_{600} of ~0.5 to 0.8. Then YdcI levels stay high for several hours, but are low again in stationary growth phase (overnight growth). It has been shown that growth in LB is carbon limited and that bacteria grow exponentially for a short time only (early exponential growth) followed by a gradual decline after ~2 h (at OD_{600} of ~0.3) (late exponential growth) presumably because the bacteria then shift to catabolism of peptides, an ample component of LB and tryptone media (Sezonov et al., 2007). It has also been reported that the shift from early to the late exponential growth is accompanied by a short lag phase (Wang and Koch, 1978). Interestingly, this short lag phase was also observed in this work (Fig. 11) with a slightly prolonged lag phase in the *ydcI* mutant indicating that YdcI has a role in nutritional shifts. In line with this, all the identified YdcI target loci have a range of roles in different metabolic pathways (Fig. 14). The gene *gltA* encodes for enzyme citrate synthase and this enzyme's activity was increased in a *ydcI* deletion mutant confirming the repression of *gltA* by YdcI (Nishio et al., 2013). Gene *gltA* maps divergent to *sdhCDAB* encoding succinate:quinone oxidoreductase. In this work, repression of *mqo* by YdcI was shown by using a P_{mqo} *lacZ* reporter and repression is evident from RNA-seq data. The gene mqo encodes for TCA cycle enzyme malate:quinone oxidoreductase. In addition, another identified

target loci of YdcI is the *yobA-yebZ-yebY* operon, which is required for copper binding and delivery to NADH dehydrogenase-II (NDH-2) (Hadley et al., 2022). The NADH produced from the TCA cycle could be further utilized by enzymes like NDH-2 of the electron transport chain. The repression of *mqo, gltA* and possibly its divergent *sdhCDAB* operon (coding for enzymes of the TCA cycle) and *yobA-yebZ-yebY* (associated with electron transport chain) by YdcI suggest a potential role of YdcI in down-regulating the carbon flux operating in the TCA cycle upon a nutritional shift from early to late exponential growth. Gene *yobF*, another target gene of YdcI, is induced upon heat shock and important for acid stress response in *E. coli* (Hobbs et al., 2010), which emphasizes on the potential role of YdcI in different stress conditions. Additional identified targets of YdcI include presumptive charged/polar amino acid ABC transporter (*yhdWXYZ* and lipoprotein YhdV), several symporters (*lldP*, *gltP*), inner membrane glycine transporter (*yicG*) and a pH-regulated NhaA H+/Na+ antiporter (implicated in maintaining cellular pH homeostasis during alkali challenge) suggesting the possible role of YdcI in transport of metabolites and pH homestasis (Fig. 14).

In this study, the absence of detectable regulation of the targets of YdcI, other than *iraP* and *mqo*, suggest that the specific growth conditions or a small molecule inducer yet remain to be identified for YdcI. Data of this work show induction of YdcI after 2 hours of growth (as the cells transit from early to mid-exponential growth phase) and this induction was independent of acid stress (Fig. 9). Additionally, the promoter *lacZ* reporter expression analysis showed a tendency of regulation of the the targets of YdcI in a time dependent manner (Fig. 13). Further, data of this work suggests that the previously reported finding of acid stress (pH 4.5) induced translation of YdcI (Schumacher et al., 2023) was caused by the growth phase and not by acid stress.

Our results validate YdcI as activator of *iraP* and show repression of *iraP* by H-NS (Fig. 14). YdcI presumably activates *iraP* by acting as H-NS antagonist, since activity of the *iraP* promoter is similarly high in the *hns* single and the *hns ydcI* double mutant. Activation by derepression of H-NS repressed promoters is well known principle of regulation (Navarre et al., 2007). H-NS is an abundant global repressor of stress and pathogenicity related genes, and H-NS mediated repression of *iraP* extends to list of stress-repressed genes regulated by H-NS. IraP associates with RssB and prevents ClpXP mediated degradation of the stress sigma factor σ^S (Bougdour et al., 2006). Thus, YdcI-mediated activation of *iraP* expression might have a role in eliciting the RpoS mediated general stress response in the cells. Furthermore, our data show that under phosphate starvation conditions, the activation of *ydcI* promoter takes place via ppGpp (either by phosphate stress response synthetases, RelA or SpoT), as shown before for *iraP* (Bougdour and Gottesman, 2007), demonstrating a link between elicitation of the stringent stress response and the expression of *ydcI* gene. Recently, YtfK has been shown to

promote the SpoT mediated accumulation of ppGpp and a possible interaction of SpoT with YtfK shifts the activity of SpoT from hydrolase to synthetase in response to phosphate, fatty acid and carbon starvation (Germain et al., 2019, Meyer et al., 2021). This suggests a possible role of YtfK in eliciting the ppGpp mediated activation of *iraP* and *ydcI* gene in response to phosphate stress.

Taking together the findings of this study, YdcI is hypothesized to play an important role in response to different encountered physiological stresses in *E. coli*. YdcI also potentially regulates the carbon flux operating in metabolic processes like the TCA cycle and a detailed functional characterization of this LTTR for deciphering the complex link of YdcI's role in different biological processes yet remains to be studied in the future.

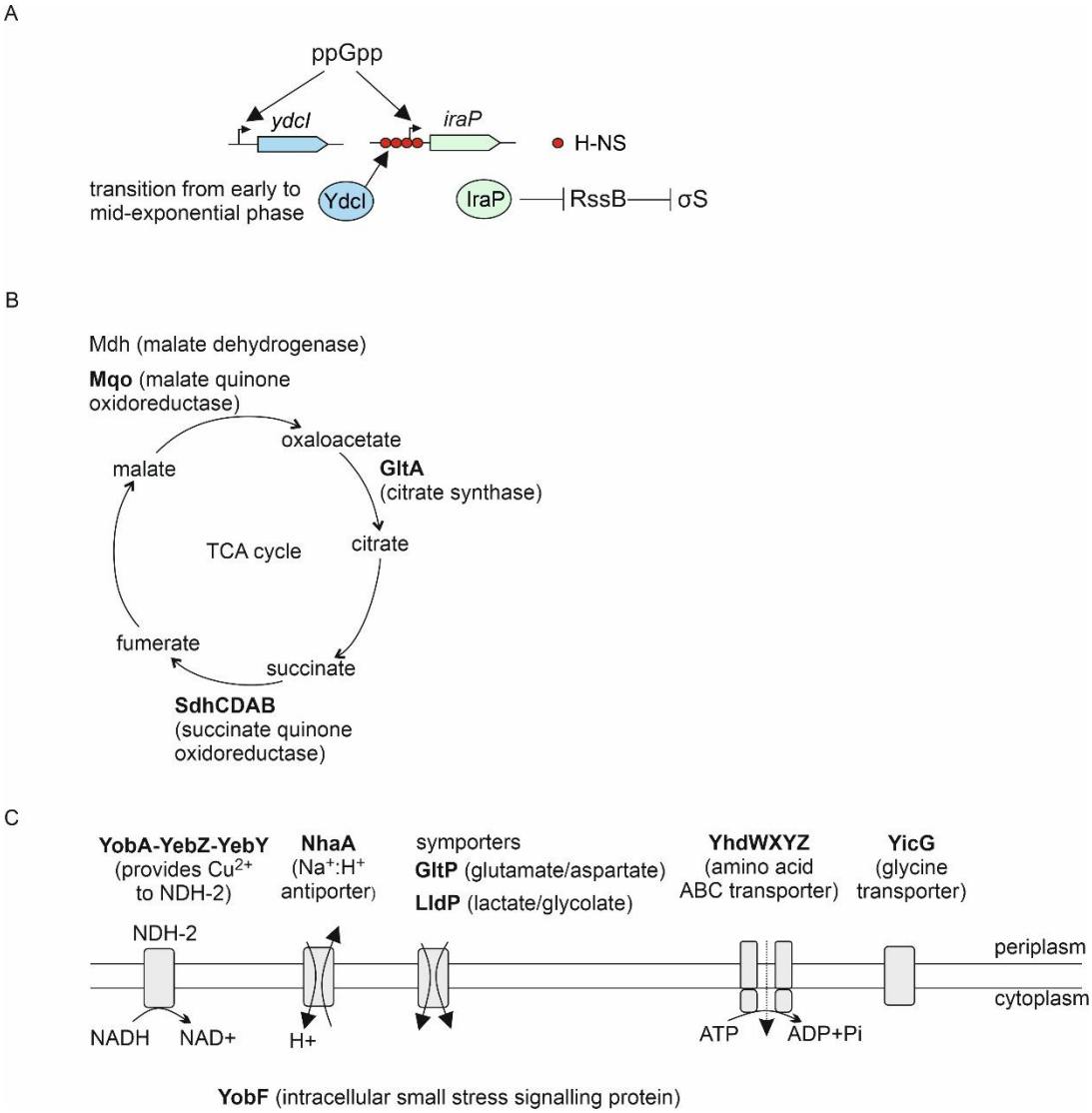


Figure 14 Model of transcriptional regulation of *ydcI* and *iraP* and function of identified targets of YdcI.

A) Under phosphate starvation growth conditions, *ydcI* and *iraP* are activated by alarmone, ppGpp. H-NS (and StpA) represses the *iraP* gene expression and YdcI activates the expression of *iraP*. B) The enzymes of the TCA cycle namely Mqo, GltA and SdhCDAB are potentially regulated by YdcI. C) Other potential target genes of YdcI encode for membrane proteins and transporters. The targets of YdcI are marked in bold and their associated functions are shown.

Material and Methods

Bacterial strains and growth conditions

E. coli K-12 strains are summarized in Table S2. Strains were constructed by transduction using phage P1*vir* (Miller, 1992), and by Lambda-Red-mediated recombineering, as described (Datsenko and Wanner, 2000). Bacteria were cultured in in LB medium (10 g/l tryptone, 5 g/l yeast extract, and 5 g/l NaCl), tryptone medium (10 g/l tryptone, 5 g/l NaCl), SOB medium (20 g/l tryptone, 5 g/l yeast extract, 0.5 g/l NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, pH7.0), SOC medium (SOB with 0.4 % glucose),
MOPS minimal 0.5% glucose medium composed of 0.1 mM MOPS pH 7.4, 0.1 mM Tricine pH 7.4, 1 mM $FeSO_4$, 0.19 M NH_4Cl, 27.6 mM K_2SO_4, 0.5 mM $CaCl_2$, 52.8 mM $MgCl_2$, 0.5 M NaCl, micronutrient stock (30 µM $(NH_4)_6Mo_7O_{24}$•4H2O, 4 µM H_3BO_3, 300 µM $CoCl_2$, 100 µM $CuSO_4$, 800 µM $MnCl_2$, 100 µM $ZnSO_4$) supplemented with 0.1 mM or 1.32 mM of K_2HPO_4. The MOPS minimal medium for phosphate starvation experiments were made according to procedure described previously (Bougdour et al., 2006, Neidhardt et al., 1974, Rice et al., 2009). Antibiotics were added to a final concentration of 50 µg/ml ampicillin, 15 µg/ml chloramphenicol, and 15 µg/ml kanamycin, as required.

For phosphate starvation experiments, all cells were grown overnight in MOPS glu 1.32 mM Pi (high phosphate) medium and inoculated at an OD_{600} of 0.08 in low (0.1 mM Pi) or high (1.32 mM Pi) phosphate MOPS glu medium. For starvation experiments with washing steps, cells were grown upto an OD_{600}~0.3-0.4 in MOPS glu (1.32 mM Pi) medium. The culture was subsequently divided in two parts (for growth in MOPS glu low and high phosphate medium). The cells were further washed twice by centrifugation followed by resuspension in the respective prewarmed MOPS glu low (0.1 mM Pi) and high (1.32 mM Pi) medium. The cells were then incubated and grown at 37°C upto OD_{600}~0.5-0.8 and cultures harvested for β-galactosidase assays. For starvation experiments without washing steps, cells were inoculated to an OD_{600} of 0.08 using an overnight culture grown in MOPS glu high (1.32 mM Pi) medium and grown at 37°C in low and high phosphate MOPS glucose medium, respectively. The cells were then grown upto mid or late exponential phase (OD_{600}~0.5-0.8) in the respective medium.

Plasmid construction

Plasmids were constructed using standard DNA cloning techniques, using oligonucleotides and gene fragments as listed in table S4. As a principle, all segments of plasmids that were generated by PCR and all cloning fusion sites were sequenced. In addition, the integrity of the sizes and structure of each plasmid was validated by restriction enzyme analyses.

β-galactosidase expression analysis

For expression analyses of promoter *lacZ* fusions bacteria were grown as described in the figure legends. The β-galactosidase assays were performed as described (Miller, 1992) of at least 3 independent biological replicates and standard deviations were calculated. Student *t*-Test was performed for pair of samples as indicated with the asterisks in the bar graphs.

Growth curve analysis

To measure the growth of relevant strains over time, the *Tecan Spark multimode reader* was used. Cultures were assayed in 96-well clear bottom microplates. The strains of interest were first grown over night at 37°C in 4 ml Tryptone. 4 ml of the culture was harvested by centrifugation and the pellet was resuspended in 4 ml MG saline followed by OD_{600} measurement for each strain. The required volume of each culture was taken to achieve OD_{600}~0.05 in 250 µl tryptone in each well. The measurements were taken every 5 min for 1 hour 25 min, every 2 min for the next 50 min and then for every 5 min for the next 1 hour 35 min.

Western blotting

For comparison of chromosomal levels of $YdcI_{3F}$, strain U584 (U4 *ydcI-3xFLAG*) coding for epitope-tagged $YdcI_{3F}$ under the control of its native promoter was used (Table S2). Cultures grown overnight were used to inoculate 20 ml of LB medium, tryptone and tryptone containing 0.2% glucose to an OD_{600} of 0.08 and grown upto 24 hours at 37°C. Samples were at the time indicated, pelleted for 1 min at 13,000 rpm.

SDS-PAGE and Western blot procedures were performed as described previously (Sambrook, 2001). Per lane, cells at an OD_{600} of 0.08 were loaded, and equal loading was verified by 2,2,2-trichloroethane staining of total protein (Ladner et al., 2004). Epitope-tagged $YdcI_{3F}$ was detected using primary antibody anti-FLAG M2 from mouse (diluted 1:2000) (catalogue number F3165; Sigma) and Alexa Fluor 680 fluorescent dye-labeled secondary anti-mouse antibody from goat (diluted 1:10000) (catalogue number A21057; Thermo Scientific) and anti-GroEL antibody (diluted 1:5000). Quantification of protein bands was performed with Odyssey V3.0 software.

Electrophoretic mobility shift assays for DNA-binding analyses

DNA-binding analyses were performed by electrophoretic mobility shift assays (EMSA). DNA fragments were amplified by PCR using oligonucleotides specified in Table S4. After clean-up of the fragments (NucleoSpin Gel and PCR Clean-up Mini kit, Macherey-Nagel, Germany), the concentration of the DNA samples were estimated by agarose gel electrophoresis and the samples stored at -20°C.

DNA-binding reactions with purified YdcI$_{His6}$ were set up on ice in 10 µl binding buffer (20 mM Tris-HCl [pH 7.5], 100 mM KCl, 2 mM DTT, 10% glycerol, 1.25 ng/µl poly[d(I-C)]) with 2.5 nM of 5' 6-FAM (Fluorescein) labelled DNA fragments, and YdcI$_{His6}$ protein at the concentration specified in the figures. Non-labelled PCR fragments were added as competitors at a concentration of 50 nM (65 ng per 10 µl assay), where indicated. After set-up of the EMSA binding samples on ice, they were incubated at 30°C for 20 minutes and then rapidly loaded onto an 8% polyacrylamide (acrylamide: bis-acrylamide 29:1), 0.5x TBE (45mM Tris-base, 45 mM boric acid, 1 mM EDTA) gel. The gel was run in the cold room for 10 min at 50 V, followed by 55 min at 200 V. Imaging of gels with the FAM-labelled fragments (Absorbance max: 495 nm, Emission max: 520 nm) was performed using a BioRad Chemidoc MP system (settings blue EPI light illumination, filter 530/28). Gel pictures are shown with inverted colors.

YdcI$_{His6}$ protein purification

For purification of C-terminally histidine-tagged YdcI$_{His6}$, strain C41(DE3) was transformed with plasmid pKESL317 carrying the *ydcI$_{His6}$* under control of the *T7* promoter (Table S3). For protein purification, overnight cultures of the transformant were used to inoculate 400 ml of LB medium containing 50 µg/ml ampicillin to an OD_{600} of 0.1, and the cultures were grown at 37°C for 1.5 h to an OD_{600} of 0.6. Protein expression was induced by adding IPTG to a final concentration of 200 µM, and the cultures were grown further for 1 h at 37°C. The cultures were harvested by centrifugation, cells were resuspended in resuspension buffer (20 mM Tris-HCl [pH 8.0], 300 mM KCl), and pelleted for storage at -80°C. Lysate preparation was performed at 4°C; the cell pellets were resuspended in lysis buffer (4 ml/g of bacterial cells) (20mM Tris-HCl [pH 8.0], 300 mM KCl, 20 mM imidazole, 20mg/ml DNase I [5,000 U/ml] [New England BioLabs, USA]) and lysed by sonication (40% amplitude for 4 min with a 2-s pulse/2-s pause) (Sonics Vibracell VCX750 High-Volume Ultrasonic Cell Disrupter). The lysate was cleared by centrifugation (two times; 40,000 rpm, 30 min each) (Sorvall SS34 rotor) and the supernatant was loaded onto Polypropylene column. The column was washed with wash buffer (20 mM Tris-HCl [pH 8.0], 300 mM KCl) with increasing imidazole concentrations (20 and 60 mM) with 10 times the resin volume at each step. The proteins were eluted with elution buffer (20 mM Tris-HCl [pH 8.0], 300 mM KCl, 250 mM imidazole), and 500 µl fractions were collected and analysed by SDS-PAGE. The eluted fractions of purified protein were stored in 20 mM Tris-HCl [pH 8.0], 300 mM KCl containing 50 mM NDSB-256. Protein concentrations were measured by Pierce bicinchoninic acid (BCA) protein assay (Thermo Scientific, Germany). Aliquots of the proteins were stored at -80°C.

RNA-seq and data analysis

For RNA-seq, total RNA was isolated from strains (U4, U456 Δ*yafC*, U505 Δ*ydcI*) with U4 taken as control. These strains were inoculated to an OD_{600} of 0.05 in 10 ml tryptone medium and grown for

2.5 h to an OD_{600} of approximately 0.8 at 37°C with shaking and 1.5 ml samples of the cultures were harvested by adding RNAprotect bacterial reagent (Qiagen, Germany) to them followed by total RNA isolation using the RNeasy minikit (Qiagen, Germany), with on-column DNA digestion by DNase I treatment, according to the manufacturer's instructions (Qiagen, Germany). The integrity of the RNA was validated by separating 1 µg of RNA on 8% denaturing acrylamide gels (7 M urea, 45 mM Tris base, 45 mM boric acid, 1 mM EDTA, 8% acrylamide-bisacrylamide [19:1, wt/wt]) and visual inspection of the 23S rRNA and 16S rRNA bands. The preparation of RNA-seq libraries was performed according to methods described previously (Borries et al., 2011), with minor adaptations. The synthesis of cDNA libraries was performed by the Cologne Center for Genomics (CCG) (http://portal.ccg.uni-koeln.de/ccg/) using an Illumina TruSeq RNA library kit according to the manufacturer's instructions. Three biological replicates of each strain were sequenced, with 20 million reads each, using an Illumina HiSeq 4000 instrument. For data analysis, the sequencing raw data files were uploaded to the public Galaxy Web platform (https://usegalaxy.org/) (Afgan et al., 2016). First, raw sequencing reads were uploaded in FASTQ file format and aligned to the E. coli K-12 MG1655 genome (GenBank assembly accession number GCA_000005845.2) using Bowtie2 (Langmead and Salzberg, 2012) with default settings for paired-end libraries. These generated BAM files containing the aligned reads were used to calculate individual gene abundances. For this, HTSeq (Anders et al., 2015) was used in union mode for stranded data sets with a minimum alignment quality of 10. Tabular lists of gene abundances for each sample were used for differential expression analysis using DEseq2 (Love et al., 2014) with default settings. In addition, the aligned reads compiled in the BAM files of the individual samples were inspected using IGV (Integrative Genomics Viewer) (Robinson et al., 2011).

Data availability

Raw sequencing files (FASTQ file format) and gene abundance tables (TXT file format) have been deposited and are available in NCBI's Gene Expression Omnibus database under accession number GSE304471.

Identification of YdcI binding motif

The DNA sequence of the YdcI targets (Table S1) showing strong binding to YdcI from the competitive EMSA were submitted to MEME-suite (Grant et al., 2011) to generate a motif.

Supplementary Information

The supplementary information includes fig S1 for EMSA of YdcI targets,
Table S1 for input DNA sequences for MEME-suite, Table S2, for strains, Table S3, for plasmids and Table S4, for the oligonucleotides used in this study.

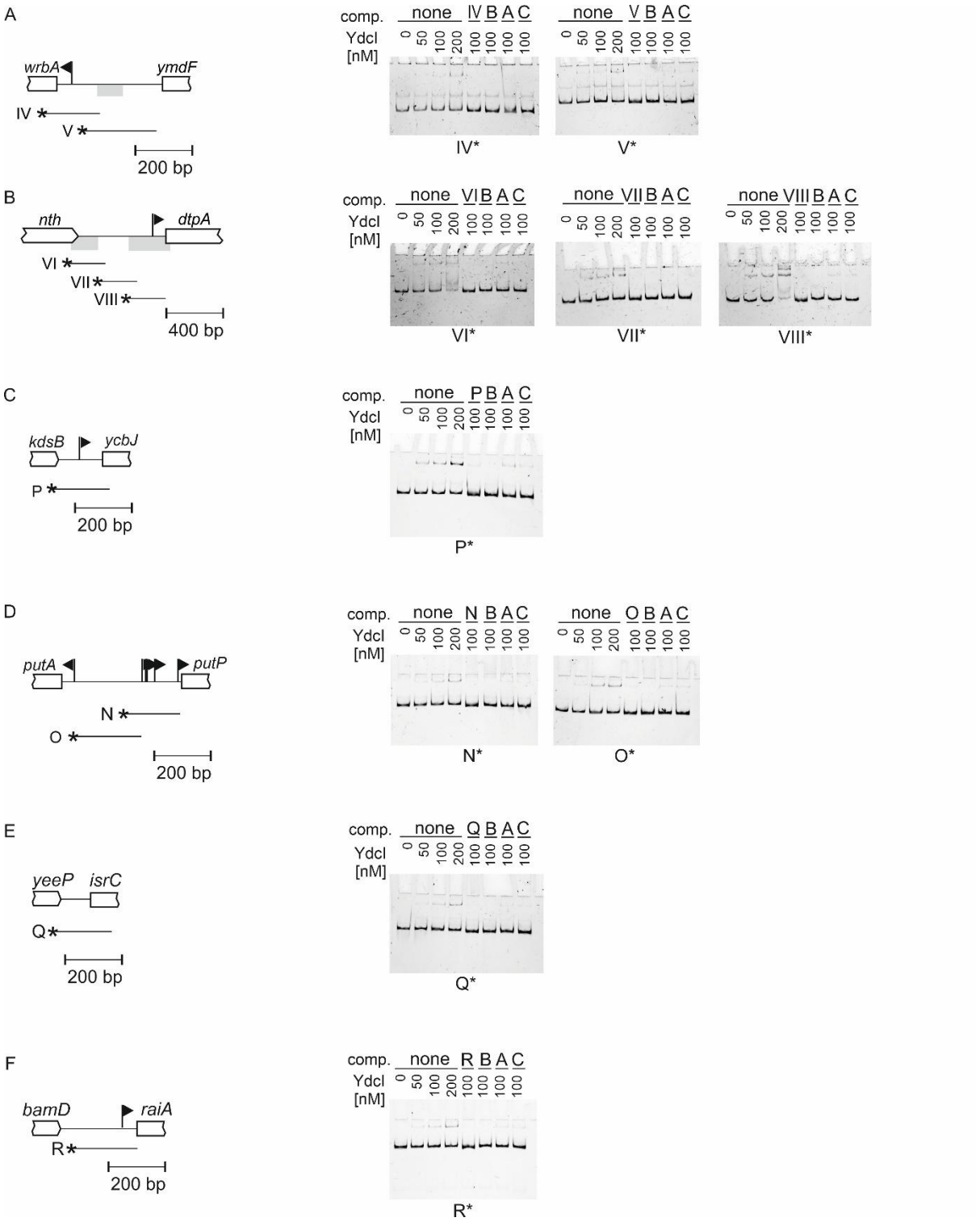
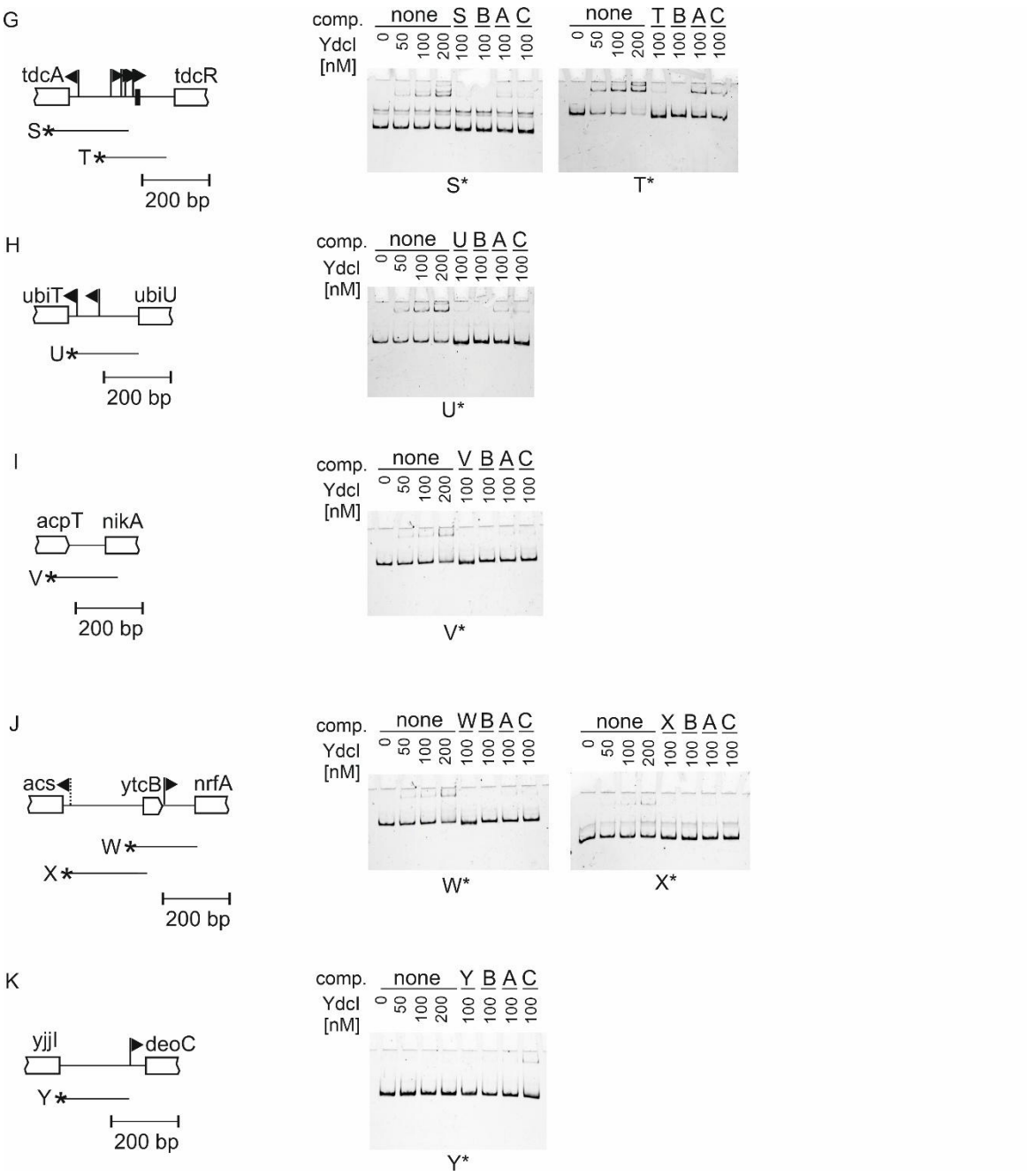


Figure S1 DNA YdcI-binding assays of non-bound fragments. For DNA-binding analyses of $YdcI_{His6}$, electrophoretic mobility shift assays (EMSAs) were performed using (A) *wrbA*, (B) *dtpA,* (C) *ycbJ*, (D) *putP*, (E) *isrC,* (F) *raiA,* (G) *tdcA*, (H) *ubiU,* (I) *nikA*, (J) *nrfA* and (K) *yjjI* promoter DNA fragments. Schematic of the upstream region of (A) *wrbA*, (B) *dtpA,* (C) *ycbJ*, (D) *putP*, (E) *isrC,* (F) *raiA,* (G) *tdcA*, (H) *ubiU,* (I) *nikA*, (J) *nrfA* and (K) *yjjI* promoter DNA fragments. All the fragments are depicted as solid lines and and were amplified by PCR using oligonucleotides listed in Table S4. 5' 6-FAM labeled fragments are marked with an asterisk. Predicted binding sites of YdcI (Fig. 1) is depicted as grey boxes for (A) *wrbA* and (B) *dtpA*. Electrophoretic mobility shift assays with increasing concentrations of purified $YdcI_{His6}$ using the (A) *wrbA* fragments IV and V*,* (B) *dtpA* fragments VI, VII

and VIII, (C) *ycbJ* fragment P, (D) *putP* fragments N and O, (E) *isrC* fragment Q, (F) *raiA* fragment R, (G) *tdcA* fragments S and T, (H) *ubiU* fragment U, (I) *nikA* fragment V, (J) nrfA fragments W and X and (K) *yjjI* fragment Y. Fragment A and C are used as non-specific competitors and fragment B is used as a specific competitor for the EMSA. Figure continued on next page.



Continuation of Figure S1

Tables

Table S1: Input sequence for MEME-suite		
Loci	Coordinates	Sequence
nhaA	17121-17330	TTTTTGTGGCGTAATTTTTTAATGATTTAATTATTTAACTTTAATTTATCTCTTCATCG CAATTATTGACGACAAGCTGGATTATTTTTGAAATATTGGCCTAACAAGCATCGCCG ACTGACAACAAATTAATTATTACTTTTCCTAATTAATCCCTCAGGAATCCTCACCTTA AGCTATGATTATCTAGGCTTAGGGTCACTCGTGAGC
iraP	401098-401202	AGCTAATTTATTCCTGGCTTAAATGGCCATGCGGTGAGTTTTTTTCTCTTAATTATAA GTTAACGAAGAGAATATATTTCATAACTTTTATTTATAATAAAGGTT
gltA	754588-754807	CTGGGTACAGAGCTCTGGGCGCTTGCAGGTAAAGGATCCATTGATGACGAATAAA TGGCGAATCAAGTACTTAGCAATCCGAATTATTAAACTTGTCTACCACTAATAACTG TCCCGAATGAATTGGTCAATACTTCCACACTGTTACATAAGTTAATCTTAGGTGAAA TACCGACTTCATAACTTTTACGCATTATATGCTTTTCCTGGTAATGTTTGT
yobF	1907821-1908100	GTTGTTCCAGGCCAGTGGGAGAAGTTATTACATAGTGCGTGCAATATCACATTTTTT GCTATGCAATGAATAAAAAGTTATATCACTTTTTCTCATAAAACAGTCAGTTAACGG CTATTAATTACCCTAAAGAGAAGTCAATCCCCAAAGGGATTGTAAATTTAAAATAA GAAAAATTGATGAATGAGCAAAAAAATCAAGAGAGAAACGTTTCGCTAATAAATA ATACCACGCTTAATCACAAATCCGATGTAAACATCCTAATAAATTAATGGGGTAA
yobA	1924941-1925150	TAGCGAAGGGAGCGTGCAGTTGAAGCCATATTATCTATTCCTTTTTGTAATAACTTT TTTACAGAGCATAACCTTGTCTAATGTCTGAGTCGAGGATCATCAATTCCGGCTTGC CATCCTGGCTCACTCTTAGTAACTTTTGCCCGCGAATGATGAGGAGATTAAGAATG CTGAAGAATCTGGCTAAACTGGATCAAACAGAAATGGATA
mqo	2307145-2307424	CCTGATAAAACTCACGACGGAAGTGTGGATCCTGATCCGCCGCCCATTCATCCAGC AGGATAATATCGCGTTCTTCTGCCAGCGCCAGCAACAGCGCCACGCGTTTTTTCTGC CCTTTTGATAACTTCAGGTTAACAATACGCCCGTTGCTTAACTCAAGCTTATGAGCC ATTTTCAGCTGCGCCAGCCACTTCTCAACCAGTTGCGGGTTAGCGGGTTTACCCTCC GGCCCCAGCAGTTGATCAAACAGCCAGACATCGGTAAACACTGCCGAAAACAG
yhdV	3418135-3418394	TAAATCAGAAACATAAAGGCGCTTTCGGGTGCCTTTATTATTTCCAGTGAAACCCAT AAAAATTAATAAGATATTCTTCTGCTCACTCTTTAAAAGCTTTCTATAGTTCCCGCTC CCTTCACTATTTTTACAATTCACATAATTTGTCATTATTCATTCCGCAAGATTTATAAC CTCCGGTGATAAAATGGCATTCAGCTCGTTAATAAGAGAGTTAACTTATTAAGCGT TAGCGTTTATTACTGAGGTAACACCATGAA
lldP	3777164-3777363	CAGACATCTCCCCCCACAAGAATTGGCCCTACCAATTCTTCGCTTATCTGACCTCTG GTTCACAATTTCCCAATTAAAACTCACATCAATGTTGCCAATACATAACATTTAGTTA ACCATTCATTGTCATTATCCCTACACAACACAATTGGCAGTGCCACTTTTACACAAC GTGTGACAAGGAGATGAGCAACAGACTC
yicG	3818587-3818866	GAAACTAAAAACTCTTTTTGTTGATTGAGACACCCGATGCGTAAGCCAAGGTCCAG GTGCAGTGAACACAATGGCTAAATATTGCACCTTTCTTTTCCCCCTCAGTTTTAACCT ATTTTTTCTTATGCATTTTCTCAGACAAGAAGTCAGAAGAATGCATCTCTGCTACAG AAAATAGCGATTTCACATAACTACAAGTTATCAATTTCCCCTCCCCTTAAAAAATCTC AATCGTGACAATGCGCACAAATCGCTACCCTGCCAGACAGATTTTTAGGGA
gltP	4294148-4294367	TGATGAAGATTATGTCATTGAGCTGCATAAAAAATAATCGAATGAACATATGCCAA AAATAATCACTAATCAGTATTATTGCAGATTAAACAAATAAAAATCTTTCCATAACA AATGGTTATTCATTAATCCTGCCATAAACTCGCGTTATTATGCATTAATGCAGCGAA AAGCTCTGTTGTTAAAGGGTTGCGCAACATACCGCGCAAATGATACTGAT

Table S2: Bacterial strain and relevant genotype		
Strains	Genotype / description[a]	Construction /reference[b]
CE41(DE3)	BL21(DE3) derivate for toxic protein expression	(Miroux and Walker, 1996)
M182	Δ(*lacIPOZYA)74 galU galK strA stpA::tet* (lab storage number S159)	(Zhang et al., 1996)
S3754	MG1655 Δ*hns*$_{KD4\text{-}kan}$	(Rangarajan and Schnetz, 2018)
T1241	BW30270 *ilvG*$^{+}$, *fhlDC*(IS1) highly motile, *dgcJ*::IS1	(Pannen et al., 2016)
AB036	MG1655 Δ*relA* $_{kan}$ *spoT*$_{E319Q}$ zib563::Tn10 Lab number T2985	(Bougdour and Gottesman, 2007)
AB037	MG1655 *spoT*$_{E319Q}$ zib563::Tn10 Lab number T2986	(Bougdour and Gottesman, 2007)
JW5226-1	BW25113 Δ*ydcI783*$_{kan}$ Lab number T2034	(Baba et al., 2006)
N4235	*rpoB*35* Δ*relA251*$_{kan}$ Δ*spoT207*$_{cm}$ Lab number T1098	(Trautinger et al., 2005)
JW0198-1	BW25113 Δ*yafC727*$_{kan}$ Lab number T2791	(Baba et al., 2006)
JW0389-1	BW25113 Δ*phoB763*$_{kan}$ Lab number T2034	(Baba et al., 2006)
JW0390-2	BW25113 Δ*phoR764*$_{kan}$ Lab number T2034	(Baba et al., 2006)
U255	T1241 *ydcI-3xFLAG*$_{kan}$	T1241/pKD46 x OA824/OA825 (pSUB11)
U4	T1241 Δ*lacZ*	(Yilmaz et al., 2020)
U261	U4 Δ*hns*$_{FRT}$	lab collection
U451	U4 Δ*yafC*$_{kan}$	U4 x P1vir [JW0198-1 Δ*yafC*$_{kan}$]
U456	U4 Δ*yafC*$_{FRT}$	U451 x pCP20
U503	U4 Δ*ydcI*$_{kan}$	U4 x P1vir [JW5226-1 Δ*ydcI*$_{kan}$]
U505	U4 Δ*ydcI*$_{FRT}$	U503 x pCP20
U556	U505 Δ*hns*$_{kan}$	U505 x P1vir [S3754 Δ*hns*$_{KD4\text{-}kan}$]
U558	U505 Δ*hns*$_{FRT}$	U556 x pCP20
U559	U4 Δ*stpA::tet*	U4 x P1vir [M182 *stpA::tet*]
U560	U505 Δ*stpA::tet*	U505 x P1vir [M182 *stpA::tet*]
U561	U4 Δ*stpA::tet* Δ*hns*$_{kan}$	U559 x P1vir [S3754 Δ*hns*$_{KD4\text{-}kan}$]
U562	U505 Δ*stpA::tet* Δ*hns*$_{kan}$	U560 x P1vir [S3754 Δ*hns*$_{KD4\text{-}kan}$]
U563	U4 Δ*stpA::tet* Δ*hns*$_{FRT}$	U561 x pCP20
U564	U505 Δ*stpA::tet* Δ*hns*$_{FRT}$	U562 x pCP20
U571	U4 Δ*relA*$_{kan}$	U4 x P1vir [N4235 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$]
U572	U505 Δ*relA*$_{kan}$	U505 x P1vir [N4235 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$]
U573	U4 Δ*phoB*$_{kan}$	U4 x P1vir [JW0389-1 Δ*phoB*$_{kan}$]
U574	U505 Δ*phoB*$_{kan}$	U505 x P1vir [JW0389-1 Δ*phoB*$_{kan}$]
U575	U4 Δ*phoR*$_{kan}$	U4 x P1vir [JW0390-2 Δ*phoR*$_{kan}$]
U576	U505 Δ*phoR*$_{kan}$	U505 x P1vir [JW0390-2 Δ*phoR*$_{kan}$]
U577	U4 *ydcI-3xFLAG*$_{kan}$	U4 x P1vir [U255 *ydcI-3xFLAG*$_{kan}$]
U578	U4 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$	U571 x P1 vir [N4235 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$]
U579	U505 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$	U572 x P1 vir [N4235 Δ*relA*$_{kan}$ Δ*spoT*$_{cm}$]
U580	U4 Δ*phoB*$_{FRT}$	U573 x pCP20
U581	U505 Δ*phoB*$_{FRT}$	U574 x pCP20
U582	U4 Δ*phoR*$_{FRT}$	U575 x pCP20

Table S2: Bacterial strain and relevant genotype		
Strains	Genotype / description[a]	Construction /reference[b]
U583	U505 Δ*phoR*$_{FRT}$	U576 x pCP20
U584	U4 *ydcI-3xFLAG*$_{FRT}$	U577 x pCP20
U613	U4 *spoT*$_{E319Q}$ zib563::Tn10	U4 x P1vir [AB037 *spoT*$_{E319Q}$ zib563::Tn10]
U614	U505 *spoT*$_{E319Q}$ zib563::Tn10	U505 x P1vir [AB037 *spoT*$_{E319Q}$ zib563::Tn10]
U615	*U4* Δ*relA*$_{kan}$ *spoT*$_{E319Q}$ zib563::Tn10	U571 x P1vir [AB037 *spoT*$_{E319Q}$ zib563::Tn10]
U616	U505 Δ*relA*$_{kan}$ *spoT*$_{E319Q}$ zib563::Tn10	U572 x P1vir [AB037 *spoT*$_{E319Q}$ zib563::Tn10]
U621	T1241 Δ*phoB*-Δ*phoR*$_{kan}$	T1241/pKD46 x OB483/OB484 (pKD13)
U622	U4 Δ*phoB*-Δ*phoR*$_{kan}$	U4 x P1vir [U621 Δ*phoB*-Δ*phoR*$_{kan}$]
U623	U505 Δ*phoB*-Δ*phoR*$_{kan}$	U505 x P1vir [U621 Δ*phoB*-Δ*phoR*$_{kan}$]
U624	U4 Δ*phoB*-Δ*phoR*$_{FRT}$	U622 x pCP20
U625	U505 Δ*phoB*-Δ*phoR*$_{FRT}$	U623 x pCP20

[a]Abbreviations and genetic features are as follow: kanamycin resistance (kan); chloramphenicol resistance (cm); FLP recombinase target site (FRT).

[b] Strains which were constructed by transduction is stated as "x phage [donor strain]"; λ Red recombineering, stated as "x PCR primer pair (template)" followed by Flp recombinase-catalyzed deletion of the resistance marker, "x pCP20".

Table S3: Plasmids		
Plasmid	Features*	Reference
pCP20	*cI*$_{857}$ *P*$_R$ *flp* in pSC101-ori rep^{ts} Amp	(Cherepanov and Wackernagel, 1995)
pKD4	FRT-*kan*-FRT oriRγ Amp	(Datsenko and Wanner, 2000)
pKD46	*araC Para γ-β-exo* in pSC101-ori rep^{ts} Amp	(Datsenko and Wanner, 2000)
pET22b(+)	*P*$_{T7}$ MCS His6 in pMB1-ori Amp	Novagen®
pSUB11	*3xFLAG* in pKD4	(Shimada et al., 2005)
pKESL317	*ydcI*$_{His6}$ (full-length) in pET-22b(+)	PCR OA865/OA866 of T1241 in pET22b(+) (NdeI, XhoI)
pKETS24	*lacI*^{q} *P*$_{UV5}$ MCS in pSC-ori Cm	(Fragel et al., 2019)
pKECY26	*mNeonGreen* in pSC-ori Cm	mNeonGreen fragment in pKETS24
pKECY49	MCS *lacZ* in pSC-ori Cm	MCS (OA899/OA900) *lacZ* in pKECY26
pKERBR091	*P*$_{iraP}$ *lacZ* in pSC-ori Cm	*P*$_{iraP}$ (PCR OB310/ OB311) in pKECY49
pKERBR0104	*P*$_{phoA}$ *lacZ* in pSC-ori Cm	*P*$_{phoA}$ (PCR OB310/ OB313) in pKECY49
pKERBR093	*P*$_{ydcI}$ *lacZ* in pSC-ori Cm	*P*$_{ydcI}$ (PCR OB314/ OB315) in pKECY49
pKERBR114	*P*$_{ydcI}$ *lacZ* in pSC-ori Tet	*tetR* (PCR OB601/OB602) from pSC101 in pKERBR93
pKERBR115	*P*$_{iraP}$ *lacZ* in pSC-ori Tet	*tetR* (PCR OB601/OB602) from pSC101 in pKERBR91
pKERBR117	*P*$_{ddlA}$ *lacZ* in pSC-ori Cm	*P*$_{ddlA}$ (PCR OB614/ OB615) in pKECY49
pKERBR090	*P*$_{lldP}$ *lacZ* in pSC-ori Cm	*P*$_{lldP}$ (PCR OB308/ OB309) in pKECY49
pKERBR089	*P*$_{yobA}$ *lacZ* in pSC-ori Cm	*P*$_{yobA}$ (PCR OB306/ OB307) in pKECY49
pKERBR113	*P*$_{mqo}$ *lacZ* in pSC-ori Cm	*P*$_{mqo}$ (PCR OB499/ OB500) in pKECY49
pKERBR112	*P*$_{yhdV}$ *lacZ* in pSC-ori Cm	*P*$_{yhdV}$ (PCR OB485/OB486) in pKECY49
pKERBR111	*P*$_{gltA}$ *lacZ* in pSC-ori Cm	*P*$_{gltA}$ (PCR OB422/ OB423) in pKECY49

Table S3: Plasmids		
Plasmid	Features*	Reference
pKERBR105	P_{sdhC} *lacZ* in pSC-ori Cm	P_{sdhC} (PCR OB405/ OB406) in pKECY49
pKERBR122	P_{nhaA} *lacZ* in pSC-ori Cm	P_{nhaA} (PCR OB700/ OB701) in pKECY49
pKERBR123	P_{gltP} *lacZ* in pSC-ori Cm	P_{gltP} (PCR OB702/ OB703) in pKECY49

Abbreviations and features are as follow: chloramphenicol resistance (Cm); tetracycline resistance (Tet); ampicillin resistance (Amp); MCS for multiple cloning site; rep^{ts} for temperature sensitive plasmid replication; His6 for 6x histidine tag; 3x FLAG for DYKDHDGDYKDHDIDYKDDDDK (Sigma-Aldrich, Germany).

Table S4: Oligonucleotides and DNA fragments		
Number	Sequence*	Description
OA824	CCAGTTGATGATTAACGCTATTCGAAAATCAATGCCGTTC gactacaaagaccatgacggtgattata	Cloning of *ydcI-3xFLAG*$_{kan}$
OA825	AAATGTTCAAGGAGGGATCGACAGATCCCTTCACCTTcatatga atatcctccttagttcctattcc	Cloning of *ydcI-3xFLAG*$_{kan}$
OB483	GACGTCGCATTAATGATCGCAACCTATTTATTACAACAGGG CAAATCATGattccggggatccgtcgacc	Cloning of Δ*phoB*-Δ*phoR*$_{kan}$
OB484	TTATGGCAATAAAAGATGACAAAGGCGGATTAATCGCTGTTT TTGGCAATtgtaggctggagctgcttcga	Cloning of Δ*phoB*-Δ*phoR*$_{kan}$
OA865	tcaccatATGGAAAAAAATAGTCTGTTTAGTCAGCG	Cloning *ydcI*$_{His6}$ under P_{tac}
OA866	tcacctcgagGAACGGCATTGATTTTCGAATAGCG	Cloning *ydcI*$_{His6}$ under P_{tac}
OA899	5'P-TCGACATCTGCAGACT	Cloning MCS in pKECY26
OA900	5'P-CTAGAGTCTGCAGATG	Cloning MCS in pKECY26
OB310	gcaggtcgacTGATTTACCACCAAAAACGATTCCTA	*iraP, phoA* promoter *lacZ* fusion
OB311	gcagtctagaAAATAACAACTCAGCAATGAGATTTTTCAT	*iraP* promoter *lacZ* fusion
OB313	gcagtctagaGAGTGCCAGTGCAATAGTGCTTTG	*phoA* promoter *lacZ* fusion
OB314	gcagtctagaGCGCTGACTAAACAGACTATTTTTTTCC	*ydcI* promoter *lacZ* fusion
OB315	gcaggtcgacAATCTCATCCGCCGTGATGCT	*ydcI* promoter *lacZ* fusion
OB601	gcagagatctTTCTCATGTTTGACAGCTTATCATCG	*tetR* insertion in pKERBR91 and pKERBR93
OB602	gcagggcgcgccATGGCGGACGCGATGGA	*tetR* insertion in pKERBR91 and pKERBR93
OB614	gcagtctagaTGATTTACCACCAAAAACGATTCCTA	*ddlA* promoter *lacZ* fusion
OB615	gcaggtcgacAAATAACAACTCAGCAATGAGATTTTTCAT	*ddlA* promoter *lacZ* fusion
OB306	gcagtctagaGTAGCGAAGGGAGCGTGCAGT	*yobA* promoter *lacZ* fusion
OB307	gcaggtcgacTGTTTGATCCAGTTTAGCCAGATTCTT	*yobA* promoter *lacZ* fusion
OB308	gcaggtcgacCAGGAGATGACTGGGTTTTCAAAG	*lldP* promoter *lacZ* fusion
OB309	gcagtctagaGGGATCGTAGTTTTGTTGCCAGAGAT	*lldP* promoter *lacZ* fusion
OB499	gcagtctagaCGAGAAGAGCATGGCAGTCACTTT	*mqo* promoter *lacZ* fusion
OB500	gcaggtcgacTGCTGTTTCTGATTGGCGGC	*mqo* promoter *lacZ* fusion
OB485	gcaggtcgacTAAATCAGAAACATAAAGGCGCTTTC	*yhdV* promoter *lacZ* fusion
OB486	gcagtctagaAGCCAGTGTGGCTATCATCATCTT	*yhdV* promoter *lacZ* fusion
OB422	gcaggtcgacAGGTCTTTGTTTTTTCACATTTCTTATCAT	*gltA* promoter *lacZ* fusion
OB423	gcagtctagaGAGGGTGAGTTTTGCTTTTGTATCAG	*gltA* promoter *lacZ* fusion
OB405	GCAGgtcgacGAGGGTGAGTTTTGCTTTTGTATCAG	*sdhC* promoter *lacZ* fusion
OB406	gcagtctagaAGGTCTTTGTTTTTTCACATTTCTTATCAT	*sdhC* promoter *lacZ* fusion
OB700	gcaggtcgacAGTAAGAGGCACTCTACATGTGTTCAG	*nhaA* promoter *lacZ* fusion
OB701	gcagtctagaACTGCTAAAGAATCGATGCAGATGTTTC	*nhaA* promoter *lacZ* fusion
OB702	gcaggtcgacGCGAAATTGTTGCAGCGGAGATTGGA	*gltP* promoter *lacZ* fusion
OB703	gcagtctagaCCAGGCCAGGCTGAATTTTATATTTTT	*gltP* promoter *lacZ* fusion
OB324	6-FAM-CGGGGGATTCGCGGCCTT	*iraP* EMSA frag. A*
OB325	CGGGGGATTCGCGGCCTT	*iraP* EMSA frag. A

Table S4: Oligonucleotides and DNA fragments		
Number	Sequence*	Description
OB326	TTCTCTTCGTTAACTTATAATTAAGAGAAAAAAAC	*iraP* EMSA frag. A*/A
OB327	6-FAM-AGCTAATTTATTCCTGGCTTAAATGGC	*iraP* EMSA frag. B*/B1*/ B1.1*/M1*/M2*/M1-M2*
OB328	AGCTAATTTATTCCTGGCTTAAATGGC	*iraP* EMSA frag. B/B1/ B1.1/M1/M2/M1-M2
OB329	ACTTTGCGCAAATATTTTTTGTTTGATTACCA	*iraP* EMSA frag. B*/B/B3
OB330	6-FAM-TTTTGTGTGGGTAATTATTTAAATAAGAGAAAC	*iraP* EMSA frag. C*
OB331	TTTTGTGTGGGTAATTATTTAAATAAGAGAAAC	*iraP* EMSA frag. C
OB332	AAATAACAACTCAGCAATGAGATTTTTCAT	*iraP* EMSA frag. C*/C
OB334	AACCTTTATTATAAATAAAAGTTATGAAATATATTCTCT	*iraP* EMSA frag. B1
OB335	6-FAM-CTCTTAATTATAAGTTAACGAAGAGAATATATTTCA	*iraP* EMSA frag. B2*
OB336	CTCTTAATTATAAGTTAACGAAGAGAATATATTTCA	*iraP* EMSA frag. B2
OB337	CTTATTTAAATAATTACCCACACAAAATATAGGC	*iraP* EMSA frag. B2*/B2
OB338	6-FAM-TTTATAATAAAGGTTGATAATTAAAAGCCTATATTT	*iraP* EMSA frag. B3*
OB339	TTTATAATAAAGGTTGATAATTAAAAGCCTATATTT	*iraP* EMSA frag. B3
OB618	CACAAAATATAGGCTTTTAATTATCAACCTTTATTA	*iraP* EMSA frag. B1.1/M1/M2/M1-M2
OB321	6-FAM-TAGCGAAGGGAGCGTGCAGT	*yobA* EMSA frag. D*
OB340	TAGCGAAGGGAGCGTGCAGT	*yobA* EMSA frag. D
OB307	GCAGgtcgacTGTTTGATCCAGTTTAGCCAGATTCTT	*yobA* EMSA frag. D*/D
OB343	6-FAM-CTGGGTACAGAGCTCTGGGCG	*gltA* EMSA frag. K*
OB344	CTGGGTACAGAGCTCTGGGCG	*gltA* EMSA frag. K
OB345	ACAAACATTACCAGGAAAAGCATATAATG	*gltA* EMSA frag. K*/K/K1*/K1/L*/L
OB616	6-FAM-CGAATGAATTGGTCAATACTTCCACA	*gltA* EMSA frag. K1*
OB617	CGAATGAATTGGTCAATACTTCCACA	*gltA* EMSA frag. K1
OB346	6-FAM-TAGGTGAAATACCGACTTCATAACTTTTAC	*gltA* EMSA frag. L*
OB347	TAGGTGAAATACCGACTTCATAACTTTTAC	*gltA* EMSA frag. L
OB348	GTTCCGGAGACCTGGCGG	*gltA* EMSA frag. L*/L
OB501	6-FAM-CGAGAAGAGCATGGCAGTCACTT	*mqo* EMSA frag. 12*
OB502	CGAGAAGAGCATGGCAGTCACTT	*mqo* EMSA frag. 12
OB503	CCGGACGGCATAATATTACGAC	*mqo* EMSA frag. 12*/12
OB504	6-FAM-AGCATATTGACCTGACGGCAGC	*mqo* EMSA frag. 14*
OB505	AGCATATTGACCTGACGGCAGC	*mqo* EMSA frag. 14
OB506	TCCTGCTGGATGAATGGGC	*mqo* EMSA frag. 14*/14
OB507	6-FAM-CCTGATAAAACTCACGACGGAAGTG	*mqo* EMSA frag. 16*
OB508	CCTGATAAAACTCACGACGGAAGTG	*mqo* EMSA frag. 16
OB509	CTGTTTTCGGCAGTGTTTACCGAT	*mqo* EMSA frag. 16*/16
OB510	6-FAM-ACCCTCCGGCCCCAGC	*mqo* EMSA frag. 18*
OB511	ACCCTCCGGCCCCAGC	*mqo* EMSA frag. 18
OB512	TGCTGTTTCTGATTGGCGGC	*mqo* EMSA frag. 18*/18
OB487	6-FAM-TAAATCAGAAACATAAAGGCGCTTTC	*yhdV* EMSA frag. 3*
OB488	TAAATCAGAAACATAAAGGCGCTTTC	*yhdV* EMSA frag. 3
OB489	TTCATGGTGTTACCTCAGTAATAAACGC	*yhdV* EMSA frag. 3*/3
OB490	6-FAM-CGTTAATAAGAGAGTTAACTTATTAAGCGTTAGC	*yhdV* EMSA frag. 5*
OB491	CGTTAATAAGAGAGTTAACTTATTAAGCGTTAGC	*yhdV* EMSA frag. 5
OB492	GGCAATAGCCCCAGCCGT	*yhdV* EMSA frag. 5*/5
OB493	6-FAM-ACCAATACCTGTATGAAAGGCACGA	*yhdV* EMSA frag. 7*
OB494	ACCAATACCTGTATGAAAGGCACGA	*yhdV* EMSA frag. 7
OB495	CGAAATAACGCGATGTAACATGC	*yhdV* EMSA frag. 7*/7

Table S4: Oligonucleotides and DNA fragments		
Number	Sequence*	Description
OB496	6-FAM-ATTGCTATATTCTGCTGGGTAATTCCC	*yhdV* EMSA frag. 9*
OB497	ATTGCTATATTCTGCTGGGTAATTCCC	*yhdV* EMSA frag. 9
OB498	CTGAGTGTAGACAGGAAATAATAAAAACAGC	*yhdV* EMSA frag. 9*/9
OB349	6-FAM-CAGACATCTCCCCCCACAAGA	*lldP* EMSA frag. M*
OB350	CAGACATCTCCCCCCACAAGA	*lldP* EMSA frag. M
OB351	GAGTCTGTTGCTCATCTCCTTGTCAC	*lldP* EMSA frag. M*/M
OB515	6-FAM-ATAGGAGGCCTCGGGTTGATG	*nhaA* EMSA frag. I*
OB516	GCCAATATTTCAAAAATAATCCAGCTT	*nhaA* EMSA frag. I*/I
OB517	ATAGGAGGCCTCGGGTTGATG	*nhaA* EMSA frag. I
OB518	6-FAM-TTAATTTATCTCTTCATCGCAATTATTGAC	*nhaA* EMSA frag. II*
OB519	CAGATTATAGTGTAGGAATAATTCGTTTTTTAC	*nhaA* EMSA frag. II*/II
OB520	TTAATTTATCTCTTCATCGCAATTATTGAC	*nhaA* EMSA frag. II
OB521	6-FAM-CTGATGGCGCAAATTCTTCAATAG	*nhaA* EMSA frag. III*
OB522	CAGGATAGCGGCAATGATAAGAATAA	*nhaA* EMSA frag. III*/III
OB523	CTGATGGCGCAAATTCTTCAATAG	*nhaA* EMSA frag. III
OB663	6-FAM-TTTTTGTGGCGTAATTTTTTAATGATTTAA	*nhaA* EMSA frag. II.1*/II.3*
OB664	GCTCACGAGTGACCCTAAG	*nhaA* EMSA frag. II.1*/II.1
OB665	TTTTTGTGGCGTAATTTTTTAATGATTTAA	*nhaA* EMSA frag. II.1/II.3
OB666	6-FAM-AATCCCTCAGGAATCCTCAC	*nhaA* EMSA frag. II.2*
OB667	TCTCTTTCAGGTGAATAGATCGTTT	*nhaA* EMSA frag. II.2*/II.3*/II.2/II.3
OB668	AATCCCTCAGGAATCCTCAC	*nhaA* EMSA frag. II.2
OB524	6-FAM-TTCAATATGTCCGTACATGGAATAATAAAG	*wrbA* EMSA frag. IV*
OB525	TGTGTCATTTTGCGACAAAATTACG	*wrbA* EMSA frag. IV*/IV
OB526	TTCAATATGTCCGTACATGGAATAATAAAG	*wrbA* EMSA frag. IV
OB527	6-FAM-ACCACTTATTTGATTTATAACAACTTTCACAA	*wrbA* EMSA frag. V*
OB528	TAGACATCATATTGCATAGAATTAACGCG	*wrbA* EMSA frag. V*/V
OB529	ACCACTTATTTGATTTATAACAACTTTCACAA	*wrbA* EMSA frag. V
OB530	6-FAM-GCCCCGCTGTGGCTCTTG	*dtpA* EMSA frag. VI*
OB531	CGACAGCATCGTGATCTTGATCAC	*dtpA* EMSA frag. VI*/VI
OB532	GCCCCGCTGTGGCTCTTG	*dtpA* EMSA frag. VI
OB533	6-FAM-AATTGTTTTTTTAATAGCGAAATTTTCTATTC	*dtpA* EMSA frag. VII*
OB534	ATCCAGCGTTAAATTTTTCAAAAAAC	*dtpA* EMSA frag. VII*/VII
OB535	AATTGTTTTTTTAATAGCGAAATTTTCTATTC	*dtpA* EMSA frag. VII
OB536	6-FAM-CCCCTCTCAGTGCAGTGAAAAAA	*dtpA* EMSA frag. VIII*
OB537	GTTGGTTTTTGGTTTGCAGTGGA	*dtpA* EMSA frag. VIII*/VIII
OB538	CCCCTCTCAGTGCAGTGAAAAAA	*dtpA* EMSA frag. VIII
OB539	6-FAM-TGTACCTCGTTTAACCCGAAATCTGT	*yobF* EMSA frag. IX*
OB540	AGCAAAAAATGTGATATTGCACGC	*yobF* EMSA frag. IX*/IX
OB541	TGTACCTCGTTTAACCCGAAATCTGT	*yobF* EMSA frag. IX
OB542	6-FAM-GTTGTTCCAGGCCAGTGGGA	*yobF* EMSA frag. X*
OB543	TTACCCCATTAATTTATTAGGATGTTTACATC	*yobF* EMSA frag. X*/X
OB544	GTTGTTCCAGGCCAGTGGGA	*yobF* EMSA frag. X
OB545	6-FAM-AAATAATACCACGCTTAATCACAAATCC	*yobF* EMSA frag. XI*
OB546	TGAAGACGAAGATATTATTCGTCTGGT	*yobF* EMSA frag. XI*/XI
OB547	AAATAATACCACGCTTAATCACAAATCC	*yobF* EMSA frag. XI
OB548	6-FAM-GAAACTAAAAACTCTTTTTGTTGATTGAGACA	*yicG* EMSA frag. XII*
OB549	TCCCTAAAAATCTGTCTGGCAGG	*yicG* EMSA frag. XII*/XII
OB550	GAAACTAAAAACTCTTTTTGTTGATTGAGACA	*yicG* EMSA frag. XII
OB551	6-FAM-TGAAATCACCGCCCTGATGC	*gltP* EMSA frag. XIII*

Table S4: Oligonucleotides and DNA fragments		
Number	Sequence*	Description
OB552	ATTTTTGGCATATGTTCATTCGATTAT	*gltP* EMSA frag. XIII*/XIII
OB553	TGAAATCACCGCCCTGATGC	*gltP* EMSA frag. XIII
OB554	6-FAM-TGATGAAGATTATGTCATTGAGCTGC	*gltP* EMSA frag. XIV*
OB555	ATCAGTATCATTTGCGCGGTATGT	*gltP* EMSA frag. XIV*/XIV
OB556	TGATGAAGATTATGTCATTGAGCTGC	*gltP* EMSA frag. XIV
OB557	6-FAM-ATGCAGCGAAAAGCTCTGTTGTTA	*gltP* EMSA frag. XV*
OB558	GGCCAGGCTGAATTTTATATTTTTCA	*gltP* EMSA frag. XV*/XV
OB559	ATGCAGCGAAAAGCTCTGTTGTTA	*gltP* EMSA frag. XV
OB446	6-FAM-TCGCGCTGAAATGCGCTA	*ycbJ* EMSA frag. P*
OB447	GGCTTAATTCGGCACGCAG	*ycbJ* EMSA frag. P*/P
OB448	TCGCGCTGAAATGCGCTA	*ycbJ* EMSA frag. P
OB440	6-FAM-TTATCGTGAAATATCCATGATGTTGC	*putP* EMSA frag. N*
OB441	ATTTACATTAAATTTTAGTGCTCAGCGAC	*putP* EMSA frag. N*/N
OB442	TTATCGTGAAATATCCATGATGTTGC	*putP* EMSA frag. N
OB443	6-FAM-TGTGAGAGAGTGCAACCTGATGAA	*putP* EMSA frag. O*
OB444	AGTCTCCAAAAAATTATTATCGGCAA	*putP* EMSA frag. O*/O
OB445	TGTGAGAGAGTGCAACCTGATGAA	*putP* EMSA frag. O
OB449	6-FAM-GGTATGGAACTGGATTTTCTTCTGAA	*isrC* EMSA frag. Q*
OB450	GCATGTTACTGACAAAGGTTATTATTAGCTT	*isrC* EMSA frag. Q*/Q
OB451	GGTATGGAACTGGATTTTCTTCTGAA	*isrC* EMSA frag. Q
OB452	6-FAM-GAGCTGCCGTTTTTTTATTCTGTCA	*raiA* EMSA frag. R*
OB453	TAAATTTTACCTCTTGTCTTCCCGTCT	*raiA* EMSA frag. R
OB454	GAGCTGCCGTTTTTTTATTCTGTCA	*raiA* EMSA frag. R*/R
OB455	6-FAM-CCTGAAAGACTACCAGGTGCTGC	*tdcA* EMSA frag. S*/S
OB456	TACTTTAACTTGTTGATATTTAAAGGTATTTAATTGTAAT	*tdcA* EMSA frag. S*
OB457	CCTGAAAGACTACCAGGTGCTGC	*tdcA* EMSA frag. S
OB458	6-FAM-ATACTTTCCAGAGTATCGTTATTACAATTAAATACC	*tdcA* EMSA frag. T*
OB459	GGTAATTATCCGGTAATTGCTTGAAA	*tdcA* EMSA frag. T*/T
OB460	ATACTTTCCAGAGTATCGTTATTACAATTAAATACC	*tdcA* EMSA frag. T
OB461	6-FAM-CTCCCTGTTTCAACAATCATCCTATTT	*ubiU* EMSA frag. U*
OB462	AATTTTTCCTGAAGGTTGCGCTA	*ubiU* EMSA frag. U*/U
OB463	CTCCCTGTTTCAACAATCATCCTATTT	*ubiU* EMSA frag. U
OB464	6-FAM-CTACTCCCTTTACGCTCACCGC	*nikA* EMSA frag. V*
OB465	GCGCAAATAGAGTGCGGCG	*nikA* EMSA frag. V*/V
OB466	CTACTCCCTTTACGCTCACCGC	*nikA* EMSA frag. V
OB473	6-FAM-TTTCCTTTTATGCCGAAGGTGATG	*yjjI* EMSA frag. Y*
OB474	TTCTAACATCTACTCCGCAACACACTTC	*yjjI* EMSA frag. Y*/Y
OB475	TTTCCTTTTATGCCGAAGGTGATG	*yjjI* EMSA frag. Y
OB467	6-FAM-CTTGTAGGATGTTAATAATATGTGGCATAAGC	*nrfA* EMSA frag. W*
OB468	ATTGTAAGTGCATGTAAAATACCACTTTAGAG	*nrfA* EMSA frag. W*/W
OB469	CTTGTAGGATGTTAATAATATGTGGCATAAGC	*nrfA* EMSA frag. W
OB470	6-FAM-TGTTGAAAAAGAGGAAGATACTGACTAACTC	*nrfA* EMSA frag. X*
OB471	AGGGGCTTCATCCGAATTGC	*nrfA* EMSA frag. X*/X
OB472	TGTTGAAAAAGAGGAAGATACTGACTAACTC	*nrfA* EMSA frag. X
Double stranded fragments		
M1	AGCTAATTTATTCCTGGCTTAAATGGCCATGCGGTGAGTTTTT TTCTCTCGGTTATAAGCCGACGAAGAGAATATATTTCATAACT TTTATTTATAATAAAGGTTGATAATTAAAAGCCTATATTTTGT GGCCCG	mutant M1, upstream of *iraP*

Table S4: Oligonucleotides and DNA fragments		
Number	Sequence*	Description
M2	AGCTAATTTATTCCTGGCTTAAATGGCCATGCGGTGAGTTTTTTTCTCTTAATTATAAGTTAACGAAGAGAATATATTTCACGGCTTTTATCCGTAATAAAGGTTGATAATTAAAAGCCTATATTTTGTG	mutant M2, upstream of *iraP*
M1-M2	AGCTAATTTATTCCTGGCTTAAATGGCCATGCGGTGAGTTTTTTTCTCTCGGTTATAAGCCGACGAAGAGAATATATTTCACGGCTTTTTATCCGTAATAAAGGTTGATAATTAAAAGCCTATATTTTGTG	mutant M1-M2, upstream of *iraP*

Oligonucleotide sequences are given in 5' to 3' direction. Homologous sequences to the aligned and indicated regions cloned are shown in capital letters and the sites for restriction endonucleases are underlined. The mutations of the fragments are indicated with red letters.

Main Discussion

In this thesis, the regulatory and functional characterization of LysR-type transcription regulators in *E. coli* was elucidated with a special focus on YdcI. The approaches undertaken in this work, as decribed in the next section, have helped in the identification of autoregulation, binding specificity and expression pattern for a subset of LTTRs. Additionally, the complex regulatory network and functional role of YdcI was deciphered using previously available ChIP-exo data (Gao et al., 2018), comparative RNA-seq analysis, *in vitro* DNA binding analyses and *in-vivo lacZ* reporter expression analysis as described in the later section. Taken together, this work has revealed the complexity of characterizing the unknown LTTRs and added novel insights into understanding the regulatory and functional role of YdcI in *E. coli*.

Approaches to characterize LTTRs

The approaches to characterize the partially characterized or uncharacterized LTTRs are not well streamlined or classified till date. This entails the scope of understanding the regulatory and functional role of the LTTRs by using a systematic characterization approach, starting with studying the expression pattern of the LTTRs in response to different conditions and narrowing down the optimal growth conditions needed for their expression. In this work, Western blot analysis has revealed the expression of 4 out of 17 LTTRs under standard growth conditions (Figure 2, Chapter 1). A combination of knowledge available from EcoCyc, UniProt, functional genomics and transcriptomics databases can be helpful to predict the conditions for expression of the different LTTRs. After gaining an idea about the potential conditions inducing the expression of a LTTR, it could be plausible to test if the identified growth conditions or abiotic stresses (pH, temperature, osmotic, nutrient stresses) could modulate the expression of the LTTRs. In addition, *lacZ* reporter screen of previously identified target loci (for partially characterized LTTRs) could be used to identify the different conditions (pH, temperature) which can induce the expression and/or activity of the LTTR.

Another approach to characterize the LTTRs would be to identify the target genes to further shed light on the involvement of LTTRs in complex gene regulatory networks and important metabolic pathways in *E. coli*, using physiological conditions under which the LTTR is expressed. The unique tetrameric structure of the LTTRs with the dimeric N-terminal DNA-binding and the dimeric C-terminal regulatory domain connected via a helical linker helix and a flexible hinge (Maddocks and Oyston, 2008, Momany and Neidle, 2012) confers an added advantage to study the binding and interaction of the LTTRs with their target genes. This makes it promising to use techniques like ChIP-seq and ChIP-exo, which allow identification of all possible binding locations and thus, the potential targets of these DNA-binding

transcription regulators. Despite the advantage conferred by these genome wide high-throughput sequencing techniques, the functional characterization of the LTTRs without prior knowledge about the biological function of the LTTRs till date remains experimentally challenging as evident by the identification of the biological function of only 3 out of 16 and 10 out of 34 transcription regulators (Gao et al., 2021, Gao et al., 2018).

The use of candidate approach of studying autoregulation of the LTTRs as shown in this work can further shed light on the regulatory role of the LTTRs. Identification of the LTTRs which function as autoregulators resolves the problem of characterising LTTRs without a known target. The candidate approach could be followed up by *in vitro* DNA binding analyses like EMSA as shown in this work to further study the direct DNA binding mode of regulation of the LTTR acting as autoregulators. Additionally, a reporter screen using the *lttr* gene could be established to identify the growth conditions, which in turn induce the expression of the LTTRs. Furthermore, a genome library screening with a promoter *lacZ* reporter of the autoregulatory gene could as well help in identification of additional regulators of these LTTRs.

Other approaches to characterise LTTRs may include identification of the small molecule effectors, which can modulate the expression of the LTTRs. The identification of specific amino acid mutations which mimics an effector bound form has identified constitutively active mutants of LeuO (Fragel et al., 2019). Hence, identification of constitutively active mutants could be useful to screen for other target loci or even identify a metabolite or a structural modification which could modulate the activity of the partially characterized LTTRs.

In this work, 8 out of the 17 studied LTTRs have been shown to act as autoregulators and the binding of 4 of these LTTRs to their own promoter fragments has been demonstrated by EMSA (Table 1, Chapter 1). Additionally, the target genes of the 4 LTTRs (LrhA, YafC, YdcI and YhaJ) were identified by a comparative RNA-seq analysis (Chapter 1, section 1.4). Nevertheless, RNA-seq is an indirect approach of target identification and could lead to identification of non-specific targets of the LTTRs. Hence, the combinational approach of ChIP-exo and comparative RNA-seq analysis is a better alternative to identify the target genes and study the regulatory and functional role of the LTTRs as elucidated for the LTTR, YdcI (Chapter 2).

The regulatory and functional role of YdcI

In order to decipher the regulatory and functional role of YdcI in *E. coli*, a combination of ChIP-exo and comparative RNA-seq analysis was used in this work. The validation of 10 out of 16 ChIP-exo predicted targets of YdcI (Gao et al., 2018) by *in vitro* DNA binding EMSA analyses was followed by identification of a consensus DNA-binding motif for YdcI (Fig. 4, Chapter 2). This approach highlights the importance of further experimental studies needed for evaluation of the existing whole genome sequencing data to narrow down the specific targets and then only further characterize such partially characterized LTTRs. Another important step to study the regulatory function of YdcI, lies in the understanding of the conditions under which the protein is expressed. YdcI-3xFLAG expression analysis showed that the protein levels of YdcI increase approximately after 2 hours of exponential growth in LB and tryptone media (Fig. 10, Chapter 2). This work also showed that YdcI expression goes up upon a decrease in pH of the media, but the induction is independent of acid stress (as suggested by Schumacher et al., 2023) and is rather dependent on the growth phase (Fig. 9, Chapter 2). The time point of expression of the protein correlates with the entry of the cells from early to mid-exponential phase. In addition, the detected levels of YdcI-3xFLAG were very stable and constant under standard laboratory growth conditions. Without any degradation of the protein within cells, stable amount of this protein would be sufficient to regulate the expression of its target genes. However, the absence of detectable regulation for most of the identified targets of YdcI (other than *iraP* and *mqo*) suggests that suitable growth conditions or small molecule inducer for studying the regulation by YdcI yet remains to be identified.

The regulatory role of YdcI for one of its target, *iraP* shows the involvement of YdcI in the bacterial stress response. This work has additionally shown that YdcI presumably activates *iraP* by acting as an H-NS antagonist. In another study, *ydcI* expression has been reported to be controlled by σS in *Salmonella enterica* servar Typhimurium, a similar mechanism of σS mediated regulation of *ydcI* expression may exist in *E. coli*, but studies using P_{ydcI} *lacZ* reporter fusions in WT and Δ*rpoS* mutant remains to be performed to address this open question. Intriguingly, σS mediated regulation of *ydcI* gene, where YdcI in turn enhances *iraP* expression and increases levels of IraP stabilizing RpoS might unravel the existence of a novel feedback regulation in response to different physiological stresses in the cells. Additionally, this study showed RelA or SpoT mediated ppGpp induction of P_{ydcI} expression under phosphate starvation conditions as shown before for *iraP* (Bougdour and Gottesman, 2007), demonstrating a link between elicitation of the stringent stress response and/or the role of ppGpp as second messenger and the expression of *ydcI* gene. There is a reported increase in ppGpp levels in response to glucose starvation (Metzger et al., 1989, Xiao et al., 1991). Moreover, the transcription

activation by ppGpp requires DksA as a coregulator upon amino acid starvation (Paul et al., 2005). Hence, studies aimed to identify other starvation conditions or co-regulators that might result in similar ppGpp mediated control of the *iraP* and *ydcI* genes could possibly be helpful to further understand other stressors dependent transcriptional regulation of YdcI and its target gene, *iraP*.

The functional role of YdcI seems to be additionally linked to regulation of important metabolic pathways in *E. coli*. The target genes of *mqo*, *gltA* and its divergent *sdhCDAB* operon encode enzymes of the TCA cycle, *yobA-yebZ-yebY* is associated with copper delivery function to NADH dehydrogenase-II (NDH-2) of the bacterial electron transport chain. These key targets suggest a potential role of YdcI in regulating the carbon flux operating in the TCA cycle under specific physiological conditions. Another subset of targets identified for YdcI include several symporters or proton transporters (*lldP*, *gltP*), inner membrane glycine transporter (*yicG*) which indicates a potential involvement of YdcI in regulating the transport of metabolites or ions across the cells. Another target gene, *yobF*, is induced upon heat shock and important for acid stress response in *E. coli* (Hobbs et al., 2010), which emphasizes on the potential role of YdcI in heat and acid stress. A possible link between the involvement of YdcI in acid stress reponse and regulation of biochemical pathway (TCA cycle) could be hypothesized as a response of the cells to severe acid or alkaline stress condition by reducing the metabolic activity within the cells to help them enter into a dormant state. Therefore, different *in vivo* or *in vitro* biochemical assays and functional studies could be performed in the future to unravel the connection of the functional role of YdcI in different stress responses and metabolic pathways and decipher the physiological relevance of this LTTR.

Alternative approaches to characterise the LTTRs

The alternative approaches which can be used to characterize the LTTRs in the future could focus on undertaking a forward genetics approach, rather than the reverse genetic approach used in this work. This approach could involve generating a LTTR mutant library consisting of deletion strains of different *lttr* alleles. The growth analysis of the wide array of mutants along with the wild-type as a control in different growth conditions (modulating the nutrient availability) or on exposure to different abiotic stresses (heat, oxidative, pH, temperature) could help in the identification of a possible phenotype of some of the partially or uncharacterized LTTRs in *E. coli*. This screen can be correlated to the available transcriptome data of *E. coli* K-12 in response to different stressors (heat, cold, oxidative stress, pH, nitrosative stress, phosphate stress), which would in turn give a comprehensive overview of the genes involved in regulating important biochemical pathways or deciphering the complex transcription regulatory network in *E. coli*. The advancement in mutagenic screening tools, genomics technology and availability of genome wide high throughput sequencing (ChIP-exo/ ChIP seq) and transcriptome profile (RNA-seq) has opened up a new dimension of techniques for the ease of characterization of the LTTRs. Hence, prior knowledge about the possible phenotype of a LTTR or biochemical pathways where a LTTR plays a key role is a prerequisite and would help to resolve the challenges encountered in characterizing the LTTRs in *E. coli* in the future.

References

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., et al. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44, W3-W10. 10.1093/nar/gkw343.

Alanazi, A.M., Neidle, E.L., and Momany, C. (2013). The DNA-binding domain of BenM reveals the structural basis for the recognition of a T-N11-A sequence motif by LysR-type transcriptional regulators. Acta Crystallogr D Biol Crystallogr 69, 1995-2007. 10.1107/S0907444913017320.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166-169. 10.1093/bioinformatics/btu638.

Anderson, L.A., Palmer, T., Price, N.C., Bornemann, S., Boxer, D.H., and Pau, R.N. (1997). Characterisation of the molybdenum-responsive ModE regulatory protein and its binding to the promoter region of the modABCD (molybdenum transport) operon of Escherichia coli. Eur J Biochem 246, 119-126. 10.1111/j.1432-1033.1997.00119.x.

Aravind, L., Anantharaman, V., Balaji, S., Babu, M.M., and Iyer, L.M. (2005). The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev 29, 231-262. 10.1016/j.femsre.2004.12.008.

Aravind, L., and Koonin, E.V. (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. Trends Biochem Sci 23, 469-472. 10.1016/s0968-0004(98)01293-6.

Atkinson, G.C., Tenson, T., and Hauryliuk, V. (2011). The RelA/SpoT homolog (RSH) superfamily: distribution and functional evolution of ppGpp synthetases and hydrolases across the tree of life. PLoS One 6, e23479. 10.1371/journal.pone.0023479.

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. 10.1038/msb4100050.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2, 28-36.

Bange, G., Brodersen, D.E., Liuzzi, A., and Steinchen, W. (2021). Two P or Not Two P: Understanding Regulation by the Bacterial Second Messengers (p)ppGpp. Annu Rev Microbiol 75, 383-406. 10.1146/annurev-micro-042621-122343.

Banos, R.C., Martinez, J., Polo, C., Madrid, C., Prenafeta, A., and Juarez, A. (2011). The yfeR gene of Salmonella enterica serovar Typhimurium encodes an osmoregulated LysR-type transcriptional regulator. FEMS Microbiol Lett 315, 63-71. 10.1111/j.1574-6968.2010.02171.x.

Battesti, A., and Bouveret, E. (2006). Acyl carrier protein/SpoT interaction, the switch linking SpoT-dependent stress response to fatty acid metabolism. Mol Microbiol 62, 1048-1063. 10.1111/j.1365-2958.2006.05442.x.

Battesti, A., Majdalani, N., and Gottesman, S. (2011). The RpoS-mediated general stress response in Escherichia coli. Annu Rev Microbiol 65, 189-213. 10.1146/annurev-micro-090110-102946.

Baugh, A.C., Momany, C., and Neidle, E.L. (2023). Versatility and Complexity: Common and Uncommon Facets of LysR-Type Transcriptional Regulators. Annu Rev Microbiol 77, 317-339. 10.1146/annurev-micro-050323-040543.

Bennison, D.J., Irving, S.E., and Corrigan, R.M. (2019). The Impact of the Stringent Response on TRAFAC GTPases and Prokaryotic Ribosome Assembly. Cells 8. 10.3390/cells8111313.

Blumer, C., Kleefeld, A., Lehnen, D., Heintz, M., Dobrindt, U., Nagy, G., Michaelis, K., Emody, L., Polen, T., Rachel, R., et al. (2005). Regulation of type 1 fimbriae synthesis and biofilm formation by the transcriptional regulator LrhA of Escherichia coli. Microbiology (Reading) 151, 3287-3298. 10.1099/mic.0.28098-0.

Borries, A., Vogel, J., and Sharma, C.M. (2011). Differential RNA Sequencing (dRNA-Seq): Deep-Sequencing-Based Analysis of Primary Transcriptomes. In Tag-Based Next Generation Sequencing, pp. 109-121. 10.1002/9783527644582.ch7.

Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., and Rimsky, S. (2007). H-NS cooperative binding to high-affinity sites in a regulatory element results in transcriptional silencing. Nat Struct Mol Biol 14, 441-448. 10.1038/nsmb1233.

Bougdour, A., Cunning, C., Baptiste, P.J., Elliott, T., and Gottesman, S. (2008). Multiple pathways for regulation of sigmaS (RpoS) stability in Escherichia coli via the action of multiple anti-adaptors. Mol Microbiol 68, 298-313. 10.1111/j.1365-2958.2008.06146.x.

Bougdour, A., and Gottesman, S. (2007). ppGpp regulation of RpoS degradation via anti-adaptor protein IraP. Proc Natl Acad Sci U S A 104, 12896-12901. 10.1073/pnas.0705561104.

Bougdour, A., Wickner, S., and Gottesman, S. (2006). Modulating RssB activity: IraP, a novel regulator of sigma(S) stability in Escherichia coli. Genes Dev 20, 884-897. 10.1101/gad.1400306.

Bouillet, S., Hamdallah, I., Majdalani, N., Tripathi, A., and Gottesman, S. (2024). A negative feedback loop is critical for recovery of RpoS after stress in Escherichia coli. PLoS Genet 20, e1011059. 10.1371/journal.pgen.1011059.

Breddermann, H., and Schnetz, K. (2016). Correlation of Antagonistic Regulation of leuO Transcription with the Cellular Levels of BglJ-RcsB and LeuO in Escherichia coli. Frontiers in cellular and infection microbiology 6, 106. 10.3389/fcimb.2016.00106.

Breddermann, H., and Schnetz, K. (2017). Activation of leuO by LrhA in Escherichia coli. Mol Microbiol 104, 664-676. 10.1111/mmi.13656.

Brugger, C., Schwartz, J., Novick, S., Tong, S., Hoskins, J.R., Majdalani, N., Kim, R., Filipovski, M., Wickner, S., Gottesman, S., et al. (2023). Structure of phosphorylated-like RssB, the adaptor delivering sigma(s) to the ClpXP proteolytic machinery, reveals an interface switch for activation. J Biol Chem 299, 105440. 10.1016/j.jbc.2023.105440.

Bundy, B.M., Collier, L.S., Hoover, T.R., and Neidle, E.L. (2002). Synergistic transcriptional activation by one regulatory protein in response to two metabolites. Proc Natl Acad Sci U S A 99, 7693-7698. 10.1073/pnas.102605799.

Busby, S.J.W. (2019). Transcription activation in bacteria: ancient and modern. Microbiology (Reading) 165, 386-395. 10.1099/mic.0.000783.

Bykowski, T., van der Ploeg, J.R., Iwanicka-Nowicka, R., and Hryniewicz, M.M. (2002). The switch from inorganic to organic sulphur assimilation in Escherichia coli: adenosine 5'-phosphosulphate (APS) as a signalling molecule for sulphate excess. Mol Microbiol 43, 1347-1358. 10.1046/j.1365-2958.2002.02846.x.

Byrne, R.T., Chen, S.H., Wood, E.A., Cabot, E.L., and Cox, M.M. (2014). Escherichia coli genes and pathways involved in surviving extreme exposure to ionizing radiation. J Bacteriol 196, 3534-3545. 10.1128/JB.01589-14.

Camarena, L., Poggio, S., Garcia, N., and Osorio, A. (1998). Transcriptional repression of gdhA in Escherichia coli is mediated by the Nac protein. FEMS Microbiol Lett 167, 51-56. 10.1111/j.1574-6968.1998.tb13206.x.

Carmel, O., Rahav-Manor, O., Dover, N., Shaanan, B., and Padan, E. (1997). The Na+-specific interaction between the LysR-type regulator, NhaR, and the nhaA gene encoding the Na+/H+ antiporter of Escherichia coli. EMBO J 16, 5922-5929. 10.1093/emboj/16.19.5922.

Cashel, M., and Kalbacher, B. (1970). The control of ribonucleic acid synthesis in Escherichia coli. V. Characterization of a nucleotide associated with the stringent response. J Biol Chem 245, 2309-2318.

Chen, C.C., and Wu, H.Y. (2005). LeuO protein delimits the transcriptionally active and repressive domains on the bacterial chromosome. J Biol Chem 280, 15111-15121. 10.1074/jbc.M414544200.

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. 10.1016/0378-1119(95)00193-a.

Cho, B.K., Kim, D., Knight, E.M., Zengler, K., and Palsson, B.O. (2014). Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. BMC Biol 12, 4. 10.1186/1741-7007-12-4.

Choi, H., Kim, S., Mukhopadhyay, P., Cho, S., Woo, J., Storz, G., and Ryu, S.E. (2001). Structural basis of the redox switch in the OxyR transcription factor. Cell 105, 103-113. 10.1016/s0092-8674(01)00300-2.

Christman, M.F., Storz, G., and Ames, B.N. (1989). OxyR, a positive regulator of hydrogen peroxide-inducible genes in Escherichia coli and Salmonella typhimurium, is homologous to a family of bacterial regulatory proteins. Proc Natl Acad Sci U S A 86, 3484-3488. 10.1073/pnas.86.10.3484.

Connolly, J.P., Gabrielsen, M., Goldstone, R.J., Grinter, R., Wang, D., Cogdell, R.J., Walker, D., Smith, D.G., and Roe, A.J. (2016). A Highly Conserved Bacterial D-Serine Uptake System Links Host Metabolism and Virulence. PLoS Pathog 12, e1005359. 10.1371/journal.ppat.1005359.

Connolly, J.P.R., O'Boyle, N., and Roe, A.J. (2020). Widespread Strain-Specific Distinctions in Chromosomal Binding Dynamics of a Highly Conserved Escherichia coli Transcription Factor. mBio 11. 10.1128/mBio.01058-20.

Connolly, J.P.R., O'Boyle, N., Turner, N.C.A., Browning, D.F., and Roe, A.J. (2019). Distinct intraspecies virulence mechanisms regulated by a conserved transcription factor. Proc Natl Acad Sci U S A 116, 19695-19704. 10.1073/pnas.1903461116.

Corrigan, R.M., Bellows, L.E., Wood, A., and Grundling, A. (2016). ppGpp negatively impacts ribosome assembly affecting growth and antimicrobial tolerance in Gram-positive bacteria. Proc Natl Acad Sci U S A 113, E1710-1719. 10.1073/pnas.1522179113.

Dame, R.T., Luijsterburg, M.S., Krin, E., Bertin, P.N., Wagner, R., and Wuite, G.J. (2005). DNA bridging: a property shared among H-NS-like proteins. J Bacteriol 187, 1845-1848. 10.1128/JB.187.5.1845-1848.2005.

Datsenko, K.A., and Wanner, B.L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A 97, 6640-6645. 10.1073/pnas.120163297.

De la Cruz, M.A., Fernandez-Mora, M., Guadarrama, C., Flores-Valdez, M.A., Bustamante, V.H., Vazquez, A., and Calva, E. (2007). LeuO antagonizes H-NS and StpA-dependent repression in Salmonella enterica ompS1. Mol Microbiol 66, 727-743. 10.1111/j.1365-2958.2007.05958.x.

Diaz, E., Ferrandez, A., and Garcia, J.L. (1998). Characterization of the hca cluster encoding the dioxygenolytic pathway for initial catabolism of 3-phenylpropionic acid in Escherichia coli K-12. J Bacteriol 180, 2915-2923. 10.1128/JB.180.11.2915-2923.1998.

Diez, S., Ryu, J., Caban, K., Gonzalez, R.L., Jr., and Dworkin, J. (2020). The alarmones (p)ppGpp directly regulate translation initiation during entry into quiescence. Proc Natl Acad Sci U S A 117, 15565-15572. 10.1073/pnas.1920013117.

Dorman, C.J. (2004). H-NS: a universal regulator for a dynamic genome. Nat Rev Microbiol 2, 391-400. 10.1038/nrmicro883.

Dorman, C.J. (2014). Function of nucleoid-associated proteins in chromosome structuring and transcriptional regulation. J Mol Microbiol Biotechnol 24, 316-331. 10.1159/000368850.

Dover, N., Higgins, C.F., Carmel, O., Rimon, A., Pinner, E., and Padan, E. (1996). Na+-induced transcription of nhaA, which encodes an Na+/H+ antiporter in Escherichia coli, is positively regulated by nhaR and affected by hns. J Bacteriol 178, 6508-6517. 10.1128/jb.178.22.6508-6517.1996.

Drazic, A., Gebendorfer, K.M., Mak, S., Steiner, A., Krause, M., Bepperling, A., and Winter, J. (2014). Tetramers are the activation-competent species of the HOCl-specific transcription factor HypT. J Biol Chem 289, 977-986. 10.1074/jbc.M113.521401.

Drazic, A., Miura, H., Peschek, J., Le, Y., Bach, N.C., Kriehuber, T., and Winter, J. (2013). Methionine oxidation activates a transcription factor in response to oxidative stress. Proc Natl Acad Sci U S A 110, 9493-9498. 10.1073/pnas.1300578110.

Dubbs, J.M., and Mongkolsuk, S. (2012). Peroxide-sensing transcriptional regulators in bacteria. J Bacteriol 194, 5495-5503. 10.1128/JB.00304-12.

Eguchi, Y., Ishii, E., Hata, K., and Utsumi, R. (2011). Regulation of acid resistance by connectors of two-component signal transduction systems in Escherichia coli. J Bacteriol 193, 1222-1228. 10.1128/JB.01124-10.

Espinosa, E., and Casadesus, J. (2014). Regulation of Salmonella enterica pathogenicity island 1 (SPI-1) by the LysR-type regulator LeuO. Mol Microbiol 91, 1057-1069. 10.1111/mmi.12500.

Fernandez-Coll, L., and Cashel, M. (2020). Possible Roles for Basal Levels of (p)ppGpp: Growth Efficiency Vs. Surviving Stress. Front Microbiol 11, 592718. 10.3389/fmicb.2020.592718.

Fragel, S.M., Montada, A., Heermann, R., Baumann, U., Schacherl, M., and Schnetz, K. (2019). Characterization of the pleiotropic LysR-type transcription regulator LeuO of Escherichia coli. Nucleic Acids Res 47, 7363-7379. 10.1093/nar/gkz506.

Fritsch, P.S., Urbanowski, M.L., and Stauffer, G.V. (2000). Role of the RNA polymerase alpha subunits in MetR-dependent activation of metE and metH: important residues in the C-terminal domain and orientation requirements within RNA polymerase. J Bacteriol 182, 5539-5550. 10.1128/JB.182.19.5539-5550.2000.

Fujii, T., Aritoku, Y., Agematu, H., and Tsunekawa, H. (2002). Increase in the rate of L-pipecolic acid production using lat-expressing Escherichia coli by lysP and yeiE amplification. bioscience biotechnology and biochemistry 66, 1981-1984. 10.1271/bbb.66.1981.

Ganduri, Y.L., Sadda, S.R., Datta, M.W., Jambukeswaran, R.K., and Datta, P. (1993). TdcA, a transcriptional activator of the tdcABC operon of Escherichia coli, is a member of the LysR family of proteins. Mol Gen Genet 240, 395-402. 10.1007/BF00280391.

Gao, Y., Lim, H.G., Verkler, H., Szubin, R., Quach, D., Rodionova, I., Chen, K., Yurkovich, J.T., Cho, B.K., and Palsson, B.O. (2021). Unraveling the functions of uncharacterized transcription factors in Escherichia coli using ChIP-exo. Nucleic Acids Res 49, 9696-9710. 10.1093/nar/gkab735.

Gao, Y., Yurkovich, J.T., Seo, S.W., Kabimoldayev, I., Drager, A., Chen, K., Sastry, A.V., Fang, X., Mih, N., Yang, L., et al. (2018). Systematic discovery of uncharacterized transcription factors in Escherichia coli K-12 MG1655. Nucleic Acids Res 46, 10682-10696. 10.1093/nar/gky752.

Gardner, S.G., and McCleary, W.R. (2019). Control of the phoBR Regulon in Escherichia coli. EcoSal Plus 8. 10.1128/ecosalplus.ESP-0006-2019.

Gebendorfer, K.M., Drazic, A., Le, Y., Gundlach, J., Bepperling, A., Kastenmuller, A., Ganzinger, K.A., Braun, N., Franzmann, T.M., and Winter, J. (2012). Identification of a hypochlorite-specific transcription factor from Escherichia coli. J Biol Chem 287, 6892-6903. 10.1074/jbc.M111.287219.

Geiger, T., Francois, P., Liebeke, M., Fraunholz, M., Goerke, C., Krismer, B., Schrenzel, J., Lalk, M., and Wolz, C. (2012). The stringent response of Staphylococcus aureus and its impact on survival after phagocytosis through the induction of intracellular PSMs expression. PLoS Pathog 8, e1003016. 10.1371/journal.ppat.1003016.

Germain, E., Guiraud, P., Byrne, D., Douzi, B., Djendli, M., and Maisonneuve, E. (2019). YtfK activates the stringent response by triggering the alarmone synthetase SpoT in Escherichia coli. Nat Commun 10, 5763. 10.1038/s41467-019-13764-4.

Goller, C., Wang, X., Itoh, Y., and Romeo, T. (2006). The cation-responsive protein NhaR of Escherichia coli activates pgaABCD transcription, required for production of the biofilm adhesin poly-beta-1,6-N-acetyl-D-glucosamine. J Bacteriol 188, 8022-8032. 10.1128/JB.01106-06.

Gottesman, S. (2019). Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. J Biol Chem 294, 11685-11700. 10.1074/jbc.REV119.005593.

Grainger, D.C., Hurd, D., Goldberg, M.D., and Busby, S.J. (2006). Association of nucleoid proteins with coding and non-coding segments of the Escherichia coli genome. Nucleic Acids Res 34, 4642-4652. 10.1093/nar/gkl542.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. Bioinformatics 27, 1017-1018. 10.1093/bioinformatics/btr064.

Guo, Y., Li, Y., Zhan, W., Wood, T.K., and Wang, X. (2019). Resistance to oxidative stress by inner membrane protein ElaB is regulated by OxyR and RpoS. Microb Biotechnol 12, 392-404. 10.1111/1751-7915.13369.

Hadley, R.C., Zhitnitsky, D., Livnat-Levanon, N., Masrati, G., Vigonsky, E., Rose, J., Ben-Tal, N., Rosenzweig, A.C., and Lewinson, O. (2022). The copper-linked Escherichia coli AZY operon: Structure, metal binding, and a possible physiological role in copper delivery. J Biol Chem 298, 101445. 10.1016/j.jbc.2021.101445.

Hagewood, B.T., Ganduri, Y.L., and Datta, P. (1994). Functional analysis of the tdcABC promoter of Escherichia coli: roles of TdcA and TdcR. J Bacteriol 176, 6214-6220. 10.1128/jb.176.20.6214-6220.1994.

Hall, D.R., Gourley, D.G., Leonard, G.A., Duke, E.M., Anderson, L.A., Boxer, D.H., and Hunter, W.N. (1999). The high-resolution crystal structure of the molybdate-dependent transcriptional regulator (ModE) from Escherichia coli: a novel combination of domain folds. EMBO J 18, 1435-1446. 10.1093/emboj/18.6.1435.

Heil, G., Stauffer, L.T., and Stauffer, G.V. (2002). Glycine binds the transcriptional accessory protein GcvR to disrupt a GcvA/GcvR interaction and allow GcvA-mediated activation of the Escherichia coli gcvTHP operon. Microbiology (Reading) 148, 2203-2214. 10.1099/00221287-148-7-2203.

Heinemeyer, E.A., and Richter, D. (1978). Mechanism of the in vitro breakdown of guanosine 5'-diphosphate 3'-diphosphate in Escherichia coli. Proc Natl Acad Sci U S A 75, 4180-4183. 10.1073/pnas.75.9.4180.

Henikoff, S., Haughn, G.W., Calvo, J.M., and Wallace, J.C. (1988). A large family of bacterial activator proteins. Proc Natl Acad Sci U S A 85, 6602-6606. 10.1073/pnas.85.18.6602.

Higashi, K., Tobe, T., Kanai, A., Uyar, E., Ishikawa, S., Suzuki, Y., Ogasawara, N., Kurokawa, K., and Oshima, T. (2016). H-NS Facilitates Sequence Diversification of Horizontally Transferred DNAs during Their Integration in Host Chromosomes. PLoS Genet 12, e1005796. 10.1371/journal.pgen.1005796.

Hingley-Wilson, S.M., Ma, N., Hu, Y., Casey, R., Bramming, A., Curry, R.J., Tang, H.L., Wu, H., Butler, R.E., Jacobs, W.R., Jr., et al. (2020). Loss of phenotypic inheritance associated with ydcI mutation leads to increased frequency of small, slow persisters in Escherichia coli. Proc Natl Acad Sci U S A 117, 4152-4157. 10.1073/pnas.1914741117.

Hobbs, E.C., Astarita, J.L., and Storz, G. (2010). Small RNAs and small proteins involved in resistance to cell envelope stress and acid shock in Escherichia coli: analysis of a bar-coded mutant collection. J Bacteriol 192, 59-67. 10.1128/JB.00873-09.

Hommais, F., Krin, E., Laurent-Winter, C., Soutourina, O., Malpertuy, A., Le Caer, J.P., Danchin, A., and Bertin, P. (2001). Large-scale monitoring of pleiotropic regulation of gene expression by the prokaryotic nucleoid-associated protein, H-NS. Mol Microbiol 40, 20-36. 10.1046/j.1365-2958.2001.02358.x.

Honda, N., Iyoda, S., Yamamoto, S., Terajima, J., and Watanabe, H. (2009). LrhA positively controls the expression of the locus of enterocyte effacement genes in enterohemorrhagic Escherichia coli by differential regulation of their master regulators PchA and PchB. Mol Microbiol 74, 1393-1341. 10.1111/j.1365-2958.2009.06937.x.

Hou, N., Yan, Z., Fan, K., Li, H., Zhao, R., Xia, Y., Xun, L., and Liu, H. (2019). OxyR senses sulfane sulfur and activates the genes for its removal in Escherichia coli. Redox Biol 26, 101293. 10.1016/j.redox.2019.101293.

Hryniewicz, M.M., and Kredich, N.M. (1991). The cysP promoter of Salmonella typhimurium: characterization of two binding sites for CysB protein, studies of in vivo transcription initiation, and demonstration of the anti-inducer effects of thiosulfate. J Bacteriol 173, 5876-5886. 10.1128/jb.173.18.5876-5886.1991.

Hussein, M.J., Green, J.M., and Nichols, B.P. (1998). Characterization of mutations that allow p-aminobenzoyl-glutamate utilization by Escherichia coli. J Bacteriol 180, 6260-6268. 10.1128/JB.180.23.6260-6268.1998.

Hustmyer, C.M., Wolfe, M.B., Welch, R.A., and Landick, R. (2022). RfaH Counter-Silences Inhibition of Transcript Elongation by H-NS-StpA Nucleoprotein Filaments in Pathogenic Escherichia coli. mBio 13, e0266222. 10.1128/mbio.02662-22.

Huttener, M., Prieto, A., Espelt, J., Bernabeu, M., and Juarez, A. (2018). Stringent Response and AggR-Dependent Virulence Regulation in the Enteroaggregative Escherichia coli Strain 042. Front Microbiol 9, 717. 10.3389/fmicb.2018.00717.

Ishihama, A., Shimada, T., and Yamazaki, Y. (2016). Transcription profile of Escherichia coli: genomic SELEX search for regulatory targets of transcription factors. Nucleic Acids Res 44, 2058-2074. 10.1093/nar/gkw051.

Iwanicka-Nowicka, R., and Hryniewicz, M.M. (1995). A new gene, cbl, encoding a member of the LysR family of transcriptional regulators belongs to Escherichia coli cys regulon. Gene 166, 11-17. 10.1016/0378-1119(95)00606-8.

Jennings, M.E., Quick, L.N., Soni, A., Davis, R.R., Crosby, K., Ott, C.M., Nickerson, C.A., and Wilson, J.W. (2011). Characterization of the Salmonella enterica serovar Typhimurium ydcI gene, which encodes a conserved DNA binding protein required for full acid stress resistance. J Bacteriol 193, 2208-2217. 10.1128/JB.01335-10.

Jo, I., Chung, I.Y., Bae, H.W., Kim, J.S., Song, S., Cho, Y.H., and Ha, N.C. (2015). Structural details of the OxyR peroxide-sensing mechanism. Proc Natl Acad Sci U S A 112, 6443-6448. 10.1073/pnas.1424495112.

Jorgensen, C., and Dandanell, G. (1999). Isolation and characterization of mutations in the Escherichia coli regulatory protein XapR. J Bacteriol 181, 4397-4403. 10.1128/JB.181.14.4397-4403.1999.

Jourdan, A.D., and Stauffer, G.V. (1998). Mutational analysis of the transcriptional regulator GcvA: amino acids important for activation, repression, and DNA binding. J Bacteriol 180, 4865-4871. 10.1128/JB.180.18.4865-4871.1998.

Jovanovic, M., Lilic, M., Savic, D.J., and Jovanovic, G. (2003). The LysR-type transcriptional regulator CysB controls the repression of hslJ transcription in Escherichia coli. Microbiology (Reading) 149, 3449-3459. 10.1099/mic.0.26609-0.

Justino, M.C., Vicente, J.B., Teixeira, M., and Saraiva, L.M. (2005). New genes implicated in the protection of anaerobically grown Escherichia coli against nitric oxide. J Biol Chem 280, 2636-2643. 10.1074/jbc.M411070200.

Kahramanoglou, C., Seshasayee, A.S., Prieto, A.I., Ibberson, D., Schmidt, S., Zimmermann, J., Benes, V., Fraser, G.M., and Luscombe, N.M. (2011). Direct and indirect effects of H-NS and Fis on global gene expression control in Escherichia coli. Nucleic Acids Res 39, 2073-2091. 10.1093/nar/gkq934.

Karp, P.D., Weaver, D., Paley, S., Fulcher, C., Kubo, A., Kothari, A., Krummenacker, M., Subhraveti, P., Weerasinghe, D., Gama-Castro, S., et al. (2014). The EcoCyc Database. EcoSal Plus 6. 10.1128/ecosalplus.ESP-0009-2013.

Keseler, I.M., Mackie, A., Santos-Zavaleta, A., Billington, R., Bonavides-Martinez, 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. 10.1093/nar/gkw1003.

Kim, S.O., Merchant, K., Nudelman, R., Beyer, W.F., Jr., Keng, T., DeAngelo, J., Hausladen, A., and Stamler, J.S. (2002). OxyR: a molecular code for redox-related signaling. Cell 109, 383-396. 10.1016/s0092-8674(02)00723-7.

Knapp, G.S., and Hu, J.C. (2009). The oligomerization of CynR in Escherichia coli. Protein Sci 18, 2307-2315. 10.1002/pro.241.

Knapp, G.S., Tsai, J.W., and Hu, J.C. (2009). The oligomerization of OxyR in Escherichia coli. Protein Sci 18, 101-107. 10.1002/pro.5.

Ko, M., and Park, C. (2000). H-NS-Dependent regulation of flagellar synthesis is mediated by a LysR family protein. J Bacteriol 182, 4670-4672. 10.1128/JB.182.16.4670-4672.2000.

Koentjoro, M.P., Adachi, N., Senda, M., Ogawa, N., and Senda, T. (2018). Crystal structure of the DNA-binding domain of the LysR-type transcriptional regulator CbnR in complex with a DNA fragment of the recognition-binding site in the promoter region. FEBS J 285, 977-989. 10.1111/febs.14380.

Krol, J.E. (2018). Regulatory loop between the CsrA system and NhaR, a high salt/high pH regulator. PLoS One 13, e0209554. 10.1371/journal.pone.0209554.

Kullik, I., Stevens, J., Toledano, M.B., and Storz, G. (1995). Mutational analysis of the redox-sensitive transcriptional regulator OxyR: regions important for DNA binding and multimerization. J Bacteriol 177, 1285-1291. 10.1128/jb.177.5.1285-1291.1995.

Ladner, C.L., Yang, J., Turner, R.J., and Edwards, R.A. (2004). Visible fluorescent detection of proteins in polyacrylamide gels without staining. Anal Biochem 326, 13-20. 10.1016/j.ab.2003.10.047.

Laishram, R.S., and Gowrishankar, J. (2007). Environmental regulation operating at the promoter clearance step of bacterial transcription. Genes Dev 21, 1258-1272. 10.1101/gad.1520507.

Lamblin, A.F., and Fuchs, J.A. (1993). Expression and purification of the cynR regulatory gene product: CynR is a DNA-binding protein. J Bacteriol 175, 7990-7999. 10.1128/jb.175.24.7990-7999.1993.

Lamblin, A.F., and Fuchs, J.A. (1994). Functional analysis of the Escherichia coli K-12 cyn operon transcriptional regulation. J Bacteriol 176, 6613-6622. 10.1128/jb.176.21.6613-6622.1994.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357-359. 10.1038/nmeth.1923.

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. 10.1046/j.1365-2958.2002.03032.x.

Lerche, M., Dian, C., Round, A., Lonneborg, R., Brzezinski, P., and Leonard, G.A. (2016). The solution configurations of inactive and activated DntR have implications for the sliding dimer mechanism of LysR transcription factors. Sci Rep 6, 19988. 10.1038/srep19988.

Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell 157, 624-635. 10.1016/j.cell.2014.02.033.

Lim, C.J., Lee, S.Y., Teramoto, J., Ishihama, A., and Yan, J. (2013). The nucleoid-associated protein Dan organizes chromosomal DNA through rigid nucleoprotein filament formation in E. coli during anoxia. Nucleic Acids Res 41, 746-753. 10.1093/nar/gks1126.

Lim, C.J., Whang, Y.R., Kenney, L.J., and Yan, J. (2012). Gene silencing H-NS paralogue StpA forms a rigid protein filament along DNA that blocks DNA accessibility. Nucleic Acids Res 40, 3316-3328. 10.1093/nar/gkr1247.

Liu, K., Bittner, A.N., and Wang, J.D. (2015). Diversity in (p)ppGpp metabolism and effectors. Curr Opin Microbiol 24, 72-79. 10.1016/j.mib.2015.01.012.

Lochowska, A., Iwanicka-Nowicka, R., Plochocka, D., and Hryniewicz, M.M. (2001). Functional dissection of the LysR-type CysB transcriptional regulator. Regions important for DNA binding, inducer response, oligomerization, and positive control. J Biol Chem 276, 2098-2107. 10.1074/jbc.M007192200.

Lorenz, E., and Stauffer, G.V. (1996). RNA polymerase, PurR and MetR interactions at the glyA promoter of Escherichia coli. Microbiology (Reading) 142 (Pt 7), 1819-1824. 10.1099/13500872-142-7-1819.
Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. 10.1186/s13059-014-0550-8.
Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., and Hinton, J.C. (2006). H-NS mediates the silencing of laterally acquired genes in bacteria. PLoS Pathog 2, e81. 10.1371/journal.ppat.0020081.
Luijsterburg, M.S., Noom, M.C., Wuite, G.J., and Dame, R.T. (2006). The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. J Struct Biol 156, 262-272. 10.1016/j.jsb.2006.05.006.
Lukas, H., Reimann, J., Kim, O.B., Grimpo, J., and Unden, G. (2010). Regulation of aerobic and anaerobic D-malate metabolism of Escherichia coli by the LysR-type regulator DmlR (YeaT). J Bacteriol 192, 2503-2511. 10.1128/JB.01665-09.
Maddocks, S.E., and Oyston, P.C.F. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology (Reading) 154, 3609-3623. 10.1099/mic.0.2008/022772-0.
Makino, K., Shinagawa, H., Amemura, M., and Nakata, A. (1986). Nucleotide sequence of the phoB gene, the positive regulatory gene for the phosphate regulon of Escherichia coli K-12. J Mol Biol 190, 37-44. 10.1016/0022-2836(86)90073-2.
Marbaniang, C.N., and Gowrishankar, J. (2011). Role of ArgP (IciA) in lysine-mediated repression in Escherichia coli. J Bacteriol 193, 5985-5996. 10.1128/JB.05869-11.
Maurer, S., Fritz, J., and Muskhelishvili, G. (2009). A systematic in vitro study of nucleoprotein complexes formed by bacterial nucleoid-associated proteins revealing novel types of DNA organization. J Mol Biol 387, 1261-1276. 10.1016/j.jmb.2009.02.050.
McNicholas, P.M., and Gunsalus, R.P. (2002). The molybdate-responsive Escherichia coli ModE transcriptional regulator coordinates periplasmic nitrate reductase (napFDAGHBC) operon expression with nitrate and molybdate availability. J Bacteriol 184, 3253-3259. 10.1128/JB.184.12.3253-3259.2002.
Medina-Aparicio, L., Rebollar-Flores, J.E., Gallego-Hernandez, A.L., Vazquez, A., Olvera, L., Gutierrez-Rios, R.M., Calva, E., and Hernandez-Lucas, I. (2011). The CRISPR/Cas immune system is an operon regulated by LeuO, H-NS, and leucine-responsive regulatory protein in Salmonella enterica serovar Typhi. J Bacteriol 193, 2396-2407. 10.1128/JB.01480-10.
Metzger, S., Schreiber, G., Aizenman, E., Cashel, M., and Glaser, G. (1989). Characterization of the relA1 mutation and a comparison of relA1 with new relA null alleles in Escherichia coli. J Biol Chem 264, 21146-21152.
Meyer, L., Germain, E., and Maisonneuve, E. (2021). Regulation of ytfK by cAMP-CRP Contributes to SpoT-Dependent Accumulation of (p)ppGpp in Response to Carbon Starvation YtfK Responds to Glucose Exhaustion. Front Microbiol 12, 775164. 10.3389/fmicb.2021.775164.
Miller, J.H. (1992). A short course in bacterial genetics. A laboratory manual and handbook for Escherichia coli and related bacteria (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).
Miroux, B., and Walker, J.E. (1996). Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260, 289-298. 10.1006/jmbi.1996.0399.
Momany, C., and Neidle, E.L. (2012). Defying stereotypes: the elusive search for a universal model of LysR-type regulation. Mol Microbiol 83, 453-456. 10.1111/j.1365-2958.2011.07960.x.
Muller, C.M., Dobrindt, U., Nagy, G., Emody, L., Uhlin, B.E., and Hacker, J. (2006). Role of histone-like proteins H-NS and StpA in expression of virulence determinants of uropathogenic Escherichia coli. J Bacteriol 188, 5428-5438. 10.1128/JB.01956-05.

Murray, K.D., and Bremer, H. (1996). Control of spoT-dependent ppGpp synthesis and degradation in Escherichia coli. J Mol Biol 259, 41-57. 10.1006/jmbi.1996.0300.
Muse, W.B., and Bender, R.A. (1998). The nac (nitrogen assimilation control) gene from Escherichia coli. J Bacteriol 180, 1166-1173. 10.1128/JB.180.5.1166-1173.1998.
Muse, W.B., and Bender, R.A. (1999). The amino-terminal 100 residues of the nitrogen assimilation control protein (NAC) encode all known properties of NAC from Klebsiella aerogenes and Escherichia coli. J Bacteriol 181, 934-940. 10.1128/JB.181.3.934-940.1999.
Nandineni, M.R., and Gowrishankar, J. (2004). Evidence for an arginine exporter encoded by yggA (argO) that is regulated by the LysR-type transcriptional regulator ArgP in Escherichia coli. J Bacteriol 186, 3539-3546. 10.1128/JB.186.11.3539-3546.2004.
Nandineni, M.R., Laishram, R.S., and Gowrishankar, J. (2004). Osmosensitivity associated with insertions in argP (iciA) or glnE in glutamate synthase-deficient mutants of Escherichia coli. J Bacteriol 186, 6391-6399. 10.1128/JB.186.19.6391-6399.2004.
Navarre, W.W., McClelland, M., Libby, S.J., and Fang, F.C. (2007). Silencing of xenogeneic DNA by H-NS-facilitation of lateral gene transfer in bacteria by a defense system that recognizes foreign DNA. Genes Dev 21, 1456-1471. 10.1101/gad.1543107.
Neidhardt, F.C., Bloch, P.L., and Smith, D.F. (1974). Culture medium for enterobacteria. J Bacteriol 119, 736-747. 10.1128/jb.119.3.736-747.1974.
Nishino, K., and Yamaguchi, A. (2001). Analysis of a complete library of putative drug transporter genes in Escherichia coli. J Bacteriol 183, 5803-5812. 10.1128/JB.183.20.5803-5812.2001.
Nishio, Y., Suzuki, T., Matsui, K., and Usuda, Y. (2013). Metabolic Control of the TCA cycle by the YdcI Transcriptional Regulator in Escherichia coli. J Microb Biochem Technol 5, 059-067. 10.4172/1948-5948.100010.
Norregaard-Madsen, M., McFall, E., and Valentin-Hansen, P. (1995). Organization and transcriptional regulation of the Escherichia coli K-12 D-serine tolerance locus. J Bacteriol 177, 6456-6461. 10.1128/jb.177.22.6456-6461.1995.
Ogasawara, H., Ishizuka, T., Hotta, S., Aoki, M., Shimada, T., and Ishihama, A. (2020). Novel regulators of the csgD gene encoding the master regulator of biofilm formation in Escherichia coli K-12. Microbiology (Reading) 166, 880-890. 10.1099/mic.0.000947.
Ono, S., Goldberg, M.D., Olsson, T., Esposito, D., Hinton, J.C., and Ladbury, J.E. (2005). H-NS is a part of a thermally controlled mechanism for bacterial gene regulation. Biochem J 391, 203-213. 10.1042/bj20050453.
Onufryk, C., Crouch, M.L., Fang, F.C., and Gross, C.A. (2005). Characterization of six lipoproteins in the sigmaE regulon. J Bacteriol 187, 4552-4561. 10.1128/JB.187.13.4552-4561.2005.
Oshima, T., and Biville, F. (2006). Functional identification of ygiP as a positive regulator of the ttdA-ttdB-ygjE operon. Microbiology (Reading) 152, 2129-2135. 10.1099/mic.0.28753-0.
Palevsky, N., Shemer, B., Connolly, J.P., and Belkin, S. (2016). The Highly Conserved Escherichia coli Transcription Factor YhaJ Regulates Aromatic Compound Degradation. Front Microbiol 7, 1490. 10.3389/fmicb.2016.01490.
Pannen, D., Fabisch, M., Gausling, L., and Schnetz, K. (2016). Interaction of the RcsB Response Regulator with Auxiliary Transcription Regulators in Escherichia coli. J Biol Chem 291, 2357-2370. 10.1074/jbc.M115.696815.
Parsek, M.R., Ye, R.W., Pun, P., and Chakrabarty, A.M. (1994). Critical nucleotides in the interaction of a LysR-type regulator with its target promoter region. catBC promoter activation by CatR. J Biol Chem 269, 11279-11284.
Patacq, C., Chaudet, N., and Letisse, F. (2020). Crucial Role of ppGpp in the Resilience of Escherichia coli to Growth Disruption. mSphere 5. 10.1128/mSphere.01132-20.
Paul, B.J., Berkmen, M.B., and Gourse, R.L. (2005). DksA potentiates direct activation of amino acid promoters by ppGpp. Proc Natl Acad Sci U S A 102, 7823-7828. 10.1073/pnas.0501170102.

Peano, C., Wolf, J., Demol, J., Rossi, E., Petiti, L., De Bellis, G., Geiselmann, J., Egli, T., Lacour, S., and Landini, P. (2015). Characterization of the Escherichia coli sigma(S) core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis. Sci Rep 5, 10469. 10.1038/srep10469.

Peterson, C.N., Carabetta, V.J., Chowdhury, T., and Silhavy, T.J. (2006). LrhA regulates rpoS translation in response to the Rcs phosphorelay system in Escherichia coli. J Bacteriol 188, 3175-3181. 10.1128/JB.188.9.3175-3181.2006.

Ptashne, M., Backman, K., Humayun, M.Z., Jeffrey, A., Maurer, R., Meyer, B., and Sauer, R.T. (1976). Autoregulation and function of a repressor in bacteriophage lambda. Science 194, 156-161. 10.1126/science.959843.

Pul, U., Wurm, R., Arslan, Z., Geissen, R., Hofmann, N., and Wagner, R. (2010). Identification and characterization of E. coli CRISPR-cas promoters and their silencing by H-NS. Mol Microbiol 75, 1495-1512. 10.1111/j.1365-2958.2010.07073.x.

Punekar, A.S., Porter, J., Carr, S.B., and Phillips, S.E. (2016). Structural basis for DNA recognition by the transcription regulator MetR. Acta Crystallogr F Struct Biol Commun 72, 417-426. 10.1107/S2053230X16006828.

Qin, L., Bdira, F.B., Sterckx, Y.G.J., Volkov, A.N., Vreede, J., Giachin, G., van Schaik, P., Ubbink, M., and Dame, R.T. (2020). Structural basis for osmotic regulation of the DNA binding properties of H-NS proteins. Nucleic Acids Res 48, 2156-2172. 10.1093/nar/gkz1226.

Rahav-Manor, O., Carmel, O., Karpel, R., Taglicht, D., Glaser, G., Schuldiner, S., and Padan, E. (1992). NhaR, a protein homologous to a family of bacterial regulatory proteins (LysR), regulates nhaA, the sodium proton antiporter gene in Escherichia coli. J Biol Chem 267, 10433-10438.

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. 10.1111/mmi.13926.

Rhee, K.Y., Senear, D.F., and Hatfield, G.W. (1998). Activation of gene expression by a ligand-induced conformational change of a protein-DNA complex. J Biol Chem 273, 11257-11266. 10.1074/jbc.273.18.11257.

Rice, C.D., Pollard, J.E., Lewis, Z.T., and McCleary, W.R. (2009). Employment of a promoter-swapping technique shows that PhoU modulates the activity of the PstSCAB2 ABC transporter in Escherichia coli. Appl Environ Microbiol 75, 573-582. 10.1128/AEM.01046-08.

Rimsky, S., Zuber, F., Buckle, M., and Buc, H. (2001). A molecular mechanism for the repression of transcription by the H-NS protein. Mol Microbiol 42, 1311-1323. 10.1046/j.1365-2958.2001.02706.x.

Rintoul, M.R., Cusa, E., Baldoma, L., Badia, J., Reitzer, L., and Aguilar, J. (2002). Regulation of the Escherichia coli allantoin regulon: coordinated function of the repressor AllR and the activator AllS. J Mol Biol 324, 599-610. 10.1016/s0022-2836(02)01134-8.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24-26. 10.1038/nbt.1754.

Romiyo, V., and Wilson, J.W. (2020). Phenotypes, transcriptome, and novel biofilm formation associated with the ydcI gene. Antonie van Leeuwenhoek 113, 1109-1122. 10.1007/s10482-020-01412-7.

Ruiz, J., Haneburger, I., and Jung, K. (2011). Identification of ArgP and Lrp as transcriptional regulators of lysP, the gene encoding the specific lysine permease of Escherichia coli. J Bacteriol 193, 2536-2548. 10.1128/JB.00815-10.

Sambrook, J. (2001). Molecular cloning : a laboratory manual (Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001).

Sanchez-Vazquez, P., Dewey, C.N., Kitten, N., Ross, W., and Gourse, R.L. (2019). Genome-wide effects on Escherichia coli transcription from ppGpp binding to its two sites on RNA polymerase. Proc Natl Acad Sci U S A 116, 8310-8319. 10.1073/pnas.1819682116.

Schellhorn, H.E. (2020). Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. Front Microbiol 11, 560099. 10.3389/fmicb.2020.560099.

Schumacher, K., Gelhausen, R., Kion-Crosby, W., Barquist, L., Backofen, R., and Jung, K. (2023). Ribosome profiling reveals the fine-tuned response of Escherichia coli to mild and severe acid stress. mSystems 8, e0103723. 10.1128/msystems.01037-23.

Seeger, C., Poulsen, C., and Dandanell, G. (1995). Identification and characterization of genes (xapA, xapB, and xapR) involved in xanthosine catabolism in Escherichia coli. J Bacteriol 177, 5506-5516. 10.1128/jb.177.19.5506-5516.1995.

Sezonov, G., Joseleau-Petit, D., and D'Ari, R. (2007). Escherichia coli physiology in Luria-Bertani broth. J Bacteriol 189, 8746-8749. 10.1128/JB.01368-07.

Shi, A., Zheng, H., Yomano, L.P., York, S.W., Shanmugam, K.T., and Ingram, L.O. (2016). Plasmidic Expression of nemA and yafC* Increased Resistance of Ethanologenic Escherichia coli LY180 to Nonvolatile Side Products from Dilute Acid Treatment of Sugarcane Bagasse and Artificial Hydrolysate. Appl Environ Microbiol 82, 2137-2145. 10.1128/AEM.03488-15.

Shimada, T., Bridier, A., Briandet, R., and Ishihama, A. (2011). Novel roles of LeuO in transcription regulation of E. coli genome: antagonistic interplay with the universal silencer H-NS. Mol Microbiol 82, 378-397. 10.1111/j.1365-2958.2011.07818.x.

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. 10.1111/j.1365-2443.2005.00888.x.

Shimada, T., Yamamoto, K., and Ishihama, A. (2009). Involvement of the leucine response transcription factor LeuO in regulation of the genes for sulfa drug efflux. J Bacteriol 191, 4562-4571. 10.1128/JB.00108-09.

Shimada, T., Yamazaki, K., and Ishihama, A. (2013). Novel regulator PgrR for switch control of peptidoglycan recycling in Escherichia coli. Genes Cells 18, 123-134. 10.1111/gtc.12026.

Solomon, L., Shah, A., Hannagan, S., and Wilson, J.W. (2014). Bacterial genus-specific tolerance for YdcI expression. Curr Microbiol 69, 640-648. 10.1007/s00284-014-0631-7.

Sonnenfield, J.M., Burns, C.M., Higgins, C.F., and Hinton, J.C. (2001). The nucleoid-associated protein StpA binds curved DNA, has a greater DNA-binding affinity than H-NS and is present in significant levels in hns mutants. Biochimie 83, 243-249. 10.1016/s0300-9084(01)01232-9.

Spira, B., Silberstein, N., and Yagil, E. (1995). Guanosine 3',5'-bispyrophosphate (ppGpp) synthesis in cells of Escherichia coli starved for Pi. J Bacteriol 177, 4053-4058. 10.1128/jb.177.14.4053-4058.1995.

Srinivasan, R., Chandraprakash, D., Krishnamurthi, R., Singh, P., Scolari, V.F., Krishna, S., and Seshasayee, A.S. (2013). Genomic analysis reveals epistatic silencing of "expensive" genes in Escherichia coli K-12. Mol Biosyst 9, 2021-2033. 10.1039/c3mb70035f.

Srivatsan, A., Han, Y., Peng, J., Tehranchi, A.K., Gibbs, R., Wang, J.D., and Chen, R. (2008). High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. PLoS Genet 4, e1000139. 10.1371/journal.pgen.1000139.

Stauffer, L.T., and Stauffer, G.V. (2012). Antagonistic Roles for GcvA and GcvB in hdeAB Expression in Escherichia coli. ISRN Microbiol 2012, 697308. 10.5402/2012/697308.

Stec, E., Witkowska-Zimny, M., Hryniewicz, M.M., Neumann, P., Wilkinson, A.J., Brzozowski, A.M., Verma, C.S., Zaim, J., Wysocki, S., and Bujacz, G.D. (2006). Structural basis of the sulphate starvation response in E. coli: crystal structure and mutational analysis of the cofactor-binding domain of the Cbl transcriptional regulator. J Mol Biol 364, 309-322. 10.1016/j.jmb.2006.06.033.

Steinchen, W., Zegarra, V., and Bange, G. (2020). (p)ppGpp: Magic Modulators of Bacterial Physiology and Metabolism. Front Microbiol 11, 2072. 10.3389/fmicb.2020.02072.

Stoebel, D.M., Free, A., and Dorman, C.J. (2008). Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. Microbiology (Reading) 154, 2533-2545. 10.1099/mic.0.2008/020693-0.

Stragier, P., Danos, O., and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in Escherichia coli. II. Nucleotide sequence of the lysA gene and its regulatory region. J Mol Biol 168, 321-331. 10.1016/s0022-2836(83)80021-7.

Stragier, P., and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in Escherichia coli. III. Nucleotide sequence and regulation of the lysR gene. J Mol Biol 168, 333-350. 10.1016/s0022-2836(83)80022-9.

Stragier, P., Richaud, F., Borne, F., and Patte, J.C. (1983). Regulation of diaminopimelate decarboxylase synthesis in Escherichia coli. I. Identification of a lysR gene encoding an activator of the lysA gene. J Mol Biol 168, 307-320. 10.1016/s0022-2836(83)80020-5.

Stratmann, T., Madhusudan, S., and Schnetz, K. (2008). Regulation of the yjjQ-bglJ operon, encoding LuxR-type transcription factors, and the divergent yjjP gene by H-NS and LeuO. J Bacteriol 190, 926-935. 10.1128/JB.01447-07.

Sung, Y.C., and Fuchs, J.A. (1992). The Escherichia coli K-12 cyn operon is positively regulated by a member of the lysR family. J Bacteriol 174, 3645-3650. 10.1128/jb.174.11.3645-3650.1992.

Sy, J., and Lipmann, F. (1973). Identification of the synthesis of guanosine tetraphosphate (MS I) as insertion of a pyrophosphoryl group into the 3'-position in guanosine 5'-diphosphate. Proc Natl Acad Sci U S A 70, 306-309. 10.1073/pnas.70.2.306.

Toesca, I., Perard, C., Bouvier, J., Gutierrez, C., and Conter, A. (2001). The transcriptional activator NhaR is responsible for the osmotic induction of osmC(p1), a promoter of the stress-inducible gene osmC in Escherichia coli. Microbiology (Reading) 147, 2795-2803. 10.1099/00221287-147-10-2795.

Toledano, M.B., Kullik, I., Trinh, F., Baird, P.T., Schneider, T.D., and Storz, G. (1994). Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: a mechanism for differential promoter selection. Cell 78, 897-909. 10.1016/s0092-8674(94)90702-1.

Trachman, J.D., and Yasmin, M. (2004). Thermo-osmoregulation of heat-labile enterotoxin expression by Escherichia coli. Curr Microbiol 49, 353-360. 10.1007/s00284-004-4282-y.

Trautinger, B.W., Jaktaji, R.P., Rusakova, E., and Lloyd, R.G. (2005). RNA polymerase modulators and DNA repair activities resolve conflicts between DNA replication and transcription. Mol Cell 19, 247-258. 10.1016/j.molcel.2005.06.004.

Turlin, E., Sismeiro, O., Le Caer, J.P., Labas, V., Danchin, A., and Biville, F. (2005). 3-phenylpropionate catabolism and the Escherichia coli oxidative stress response. Res Microbiol 156, 312-321. 10.1016/j.resmic.2004.10.012.

Tyrrell, R., Verschueren, K.H., Dodson, E.J., Murshudov, G.N., Addy, C., and Wilkinson, A.J. (1997). The structure of the cofactor-binding fragment of the LysR family member, CysB: a familiar fold with a surprising subunit arrangement. Structure 5, 1017-1032. 10.1016/s0969-2126(97)00254-2.

Uyar, E., Kurokawa, K., Yoshimura, M., Ishikawa, S., Ogasawara, N., and Oshima, T. (2009). Differential binding profiles of StpA in wild-type and h-ns mutant cells: a comparative analysis of cooperative partners by chromatin immunoprecipitation-microarray analysis. J Bacteriol 191, 2388-2391. 10.1128/JB.01594-08.

van der Ploeg, J.R., Eichhorn, E., and Leisinger, T. (2001). Sulfonate-sulfur metabolism and its regulation in Escherichia coli. Arch Microbiol 176, 1-8. 10.1007/s002030100298.

van Der Ploeg, J.R., Iwanicka-Nowicka, R., Bykowski, T., Hryniewicz, M.M., and Leisinger, T. (1999). The Escherichia coli ssuEADCB gene cluster is required for the utilization of sulfur from aliphatic sulfonates and is regulated by the transcriptional activator Cbl. J Biol Chem 274, 29358-29365. 10.1074/jbc.274.41.29358.

van der Ploeg, J.R., Iwanicka-Nowicka, R., Kertesz, M.A., Leisinger, T., and Hryniewicz, M.M. (1997). Involvement of CysB and Cbl regulatory proteins in expression of the tauABCD operon and other sulfate starvation-inducible genes in Escherichia coli. J Bacteriol 179, 7671-7678. 10.1128/jb.179.24.7671-7678.1997.

Van Dyk, T.K., Templeton, L.J., Cantera, K.A., Sharpe, P.L., and Sariaslani, F.S. (2004). Characterization of the Escherichia coli AaeAB efflux pump: a metabolic relief valve? J Bacteriol 186, 7196-7204. 10.1128/JB.186.21.7196-7204.2004.

Verstraeten, N., Knapen, W.J., Kint, C.I., Liebens, V., Van den Bergh, B., Dewachter, L., Michiels, J.E., Fu, Q., David, C.C., Fierro, A.C., et al. (2015). Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. Mol Cell 59, 9-21. 10.1016/j.molcel.2015.05.011.

Vinella, D., Albrecht, C., Cashel, M., and D'Ari, R. (2005). Iron limitation induces SpoT-dependent accumulation of ppGpp in Escherichia coli. Mol Microbiol 56, 958-970. 10.1111/j.1365-2958.2005.04601.x.

Wang, C.H., and Koch, A.L. (1978). Constancy of growth on simple and complex media. J Bacteriol 136, 969-975. 10.1128/jb.136.3.969-975.1978.

Weissbach, H., and Brot, N. (1991). Regulation of methionine synthesis in Escherichia coli. Mol Microbiol 5, 1593-1597. 10.1111/j.1365-2958.1991.tb01905.x.

Wek, R.C., and Hatfield, G.W. (1988). Transcriptional activation at adjacent operators in the divergent-overlapping ilvY and ilvC promoters of Escherichia coli. J Mol Biol 203, 643-663. 10.1016/0022-2836(88)90199-4.

Westra, E.R., Pul, U., Heidrich, N., Jore, M.M., Lundgren, M., Stratmann, T., Wurm, R., Raine, A., Mescher, M., Van Heereveld, L., et al. (2010). H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K-12 can be relieved by the transcription activator LeuO. Mol Microbiol 77, 1380-1393. 10.1111/j.1365-2958.2010.07315.x.

Will, W.R., Navarre, W.W., and Fang, F.C. (2015). Integrated circuits: how transcriptional silencing and counter-silencing facilitate bacterial evolution. Curr Opin Microbiol 23, 8-13. 10.1016/j.mib.2014.10.005.

Wilson, R.L., Urbanowski, M.L., and Stauffer, G.V. (1995). DNA binding sites of the LysR-type regulator GcvA in the gcv and gcvA control regions of Escherichia coli. J Bacteriol 177, 4940-4946. 10.1128/jb.177.17.4940-4946.1995.

Winardhi, R.S., Yan, J., and Kenney, L.J. (2015). H-NS Regulates Gene Expression and Compacts the Nucleoid: Insights from Single-Molecule Experiments. Biophys J 109, 1321-1329. 10.1016/j.bpj.2015.08.016.

Wong, G.T., Bonocora, R.P., Schep, A.N., Beeler, S.M., Lee Fong, A.J., Shull, L.M., Batachari, L.E., Dillon, M., Evans, C., Becker, C.J., et al. (2017). Genome-Wide Transcriptional Response to Varying RpoS Levels in Escherichia coli K-12. J Bacteriol 199. 10.1128/JB.00755-16.

Worthan, S.B., McCarthy, R.D.P., Delaleau, M., Stikeleather, R., Bratton, B.P., Boudvillain, M., and Behringer, M.G. (2024). Evolution of pH-sensitive transcription termination during adaptation to repeated long-term starvation. bioRxiv. 10.1101/2024.03.01.582989.

Wu, N., He, L., Cui, P., Wang, W., Yuan, Y., Liu, S., Xu, T., Zhang, S., Wu, J., Zhang, W., and Zhang, Y. (2015). Ranking of persister genes in the same Escherichia coli genetic background demonstrates varying importance of individual persister genes in tolerance to different antibiotics. Front Microbiol 6, 1003. 10.3389/fmicb.2015.01003.

Wu, X., Altman, R., Eiteman, M.A., and Altman, E. (2013). Effect of overexpressing nhaA and nhaR on sodium tolerance and lactate production in Escherichia coli. J Biol Eng 7, 3. 10.1186/1754-1611-7-3.

Xiao, H., Kalman, M., Ikehara, K., Zemel, S., Glaser, G., and Cashel, M. (1991). Residual guanosine 3',5'-bispyrophosphate synthetic activity of relA null mutants can be eliminated by spoT null mutations. J Biol Chem 266, 5980-5990.

Xu, J., Li, T., Gao, Y., Deng, J., and Gu, J. (2019). MgrB affects the acid stress response of Escherichia coli by modulating the expression of iraM. FEMS Microbiol Lett 366. 10.1093/femsle/fnz123.

Yamada, M., Talukder, A.A., and Nitta, T. (1999). Characterization of the ssnA gene, which is involved in the decline of cell viability at the beginning of stationary phase in Escherichia coli. J Bacteriol 181, 1838-1846. 10.1128/JB.181.6.1838-1846.1999.

Yamanaka, Y., Aizawa, S.I., and Yamamoto, K. (2022). The hdeD Gene Represses the Expression of Flagellum Biosynthesis via LrhA in Escherichia coli K-12. J Bacteriol 204, e0042021. 10.1128/JB.00420-21.

Yang, J., Tauschek, M., Strugnell, R., and Robins-Browne, R.M. (2005). The H-NS protein represses transcription of the eltAB operon, which encodes heat-labile enterotoxin in enterotoxigenic Escherichia coli, by binding to regions downstream of the promoter. Microbiology (Reading) 151, 1199-1208. 10.1099/mic.0.27734-0.

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. 10.1128/JB.00427-20.

Zhang, A., and Belfort, M. (1992). Nucleotide sequence of a newly-identified Escherichia coli gene, stpA, encoding an H-NS-like protein. Nucleic Acids Res 20, 6735. 10.1093/nar/20.24.6735.

Zhang, A., Rimsky, S., Reaban, M.E., Buc, H., and Belfort, M. (1996). Escherichia coli protein analogs StpA and H-NS: regulatory loops, similar and disparate effects on nucleic acid dynamics. EMBO J 15, 1340-1349.

Zheng, M., Aslund, F., and Storz, G. (1998). Activation of the OxyR transcription factor by reversible disulfide bond formation. Science 279, 1718-1721. 10.1126/science.279.5357.1718.

Zhu, M., Mu, H., and Dai, X. (2024). Integrated control of bacterial growth and stress response by (p)ppGpp in Escherichia coli: A seesaw fashion. iScience 27, 108818. 10.1016/j.isci.2024.108818.