Functional and structural characterization of LeuO, a pleiotropic LysR-type transcription regulator in *Escherichia coli*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Susann Marlies Fragel

aus Berlin

Köln, September 2018

Berichterstatter:	Prof. Dr. Karin Schnetz
(Gutachter)	Prof. Dr. Niels Gehring
Tag der mündlichen Prüfung:	05.09.2018

Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.

Marie Curie

Contents

Zusammenfassung

LeuO ist ein konservierter LysR-Typ Transkriptionsregulator (LTTR) in *Escherichia coli* und anderen Gammaproteobakterien. LeuO reguliert Gene, die unter anderem für Pathogenität und Stressantworten verantwortlich sind. Viele der LeuO-regulierten Gene werden durch das globale Repressorprotein H-NS reprimiert und co-reguliert. Daher wird LeuO als globaler Antagonist von H-NS angesehen. LTTRs regulieren ihre Zielgene typischerweise als Tetramere, die sich aus zwei Dimeren der N-terminalen DNA-Bindungsdomänen (DBDs) und zwei Dimeren der C-terminalen Effektorbindungsdomänen (EBDs) zusammensetzen. Viele der bisher beschriebenen LTTRs benötigen einen Effektor, der strukturelle Veränderungen induziert. Diese strukturellen Veränderungen führen zur Änderung der DNA-Bindung des LTTR und ermöglichen die Aktivierung oder Repression spezifischer Zielgene. Die Regulation der LeuO Aktivität und deren strukturelle Merkmale sind jedoch weitgehend unbekannt.

In dieser Arbeit wurde der CRISPR-assoziierte *cas* Promotor als ideales *Reporter* identifiziert, da er sensitiv Veränderungen der LeuO-Aktivität abbildet. Mit diesem Reporter wurden in einem genetischen *Screen* neun hyperaktive Mutanten von LeuO identifiziert, welche den Reporter bei basaler Expression voll aktivieren. Die gelösten Kristallstrukturen der EBDs des wild-typischen LeuO und der hyperaktiven Ser120Asp Mutante zeigen, dass die Mutationen, die Hyperaktivität bewirken, in der Dimerisierungsoberfläche der EBD lokalisiert sind. Eine Überlagerung der Strukturen der EBDs deutet darauf hin, dass die S120D-Mutation eine strukturelle Veränderung induziert, welche einen Effektor-gebundenen Zustand nachahmt. Diese strukturelle Veränderung kann sich dann auf die sterische Anordnung der DBDs übertragen und somit die DNA-Bindeeigenschaften beeinflussen. Eine DNase I *Footprint* Analyse der LeuO-Bindestellen in der Region des *cas* Promotors zeigt eine hochspezifische DNA-Bindung der LeuO Mutante (S120D), welche außerdem im Vergleich zum wild-typischen LeuO ein deutlicheres Bindemuster an die DNA zeigt. Eine detaillierte Analyse dieses Bindemusters deckte eine "*Core-Site*" auf, welche im Weiteren die Definition eines palindromischen DNA-Bindemotifs ermöglichte. Die Daten deuten darauf hin, dass LeuO möglicherweise eine andere Gruppe von Zielgenen reguliert, wenn ein Effektor gebunden ist, der strukturelle Veränderungen des Proteins auslöst. Die Analyse von hyperaktiven Mutanten, welche vermutlich eine Effektor-gebundene Form abbilden, kann zum Verständnis des molekularen Mechanismus der Transkriptionsregulation durch LeuO beitragen.

Abstract

LeuO is a conserved LysR-type transcriptional regulator (LTTR) of global function in *Escherichia coli* and other Gammaproteobacteria. LeuO regulates genes related to pathogenicity and stress adaptation, many of which are co-regulated by the nucleoid-associated global repressor protein H-NS. Therefore, LeuO is considered a general H-NS antagonist. LTTRs typically are tetramers that consist of two dimeric N-terminal DNA-binding domains (DBDs) and two dimeric C-terminal effector-binding domains (EBDs). Many LTTRs described so far require an effector which induces structural changes that lead to alteration of the LTTR activity in repression or activation of specific target genes. However, the regulation of LeuO activity and its structural features are poorly understood.

In this work, I characterized the CRISPR-associated *cas* promoter as the most sensitive reporter for monitoring LeuO activity. Using this reporter nine LeuO mutants were identified in a genetic screen that render LeuO hyperactive. The solved crystal structures of the EBDs of wild-type LeuO and one of the hyperactive mutants (Serine 120 to aspartic acid exchange) revealed that all mutations causing hyperactivity localize to the dimerization interface of the EBD. Comparison and superposing of the structures of the wild-type and S120D EBDs suggests that the mutation of S120D induces a structural change that mimics an effector-bound state. This structural change is presumed to transmit to the arrangement of the DBDs and changes the binding to the DNA. DNase I footprinting of the *cas* promoter region presented here revealed a highly specific DNA-binding and a more distinct DNA-binding pattern of the S120D mutant compared to the wild-type LeuO. A detailed analysis of these footprints unraveled a “core-site” of palindromic sequence, which allowed the definition of a specific DNA-binding sequence motif for LeuO. The data indicate that LeuO might regulate a different set of target genes when an effector is bound. This is relevant in understanding the molecular mechanism of transcriptional regulation by LeuO and the role of LeuO as a global regulator.

1 Introduction

Bacteria face harsh and quickly changing environmental conditions encountering various stresses and changing nutrition sources. In order to adapt to, and survive in different environments, bacteria like *E. coli* have evolved a complex gene regulation network (Lee *et al.*, 2012). The regulation of transcription by transcription factors allows an individual control of a single gene or of specific subsets of genes in response to an environmental stimulus. Activating transcription factors bind close to the transcription start site and orchestrate the formation of the transcription initiation complex. These activators can directly recruit the RNA polymerase to the promoter or they induce structural changes of the DNA promoter region, enabling the binding and initiation of transcription by RNA polymerase (Browning & Busby, 2016). Most repressive transcription factors inhibit transcription initiation by steric hindrance, in which the repressor DNA-binding denies access of the RNA polymerase to the promoter. A repressor can also block transcription initiation by targeting the activating transcription factor. Contrary, a transcriptional factor can also obstruct a potential repressor and hence indirectly activate transcription (Browning & Busby, 2016).

The family of LysR-type transcriptional regulators (LTTRs) is most prevalent in bacteria. The LysR-type transcriptional regulators activate or repress diverse sets of genes related to virulence, stress responses and motility (Maddocks & Oyston, 2008). The activity of LTTRs is controlled by an effector that binds to an “effector-binding cleft” or elsewhere, or by modification of amino acids. LeuO is a conserved LysR-type transcription regulator in *Enterobacteriaceae* regulating many targets related to stress adaptation, pathogenicity and biofilm formation (Henikoff *et al.*, 1988, Dillon *et al.*, 2012, Guadarrama *et al.*, 2014b). However, the control of LeuO activity and the mechanism of its regulation have not been studied in detail. For LeuO, no effector or condition is known that modulates its activity. Further, the intrinsic structural properties of LeuO and its DNA-binding mode are open questions. In this work, a functional and structural analysis of the LysR-type transcription regulator LeuO was conducted.

1.1 LysR-type transcription regulators

LysR-type transcription regulators (LTTRs) represent the largest family of transcription factors, resembling 16% of the overall repertoire of transcription factors in bacteria (Tropel & van der Meer, 2004). They are highly conserved and orthologues have been identified in Archaea and eukaryotic organisms (Henikoff *et al.*, 1988, Schell, 1993, Maddocks & Oyston, 2008). To date, 46 LTTRs are present in *Escherichia coli* K12 (Karp *et al.*, 2014) and these LTTRs are involved in the regulation of diverse cellular functions as oxidative stress responses, metabolic pathways, virulence and biofilm formation (Laishram & Gowrishankar, 2007, Maddocks & Oyston, 2008, Liu *et al.*, 2011). Many LysR-

type-regulators are homo-tetramers and their activity is controlled by structural changes within the tetrameric structure. These structural changes can be induced by effectors or the modification of amino acid residues. Effectors can bind to a specific “binding pocket”, in the interface of two dimers or elsewhere on the tetramer. Different effectors that bind to the LTTR can regulate different target loci or even change the LTTR from an activator to a repressor at a given target loci (Laishram & Gowrishankar, 2007). In general, LTTRs regulate expression of a target gene by binding two distinct sites in the promoter region matching the degenerated consensus binding motif T-11N-A (Maddocks & Oyston, 2008).

1.2 Structure of LysR-type transcription regulators (LTTRs)

LTTR monomers comprises of approximately 300 amino acids. They consist of a characteristic N-terminal DNA-binding domain (DBD) with a winged helix-turn-helix (wHTH) domain that connects via a helical linker and a flexible hinge to a C-terminal effector-binding domain (EBD, also called regulatory domain, RD; Figure 1)(Maddocks & Oyston, 2008, Momany & Neidle, 2012) The DNA-binding domain is highly conserved and the wHTH domain comprises of three core helices, forming a helical bundle with an almost triangular outline. The specific DNA target is determined by the α3 DNA-recognition helix and forms a DNA-protein interface by insertion into the major groove of the DNA (Aravind *et al.*, 2005). The “wing” is a β-strand hairpin unit between helix α3 and α4 and usually interacts with the minor groove of the DNA (Aravind *et al.*, 2005). Two DBDs can dimerize via their α4 helical linker and bind to a palindromic DNA sequence in an inverted arrangement of the DBDs.

The homology of the C-terminal effector-binding domain (EBD) in LTTRs is less well conserved on the amino acid sequence level, but the EBDs have a conserved domain structure. The EBDs exhibit a typical fold with two α/β subdomains (also called RD-1 and RD-2 domains) that connect by two extended, antiparallel cross-over β-strands. A cleft locates between the two subdomains of the EBD (RD-1 and RD-2) and for some LTTRs effectors are described to bind in this cleft. Two EBDs dimerize in antiparallel head-to-tail orientation and form a conserved dimerization interface (Figure 2). In this dimerization interface, likewise effector-binding was described (Jiang *et al.*, 2018).

Several structures of LTTRs are solved so far, and a typical LTTR tetramer can be described as a dimer of dimers, where a dimer comprises of two monomers, dimerizing via their EBDs and displaying an extended and a compact conformation of their DBDs, as shown by the structure of a structure model of LeuO in Figure 2. The assembly of the tetramer arises when the compact DBD subunits interact with the extended DBD subunits of the other EBD dimer and contact is made by their linker helices (Henikoff *et al.*, 1988, Maddocks & Oyston, 2008, Momany & Neidle, 2012, Taylor *et al.*, 2012).

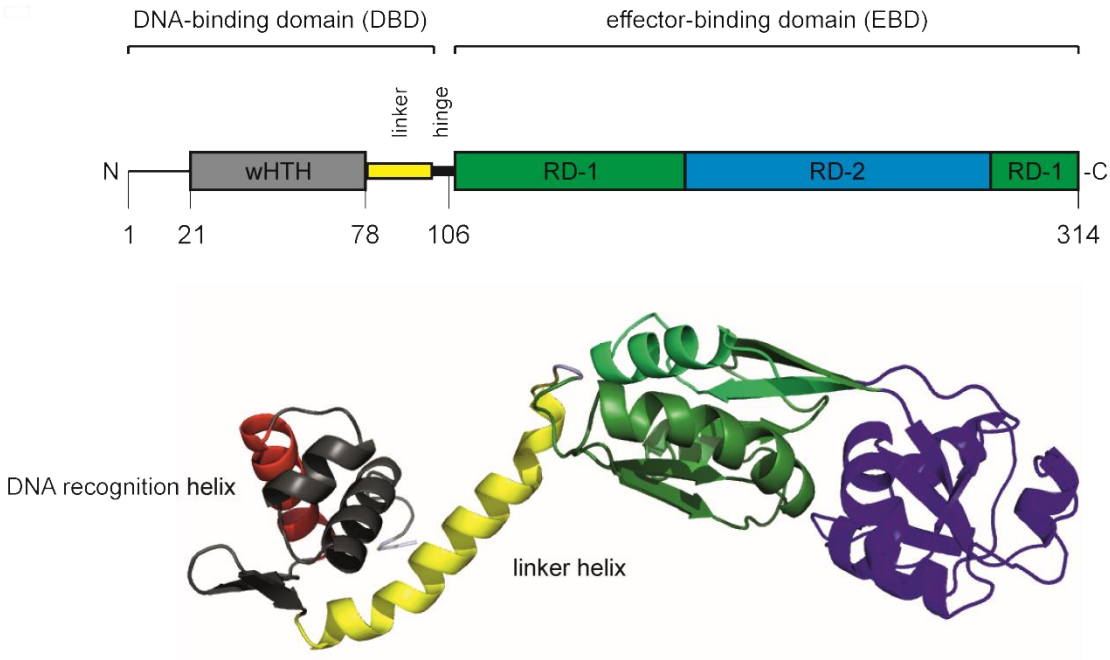

Figure 1: The domain structure of a monomeric LysR-type regulator.

LTTR monomers consist of an N-terminal winged HTH DNA-binding domain (DBD) connected via a helical linker and a hinge to a C-terminal effector-binding domain (EBD). The DNA recognition helix is shown in red. The α4 helical linker, depicted in yellow is important for dimerization of the DBDs. A "cleft" exists between the two regulatory subdomains (RD-1 and RD-2) where a putative effector binds or the modification of an amino acid residue may alter the protein structure. However, alternative effector-binding sites exist for some LTTRs in the interface of two EBD. Depicted is the monomeric structure of LeuO as example of a LysR-type transcriptional regulator (uniprot entry: P10151; https://swissmodel.expasy.org/repository/uniprot/P10151). The figure was adapted from (Lerche *et al.*, 2016) and generated using PyMOL (The PyMOL Molecular Graphics System).

The tetrameric LTTR harbors two dimers of the DBD that bind to two adjacent imperfect palindromic sites of the DNA, a high affinity DNA binding site (generally called RBS, regulatory binding site) and an alternative low-affinity DNA-binding site (ABS, activator binding site) (Toledano *et al.*, 1994, Wek & Hatfield, 1988, Rhee *et al.*, 1998). For most LTTRs, this target recognition and binding to the DNA is independent of a bound effector. However, the binding of an effector or the mutation of an amino acid residue can induce structural changes of the EBDs and subsequently change the distance or the angle of the two DBDs, as shown for OxyR (Figure 2D) (Jo *et al.*, 2015). This steric re-arrangement alters the DNA-binding pattern, the DBD dimer bound to the high affinity DNA-binding site is maintained, while the second DBD dimer "slides" along the DNA to an alternative DNA-binding site. This "sliding" along the sequence, leads to a different DNA-binding pattern, resulting in a differential regulation of gene expression. LTTRs regulate transcription of target genes by this "sliding-dimer"-mechanism, as described for example for DntR (Lerche *et al.*, 2016, Toledano *et al.*, 1994). The mechanism of the sliding dimer was also shown by DNase I footprinting for BenM, ArgP and OxyR (Bundy *et al.*, 2002, Laishram & Gowrishankar, 2007, Kullik *et al.*, 1995).

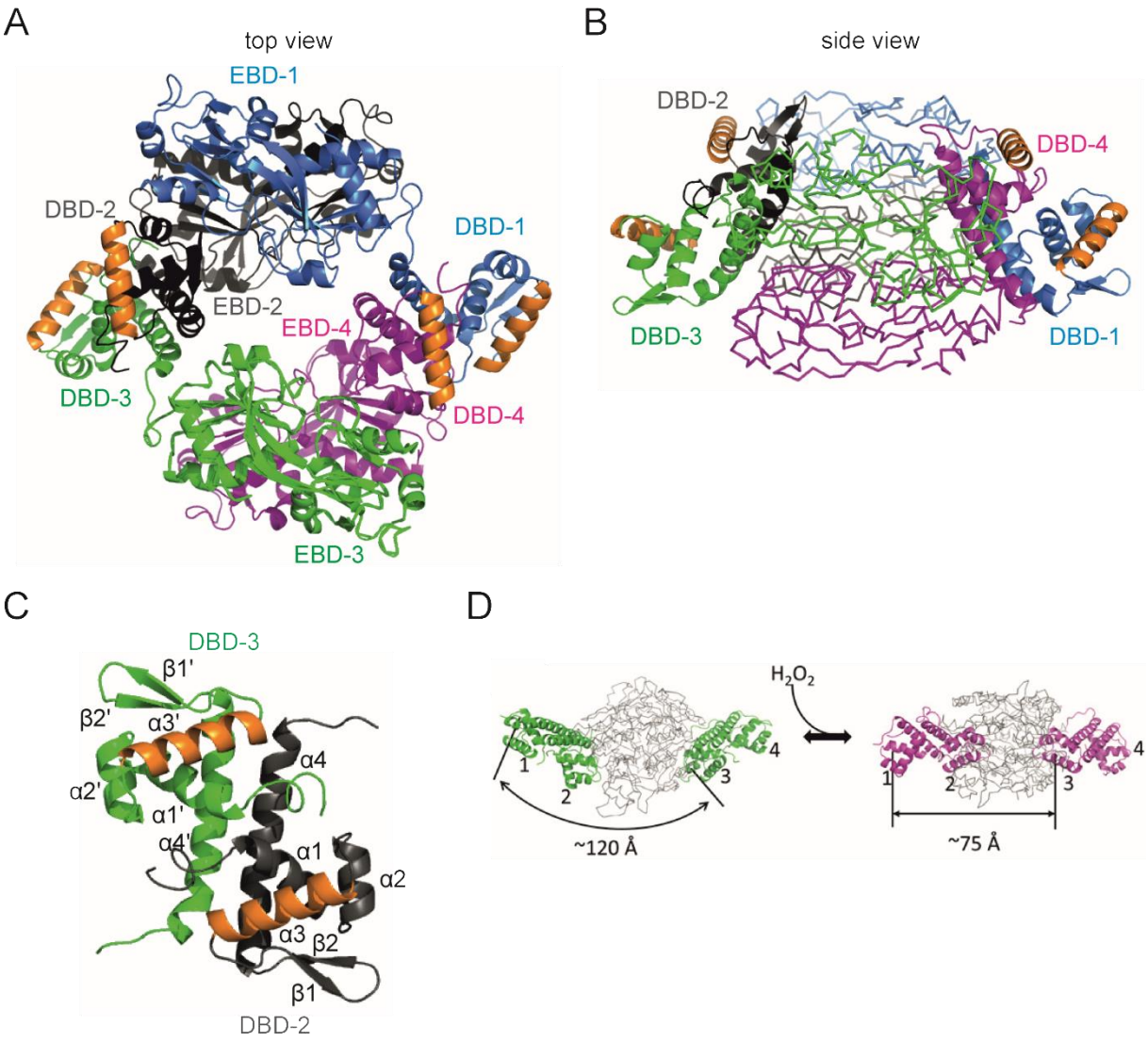

Figure 2: Predicted tetrameric structure of LeuO.

A top view (A) and a side view (B) of the predicted LeuO tetramer structure with monomeric LeuO proteins highlighted in blue, grey, green and magenta (uniprot entry: P10151; https://swissmodel.expasy.org/repository/uniprot/P10151 based on (Lunin *et al.*, 2005)). Dimerization is shown of the effector-binding domains EBD-1 (blue) with EBD-2 (grey) and dimerization of EBD-3 (green) with EBD-4 (magenta). The DNA-binding domain DBD-1 dimerizes with DBD-4 and the DBD-2 dimerizes with DBD-3. The α3 DNA-recognition helices are marked in orange. (C) A structural model of the DBD-dimer. The individual monomers of the DBDs are depicted in grey and green with the α3 DNA-recognition helices marked in orange. The dimerization of the DBDs is mediated by the α4 linker helices. (D) The reduced (left side) and oxidized form of OxyR (right side) shows the structural change transmitted to the DBDs upon oxidization of the protein (Jo *et al.*, 2015). The α3 helices of the DBDs are indicated by number and the distance between DBD-1 and DBD-3 is 120 Å in the reduced full-length protein. The DBD-1/DBD-2 dimer and the DBD-3/DBD-4 dimer should get closer by 45 Å during the transition to the oxidized state (Jo *et al.*, 2015). The figure was generated using PyMOL (The PyMOL Molecular Graphics System).

1.3 Regulation of target genes by LysR-type regulators

Structural changes in the LTTR tetramer can be induced by binding of an effector or a modification of an amino acid residue. The LTTR BenM in *Acinetobacter sp*. controls the aromatic compound metabolism by regulating genes in the β-ketoadipate pathway (Bundy *et al.*, 2002). In absence of an effector, BenM binds to the promoter of *benA* (encoding a benzoate-dioxygenase) and represses transcription. In the presence of the metabolites benzoate and muconate structural changes of BenM shift the binding to a different DNA-binding site of the *benA* promoter region and BenM synergistically activates the transcription of *benA* expression. In this way, the benzoate degradation is directly coupled to the presence of benzoate in the cell (Bundy *et al.*, 2002).

Modifications of an amino acid are possibly caused by a change of pH or by oxidation. An example for the control of an LTTR by amino acid modification is the hydroxylation of the cysteine residue (C199) in OxyR, leading to structural changes of the tetrameric OxyR. Consequently, the OxyR-regulated small RNA oxyS is activated which regulates the oxidative stress response (Toledano *et al.*, 1994, Jo *et al.*, 2015, Kullik *et al.*, 1995). The OxyR-C199D mutant described in (Jo *et al.*, 2015) mimics an oxidized form of the protein that renders the protein active. For AphB in *Vibrio cholera* the N100E mutation renders the pH-sensitive transcription regulator active at non-permissive pH 8.5. The structures of wild-type AphB and of mutant N100E revealed structural changes of the tetrameric AphB that render AphB-N100E constitutively active at its target loci (Taylor *et al.*, 2012). Mutations in the EBD of LysR-type regulators can render the LTTR constitutively active, suggesting that the mutation mimics an effector-bound state and induce the same structural change as the effector. Constitutively active mutants were also described for the LTTRs AmpR, CysB, DarR, NahR, TsaR, and XapR (Balcewich *et al.*, 2010, Colyer & Kredich, 1994, Jones *et al.*, 2018, Schell *et al.*, 1990, Monferrer *et al.*, 2010, Jørgensen & Dandanell, 1999).

Effectors are usually metabolites or ions and can bind either to the cleft between the two RD subdomains of the EBD or to a secondary binding site in the dimeric interface of two EBDs (Jiang *et al.*, 2018). In some cases, different effectors are bound by the LTTR.

These effectors can alter the transcription regulation of the LTTR as a direct read out of a stimulus in the cell. For example, to prevent toxic levels of arginine in the cell, the LTTR ArgP binds the effector arginine and activates the transcription of the arginine exporter encoding *argO.* When lysine is bound to ArgP transcription of *argO* is inhibited and also additional ArgP target genes are inhibited (Laishram & Gowrishankar, 2007). The LTTR NdhR coordinates the balance of carbon and nitrogen metabolism in cyanobacteria. For NdhR two effector binding sites were identified, where different effectors can bind (Jiang *et al.*, 2018). One binding site is located in the cleft of the effector-binding domain between two RD subdomains and binds 2-PG (2-phosphoglycolate, depicted in green; Figure

3). A secondary binding site is located in the dimeric interface of two EBDs of NdhR and binds 2-OG (2-oxoglutarate, shown in blue; Figure 3). Binding of either one of the effectors to different sites of the protein allows a fine-tuning in the regulation, depending on the incoming signal. NdhR adopts a repressor conformation upon 2-OG binding. The binding of 2-PG to the cleft induces structural changes and leads to the activation of gene expression. When the repressing effector is bound, the activating effector is no longer able to bind due to the structural changes (Jiang *et al.*, 2018).

It is noteworthy that some of the described LTTRs map next to genes they regulate, linking their location to their function. For example, BenM regulates the transcription of the divergent *benA* gene (Maddocks & Oyston, 2008, Bundy *et al.*, 2002). In the *E. coli* genome about 50% of the LTTRs map divergently to their target genes. The LTTR IlvY is located divergent to its regulated target gene *ilvC* (Wek & Hatfield, 1988). The *ilvC* gene encodes a ketol-acid reductoisomerase for branched-chain amino acid synthesis (Rhee *et al.*, 1998).

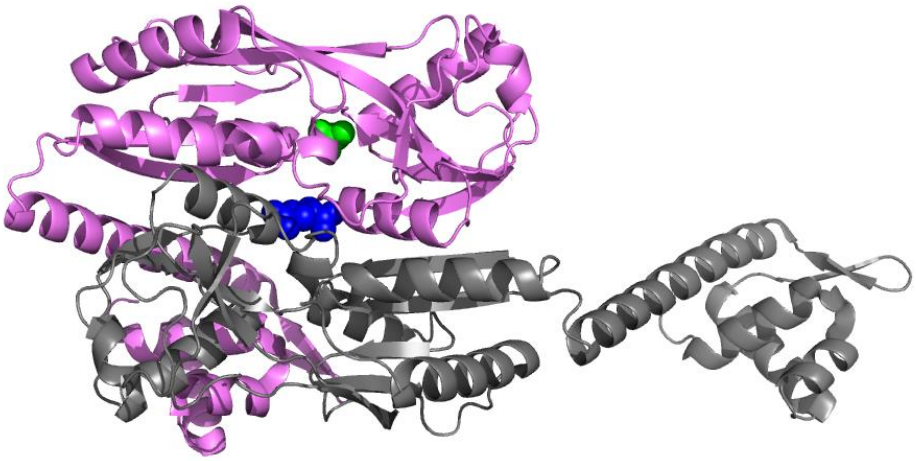

Figure 3: The binding of effectors to different sites of the LTTR NdhR.

Depicted is dimer of the NdhR tetrameric structure with two different effectors (indicated in green and blue) bound to different sites in the structure (Jiang *et al.*, 2018). The effector 2-PG (in green) binds to the cleft between the two RD subdomains of NdhR and induces gene expression. The effector 2-OG (in blue) binds to a secondary binding site in the dimer interface between two EBDs and acts as co-repressor of NdhR (Jiang *et al.*, 2018).

1.4 LeuO

LeuO is a pleiotropic LysR-type transcriptional regulator and conserved in Gammaproteobacteria (Henikoff *et al.*, 1988, Hernández-Lucas & Calva, 2012, The UniProt Consortium, 2017). LeuO has been best studied in the species *Escherichia coli* and *Salmonella enterica,* as well as in *Vibrio cholera.* The LeuO proteins are 90% identical in the *Enterobacteriaceae* and 50% identical to LeuO in *V. cholera* and LeuO is presumed to act as a tetramer, like other LTTRs (Guadarrama *et al.*, 2014b).

LeuO regulates more than 100 loci in *E. coli* and *S. enterica*, as identified by microarray transcriptome analyses and genomic SELEX screening (Dillon *et al.*, 2012, Shimada *et al.*, 2011, Stratmann *et al.*, 2012, Ishihama *et al.*, 2016). LeuO activates and represses loci related to acid-stress resistance, pathogenicity, multi-drug efflux, and biofilm formation (Henikoff *et al.*, 1988, Dillon *et al.*, 2012, Guadarrama *et al.*, 2014b).

Among many targets, LeuO activates the transcription of CRISPR-*cas* bacterial defense system, important for defense against foreign DNA and phage infection (Medina-Aparicio *et al.*, 2011, Pul *et al.*, 2010, Westra *et al.*, 2012, Hernández-Lucas & Calva, 2012). Further, LeuO plays a positive role in the regulation of the *yjjQ-bglJ* operon, encoding FixJ/NarL-type transcriptional regulators and the *bgl* operon, encoding proteins for uptake and utilization of aryl-β,D-glucosides (Ueguchi *et al.*, 1998, Stratmann *et al.*, 2008). In *Vibrio*, LeuO inhibits the expression of the cholera toxin (CT) and the toxin co-regulated pilus (TCP). Furthermore, *leuO* expression is activated by the membrane bound transcription regulator ToxR, which links *leuO* expression to various stimuli that activate the ToxR-dependent virulence regulon of *V. cholerae* (Ayala *et al.*, 2018, Bina *et al.*, 2013, Ante *et al.*, 2015a, Ante *et al.*, 2015b).

Many LeuO-activated target loci, like the CRISPR-*cas* operon, are repressed by the global repressor H-NS (heat-stable nucleoid structuring protein) and its paralogue StpA. LeuO presumably competes with H-NS for DNA-binding and delimits spreading and thus the formation of the repressing oligomeric H-NS complex (De la Cruz *et al.*, 2007, Chen & Wu, 2005, Westra *et al.*, 2010). LeuO is therefore considered an H-NS antagonist (Ueguchi *et al.*, 1998, De la Cruz *et al.*, 2007). However, in several pathways LeuO has been identified as additional repressor of H-NS-repressed gene loci and thus seems to functionally cooperate with H-NS or functions as a “back-up” of H-NS. For example, LeuO causes repression via HilE of pathogenicity island 1 (SPI-1) in Salmonella when repression by H-NS is impaired (Espinosa & Casadesús, 2014).

The *leuO* gene (GenBank Accession number U00096.3 K12: 84,368 to 85,312) is located between the divergent *leuLABCD* operon and the *ilvIH* operon, both encoding enzymes for the synthesis of branched-chain amino acids (Gemmill *et al.*, 1983, Wang & Calvo, 1993) (Figure 4). In standard laboratory conditions *leuO* is silent, but is moderately induced in response to starvation in stationary

growth phase (Majumder *et al.*, 2001, Fang *et al.*, 2000). Transcription of *leuO* is repressed by H-NS and StpA under standard laboratory conditions (Klauck *et al.*, 1997, Stratmann *et al.*, 2012). Transcription of *leuO* is activated by the transcription regulator BglJ-RcsB and by the LysR-type regulator LrhA (Stratmann *et al.*, 2012, Breddermann & Schnetz, 2017). LeuO itself shows a weak positive autoregulation, when expressed in high copy numbers (Chen & Wu, 2005). However, LeuO acts as a negative autoregulator, counteracting the activation by BglJ-RcsB as well as by LrhA, and derepression in *hns stpA* mutants (Stratmann *et al.*, 2012, Breddermann & Schnetz, 2017). Two distinct LeuO DNA-binding sites in the promoter sequence of *leuO* have been described (Stratmann *et al.*, 2012).

Although LeuO regulates many target loci, only a degenerated consensus DNA-binding motif was identified displaying the typical weak LysR-type binding motif T-11N-A (Dillon *et al.*, 2012, Maddocks & Oyston, 2008). The weak DNA-binding specificity might be increased when LeuO activity is altered by an effector or an amino acid modification, as described for other LysR-type regulators. Up to date, no effector or condition has been described that modulates the LeuO activity.

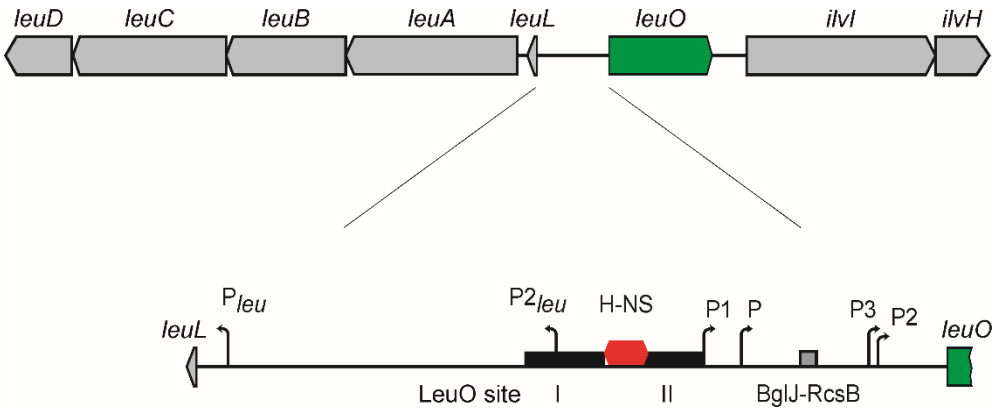

Figure 4: Schematic model of the organization of the *leuO* promoter region including the binding sites of transcriptional regulators.

The *leuO* gene is located between the operons *ilvIH* and *leuLABCD*, both encoding for enzymes involved in branched-chain amino acid synthesis. LeuO DNA-binding sites are indicated as black boxes, the H-NS nucleation site is shown as red box and the binding site for the heterodimeric activator BglJ-RcsB as grey box (Stratmann *et al.*, 2012). Transcription of *leuO* is repressed by H-NS and StpA under standard laboratory conditions. The promoter P2 is activated by BglJ-RcsB, while LrhA mainly activates P3 (and weakly P and P1). In *hns stpA* mutant strains, LeuO acts as an autorepressor, and it counteracts the activation by LrhA and BglJ-RcsB in the wild-type (Stratmann *et al.*, 2012, Breddermann & Schnetz, 2017).

1.5 Objectives of this thesis

LeuO is a pleiotropic LysR-type regulator modulating the transcription of many target loci in *Escherichia coli*. For other LysR-type regulators effectors or modifications of amino acids have been described which alter the activity of the protein. No effector is known so far for LeuO and the control of LeuO activity is poorly understood. Additionally, the mode of LeuO binding to the DNA and the mechanism of regulation of its target loci remain open questions.

In this study, I identified the *cas* promoter as the most sensitive reporter for analyzing the activity of LeuO. Using this reporter, I isolated nine LeuO mutants that are highly active causing full activation even when present at low levels. I further characterize the LeuO structure, by crystallization of the effector-binding domains (EBDs) of the wild-type LeuO and one of the hyperactive mutants (S120D). The structures revealed that all hyperactive LeuO mutants map in the dimerization interface of the EBDs and the S120D mutant causes a structural change.

In addition, I determined the DNA-binding mode of wild-type LeuO, of the hyperactive LeuO-S120D mutant and of the LeuO-DBD dimer. The DNA-binding specificity of the S120D mutant and of the DBD dimer is higher than of wild-type LeuO, which enabled the identification of a palindromic DNA-binding sequence and a specific LeuO DNA-binding motif. Furthermore, using the motif I identified and initiated the characterization of putative LeuO target loci.

2 Results

2.1 Screen for hyperactive LeuO mutants

For the LysR-type regulator LeuO no effector is known, while for other LysR-type regulators, such as AphB, XapR and OxyR, effectors have been described and amino acid mutations have been characterized that render these regulators constitutively active (Taylor *et al.*, 2012, Jørgensen & Dandanell, 1999, Jo *et al.*, 2015). Therefore, to characterize LeuO, I set up a screen for LeuO mutants that are constitutively active. As a reporter, the promoter of the *cas* (CRISPR-associated) operon, which is activated by LeuO (Pul *et al.*, 2010, Westra *et al.*, 2010) was fused to *lacZ,* and this *Pcas-lacZ* fusion was integrated into the genome (at the phage λ attachment site *attB*) in a *ΔlacZ ΔleuO* Δ(*yjjP-yjjQ-bglJ*) background (T1610; Figure 5A). The expression of the *Pcas-lacZ* reporter was strongly activated by LeuO (from 6 units to 568 units), when LeuO was expressed at high levels using a plasmid carrying *leuO* under the control of the IPTG (isopropyl-β,D-thiogalactopyranoside) inducible tac promoter (Figure: 5B and 5C), as shown before (Westra *et al.*, 2010). At basal expression levels of LeuO, without induction of *Ptac*, LeuO activated *Pcas-lacZ* moderately (compare 51 units to 6 units of control, Figure 5B). Thus, the basal expression level of LeuO is a suitable condition to screen for constitutive LeuO mutants. In this screen, promoter *lacZ* fusions of the LeuO-activated *bgl* and *yjjQ* promoters were also included as reporters (Schubert, 2013). However the *Pcas-lacZ* fusion was found to be the most sensitive reporter. For mutagenesis the *leuO* gene fragment was amplified by PCR in parallel reactions, using the non-proofreading Taq polymerase, and the PCR fragments were ligated into the *Ptac* expression plasmid (Figure 5A). Transformants of the *Pcas-lacZ* reporter strain with these ligations were screened for a Lac-positive phenotype on tryptone X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) indicator plates without IPTG (Figure 5A and 5C). Nine different mutations of single amino acid residues were isolated from 25 independent mutagenesis PCR reactions. These mutants include 3 independent isolates of LeuO-S128P as well as 2 independent isolates each of LeuO-H142R, LeuO-Q210R, LeuO-A237V and LeuO-H254R, respectively. The mutants LeuO-T127I and LeuO-R218C, respectively, were isolated once. The mutant LeuO-M244T was isolated as a single mutant and as an independent double mutant with the second mutation V230I (Figure 8A). In addition, mutant LeuO-S120D was isolated as a triple mutant with additional amino acid exchanges at residues E111D and D205N. This triple mutant was isolated using the *Pbgl-lacZ* reporter strain (Schubert, 2013) (Figure 8A). To analyze the impact of LeuO Ser120 to Asp exchange, a single mutant was constructed and used for further analysis. Multiple independent isolations of the same mutants suggest that the screen for constitutively active LeuO mutants was saturated.

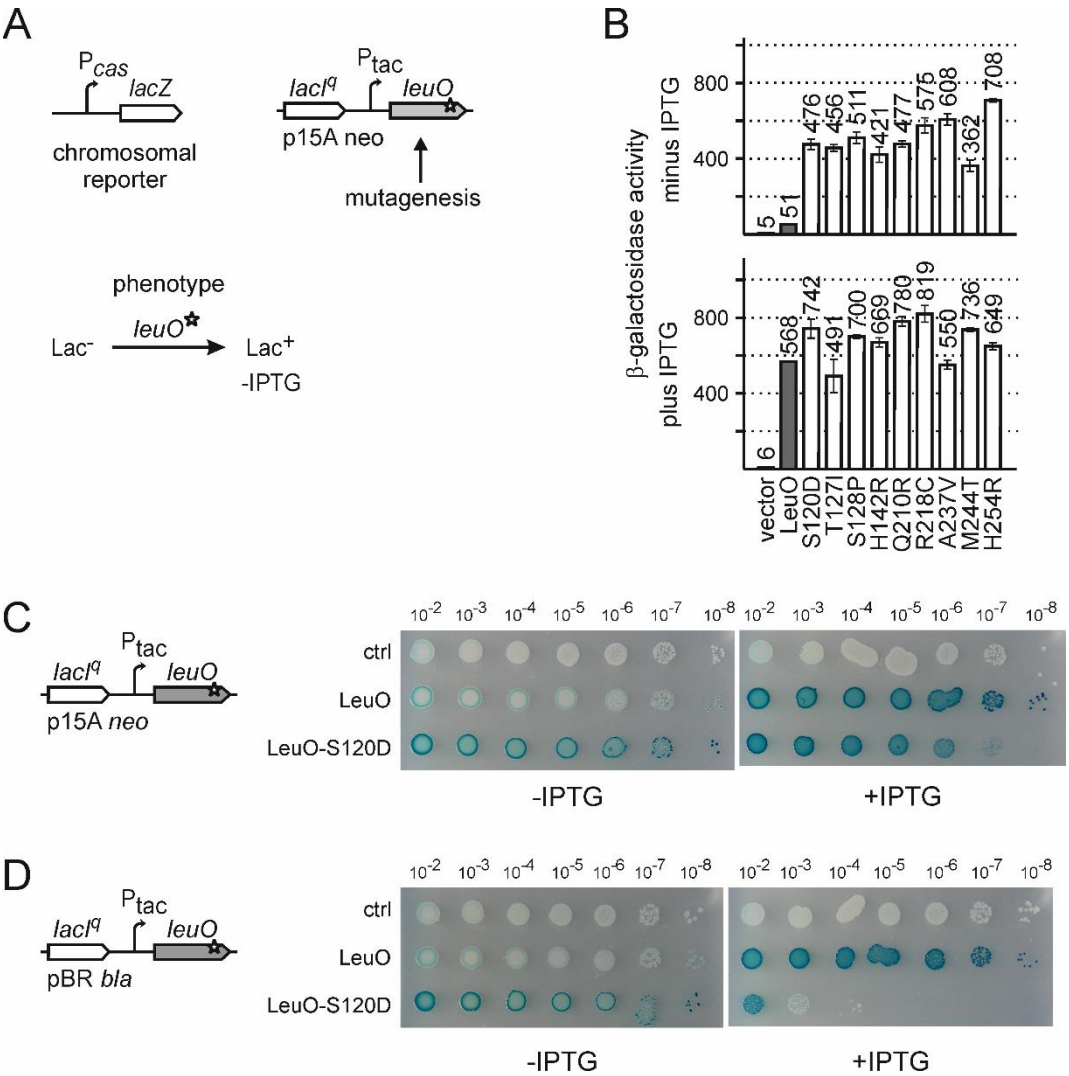

Figure 5: Screen for constitutively active LeuO mutants.

(A) To screen for active LeuO mutants, a P*cas-lacZ* fusion was integrated into the chromosomal *attB-site* of strain T1281 [Δ*lacZ* Δ*leuO* Δ(*yjjP-yjjQ-bglJ*)], and this *Pcas-lacZ* fusion was used as reporter for LeuO activity (top left side). The *leuO* gene was amplified by non-proofreading PCR and the *leuO* fragments were ligated into plasmid pKESK22 (carrying a p15A replication origin, kanamycin resistance) to place *leuO* under control of the IPTG inducible *tac* promoter (*Ptac*) (top right side). Transformants were selected on tryptone X-Gal, kanamycin indicator plates and screened for a Lac-positive phenotype (A bottom). Of a total of 17 Lac^+ clones the *leuO* fragment was sequenced, which revealed 9 different *leuO* mutants (see Table 3). (B) Activation of *Pcas* by LeuO mutants. Expression levels of *Pcas-lacZ* were determined of cultures grown without induction (top panel) and with 1 mM IPTG for induction of *leuO* expression (bottom panel). Expression was determined in absence (vector control, pKESK22) and in presence of plasmid-provided LeuO (pKETS5) and its mutants S120D (pKESL104), T127I (pKESL74), S128P (pKESL75), H142R (pKESL76), Q210R (pKESL77), R218C (pKESL78), A237V (pKESL80), M244T (pKESL73), and H254R (pKESL101). Cultures for β-galactosidase assays were inoculated from overnight cultures and grown in LB medium with kanamycin to an OD_{600} of 0.5. Where indicated 1 mM IPTG was added both to the overnight and exponential culture. Mean values of at least 3 independent biological replicates are shown as bars, and error bars indicate standard deviations. (C and D) Spotting of P*cas-lacZ* strain (T1281) carrying *leuO* on a plasmid. Over day cultures were grown in LB with antibiotics to an OD_{600} ~ 3.0 and diluted to OD_{600} = 1.0 for spotting. OD_{600} = 1.0 equals approximately 10^9 cells per ml, and 10 µl (= 10^{-2} ml) and 10 µl of the ten-fold serial dilutions (10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8})

Figure 5 continued from the previous page

were spotted on tryptone X-Gal indicator plates, with ampicillin or kanamycin. Plates were supplemented with 200 µM IPTG where indicated. The reporter strain T1281 was transformed with plasmids expressing LeuO or LeuO-S120D under the control of the IPTG inducible *lacI*^{q}*-tac* promoter system and either (C) expressed moderate levels of LeuO from a low to medium copy plasmid (p15A origin of replication; pKESK22, control; pKETS5, LeuO; pKESL104 LeuO-S120D) or (D) high levels of LeuO from a pBR derived plasmid (pKES334, control; pKESMS38, LeuO; pKESMS79 LeuO-S120D).

The LeuO mutants conferring a Lac^{+} phenotype at basal expression levels (without IPTG) were used in a quantitative assay of *Pcas-lacZ* activation (Figure 5B). To this end, transformants of the *Pcas-lacZ* reporter strain with plasmids carrying *leuO* mutants were grown without (no IPTG) and with induction of *leuO* expression (1 mM IPTG) and the β-galactosidase activities were determined. When LeuO expression was induced (1 mM IPTG), all mutants caused full activation of *Pcas-lacZ* (Figure 5B, bottom panel). At basal expression levels of *leuO*, without induction, all LeuO mutants isolated in the screen fully activated the *Pcas-lacZ* reporter (Figure 5B top panel). The expression levels ranged from 362 to 708 units referring to a 7 to 14-fold stronger activation of *Pcas-lacZ* than by wild-type LeuO (51 units, Figure 5B top panel). Thus, each of the nine mutations of single amino acid residues render LeuO hyperactive. Interestingly, all the residues locate in the effector-binding domain (EBD, amino acid residue 108 to 314), and not a single mutant mapped to the N-terminal DNA-binding domain (DBD, amino acid residue 1 to 101).

To visualize the increased activity of LeuO-S120D and to further analyze the growth of the mutant, transformants of the P*cas-lacZ* reporter strain (T1281) with plasmids expressing wild-type (pKETS5, p15A-ori or pKESMS38, pBR-ori) and mutant LeuO (pKESL104, p15A-ori or pKESMS79, pBR-ori) were spotted in serial dilutions on tryptone X-Gal indicator plates with and without 200 µM IPTG. Empty plasmids served as control (pKESK22, p15A-ori or pKES334, pBR-ori). The plasmids carrying ap15A-ori express moderate levels of LeuO (Figure 5C), while pBR-derived plasmids express high levels of LeuO (Figure 5D), under the control of the IPTG inducible *tac* promoter. Transformants were grown in LB with antibiotics to an OD_{600}~3.0 and diluted to OD_{600}=1.0 (equals approximately 10^{9} cells per ml). Ten-fold serial dilutions were made and 10 µl was pipetted per spot of each dilution (= 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, 10^{-8} per ml).

The P*cas-lacZ* strain expressing LeuO showed a blue phenotype when induced with IPTG and a light blue color when the plasmids were not induced. In agreement with the data from β-galactosidase assays, the mutant LeuO-S120D conferred a Lac^{+} phenotype even when no IPTG was present. Further the cells expressing wild-type LeuO showed a normal growth (up to dilution 10^{-8}), while the hyperactive LeuO-S120D mutant conferred a growth defect when the expression was induced from the high-copy number pBR-ori plasmid (no growth as of dilution 10^{-4}). This suggests that high levels of

hyperactive LeuO-S120D are toxic for *E. coli* cells. Further, all colonies of the dilution 10^{-3} and 10^{-4} were Lac^-. The sequencing of *leuO* in several of these clones revealed an additional mutation W303*. The additional mutation leads to a truncated protein with lost hyperactivity.

2.2 Crystal structure of the LeuO effector-binding domain (EBD)

The nine mutations causing hyperactivity of LeuO, map in the effector-binding domain (EBD) that extends from residue Ala108 to Arg314, according to a structural model of *E. coli* LeuO (https://swissmodel.expasy.org/repository/uniprot/P10151) and an alignment of *S. enterica* LeuO with other LysR-type regulators (Figure 6) (Guadarrama *et al.*, 2014a). To characterize the functional relevance of these residues, the crystal structure of the LeuO-EBD was determined in cooperation with the Lab of Prof. Dr. Ulrich Baumann, Institute for Biochemistry, Cologne (Figure 7). In addition, the structure of the EBD of hyperactive mutant LeuO-S120D was solved in an orthorhombic (space group I222) and a monoclinic (space group C2) crystal form (Figure 7A). Most residues are well resolved in the electron density maps with the exception of the loop connecting strands β4 and β5 (residues 151-158 including helix α6; secondary structures are numbered according to the prediction for the full-length protein), which has weak electron density in the wild-type and in the monoclinic crystal form of the S120D structure, but it could be traced in the orthorhombic S120D crystal form. Furthermore, the loop between β5 and β6 residues (R173 to E175) is not resolved in the wild-type and monoclinic mutant structures and possesses only weak density in the orthorhombic S120D structure.

The EBD exhibits the typical fold with two α/β subdomains (also denoted as RD domains) connected by two extended, antiparallel cross-over β-strands β6 and β11. DALI and PDB FOLD searches revealed the effector domain of DntR of *Burckholderia sp.* as the structurally most similar protein with known structure (PDB ID: 5AE5) (Lerche *et al.*, 2016).

In all three crystal structures, the EBDs form a dimer (Figure 7A) that is predicted to be stable in solution under standard conditions as judged by the PISA server (Krissinel & Henrick, 2007). This dimer is similar to those observed in structures of other isolated EBDs as well as in full-length structures of LysR-type regulators, e.g. in the structures of AphB (Taylor *et al.*, 2012). As anticipated, there is no sign of a tetramer in all three crystal structures presented here, because tetramerization requires the α4 linker helix of the DNA-binding domain that is not present in the constructs.

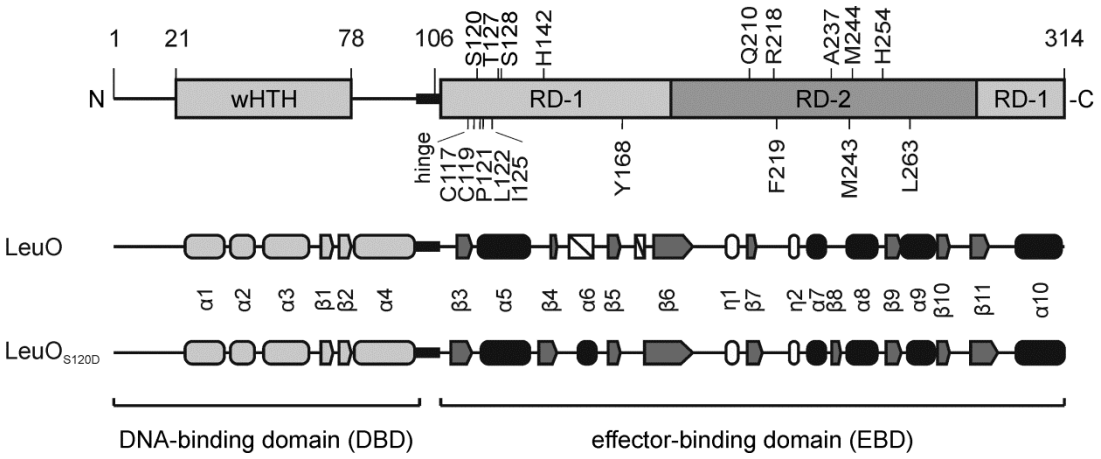

Figure 6: Domain architecture and secondary structure elements of LeuO and LeuO-S120D.

(A) Schematic presentation of the LeuO domain and secondary structure, with the N-terminal DNA-binding domain (DBD) carrying a wHTH motif, and the C-terminal effector-binding domain (EBD), connected by a hinge region. Mutations of residues, which render the protein hyperactive, are shown atop and additional mutations are depicted below the scheme. The secondary structure of the solved crystal structures is depicted in a schematic representation showing α helices as round filled cylinders, η helices as round unfilled cylinders, β sheets as arrows and non-solved areas as white boxes. The secondary structure of the DBD (shown in light grey) is based on a structural model (https://swissmodel.expasy.org/repository/uniprot/P10151).

Detailed analysis of the crystal structures by Prof. Dr. Baumann revealed the following features: The three-dimensional structures of wild-type and S120D mutant are similar, however with some important differences (Figure 7D). The root-mean-square (RMS) deviation of all 192 equivalent Cα pairs is about 1.7 Å for the superposition of wild-type and each of the two S120D structures, while superposing the latter results in a significantly lower RMS deviation of 1.2 Å. Flexible alignment by the RAPIDO web server (Mosca & Schneider, 2008) reduced the RMS deviation to about 0.8 Å for 175 Cα pairs, revealing two rigid bodies. The first rigid body essentially consists of β-strands 3 and 4 and helices α5 and α10. This becomes more apparent when superposing only domains RD-2 (residues 185 to 285), which results in an RMS deviation of the 101 Cα pairs of 0.4 Å between wild-type and each mutant structure. Superposing in this way the entire structures lead to a good alignment of both S120D mutant structures, while the wild-type differs especially in the position of β-strands β3 and β4 as well as helices α5 and α10 of domain RD-1 (Figure 7D). Remarkably, the Cα atoms of the first N-terminally resolved residue (Arg112) have a distance of about 4 Å between wild-type and mutant (Figure 7D). Similar but less pronounced differences between inactive and active EBDs have been observed previously, e.g. in the structures of DntR (PDB ID: 2y7w, 2y7r) (Devesse *et al.*, 2011). Interestingly, the largest deviations of the three structures in the RD-2 domain occur at the "arginine elbow" at residue 218, of which a mutation to glutamic acid, cysteine or alanine results in hyperactivation.

Intriguingly, the amino acid whose mutation causes hyperactivity, are surface exposed and all map in the putative dimerization interface (Figure 7A). The residues Ser120, Arg218, and Met244 are located in the vicinity of the putative effector-binding cleft, which is located in between the two α/β subdomains. Met244 is close to its counterpart of the second monomer (Figure 7C). Arg218 is reasonably well defined in the wild-type structure and points towards the interior of the putative effector-binding pocket. The mutation of residue Ser120 to Asp induces a reorientation of the side-chain of Arg218 (Figure 7D) with much weaker electron density. Further, in the wild-type protein structure there appears to be a chloride ion bound close to the position in which the carboxylate group of Asp120 is located in the S120D mutant (Figure 7C). The assignment of the electron density peak as chloride is based on the coordination of the ion, the peak height and a weak anomalous signal. The putative chloride ion is coordinated by the backbone amide of Met243, the side-chain hydroxyl group of Ser120, and two water molecules. The vicinity is otherwise rather hydrophobic, with the side-chains of Pro121, Ala242, Met243 and Val214 lining the pocket (Figure 7E). The guanidinium group of Arg218 is about 4.5 Å apart from the putative chloride ion (Figure 7C).

Taken together, crystallization in cooperation with the Lab of Ulrich Baumann and the analysis of the crystal structure by Prof. Dr. Ulrich Baumann suggests the typical fold of the EBD with the two α/β subdomains. The structure is quite similar to LysR-type regulator DntR (Lerche *et al.*, 2016) and the structural change between the EBDs of LeuO wild-type and the hyperactive S120D mutant is similar to structural changes described for other LysR-type regulators. For LysR-type regulators DntR, AphB and BenM, the structural changes are induced by an effector binding to the EBD. In the crystal structures of the LeuO-EBD and the hyperactive S120D the side-chain of Arg218 is reoriented at the edge of the presumptive effector-binding cleft, hinting to an effector regulating the LeuO activity. This putative LeuO effector is also supported by the putative chloride ion binding to the rather hydrophobic effector binding pocket. The change in the structure, caused by an effector or a hyperactive mutant, may transmit to the DNA-binding domains and alter for example the angle of the DNA-binding.

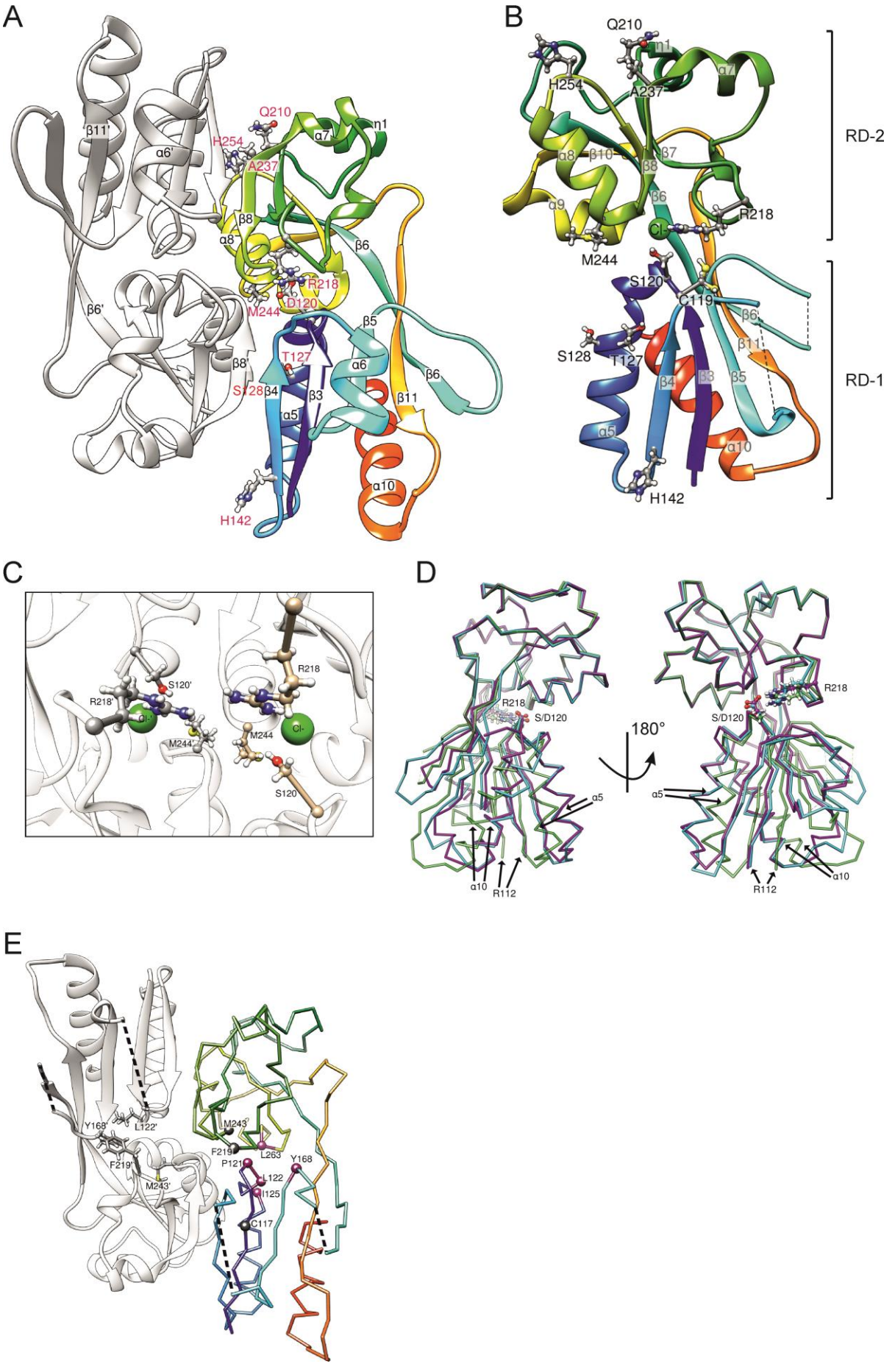

Figure legend next page

Figure 7: Crystal structures of LeuO-EBD and LeuO-S120D-EBD monomer and homodimers by Prof. Dr. Ulrich Baumann.

(A) Cartoon representation of LeuO-S120D-EBD dimer with residues causing hyperactivity when mutated. (B) LeuO-EBD monomer shown in gradient coloring (blue, N-terminus; red C-terminus), with secondary structure elements labeled. A bound chloride ion is shown in green and activating residues are indicated as sticks. (C) Close-up view of dimer LeuO-EBD with bound chloride ion and residues Ser120, Met244 and Arg218 shown as sticks. (D) Superposition of LeuO-EBD in green and LeuO-S120D-EBD crystallized in different space groups in cyan (I222) and purple (C2) in ribbon representation. The Cα atoms of residue Arg112 have a distance of about 4 Å between wild-type and mutant. (E) Dimer of the LeuO-EBD. The monomeric EBD is shown in gray ribbon (left), with amino acids Pro121, Leu122, Tyr168, Phe219, Met243 (depicted as sticks) lining the pocket that holds the putative chloride ion. The second monomeric EBD is shown in tubes in green, blue, yellow, and orange. Residues that render LeuO inactive are indicated as magenta spheres (Pro121, Leu122, Ile125, Tyr168, and Leu263). Residues that are neutral when mutated compared to wild-type LeuO are depicted in gray spheres (Met243, Phe219, and Cys117). Numbering of residues is according to the full-length protein sequence of LeuO. The Figure was prepared using Chimera (Pettersen *et al.*, 2004).

2.3 Characterization of additional LeuO mutants

Residues Ser120, Arg218 and Met244 map close to the putative effector binding pocket, and are involved in intramolecular and intermolecular contacts. To further characterize these residues, they were mutated to alanine, and Arg218 was in addition mutated to glutamic acid. Further, a LeuO-S120D/M244T double mutant was constructed. The LeuO mutants were constructed in plasmid pKESK22 as before, carrying leuO under control of the IPTG inducible *tac* promoter (*Ptac*), a p15A replication origin and a kanamycin resistance. The activity of these LeuO mutants was analyzed using the *Pcas-lacZ* reporter (strain T1281). LeuO mutants S120A and M244A showed a reduced activation of *Pcas* compared to wild-type LeuO at basal expression levels (no IPTG, Figure 8A, top panel). Upon induction of their expression (1 mM IPTG) the mutants fully activated *Pcas* (Figure 8A, bottom panel). These results show that the hyperactivity is specific for the Ser120 to Asp and Met244 to Thr amino acid exchanges, but not caused by the exchange to alanine (S120A and M244A, Figure 8A). The double mutant LeuO-S120D/M244T remained hyperactive, causing full activation of the *Pcas* (Figure 8A). Interestingly, mutation of Arg218 to alanine or glutamic acid rendered the protein similarly hyperactive as the Arg218 to cysteine mutation, which was isolated in the screen and caused full activation of *Pcas* at basal expression levels (Figure 8A).

LeuO harbors a cysteine residue at position 119, located in the cleft of the EBD. A second cysteine at position 117 is located in the dimer interface of the LeuO-EBDs (Figure 7B and 7E). To test the relevance of Cys117 and Cys119 for LeuO activity, these residues were mutated to a serine and aspartic acid, respectively. The mutant C119D showed increased activation of *Pcas* at basal expression levels compared to the wild-type (increase from 51 units to 148 units; Figure 8A), and caused full activation of *Pcas* at high expression levels. In contrast the other three mutants C117S,

C117D and C119S had a slightly reduced activity even when expressed at high levels (Figure 8A). Thus, introduction of the negatively charged aspartic acid at residue 119 enhances LeuO activity, which may have a similar effect as the mutation of residue 120 from serine to the negatively charged aspartic acid in the hyperactive mutants S120D.

In the solved LeuO-EBD structure several loops showed a weak electron density, as e.g. between β4 and β5 (residues Q151 to R158) and between β5 and β6 (residues R173 to E175; Figure 7B and 8C). To analyze these unstructured regions, amino acids in these loops (Asn150, Gln151, Asn152, His155, Gln156, Arg158, His172, Arg173) and additional positively charged residues His204, Trp226 and Lys232, were mutated to alanine and to glutamic acid. The mutants Asn150, Gln151, Asn152, His155, Gln156, Arg158, His172, Arg173, His204 showed full activation of *Pcas* when their expression was induced with IPTG, irrespective of being mutated to alanine or glutamic acid (Figure 8A). At basal expression levels the LeuO mutants N150E, Q151E, Q156E and H204E showed a slightly increased activation (2 to 3-fold) of the reporter compared to wild-type LeuO. The LeuO mutants N150A, Q151A, N152A, N152E, H155A, H155E, Q156A, R158A and R158E showed a similar activation at basal expression levels as the wild-type (Figure 8A). The activation of *Pcas* was reduced for LeuO mutants W226A and W226E with induction and abolished without induction possibly due to protein instability or structural changes. The LeuO mutants K232A and K232E had a slightly reduced activity even when expressed at high levels and showed no activation at basal expression levels (Figure 8A).

In the screen for constitutively active LeuO mutants a triple mutant LeuO-S120D/E111D/D205N and a double mutant LeuO-V230I/M244T were isolated. In addition, several double mutants of LeuO (S120D/E111D, S120D/C119S, M244A/M34V, M244A/G240V) were isolated in the process of cloning. The activation of the *Pcas-lacZ* by these LeuO mutants was analyzed as well. All LeuO mutant variants of S120D showed full activation of *Pcas-lacZ* with and without induction and irrespectively of the second or third additional mutation (LeuO-S120D/E111D/D205N, LeuO-S120D/E111D, LeuO-S120D/C119S; Figure 8A). LeuO mutant M244A/M34V showed a similar reduced activation of *Pcas* as the single mutant of LeuO-M244A, while the activation by LeuO-M244A/G240V was abolished. The additional mutant LeuO-M244T/V230I showed full activation of *Pcas-lacZ* just as the LeuO-M244T single mutant. Taken together, the additional secondary LeuO mutations were rather negligible, as the activation does not differ significantly from their corresponding single mutant (Figure 8A).

In order to further characterize residues in the cleft of the EBD, several additional residues were mutated (Figure 6 and 7E). Pro121 was mutated to the negatively charged aspartic acid and Leu122, Ile125, Tyr168, Phe219, Met243, and Leu263 to negatively charged glutamic acid. The activity of these mutants was analyzed using a low copy number (pSC origin) expression plasmid carrying *leuO* under control of the moderate lacUV5 promoter (Figure 8B). The LeuO mutant L122E was

additionally analyzed expressing the mutant from the Ptac, p15A-ori kanR plasmid. Upon induction of *leuO* expression mutants F219E and M243E showed a similar activation of the *Pcas-lacZ* reporter as the wild-type LeuO (Figure 8B, plus IPTG). In contrast, activation of *Pcas* was strongly decreased for LeuO mutants P121D, L122E, I125E, Y168E and L263E (Figure 8B, plus IPTG). Expressing the LeuO mutant L122E from a medium copy number plasmid (p15A-ori) does not activate the P*cas-lacZ* reporter without IPTG and showed reduced activation upon induction, compared to the wild-type (Figure 8A). Thus, these mutants are inactive possibly due to a structural change or protein instability.

Taken together, the analysis of additional LeuO mutants revealed that hyperactivity is specific for the particular amino acid exchanges S120D and M244T, while the activities of mutants C119D, N150E, Q151E, Q156E and H204E are moderately increased. Arg218 may be inhibitory since mutation of this residue to alanine, cysteine, or glutamic acid, causes hyperactivity.

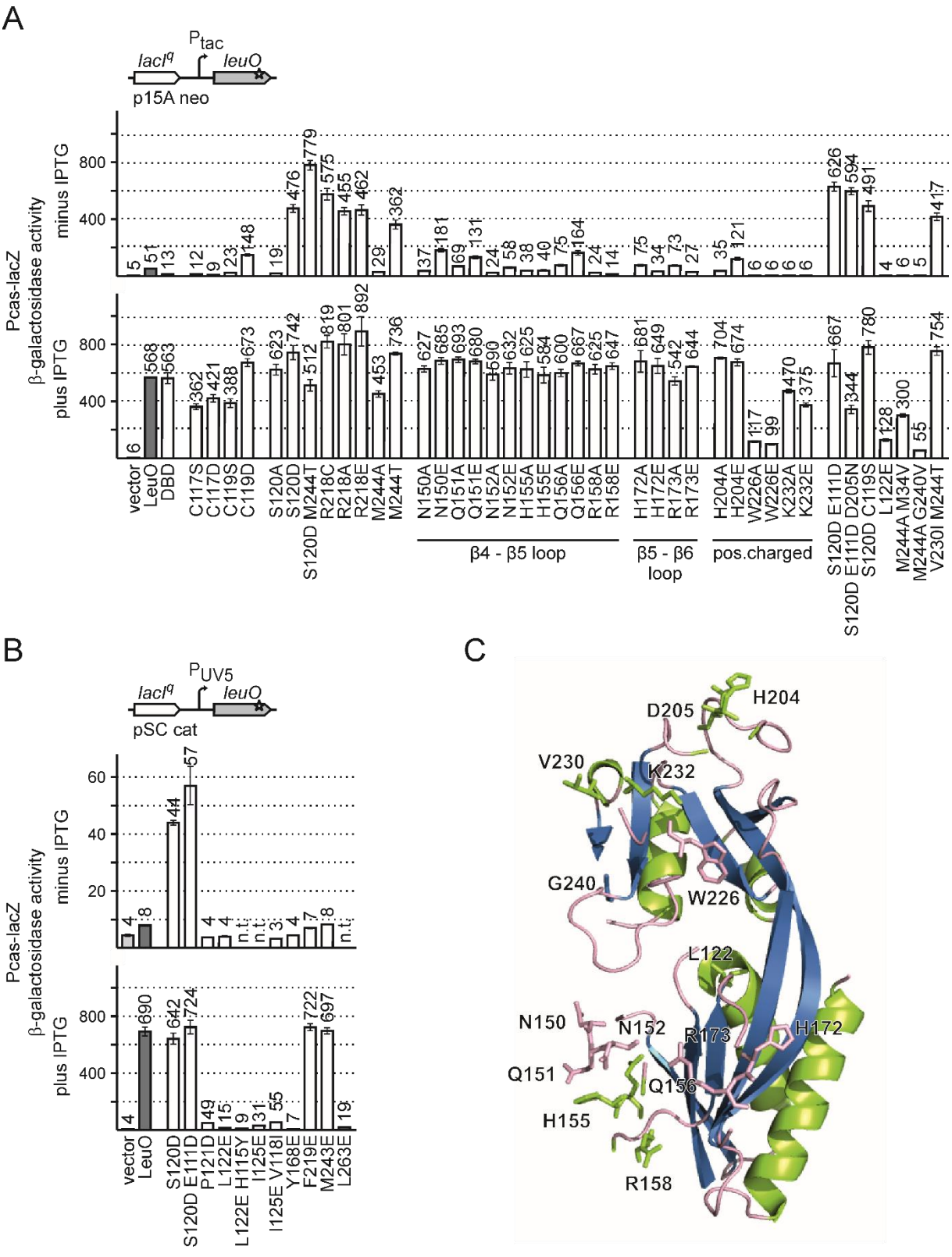

Figure 8: Characterization of additional LeuO mutants.

Activation of the *Pcas-lacZ* by LeuO mutants was determined in reporter strain T1281. Expression levels were determined of cultures grown without induction (top panel) or with 1 mM IPTG for induction of *leuO* expression (bottom panel). (A) *Pcas-lacZ* expression was analyzed of transformants carrying *Ptac* plasmid for directing expression of wild-type LeuO (pKETS5), LeuO-DBD (pKESMS63), as well as mutants S120D (pKESL104), S120A (pKESL108), R218C (pKESL78), R218A (pKESL238), R218E (pKESL243), M244T (pKESL73), M244A (pKESMS1), S120D/M244T (pKESL107). C117S (pKESMS64), C117D (pKESMS67), C119S (pKESMS65),

Figure 8 continued from the previous page

C119D (pKESMS66), L122E (pKESL106), N150A (pKESL221), N150E (pKESL231), Q151A (pKESL229), Q151E (pKESL232), N152A (pKESL222), N152E (pKESL233), H155A (pKESL223), H155E (pKESL234), Q156A (pKESL224), Q156E (pKESL241), R158A (pKESL228), R158E (pKESL235), H172A (pKESL225), H172E (pKESL236), R173A (pKESL226), R173E (pKESL242), H204A (pKESL230), H204E (pKESL237), W226A (pKESL239), W226E (pKESL240), K232A (pKESL227), K232E (pKESL244), S120D/E111D (pKESL71), S120D/E111D/D205N (pKESL72), S120D/C119D (pKESL105), M244A/M34V(pKESL110), M244A/G240V(pKESL109), V230I/M244T (pKESL79), empty vector pKESK22 served as a control (vector). (B) *Pcas-lacZ* expression was analyzed in transformants carrying low copy plasmids (pSC-ori) with the *lacUV5* promoter (P_{UV5}) directing expression of wild-type LeuO (pKESL39), and mutants S120D (pKESL102), S120D/E111D (pKESL40) P121D (pKESL69), L122E (pKESL103), L122E/H115Y (pKESL68), I125E (pKESL64), I125E/V118I (pKESL65), Y168E (pKESL70), F219E (pKESL67), M243E (pKESL61), and L263E (pKESL62). The empty parent vector of these plasmids served as control (vector). Cultures for β-galactosidase assays were grown to in LB medium with appropriate antibiotics to an OD_{600} of 0.5. Mean values of at least 3 independent biological replicates are shown as bars; error bars indicate standard deviations. (C) Mutated amino acid residues of unstructured loops of the LeuO protein. The Figure was created using PyMOL (The PyMOL Molecular Graphics System).

2.4 The LeuO DNA-binding domain (DBD) is functional

The LeuO protein in *E. coli* is presumably a tetramer, like *S. enterica* LeuO and other LysR-type regulators (Guadarrama *et al.*, 2014a) (predicted structure model from https://swissmodel.expasy.org/repository/uniprot/P10151 shown in Figure 2A and 2B). Tetramers of LysR-type regulators contain two pairs of dimeric effector binding domains (EBD) and two pairs of dimeric DNA-binding domains (DBD; Figure 2A and 2B). Dimerization of the N-terminal DBD is mediated by the α4 helical linker (Figure 2), while the EBDs provide an additional dimerization interface (Figure 7A and Figure 2). In Figure 2A the four monomeric LeuO proteins are highlighted in blue, grey, green and magenta. The EBD-1 (blue) dimerizes with EBD-2 (grey) and the EBD-3 (green) dimerizes with EBD-4 (magenta). The tetrameric complex is completed by the dimerization of DBD-1 with DBD-4 and the dimerization of DBD-2 with DBD-3 (shown in Figure 2B and 2C). The α3 helices (marked in orange) are the DNA recognition helices, forming the DNA-protein interface by insertion into the major groove of the DNA (Aravind *et al.*, 2005) (Figure 2C).

The two DBD dimers within one tetramer presumably bind to two adjacent sites on the DNA. Therefore a single dimeric DBD may be capable of specifically binding to the DNA as well. To address this, I tested whether a LeuO DBD-dimer is sufficient to activate the *Pcas-lacZ* reporter. The DBD (aa1 to aa101) was expressed using the *Ptac* expression plasmid pKESMS63, and the capability of the DBD to activate the *Pcas-lacZ* reporter was analyzed. Upon induction of the DBD expression, full activation of *Pcas-lacZ* was observed (Figure 8A). At basal expression levels of the DBD (no IPTG), *Pcas* activation was 4-fold lower by the DBD than by wild-type LeuO, but still detectable (Figure 8A, compare 13 and 51 units). These data show that the DBD alone is sufficient to activate *Pcas*, suggesting specific DNA binding by the presumptive LeuO DBD dimer.

2.5 LeuO binds to the promoter regions of *cas, yjjQ* and *leuO*

LeuO is an activator of *Pcas* and *PyjjQ* and represses its own promoter (*PleuO*). DNase I footprinting showed two distinct LeuO DNA-binding sites for the promoter regions of *cas* and *leuO* and three putative LeuO DNA-binding sites for *PyjjQ* (Westra *et al.*, 2010, Stratmann *et al.*, 2012) (and unpublished lab data). The LeuO DNA-binding sites are indicated in the schematic view of the promoter regions in Figure 9. To further characterize the kinetics of the LeuO binding to the individual DNA-binding sites, I performed electrophoretic mobility shift assays (EMSA). 80 bp DNA fragments were amplified by PCR, covering the different possible LeuO DNA-binding sites of the promoter regions of *Pcas* (I and II), *PleuO* (I and II) and *PyjjQ* (I, II and III; see schemes in Figure 9). C-terminally His-tagged LeuO was purified using the ÄKTA fast protein liquid chromatography (FPLC) system. The DNA fragments (20 nM) were incubated with increasing concentrations of LeuO and separated by gel electrophoresis on 8% native polyacrylamide gels.

A distinct band is visible for the DNA fragments when no LeuO protein was added. For the fragment of LeuO DNA binding site II of the *cas* locus the intensity of the band decreases at a LeuO concentration of 100 nM and a shift is visible. The band intensity of the shift increases at 250 nM and is completely shifted at 500 nM and 1000 nM LeuO (Figure 9A). The DNA fragment of DNA-binding site I of the *cas* locus shows a decrease of the band intensity at 250 nM, and the shift is visible at protein concentrations 500 nM and at 1000 nM (Figure 9A). The band intensity of both DNA fragments of the DNA-binding sites I and II at the leuO locus decreases at LeuO concentration of 250 nM and is completely shifted at 500 nM and 1000 nM LeuO (Figure 9B). The promoter region of *yjjQ* contains 3 LeuO DNA-binding sites characterized by DNase I footprinting (unpublished lab data; Figure 9C). The independent analysis of the single DNA-binding sites shows an unspecific shift for DNA-binding site I, where the band intensity starts to decrease at high protein concentrations of 500 nM and 1000 nM LeuO. Similarly, for DNA-binding site II the band intensity starts to decrease at 250 nM and 500 nM and is shifted at 1000 nM LeuO. The intensity of the band of DNA-binding site III is highly decreased at 250 nM and shifted at 500 nM and 1000 nM LeuO (Figure 9C). For DNA-binding site I and II at the *yjjQ* locus the band is shifted at rather high protein concentrations indicating a low unspecific binding to these sites (Figure 9C).

Taken together, LeuO binds to two distinct DNA-binding sites at the *cas* and *leuO* locus with high specificity, as shown before (Westra *et al.*, 2010, Stratmann *et al.*, 2012). LeuO binding at the *yjjQ* locus was shown only for binding site III at high protein concentrations. However, a more specific characterization of the DNA-binding sites seems not possible, since the EMSA is not sensitive enough to characterize differences in the DNA-binding sites.

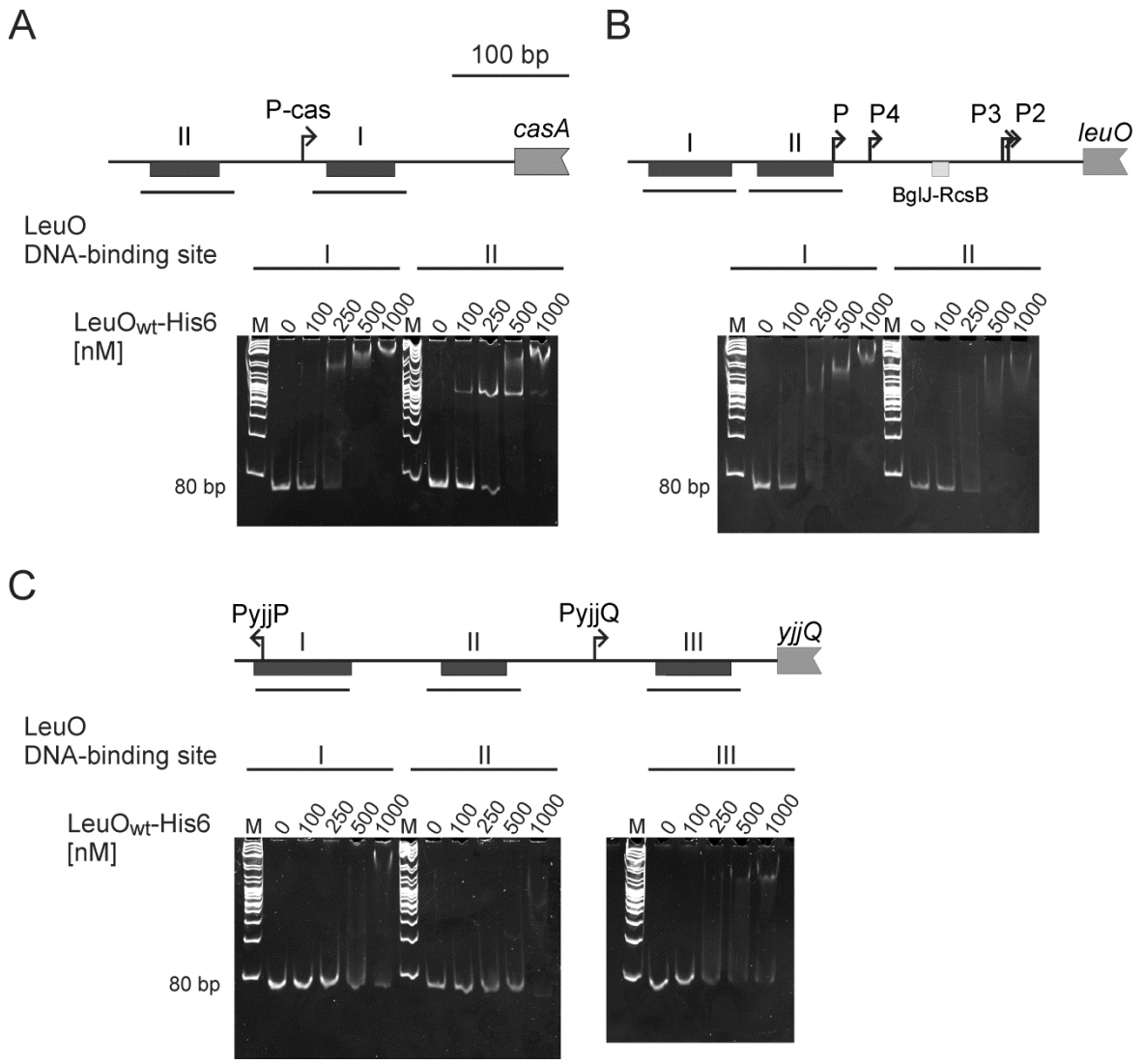

Figure 9: LeuO binding to DNA-binding sites at the *cas, leuO* and *yjjQ* loci.

Binding of LeuO to DNA-binding sites at *cas*, *leuO* and *yjjQ* loci was tested by electrophoretic mobility shift assays (EMSA). A schematic view of the promoter regions of *Pcas* (A), *PleuO* (B) and *PyjjQ* (C) is depicted on top of each panel, with the LeuO DNA-binding sites indicated (I, II and III) (Westra *et al.*, 2010, Stratmann *et al.*, 2012) (and unpublished lab data). The gel pictures of the EMSAs are depicted below each scheme with the DNA-binding sites indicated. The 80 bp DNA fragments were amplified by PCR and 20 nM were incubated without protein or with increasing concentrations of LeuO as indicated (0, 100, 250, 500 and 1000 nM). The samples were separated on 8% native polyacrylamide gels and stained with ethidium bromide.

2.6 DNase I footprinting by LeuO, LeuO-S120D and the LeuO-DBD

Two extended AT-rich DNA-binding sites for LeuO at the *cas* promoter region (LeuO sites I and II) have been identified by DNase I footprinting (Westra *et al.*, 2010). In these footprints a rather high LeuO concentration was used, which corresponds to high levels of wild-type LeuO that is required for activation of *Pcas in vivo*. In contrast, only low levels of the hyperactive LeuO-S120D are required for *Pcas* activation. Therefore, DNA-binding by S120D may be more specific than by wild-type LeuO. To test this, DNA-binding of LeuO and its S120D mutant was analyzed by DNase I footprinting using a

broad range of protein concentrations. In addition, to narrow down on the DNA-binding sites, LeuO-DBD was included in the footprinting analysis. The rational of this was that the full-length protein carrying two DBD dimers presumably contacts DNA at two sites and therefore yields an extended footprint. The DBD presumably occupies only half of the full-length LeuO DNA-binding site, which may allow defining a consensus sequence.

For footprinting His-tagged LeuO, LeuO-S120D and LeuO-DBD were purified. DNA fragments covering the full-length LeuO DNA-binding sites I and II of the *cas* promoter region, were [^{32}P]-labeled at the top and bottom strand, respectively, and used for footprinting analysis (Figure 10). Footprints of DNA fragments "a" and "b" covering LeuO site II are shown in Figure 11A and 11B, while the footprints of LeuO site I (fragments "c" and "d") are shown in Figure 11C and 11D. Corresponding to the previously described LeuO site II, DNase I protection by all 3 proteins was detected (Figure 11A to D). A hypersensitive site maps in the middle of this footprint (black triangle) and therefore two half-sites IIa and IIb are labeled in the footprint (IIa & IIb, Figure 11 A and 11B). Strikingly, DNase I protection by LeuO mutant S120D occurred already at very low protein concentrations (31 nM), suggesting that amino acid exchange S120D causes a higher DNA-binding affinity. This higher DNA-binding affinity is reflected by additional protection sites only detected for S120D (Figure 11A, 11C, 11D and 12A). Intriguingly, the DBD caused a DNase I protection pattern at low concentrations (31 nM) as well, but only at DNA-binding site IIa, while at higher protein concentrations (125 nM) both DNA-binding sites were protected (Figure 11B). For wild-type LeuO a footprint was apparent as of a concentration of 125 nM at both DNA-binding sites, but this footprint is more diffuse, at least at DNA-binding site IIb (Figure 11A and 11B). The DNase I footprints by LeuO-S120D and in particular by LeuO-DBD indicate that DNA-binding site IIa (in LeuO DNA-binding site II) is a high-affinity DNA-binding site, which is called "core site" in the following. This "core site" is palindromic (indicated by inverted arrows and bold letters, Figure 11A and 11B). Palindromic sequences are typical for DNA-binding of transcription regulators with dimeric DBDs (Browning & Busby, 2016). Thus the palindromic sequence of DNA-binding site IIa may represent one of the two DNA contact sites of the tetrameric LeuO. For LeuO site I (footprints of fragments "c" and "d") all three proteins protected the DNA from DNase I cleavage, confirming previous data (Westra *et al.*, 2010) (Fig footprint C and D). However, no palindromic core site is apparent in LeuO DNA-binding site I (Figure 12A). Furthermore, the DNase I footprint of LeuO-S120D showed an additional hypersensitive site, indicating the binding of at least two tetramers that might form a higher order complex (open arrowhead, Figure 11A).

Taken together, these results confirm the LeuO DNA-binding sites at the *cas* promoter (Westra *et al.*, 2010). In addition, hyperactive mutant S120D and the DBD bind with higher affinity to LeuO site II than wild-type LeuO, Further, the footprint by S120D and DBD is more distinct, and allowed the

identification of the core DNA-binding site in DNA-binding site IIa. The palindromic sequence of the core site suggests that this represents a close to ideal DNA-binding site of LeuO.

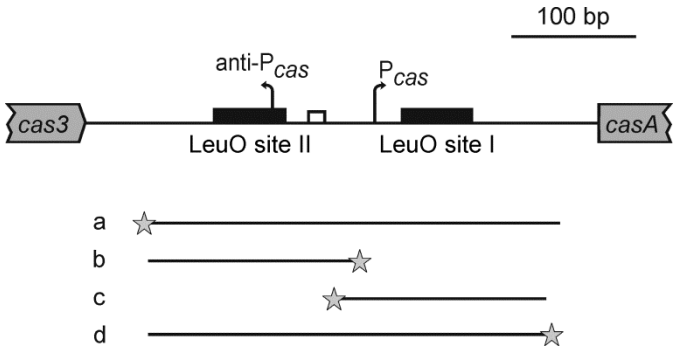

Figure 10: Schematic overview of the *cas* promoter region and DNA fragments for DNase I footprinting with $LeuO_{His6}$, $LeuO\text{-}DBD_{His6}$, and $LeuO\text{-}S120D_{His6}$.

Indicated are LeuO DNA-binding sites I and II (black boxes) (Westra *et al.*, 2012), the H-NS nucleation site (open box), and the transcription start sites (bent arrows) (Pul *et al.*, 2010). Fragments ("a" to "d") used for DNase I footprinting cover either one or both LeuO DNA-binding sites, as indicated. The fragments were generated by PCR using primer pairs of which one primer was 5' [^{32}P]-labeled, indicated by an asterisk (fragment "a": oligonucleotides [^{32}P]-OA477/OA474, fragment "b": OA477/[^{32}P]-OA475, fragment "c": [^{32}P]-OA476/OA473, fragment "d": OA477/[^{32}P]-OA473). Fragments "a" and "c" were labeled at the top strand, and fragments "b" and "d" were labeled at the bottom strand. The [^{32}P]-labeled DNA fragments were incubated with increasing protein concentrations of $LeuO_{His6}$, $LeuO\text{-}DBD_{His6}$ and $LeuO\text{-}S120D_{His6}$, as indicated.

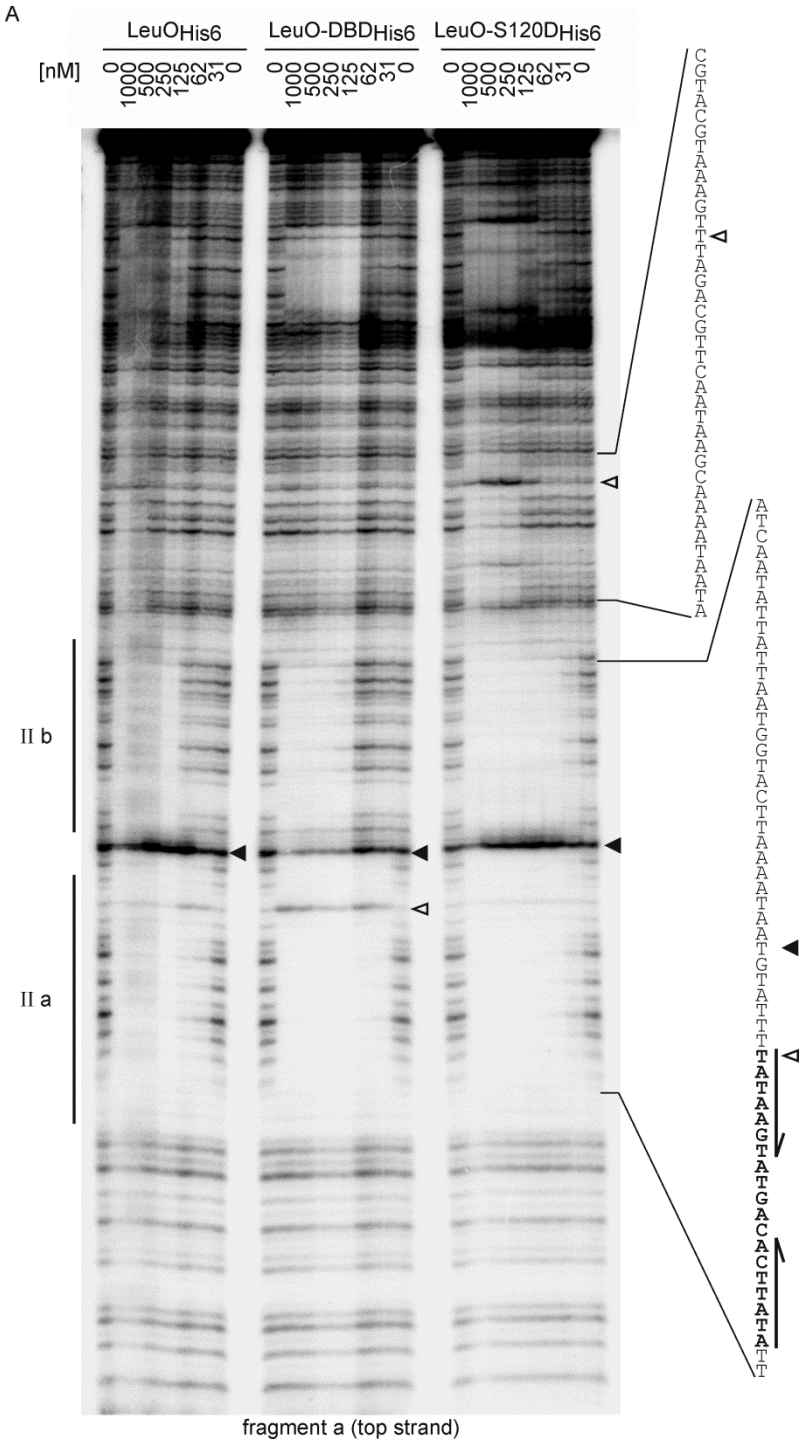

Figure 11 continued

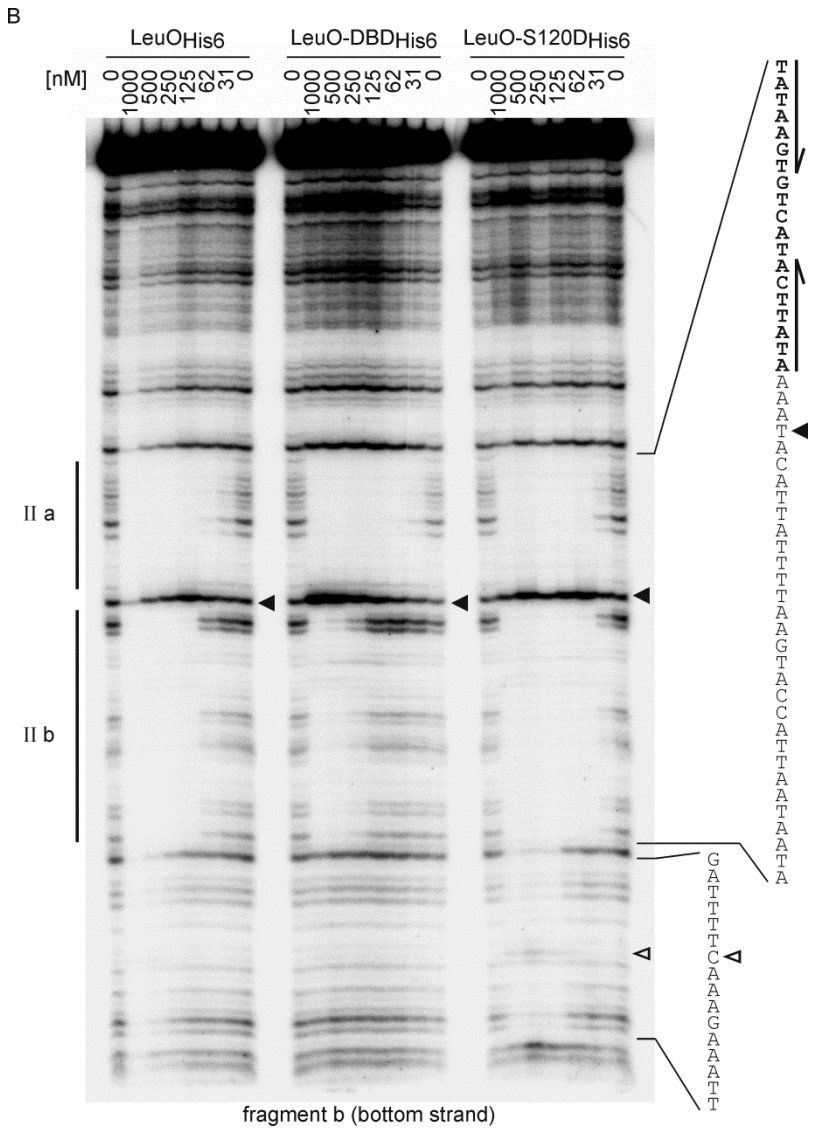

Figure 11 continued

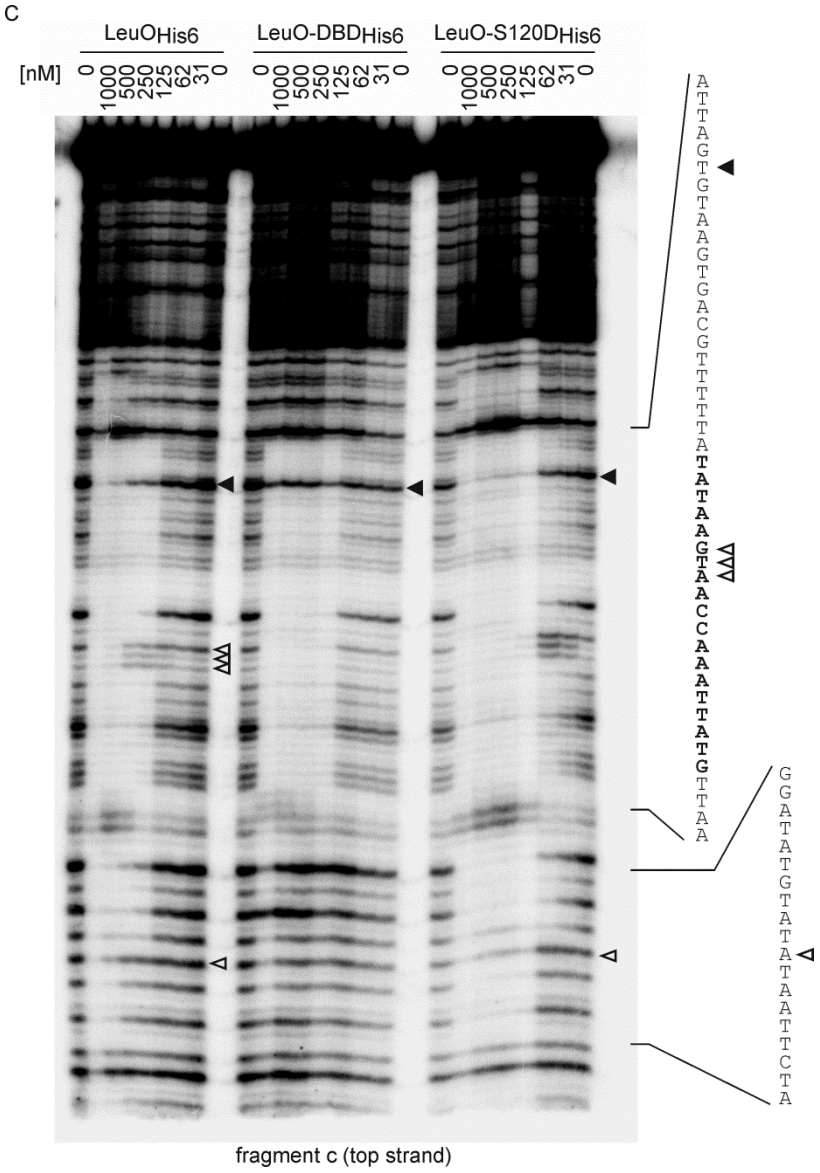

Figure 11 continued

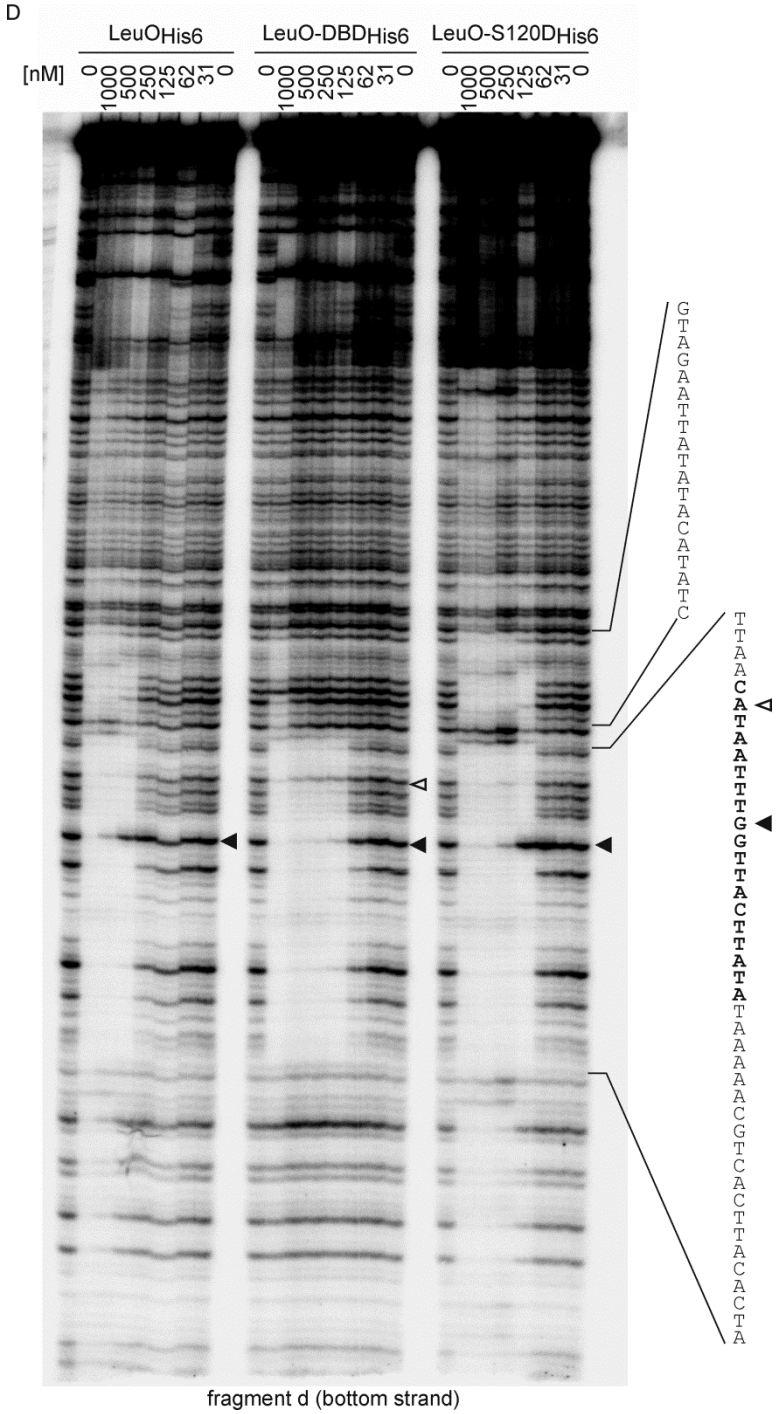

Figure 11: DNase I footprinting of the *cas* promoter region.

Shown are autoradiograms of DNase I footprints of fragment "a" (A) and fragment "b" (B) covering LeuO site II of the *cas* locus, and of fragment "c" (C) and "d" (D) covering LeuO site I. The [^{32}P]-labeled DNA fragments were incubated with increasing protein concentrations of $LeuO_{His6}$, $LeuO\text{-}DBD_{His6}$ and $LeuO\text{-}S120D_{His6}$, as indicated. Samples were separated on 6% denaturing polyacrylamide gels. Filled arrowheads indicate DNase I hypersensitive sites that were detected for all proteins, open arrowheads indicate hypersensitive sites detected for one or two protein variants. LeuO-protected sites are marked by two lines (IIa and IIb), representing the two half-sites of LeuO DNA-binding site II. The sequence of the protected DNA region is given to the right with the palindromic core sequence (at LeuO site II) indicated in bold and with inverted arrows.

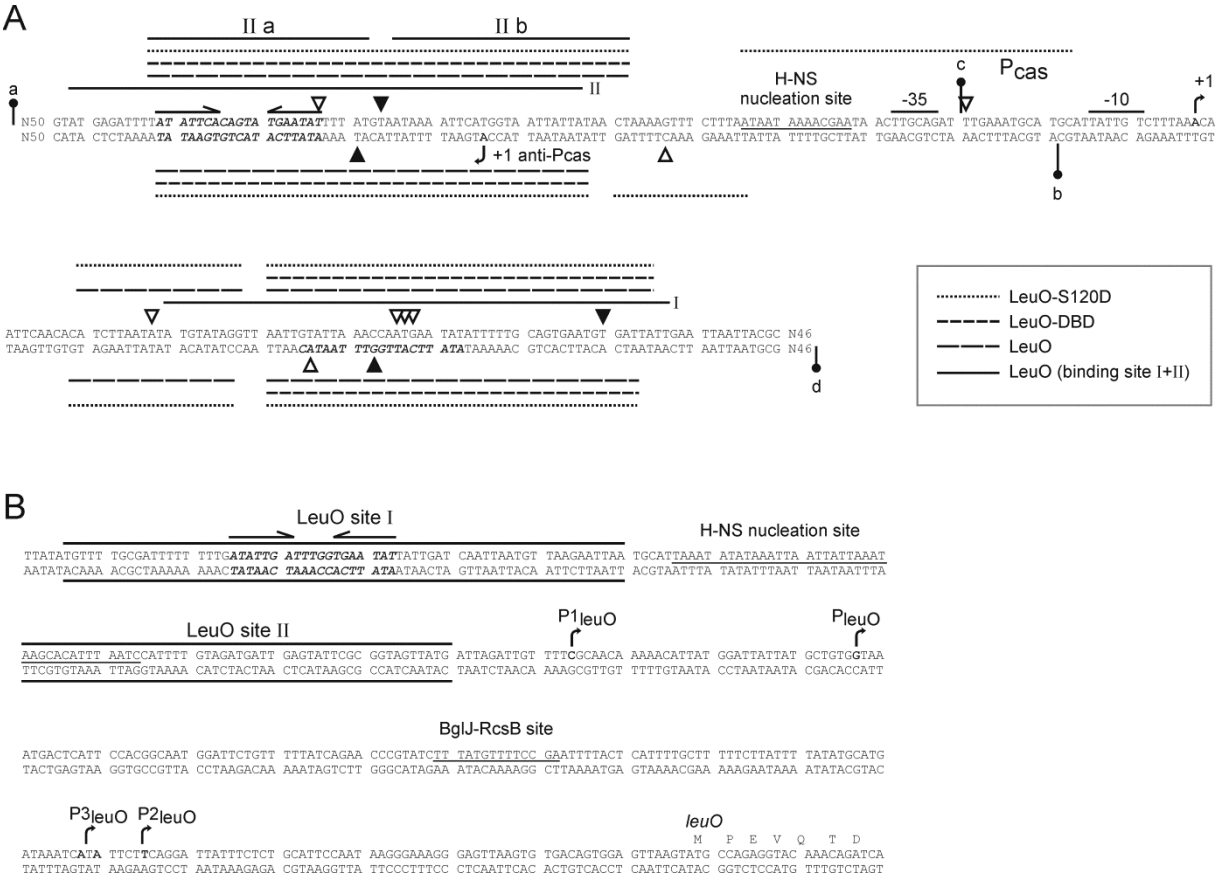

Figure 12: Compiled overview of the DNase I footprinting results at the *cas* locus.

(A) The sequence of the *cas* regulatory region with the promoters *Pcas* and *anti-Pcas* (bent arrows) and the H-NS nucleation site is shown (Pul *et al.*, 2010). The areas of protection observed in the DNase I footprints are indicated by lines above the sequence for protection of the top strand and underneath the sequence for protection of the bottom strand, respectively. Protected areas are marked: $LeuO_{His6}$ (solid line), $LeuO\text{-}DBD_{His6}$ (dashed line) and $LeuO\text{-}S120D_{His6}$ (dotted line). The 5' [^{32}P]-labeled ends of fragments "a" to "d" are indicated by lollipops. The presumptive palindromic DNA-binding core site within DNA-binding site II is shown in bold and italic, and marked by inverted arrows. (B) The sequence of the *leuO* regulatory region with promoters, the H-NS nucleation site, and BglJ-RcsB binding site indicated, as described before (Breddermann & Schnetz, 2016). LeuO DNA-binding sites I and II, which were characterized by DNase I footprinting (Stratmann *et al.*, 2012), are labeled with a black line. The presumptive palindromic DNA-binding core site within LeuO DNA-binding site I is shown in bold and italics, and marked by inverted arrows.

2.7 DNA-binding consensus sequence of LeuO

For LeuO, a 28 bp large and rather weak consensus sequence was described before (Dillon *et al.*, 2012). The DNase I footprint at the *cas* promoter region revealed the high affinity core DNA-binding site IIa harboring a 19 bp palindromic sequence (shown in bold letters and indicated by arrows in Figure 11A and 11B and Figure 12A). In the DNA-binding site I of LeuO at the autoregulated *leuO* promoter (Stratmann *et al.*, 2012), an almost identical palindromic sequence was detected within LeuO DNA-binding site I at the *leuO* locus (Figure 12B). The 33 bp sequences covering the palindromic sequences of LeuO DNA-binding site II of the *cas* locus and LeuO DNA-binding site I of the *leuO* locus

were submitted to MEME Suite (Bailey & Elkan, 1994) (Figure 18A and Figure 13, top panel motif). The resulting motif (Figure 13, top panel motif) was used to search the *E. coli* K12 MG1655 genome for putative LeuO DNA-binding sites, using FIMO (Grant *et al.*, 2011). The obtained top scores were filtered in respect to their intergenic position as well as their previous identification as presumptive LeuO targets in a microarray transcriptome analyses and in a genomic SELEX screening (Stratmann *et al.*, 2012, Shimada *et al.*, 2011) (Table 5). Then the sequences of the six top score loci (Figure 18B) were used to generate a less stringent motif using MEME Suite (Table 5; Figure 13, bottom panel motif). To validate this motif, it was again submitted to FIMO to search the *E. coli* K12 MG1655 genome for putative LeuO DNA-binding sites. By this approach we identified 37 sequences with top scores, which showed a striking correlation to LeuO targets identified in microarray and the genome scale SELEX (Stratmann *et al.*, 2012, Shimada *et al.*, 2011) (see Table 5). The LeuO DNA-binding motif (Figure 13, bottom panel motif) presumably represents the DNA-binding site of only one dimeric DBD.

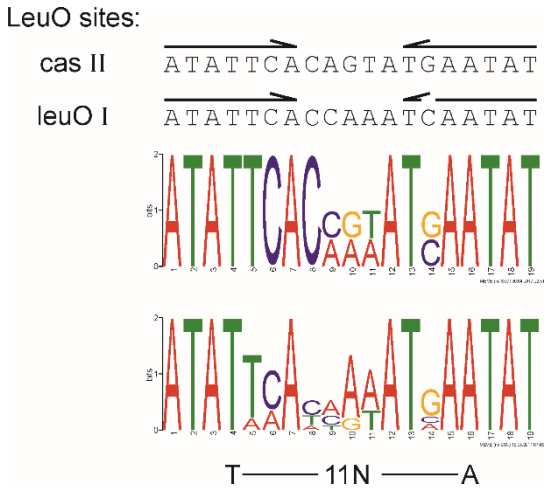

Figure 13: LeuO DNA-binding motif.

Sequences of the palindromic core DNA-binding sites of LeuO site II at *cas* and of LeuO site I at *leuO* (shown at the top) were used in an iterative approach to determine a DNA-binding consensus sequence. Arrows indicate the palindromic sequence. The top sequence logo represents the motif that is based on the LeuO DNA-binding sites at *cas* and *leuO* and was generated with MEME Suite (Bailey & Elkan, 1994). The bottom motif is based on 6 top score hits of the *E. coli* K12 genome identified by FIMO (see Table 5). The T-11N-A motif typical for DNA-binding sites of LysR-type regulators is indicated underneath the logo (Maddocks & Oyston, 2008).

2.8 Regulation of presumptive target loci by LeuO

LeuO regulates many loci in *E. coli* and *Salmonella* (Stratmann *et al.*, 2012, Shimada *et al.*, 2011, Dillon *et al.*, 2012). The palindromic DNA binding motif for LeuO shown in Figure 13 was further validated with previous data of LeuO DNA-binding and target regulation (Stratmann *et al.*, 2012, Shimada *et al.*, 2011) (Figure 13, Table 5). Many putative LeuO targets were identified that harbor a presumptive LeuO DNA-binding site and were shown to be LeuO-regulated in the DNA microarray

analyses and LeuO bound in genomic SELEX screening (Stratmann *et al.*, 2012, Shimada *et al.*, 2011) (Table 5). Nine loci were selected from these putative LeuO targets. These genes harboring a putative LeuO DNA-binding site encode proteins of predicted or unknown function. Nine additional loci were further selected from the DNA microarray that are regulated between 12-fold to 200-fold when LeuO is present and are shown to be LeuO bound targets in the genomic SELEX screening (Stratmann *et al.*, 2012, Shimada *et al.*, 2011). Some of these putative targets are interesting due to their function in *E. coli*, for others only little is known yet. Many of the putative LeuO regulated targets encode proteins for membrane structures, efflux systems or are involved in the bacterial stress response (Dillon *et al.*, 2012, Shimada *et al.*, 2011).

To analyze whether transcription of these presumptive targets is indeed regulated by LeuO, I constructed *lacZ* transcriptional fusions of the promoter regions of the putative target loci. The cloned fragments include the intergenic region of the target gene loci and the upstream gene, including the promoter and putative regulatory binding sites. The promoter *lacZ* fusions were integrated into the genome of a strain lacking LeuO (*ΔleuO*) and a strain expressing LeuO constitutively (*leuOc*, miniTn10 insertion upstream of *leuO*). This allowed determining the expression levels of the promoter *lacZ* fusions in absence and presence of LeuO via β-galactosidase assay. In addition, the strains carry a deletion of *yjjP-yjjQ-bglJ* and (*ΔPQ-J*) to avoid feedback regulation with BglJ-RcsB. The results of the expression analysis (*leuOc/ΔleuO*) are depicted in Table 1 and show no activation (<2-fold) of the promoters of *gspC, ybeQ, ybeR, yghT, uspG* and *sdiA*. LeuO slightly activates the promoters of *ompN, chiA, yafT, ybdO, yghS* and *yjfI* with a 2 to 3-fold increase, compared to strains deleted for LeuO. The small RNA encoding micC is moderately activated when LeuO is constitutively expressed. The divergent genes *envR* and *envC* are also moderately regulated by LeuO, showing 5-fold and 4-fold activation, respectively. Of all targets tested, downregulation by LeuO is shown only for the promoter *lacZ* fusion of *nmpC* (5578 units decrease to 1339 units). The promoters of *cas* and *yjjQ* are highly regulated by LeuO with 44-fold (compare 7 units and 306 units) and a 10-fold activation (compare 1.8 and 18 units), respectively (Table 1).

Taken together, LeuO regulates the *cas* promoter, and the bacterial defense system might be one of the main targets of LeuO. Most loci identified by searching the genome with the LeuO DNA-binding motif are directly regulated by LeuO, although the regulation by LeuO is rather weak. The putative LeuO targets *gspC, ybeQ, ybeR, yghT, uspG* and *sdiA* show no direct regulation by LeuO. Nevertheless, the targets might be regulated indirectly by LeuO or the regulation of some of the targets might depend on an effector that increases the effect of LeuO. Several target genes, as *ompN*, micC, *cas*, *ybdO* and *envR* are repressed by H-NS and activated by LeuO, supporting LeuO as a specific antagonist of H-NS repression.

Table 1: Regulation of presumptive targets by LeuO

P-lacZ	*ΔleuO*	*leuOc*	fold reg.	micro-array	SELEX	function[2)]	references
ompN[1)]	42	82	**2.0**	73	41	outer membrane porin	(Fàbrega *et al.*, 2012)
micC[1)]	310	1695	**5.5**	199		small regulatory RNA	(Chen *et al.*, 2004)
cas[1)]	7	306	**43.7**	65	77	bacterial defense system	(Barrangou *et al.*, 2007)
gspC	150	167	**1.1**	37	70	inner membrane protein	(Lomize *et al.*, 2017) (Francetic *et al.*, 2000b)
chiA[1)]	5.1	15.9	**3.1**	27	25	endochitinase activity	(Francetic *et al.*, 2000a)
yjjQ[1)]	1.8	18	**10.0**	n.d.	48	transcriptional repressor	(Stratmann *et al.*, 2008, Wiebe *et al.*, 2015)
yafT	37	85	**2.3**	8	6	predicted lipoprotein	(Juncker *et al.*, 2003)
ybdO[1)]	141	285	**2.0**	23	48	putative LysR-type regulator	(Higashi *et al.*, 2016)
ybeQ	16	23	**1.4**	43	10	oxidative stress response	(Krisko *et al.*, 2014)
ybeR	170	286	**1.7**	20		predicted polypeptide	
yghS	2.1	4.6	**2.2**	28	55	putative protein	
yghT	23	43	**1.9**	12		putative protein	
envR[1)]	0.1	0.5	**5.0**	48	81	(*acrS*) transcriptional regulator	(Hirakawa *et al.*, 2008)
envC[1)]	0.3	1.2	**4.0**			(*acrE*) lipoprotein, drug efflux pump	(Kobayashi *et al.*, 2001)
yjfI	50	124	**2.5**	39	54	conserved protein	
nmpC	5578	1339	**4.2**	-12	8	general bacterial porin	(Zhai & Saier, 2002)
uspG	8.7	5.2	**1.7**	-52	7	universal stress protein	(Kvint *et al.*, 2003)
sdiA	46	44	**1.0**	-17	13	transcription factor (cell division)	(García-Lara *et al.*, 1996)

[1)] H-NS-repressed

[2)] Description of protein functions were taken from the EcoCyc Database (Karp *et al.*, 2014).

2.9 Autoregulation of P*leuO* by hyperactive LeuO mutants

Transcription of *leuO* is repressed by H-NS and StpA under standard laboratory conditions, while the transcription regulator BglJ-RcsB activates *leuO* (Stratmann *et al.*, 2012). LeuO positively and negatively autoregulates the *leuO* expression. In wild-type cells, LeuO acts as a weak positive autoregulator, when provided at high levels from a plasmid (Figure 14, compare 3 to 7 units), as described (Chen & Wu, 2005). However, LeuO also acts as a negative autoregulator and inhibits the activation of *leuO* by BglJ-RcsB and derepression of *leuO* in an *hns stpA* mutant (Figure 14, compare 308 units to 19 units), as described by (Stratmann *et al.*, 2012).

To analyze whether the mutations S120D and M244T, which are hyperactive in *Pcas* activation, also affect other loci, their activity in autoregulation was analyzed. The expression of a chromosomal *PleuO-lacZ* reporter was analyzed in three different strain backgrounds (Figure 14A). In addition,

autoregulation by the LeuO-DBD was tested. The LeuO mutants were expressed from a plasmid under the control of an IPTG inducible *tac* promoter. The strain backgrounds include (1) a *ΔleuO ΔbglJ* strain to avoid activation of *PleuO* by BglJ-RcsB. Additionally, negative autoregulation was tested in a BglJ-RcsB positive strain (2), in which *bglJ* is constitutively expressed by allele *bglJc* and in an *hns stpA* mutant derivative (3). In both of these strain backgrounds *PleuO* is active (Figure 14B). High expression levels of LeuO and its variants (S120D, M244T and DBD) inhibit activation of *PleuO* by BglJ-RcsB and cause repression in the *hns stpA* background (Figure 14B, bottom panel). Further positive autoregulation remained marginal (Figure 14B). However, at basal expression levels LeuO-S120D acts as a hyperactive autorepressor in the *hns stpA* mutant, while it seems to enhance activation by BglJ-RcsB (Figure 14B, top panel). In addition, positive autoregulation of *PleuO* by low levels of LeuO-S120D is enhanced (18 units compared to 4 units, Figure 14B, top panel). Basal expression levels of M244T and of the DBD had similar effects as wild-type LeuO on *PleuO-lacZ* in all three strain backgrounds. Taken together, LeuO-S120D is hyperactive in autoregulation, but not LeuO-M244T. Surprisingly, LeuO-DBD is sufficient for autorepression when expressed at high levels.

The regulation of *PleuO-lacZ* was also tested for the LeuO mutants that were shown hyperactive at *Pcas* (T127I, S128P, H142R, Q210R, R218C, A237V, H254R; Figure 5). In addition, LeuO mutants S120A, M244A and the double mutant S120D/M244T were included and tested as described before at high expression levels (plus IPTG). Under non-activating conditions (*ΔleuO ΔbglJ)* the additional LeuO mutants that were shown hyperactive at *Pcas* show slightly reduced activity compared to wild-type LeuO, except LeuO-A237V showing similar activation as the wild-type and the LeuO-S120A mutant showing a slightly increased activation of *PleuO* (Figure 14C, top panel). All tested LeuO mutants negatively autoregulated *PleuO* in the presence of BglJ-RcsB (Figure 14C, middle panel). The hyperactive LeuO mutants inhibited activation of *PleuO* by BglJ-RcsB except LeuO mutants S120A, T127I, A237V, M244A and H254R. The autorepression by these LeuO mutants is weaker than the wild-type LeuO.

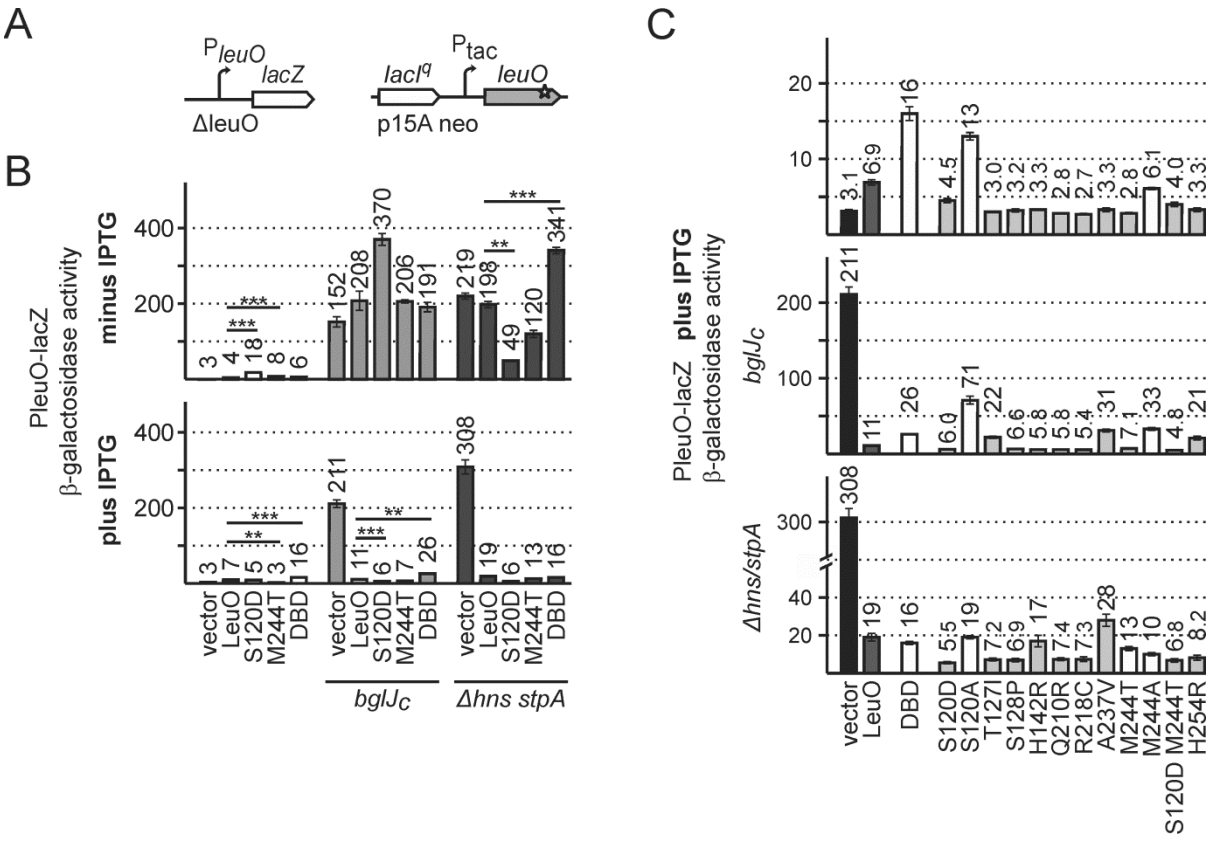

Figure 14: Autoregulation of *leuO* by LeuO and LeuO mutants.

(A) Autoregulation of *leuO* was analyzed using a chromosomal *PleuO-lacZ* reporter (Stratmann *et al.*, 2012) and LeuO and LeuO mutants were expressed from a plasmid carrying an IPTG inducible *tac* promoter (*Ptac*), a p15A replication origin and a kanamycin resistance. (B) The chromosomal *PleuO-lacZ* reporter was integrated in strains *ΔleuO Δ(yjjP-yjjQ-bglJ)* (T308), T862 *ΔleuO (bglJ$_C$)*, and T352 *ΔleuO Δ(yjjP-yjjQ-bglJ) Δhns stpA* (Stratmann *et al.*, 2012). Transformants of these strains with the vector control (pKESK22), and plasmids providing LeuO (pKETS5), or LeuO mutants S120D (pKESL104), M244T (pKESL73) and LeuO-DBD (pKESMS63), were grown in LB medium with kanamycin to an OD_{600} of 0.5 without induction (top panel) or with induction of *leuO* expression (plus 1 mM IPTG; bottom panel). (C)Analysis of *PleuO* autoregulation by additional LeuO mutants S120A (pKESL108), T127I (pKESL74), S128P (pKESL75), H142R (pKESL76), Q210R (pKESL77), R218C (pKESL78), A237V (pKESL80), M244A (pKESMS1), S120D/M244T (pKESL107), H254R (pKESL101) was tested at induced LeuO expression levels (plus IPTG). LeuO mutants showing hyperactivity at *Pcas* are indicated with light grey bars. Mean values of at least 3 independent biological replicates are shown as bars, and error bars indicate standard deviations. The statistical significance was calculated by Student's t-test in comparison to wild-type LeuO with **p<0.01 and ***p<0.001.

2.10 Is LeuO activity modulated by an effector?

Many LysR-type regulators require an effector for activating or repressing transcription (Maddocks & Oyston, 2008, van Keulen *et al.*, 1998, Picossi *et al.*, 2007). The described effectors so far, are all metabolites, ions or modifications of amino acid residues. For example, arginine and lysine are effectors of ArgP (Laishram & Gowrishankar, 2007) and Na^+ is an effector of NhaR, the activator of a sodium antiporter (Dover & Padan, 2001). However, no effector is known for LeuO, although results of this work support the hypothesis that LeuO activity might be regulated by an unknown effector.

The LeuO protein harbors the typical cleft of the effector-binding domain (Figure 7) and several mutations in this cleft rendered the protein inactive (Figure 8B). The hyperactive LeuO mutants that were isolated in this work are also located either in the dimer interface of the LeuO effector-binding domain (EBD) or at the entrance of the cleft of the EBD, indicating an important role of this cleft. The hyperactive LeuO mutant S120D induces a structural change and alters the activity of LeuO. These hints support the hypothesis that the LeuO activity may indeed be modulated by an unknown effector and that binding of the effector is required for the LeuO function.

Under amino acid starvation and in stationary growth phase expression of the *leuO* gene is upregulated (Majumder *et al.*, 2001, Fang *et al.*, 2000). A putative effector might be present under the very same conditions. For some LysR-type regulators it is known that their location in the genome is closely related to their regulated targets, like operons of metabolic pathways that are located divergent from their regulators (Maddocks & Oyston, 2008). The product (or an intermediate product) of the regulated synthesis pathway can then in turn act as an effector for the LysR-type regulator as a feedback mechanism of the transcriptional regulation (Maddocks & Oyston, 2008, Picossi *et al.*, 2007). The *leuO* gene is located between two operons encoding for enzymes of the branched-chain amino acid synthesis (*leu* and *ilvIH*). This might hint to an intermediate product of the branched-chain amino acid synthesis as an effector for LeuO. The scheme of the synthesis pathway of the branched-chain amino acids L-valine, L-leucine and L-isoleucine is shown in Figure 15. The precursor of L-valine is α-ketoisovalerate, the L-leucine precursor is α-ketoisocaproate and the precursor of L-isoleucine is α-keto-β-methylvalerate. The enzyme transaminase B, encoded in *ilvE*, catalyzes the transamination between the α-keto acids of the precursors and the native amino acid (Figure 15). The *ilvG* gene, which carries a frameshift in K12 wild-type strains was restored to *ilvG*$^{+}$.

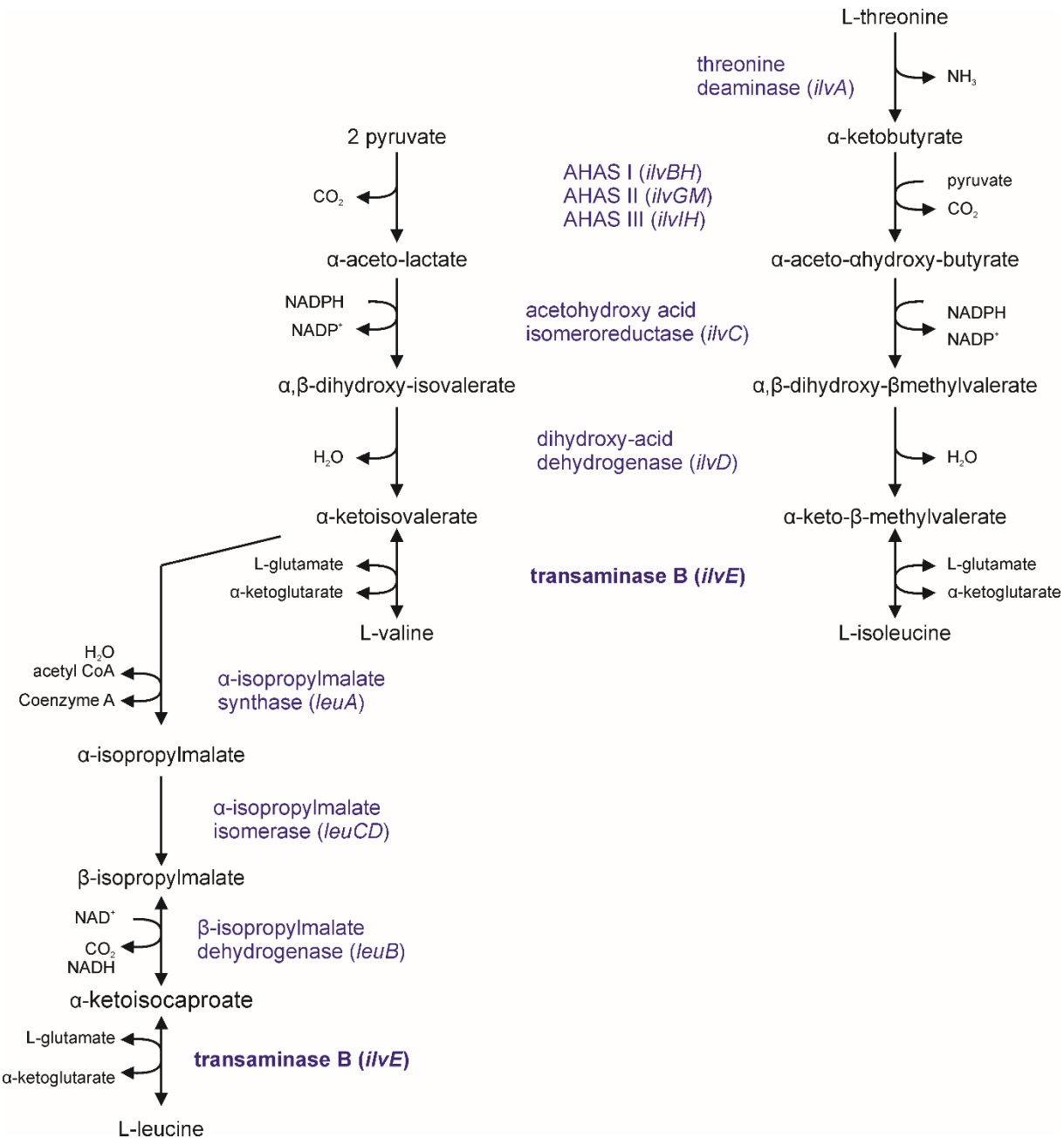

Figure 15: Metabolic pathways for the biosynthesis of the branched-chain amino acids L-isoleucine, L-valine and L-leucine.

The enzymatic reactions are indicated by arrows and the enzymes with their corresponding genes are shown in blue. Figure adapted and modified from (Salmon *et al.*, 2006).

The *Pcas-lacZ* fusion is a very sensitive reporter for monitoring changes of LeuO activity, when comparing a Δ*leuO* strain (T1281) and a strain expressing leuO constitutively (*leuOc*, T1295). In order to address whether the accumulation of precursors of the branched-chain amino acids in the cells affects activation of *Pcas* by LeuO, the *ilvE* gene was deleted in strains carrying the *Pcas* reporter and a *leuO* deletion (Δ*leuO*, T1927) or expressing *leuO* constitutively (*leuOc*, T1933; Figure 16). Cultures of these *ilvE*$^+$ and Δ*ilvE* strains were grown in M9 minimal glucose medium, supplemented with no

amino acids, or with all three branched-chain amino acids (L-valine, L-leucine and L-isoleucine) or with only one of these amino acids. The cultures were grown to OD_{600} = 0.5 and the β-galactosidase activity was measured (Figure 16B). In the Δ*leuO* strains, *Pcas-lacZ* directed the expression of 16 units β-galactosidase activation without and with amino acids supplemented. Only a slight decrease is detectable when the cells were grown in the presence of L-isoleucine (compare 12 units of +ile with 16 units when no amino acids are supplemented; Figure 16B, left panel). When LeuO is expressed constitutively, *Pcas* directed 121 units in media with no amino acids supplemented. The *Pcas* activity is slightly increased when only L-leucine or all three amino acids are supplemented (compare 198 units and 174 units; Figure 16B, right panel). When the cells were grown in the presence of L-isoleucine, the *Pcas* activation by LeuO decreases 2-fold from 121 units to 58 units (Figure 16). However, the additional deletion of *ilvE* has no effect on the *Pcas* activity in the Δ*leuO* and *leuOc* strain with and without LeuO.

To evaluate growth defects that are caused by the deletion of *ilvE*, the strains *leuOc* (T1295) and *ΔilvE leuOc* (T1933) were grown for 24 h and the OD_{600} was measured to generate a growth curve (Figure 16C). The cultures were again grown in M9 minimal glucose medium and either no, only one or all three branched-chain amino acids were supplemented. The growth of the cultures did not differ, except for the cells grown in the presence of L-isoleucine (+ile depicted in red, Figure 16C). Growth of the culture in isoleucine supplemented medium was delayed compared to cells grown in any of the other growth conditions (Figure 16C). Therefore, the 2-fold decrease in the activation of *Pcas-lacZ* may be caused by the growth-retardation. The deletion of *ilvE* causes no altered growth of the bacterial cells compared to the *leuOc* strain, when the branched-chain amino acids were supplemented. The *ΔilvE* strain needs to be supplemented with all three branched-chain amino acids.

Although several results in this work support the hypothesis that LeuO might be modulated by an unknown effector, the precursors of branched-chain amino acids seemed to have no effect. The accumulation of precursors of the branched-chain amino acids did not affect the activity of *Pcas* by LeuO, since the decrease of activation in the presence of isoleucine resulted from a growth-retardation. It is possible that the accumulation of the amino acid precursors is not sufficient to alter the activation of *Pcas-lacZ* or the β-galactosidase assay is not sensitive enough to detect changes in activation.

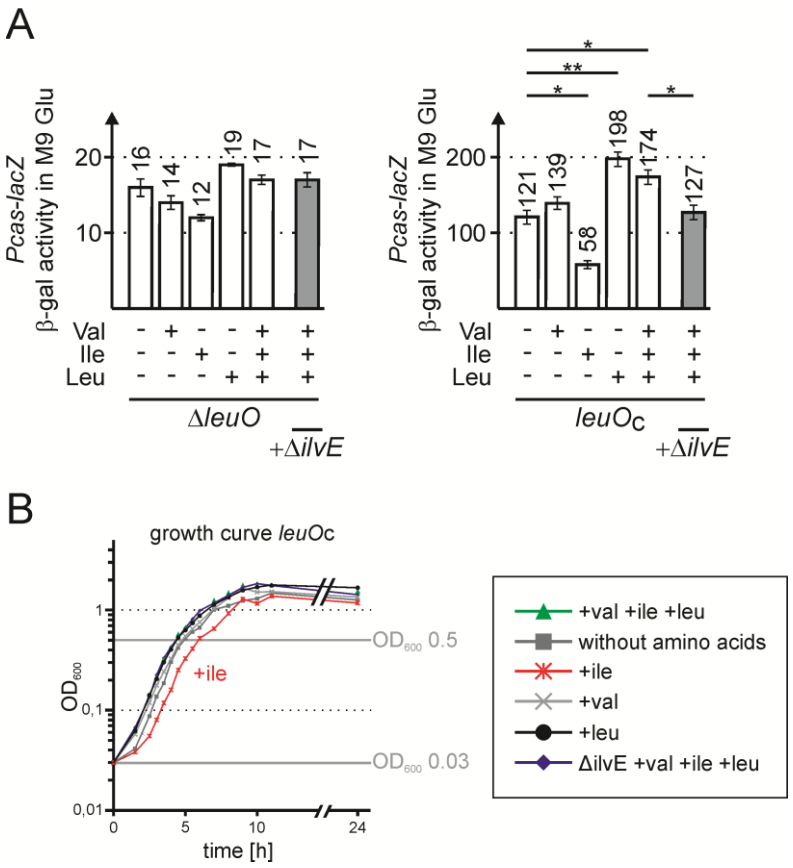

Figure 16: Effect of branched-chain amino acids on *Pcas* activation by LeuO.

(A) Reporter strains carrying *Pcas-lacZ* at the attB site include T1281 [Δ*leuO* Δ*(yjjPQ-bglJ)* Δ*lacZ],* T1927 [Δ*leuO* Δ*ilvE* Δ*(yjjPQ-bglJ)* Δ*lacZ]*, T1295 [*leuOc* Δ(*yjjPQ-bglJ*)] andT1933 [*leuOc* Δ*ilvE* Δ(*yjjPQ-bglJ*)]. For β-galactosidase assays the O/N cultures were grown in M9 minimal medium with 1% glucose, the media was supplemented with amino acids L-valine, L-leucine and L-isoleucine (300 µM), where indicated. The exponential cultures were inoculated to an OD_{600}=0.05 and grown to an OD_{600}=0.5 in M9 minimal glucose medium, supplemented with no amino acids, with all three branched-chain amino acids (300 µM each) or with only one amino acid. Mean values of at least 3 independent biological replicates are shown as bars, and error bars indicate standard deviations. The statistical significance was calculated by t-test for each culture condition compared to the condition without amino acids supplemented with *p < 0.05 and **p < 0.01. (B) Bacterial growth curve of strains T1295 (*leuOc*) and T1933 (*leuOc* Δ*ilvE*). For measurement of the growth curve an O/N culture was grown in M9 minimal glucose medium, supplemented with no amino acid or with all three branched-chain amino acids (300 µM each). Exponential cultures were grown in M9 minimal glucose medium, without amino acid supplementation (grey rectangles), with the supplementation of all three branched-chain amino acids (green triangles for *leuOc;* blue rectangles for *leuOc* Δ*ilvE*) or with only one amino acid (red asterisk for L-isoleucine; grey cross for L-valine and black circles for L-leucine). The cultures were inoculated to OD_{600}=0.03 and grown for 24 h and OD_{600} was measured every hour for 11 hours.

3 Discussion

LeuO is a conserved LysR-type transcriptional regulator (LTTR) of pleiotropic function and a co-regulator of the abundant nucleoid-associated repressor protein H-NS in *Escherichia coli* (Ueguchi *et al.*, 1998, Stoebel *et al.*, 2008, Shimada *et al.*, 2011). The signals and effectors that modulate the activity of the LeuO protein are unknown. In this work the control of LeuO activity was addressed. A screen for hyperactive LeuO mutants yielded nine single residue exchanges mapping to the presumptive cleft of the C-terminal effector-binding domain (EBD) and dimerization interface of the EBD. Low levels of these LeuO mutants are sufficient to fully activate the LeuO-controlled *cas* promoter, while high levels of wild-type LeuO are required for *Pcas* activation. The crystal structures of the EBD of one of these hyperactive mutants, LeuO-S120D, and of wild-type LeuO was solved. Comparison of the structures suggests that the mutation triggers an allosteric structural of the EBD, which is related to effector-induced structural changes of other LTTRs. This structural change transmits to the arrangement of the DNA-binding domains and is reflected in an enhanced DNA-binding specificity of the LeuO-S120D as shown by DNase I footprinting (Figure 17). A detailed analysis of the DNA-binding sites unraveled a palindromic sequence, which allowed defining a DNA-binding motif for LeuO. LeuO-S120D might mimic an effector-bound state, which regulates a putative different set of target loci than the wild-type LeuO. The LeuO mutants can be used as a tool to study the mechanism of transcriptional regulation and the H-NS co-regulation.

3.1 DNA-binding of LeuO and LeuO-S120D

DNase I footprinting of LeuO DNA-binding sites at the *cas* locus (Westra *et al.*, 2010) with a broad concentration range of wild-type and LeuO-S120D demonstrated that DNA-binding by LeuO-S120D is more specific and distinct than by wild-type LeuO. Intriguingly, DNase I footprinting with just the dimeric DNA-binding domain (DBD) of LeuO yielded a similarly distinct and specific pattern as LeuO-S120D. The distinct DNase I footprint pattern is in agreement with binding of a presumptive LeuO-S120D tetramer (carrying two dimeric DBDs) to a high-affinity palindromic DNA-binding-site (“core-site”) and secondary DNA-binding site (Figure 17). For this high-affinity core DNA-binding site a specific consensus sequence motif was defined. The crystal structures of the EBDs of LeuO and LeuO-S120D were solved in this work. The structural change observed for the LeuO-S120D mutant correlates with an enhanced DNA-binding specificity and distinct DNA-binding pattern at the *cas* LeuO-binding sites. It seems possible that wild-type LeuO adopts a conformation that is poorly competent in DNA-binding, while the S120D mutant is fully active. Similarly, apo-AphB may not bind DNA, while the structure of the AphB-N100E mutant suggests that this effector-independent mutant binds DNA and presumably mimics a ligand-bound AphB (Taylor *et al.*, 2012). The model that the

LeuO-S120D mutation likewise allows (or enhances) DNA-binding is supported by the fact that a low LeuO-S120D concentration is sufficient to footprint the *cas* LeuO binding-site II (Figure 11A). S120D shows additional protection from DNase I digestion between the two LeuO DNA-binding sites I and II, indicated by an additional hypersensitive site (open arrowhead, Figure 11A). These additional DNA-binding sites suggest that a higher order complex of two or more tetramers might be formed for the hyperactive LeuO-S120D. Further, the DNase I footprint pattern of LeuO-S120D is very similar to that of the isolated dimeric LeuO-DBD which binds DNA without steric hindrance (Figure 11). Taken together, the data indicate that the hyperactive mutants of LeuO, in particular S120D, may mimic an effector-induced form of LeuO with enhanced DNA-binding specificity.

However, the crystallization of the full-length proteins failed so far. The structure of full-length LeuO and hyperactive LeuO-S120D or their oligomeric state might be too flexible for crystallization. Possibly, a co-crystallization approach of LeuO-S120D with a short DNA-fragment that comprises the specific consensus sequence motif for LeuO might be successful. A complex of the purified protein with short DNA fragments representing the high-affinity core DNA-binding site might lead to a stabilized complex that can further be used for crystallization.

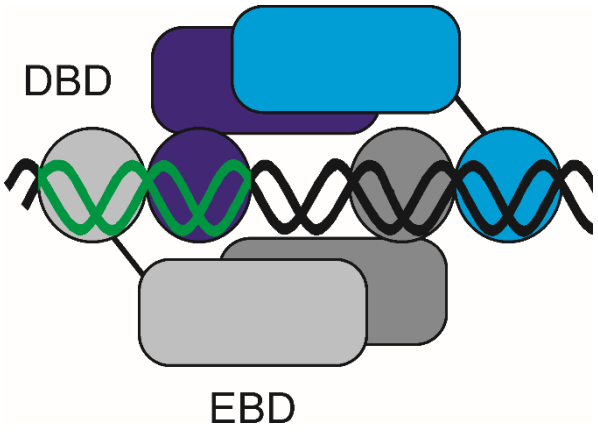

Figure 17: A schematic model for DNA-binding of a LeuO tetramer.

LeuO monomers are depicted with a spheric DNA-binding domain (DBD) and a rectangular effector-binding domain (EBD), connected by a linker. In the tetrameric structure, two EBDs dimerize in a head-to-tail orientation through a conserved dimerization interface (here: blue with light blue and grey with light grey). The DBDs dimerize via the α-helical linker (here: blue with light grey and grey with light blue). Binding of the LeuO tetramer to the DNA might be dependent on a specific angle and distance of the DBDs, which are altered by structural changes of the EBDs, either upon effector-binding or mutation of an amino acid. One DBD dimer binds to the high affinity DNA-binding site (core-site shown in green), while the other DBD dimer binds to an adjacent DNA-binding site.

3.2 Is LeuO activity regulated by an effector?

The LeuO-S120D mutant showed an enhanced DNA-binding specificity and distinct DNA-binding pattern and may mimic an effector-induced state of LeuO. Comparing the crystal structures of the

EBDs of wild-type LeuO and the S120D mutant suggest that the mutation (S120D) induces a structural change of the two EBD subdomains RD-1 and RD-2. Further, the side-chain of Arg218 is re-oriented, which resides at the presumptive cleft of the effector-binding domain. The mutation of Arg218 to Glu, Ala, and Cys also causes hyperactivity, indicating that the positively charged guanidinium group of Arginine may be crucial for an inactive LeuO conformation. These structural changes might also be triggered by binding of an effector regulating the activity of LeuO in response to a signal.

A putative effector regulating LeuO activity is also supported by the putative chloride ion binding next to the rather hydrophobic cleft of the LeuO-EBD. The putative chloride ion locates in the cleft of the EBD and is coordinated mainly by the backbone amide of Met243 and the l side-chain hydroxyl group of Ser120. This putative chloride ion is only found in the wild-type protein and not in the hyperactive mutant where this “space” is occupied by the re-oriented Arg218 (Figure 7). It might be possible that the chloride ion stabilizes the wild-type protein. To date, no effector of LeuO is known. However, a metabolite with a negative charge and a hydrophobic moiety might bind in the cleft of the EBD and therefore induce structural changes that alter the DNA-binding of LeuO. Similar effector-induced or mutation-based rearrangements have been described for other structures of LTTRs including DntR, AphB, BenM, and OxyR (Lerche *et al.*, 2016, Devesse *et al.*, 2011, Taylor *et al.*, 2012, Ruangprasert *et al.*, 2010, Ezezika *et al.*, 2007, Jo *et al.*, 2015, Kullik *et al.*, 1995).

However, some LTTRs are bound by several effectors and for some several effector-binding sites have been identified. For example, in NdhR a second effector-binding site in the dimeric interface between two EBDs is formed (Jiang *et al.*, 2018). Several hyperactive LeuO mutants as H142R, A237V, H254R and Q210R locate rather distant from the cleft of the EBD, but still in the interface between two dimerizing EBDs. These mutants may cause hyperactivity by stabilizing the dimer or by affecting a secondary effector-binding site in LeuO.

Several described LTTRs map next to the genes they regulate, for example BenM and IlvY (Maddocks & Oyston, 2008, Bundy *et al.*, 2002, Wek & Hatfield, 1988). The *leuO* gene is located between two operons related to branched-chain amino acid synthesis. The *leu* operon encodes for leucine synthesis and the *ilvIH* operon encodes for enzymes involved in the synthesis of valine and isoleucine. In addition, *leuO* is upregulated in stationary growth phase and by amino acid starvation (Fang *et al.*, 2000, Majumder *et al.*, 2001). A putative effector regulating LeuO might be present in the same conditions when LeuO is upregulated or when an intermediate product of the amino acid synthesis is present. Nonetheless, the accumulation of precursors of branched-chain amino acids had no striking effect on the LeuO activity (Figure 16). Possible reasons might be insufficient levels of the precursors or that an earlier starvation signal leads to the activation of LeuO. However, the *Pcas-lacZ*

reporter can be a valuable tool for searching an effector. Another approach to identify putative effectors regulating the LeuO activity, a fluorescence-based thermal stability assay (Thermoflour screen) can screen for possible metabolites (in solution conditions) that, when bound to the protein induces structural changes and enhance the protein stability (Boivin *et al.*, 2013). To identify an effector for LeuO, another possible approach might be a metabolite library screen using the *Pcas* reporter or another LeuO target gene reporter. LeuO strongly activates *Pcas*, associated with the bacterial defense system in *E. coli* against foreign DNA and phage invasion (Barrangou *et al.*, 2007). It is also possible that a putative effector of LeuO is present (and induces the structural changes leading to increased binding specificity like S120D) when *cas* is activated by LeuO. Components of foreign DNA or elements of invading phages might be metabolites that bind to LeuO and alter its activity as an effector.

3.3 The spectrum of LeuO target loci

Up to date, LeuO target genes and their regulation by LeuO have been identified and studied with wild-type LeuO only. These studies led to the identification of a broad spectrum of target genes and a rather weak consensus sequence for the LeuO DNA-binding site (Dillon *et al.*, 2012, Ishihama *et al.*, 2016, Shimada *et al.*, 2011). The palindromic consensus sequence defined here on the basis of the core-DNA-binding site detected by DNase I footprinting is very specific, but it is also in agreement with the consensus motif characterized before (Dillon *et al.*, 2012). The bioinformatic analysis using MEME Suite and FIMO in an iterative approach yielded a good correlation of the top LeuO target genes. However, not all LeuO targets could be identified in this approach. Assuming that LeuO activity is indeed effector-controlled the spectrum of LeuO target genes might change upon effector-binding, with some targets being up-regulated, others down-regulated or not affected at all. Such a scenario has been characterized for example for the control of *E. coli* ArgP activity by the effectors lysine (acting inhibitory) and arginine (acting positively) (Nguyen Le Minh *et al.*, 2018, Laishram & Gowrishankar, 2007). In this study, the effect of the LeuO mutants that are hyperactive in regulation of the *cas* promoter, were compared to a second LeuO target, the autoregulated *leuO* promoter. Interestingly, S120D as well as S128P, H142R, Q210R, R218C, appear hyperactive in positive and in negative *leuO* autoregulation. The other hyperactive LeuO mutants T127I, A237V, and H254R regulate the *leuO* promoter similar to wild-type LeuO. This indicates that indeed target-gene specific differences may exist. Depending on the arrangement of the core and auxiliary LeuO DNA-binding sites in a LeuO-regulated promoter, effector-binding or a mutations-causing hyperactivity in *cas* regulation may have no effect, enhance competition or co-regulation with H-NS, and allow or prevent interaction with RNA-polymerase. Thus the spectrum of LeuO target genes may vary in dependence of an effector, which could be addressed in the future for example by analyzing other

hyperactive LeuO mutants that may vary from LeuO-S120D. LeuO regulated targets analyzed here were mainly weakly regulated by wild-type LeuO (Table 5). The target loci were identified using high levels of wild-type LeuO. The LeuO-S120D mutant might mimic an effector-induced state and as a next step, analyzing the transcriptome by differential RNA seq, of cells expressing genomic LeuO-S120D instead of the endogenous LeuO, is a promising approach to identify effector-dependent targets of LeuO.

3.4 How does LeuO regulate transcription?

The relationship between the LysR-type regulator LeuO and the global silencer H-NS is quite versatile and different aspects have been analyzed in the past. Many LeuO-activated target loci are repressed by H-NS and LeuO is therefore considered an H-NS antagonist (Stoebel *et al.*, 2008, Hernández-Lucas & Calva, 2012). LeuO can also function as a "back-up" of H-NS and repress target loci when H-NS is not present (Espinosa & Casadesús, 2014). In addition, genomic SELEX data shows a significant overlap of co-regulation of LeuO target loci by H-NS (Stratmann *et al.*, 2012, Dillon *et al.*, 2012, Shimada *et al.*, 2011). However, the exact mechanisms behind the binding of both regulators to the target DNA and the resulting regulation remain elusive. The binding sites for LeuO and H-NS may determine whether LeuO acts as an activator or repressor. For the *cas* promoter it was proposed that LeuO does not compete for binding with LeuO and that LeuO might limit the spreading of H-NS into the promoter region (Westra *et al.*, 2010). The gene *ompS1* is co-regulated by LeuO and H-NS and it was suggested that activation by LeuO is presumably achieved when H-NS and LeuO compete for binding to the DNA (De la Cruz *et al.*, 2007). The spreading of the H-NS nucleocomplex can also be delimited by LeuO by transcription induced changes in DNA supercoiling (Chen & Wu, 2005). An *in vitro* transcription assay might shed light on the mechanism of the co-regulation with H-NS when the LeuO-S120D mutant alters the complex formation on the DNA or the removal of H-NS from the DNA. The nine hyperactive LeuO mutants showed differences in the regulation of *cas* and *leuO*, which might be caused by a different DNA-binding specificity, altered stabilization of the protein dimer or tetramer, or an altered interaction with H-NS. A more detailed analysis of the mutants may give new insights in the function of LeuO and its co-regulation with H-NS. Differences in the DNA-binding by the hyperactive LeuO mutants will be analyzed using surface plasmon resonance (SPR) spectroscopy to determine the binding affinity.

Another approach of characterization the interaction between H-NS and LeuO is the introduction of a non-canonical and UV-crosslinkable amino acid. The interactions with other transcription factors, with H-NS and the RNA polymerase can be analyzed, as well as the presumed dimerization and tetramer formation of LeuO.

3.5 Approaches for analyzing LysR-type transcriptional regulators

In *E. coli* 46 LysR-type transcriptional regulators (LTTRs) are described: for 17 LTTRs target genes and signals that modulate their activity are described, for 10 LTTRs some aspects of their function are known but their regulation and possible effectors are unidentified and 19 LTTRs are of unknown function. Characterizing LTTRs remains experimentally challenging due to their intrinsic structural properties and variable DNA-binding mode, but also the identification of effectors is a difficult task.

Target loci of LTTRs that were identified in the past used rather high and artificial concentrations of the wild-type protein, for example in microarray transcriptome analyses or genome scale SELEX screenings. These high protein concentrations lead to unspecific DNA-binding and target loci identification that might not be regulated in native cell conditions. To circumvent the problem of unspecific DNA-binding, the DNA-binding domain (DBD) might be a promising approach for RNA seq, due to specific DNA-binding and the possible lower protein concentrations that were needed. An already described target locus might be of good use as a reporter. Using for example a reporter lacZ fusion, hyperactive mutants can be screened, as described here for LeuO (see chapter 2.1). Mutants of the LTTR can shed light to a different set of target loci, to structural characteristics and even indicate a metabolite that alters the activity as an effector. To further analyze the function and the oligomeric state of an LTTR the protein can be crystallized to solve the structure. It is possible that putative metabolites that alter the activity of the LTTR co-crystallize. Another approach to identify effectors of LTTRs might be a metabolic library screen.

4 Materials and Methods

4.1 Media, antibiotics and bacterial cultivation

Bacterial cultures of *Escherichia coli* K12 were grown in LB medium (5 g/l Bacto Yeast extract, 10 g/l Bacto Tryptone, 5 g/l NaCl), for plates 15 g/l Bacto Agar was added. Antibiotics were added as indicated in final concentrations of 50 µg/ml ampicillin, 15 µg/ml chloramphenicol, 50 µg/ml spectinomycin, 12 µg/ml tetracyclin and 25 µg/ml kanamycin. IPTG (isopropyl-β,D-thiogalactopyranoside) was added to liquid cultures at a concentration of 1 mM and to plates at a concentration of 200 µM. X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was added to a final concentration of 40 µg/ml to tryptone plates, where indicated.

4.2 Standard molecular techniques

Standard molecular techniques like agarose gel electrophoresis, molecular cloning and PCR were carried out according to published protocols (Ausubel, 2005). Sequencing was performed by GATC Biotech AG, Konstanz, Germany.

4.3 Bacterial strains, plasmids and oligonucleotides

Escherichia coli K12 strains used and constructed in this study are listed in Table 2, plasmids are given in Table 3 and sequences of oligonucleotides are summarized in Table 4.

Table 2: *Escherichia coli* K12 strains

strain	genotype[a)]	reference/ construction
BW30270	MG1655 *rph*$^+$ (laboratory collection S3839)	CGSC#7925
C41(DE3)	BL21(DE3) derivate selected for expression of toxic proteins	(Miroux & Walker, 1996)
T1241	BW30270 *ilvG*$^+$ (motile)	(Pannen *et al.*, 2016)
S1729	*leuOc (leuO-Y1*::miniTn10-cmR; constitutive expression of leuO)	(Madhusudan *et al.*, 2005)
S4197	BW30270 ilvG$^+$ ΔlacZ (non-motile)	(Venkatesh *et al.*, 2010)
T23	S4197 Δ(*yjjP-yjjQ-bglJ*)$_{FRT}$	(Stratmann *et al.*, 2012)
T61	S4197 ΔleuO::KD4-kanR	lab collection (T. Stratmann)
T70	S4197 Δ(*yjjP-yjjQ-bglJ*)::KD3-cmR	(Venkatesh *et al.*, 2010)
T314	S4197 Δ*leuO*$_{FRT}$ Δ(*yjjP-yjjQ-bglJ*)$_{FRT}$	(Venkatesh *et al.*, 2010)
T1146	S4197 *leuO*$_C$ (=*leuO-Y1*::mTn10-cmR)	S4197 x T4GT7(S1729), lab collection (T. Stratmann)
T1148	S4197 Δ*yjjP-yjjQ-bglJ*::FRT leuO$_C$	T23 x T4GT7(S1729), lab collection (T. Stratmann)
T1610	T314 attB::(*Pcas lacZ aadA*) / F' Tn10 (TetR) proA$^+$B$^+$ lacI^q Δ(lacZ)M15	T1281 x F' (XL1-Blue) conjugation
***E. coli* K12 strains with gene deletions constructed as described (Datsenko & Wanner, 2000, Kolmsee & Hengge, 2011)**		
T1922	T23 Δ*ilvE*::KD4-kanR	T23/pKD46 x T967/T968 of pKD4
T1923	T1148 Δ*ilvE*::KD4-kanR	T1148/pKD46 x T967/T968 of pKD4

Table 2: *Escherichia coli* K12 strains		
strain	**genotype**[a)]	**reference/ construction**
T1924	T1281 Δ*ilvE*::KD4-kanR	T1281/pKD46 x T967/T968 of pKD4
T1925	T23 Δ*ilvE*::FRT	T1922 x pCP20
T1926	T1148 Δ*ilvE*::FRT	T1923 x pCP20
T1927	T1281 Δ*ilvE*::FRT	T1924 x pCP20
T1928	T1295 Δ*ilvE*::KD4-kanR	T1295/pKD46 x T967/T968 of pKD4
T1929	T314 Δ*ilvE*::KD3-cmR	T314/pKD46 x T967/T968 of pKD3
T1933	T1295 Δ*ilvE*::FRT	T1928 x pCP20
T1934	T314 Δ*ilvE*::FRT	T1929 x pCP20
T1876	T1241 Δ*leuO*::*ccdB neo*	T1241/pKD46 x T994/T995 of pKILL45
T1877	S3839 Δ*leuO*::*ccdB neo*	S3839/pKD46 x T994/T995 of pKILL45
T1921	T1146 P*leuO*c Δ*leuO*::*ccdB neo*	T1146/pKD46 x T994/T995 of pKILL45
T1930	S3839 *leuO*-S120D/M244T	T1877/pKD46 x T996/T997 of pKESL107
T1931	S3839 *leuO*-M244T	T1877/pKD46 x T996/T997 of pKESL73
T1932	S3839 *leuO*-S120D	T1877/pKD46 T996/T997 of pKESL104
T1938	S4197 P*leuO*c *leuO*-M244T	T1921/pKD46 x T996/T997 of pKESL73
***E. coli* K12 strains constructed by phage transduction**		
T1988	BW30270 Δ*leuO*::KD4-kanR	BW30270 x T4GT7 (T61)
T1989	BW30270 Δ*yjjP-yjjQ-bglJ*::KD3-cmR	BW30270 x T4GT7 (T70)
T1990	BW30270 Δ(*yjjP-yjjQ-bglJ*)::KD3-cmR, *leuO*-S120D	T1932 x T4GT7 (T70)
T1991	BW30270 Δ(*yjjP-yjjQ-bglJ*)::KD3-cmR, *leuO*-M244T	T1931 x T4GT7 (T70)
T2009	BW30270 Δ*leuO*::FRT	T1988 x pCP20
T2010	BW30270 Δ*yjjP-yjjQ-bglJ*::FRT	T1989 x pCP20
T2011	BW30270 Δ*yjjP-yjjQ-bglJ*::FRT *leuO*-S120D	T1990 x pCP20
T2012	BW30270 Δ*yjjP-yjjQ-bglJ*::FRT *leuO*-M244T	T1991 x pCP20
***E. coli* K12 strains with promoter *lacZ* fusion at attB integration site** [b)]		
T308	S4197 attB::(*PleuO lacZ aadA*) Δ$leuO_{FRT}$ Δ(*yjjP-yjjQ-bglJ*)$_{FRT}$	(Stratmann *et al.*, 2012)
T352	S4197 attB::(*PleuO lacZ aadA*) Δ$leuO_{FRT}$ Δ(*yjjP-yjjQ-bglJ*)$_{FRT}$ Δhns_{FRT} *stpA*::*tet*	(Stratmann *et al.*, 2012)
T568	T314 attB::(*Pbgl* $t1_{RAT}$ *bglG lacZ aadA*)	(Venkatesh *et al.*, 2010)
T862	S4197 attB::(*PleuO lacZ aadA*) ΔleuO$_{FRT}$ ($bglJ_C$)	(Stratmann *et al.*, 2012)
T1188	T314 attB::(*PchiA lacZ aadA*)	(Salscheider *et al.*, 2014)
T1275	T314 attB::(*PuspG lacZ aadA*)	T314/pLDR8 x pKESL31
T1277	T314 attB::(*PompN lacZ aadA*)	T314/pLDR8 x pKESL32
T1279	T314 attB::(*PmicC lacZ aadA*)	T314/pLDR8 x pKESL33
T1281	T314 attB::(*Pcas lacZ aadA*)	T314/pLDR8 x pKESL34
T1291	T1148 attB::(*PuspG lacZ aadA*)	T1148/pLDR8 x pKESL31

Table 2: *Escherichia coli* K12 strains

strain	genotype[a)]	reference/ construction
T1293	T1148 attB::(*PompN lacZ aadA)*	T1148/pLDR8 x pKESL32
T1295	T1148 attB::(*Pcas lacZ aadA*)	T1148/pLDR8 x pKESL34
T1313	T1148 attB::(*PchiA lacZ aadA*)	T1148/pLDR8 x pKESL9
T1317	T1148 attB::(*PmicC lacZ aadA*)	T1148/pLDR8 x pKESL33
T1319	T314 attB::(*PyjjQ lacZ aadA*)	T314/pLDR8 x pKES111
T1323	T1148 attB::(*PyjjQ lacZ aadA*)	T1148/pLDR8 x pKES111
T1494	T314 attB::(*PsdiA lacZ aadA*)	T314/pLDR8 x pKESL36
T1509	T314 attB::(*PnmpC lacZ aadA*)	T314/pLDR8 x pKESL35
T1511	T1148 attB::(*PnmpC lacZ aadA*)	T1148/pLDR8 x pKESL35
T1515	T314 attB::(*PgspC lacZ aadA*)	T314/pLDR8 x pKESL38
T1517	T1148 attB:: (*PgspC lacZ aadA*)	T1148/pLDR8 x pKESL38
T1527	T1148 attB::(*PsdiA lacZ aadA*)	T1148/pLDR8 x pKESL36
T2396	T314 attB::(*PyafT lacZ aadA*)	T314/pLDR8 x pKESMS70
T2397	T1148 attB::(*PyafT lacZ aadA*)	T1148/pLDR8 x pKESMS70
T2398	T314 attB::(*PybdO lacZ aadA*)	T314/pLDR8 x pKESMS71
T2399	T1148 attB::(*PybdO lacZ aadA*)	T1148/pLDR8 x pKESMS71
T2400	T314 attB::(*PybeQ lacZ aadA*)	T314/pLDR8 x pKESMS77
T2401	T1148 attB::(*PybeQ lacZ aadA*)	T1148/pLDR8 x pKESMS77
T2402	T314 attB::(*PybeR lacZ aadA*)	T314/pLDR8 x pKESMS72
T2403	T1148 attB::(*PybeR lacZ aadA*)	T1148/pLDR8 x pKESMS72
T2404	T314 attB::(*PyghS lacZ aadA*)	T314/pLDR8 x pKESMS73
T2405	T1148 attB::(*PyghS lacZ aadA*)	T1148/pLDR8 x pKESMS73
T2406	T314 attB::(*PyghT lacZ aadA*)	T314/pLDR8 x pKESMS74
T2407	T1148 attB::(*PyghT lacZ aadA*)	T1148/pLDR8 x pKESMS74
T2408	T314 attB::(*PenvR lacZ aadA*)	T314/pLDR8 x pKESMS78
T2409	T1148 attB::(*PenvR lacZ aadA*)	T1148/pLDR8 x pKESMS78
T2410	T314 attB::(*PenvC lacZ aadA*)	T314/pLDR8 x pKESMS75
T2411	T1148 attB::(*PenvC lacZ aadA*)	T1148/pLDR8 x pKESMS75
T2412	T314 attB::(*PyjfI lacZ aadA*)	T314/pLDR8 x pKESMS76
T2413	T1148 attB::(*PyjfI lacZ aadA*)	T1148/pLDR8 x pKESMS76

[a)] The following abbreviations and genetic designations are used: FRT for Flp recombinase target site, *aadA* for spectinomycin resistance, *tet* for tetracycline resistance, *neo* for kanamycin resistance. Allele $bglJ_C$ refers to allele *yjjQ/bglJ*-Y6::mTn10-cmR, directing constitutive expression of *bglJ* (Madhusudan *et al.*, 2005). Allele *leuOc* refers to leuO-Y1::mTn10-cmR, directing constitutive *leuO* expression. Chromosomal deletions and insertion of mutations were constructed by λ-Red mediated recombination (Datsenko & Wanner, 2000) or (Kolmsee & Hengge, 2011) described in chapter 4.9 and 4.10. Flipping of the FRT (Flp recombinant target site) flanked resistance cassette by Flp recombinase was performed using plasmid pCP20 (x pCP20). T4 phage transduction was performed as described (Wilson *et al.*, 1979) and is represented as "recipient strain x phage T4GT7 (donor strain)", described in chapter 4.11.

[b)] Construction of strains was performed by integration promoter *lacZ* reporter fusions into the chromosomal *attB* site (indicated as strain/pLDR8 x plasmid designation) as described (Diederich *et al.*, 1992). In brief, integration of reporter *lacZ* constructs into the phage λ *attB* site was performed using replication origin-less re-

ligated BamHI fragments of the indicated plasmids carrying a promoter *lacZ* fusion, the *attP*-site and the *aadA* gene for selection by spectinomycin. All alleles were confirmed by PCR.

Table 3: Plasmids

plasmid	features[a]	reference/ construction[b]
pCP20	cI_{857} λ-P_R flp in pSC101 rep^{ts} ampR	(Cherepanov & Wackernagel, 1995)
pET-22b(+)	T7 promoter, His6 expression vector ampR	Novagen®
pFDY157	$lacI^q$ Ptac *lacZ* rrnB-T1,T2 pBR-ori ampR	lab collection
pKD3	FRT-cmR-FRT oriRγ ampR	(Datsenko & Wanner, 2000)
pKD4	FRT-kanR-FRT oriRγ ampR	(Datsenko & Wanner, 2000)
pKD45	rhaP 3 *ccdB* (L50F) *neo* cassette inR6Kgamma	(Kolmsee & Hengge, 2011)
pKD46	*araC* P_{BAD} γ-β-exo in pSC101 rep^{ts} ampR	(Datsenko & Wanner, 2000)
pKILL45	rhaP 3 *ccdB*(wt) *neo* cassette inR6Kgamma in pKD45	cloned with (OA001/OA002)
pLDR8	cI_{857} P_R λ-*int* in pSC101 rep^{ts} kanR	(Diederich *et al.*, 1992)
pKES268	PlacUV5 MCS *lacZ* ori-p15A kanR attP *aadA*	(Salscheider *et al.*, 2014)
pKES334	$lacI^q$ Ptac MCS rrnB-T1,T2 pBR-ori ampR	pFDY157 x adaptor (OA576/OA577)
pKESK22	$lacI^q$ Ptac MCS in ori-p15A kanR	(Stratmann *et al.*, 2008)
pKETS2	P*leuO*(-846 to +45)-XbaI-TAA *lacZ* in p15A kanR attP *aadA*	lab collection, (T. Stratmann)
pKETS5	*leuO* in pKESK22	(Stratmann *et al.*, 2012)
pKESK10	PlacUV5 *bglG* ori-pSC cmR	(Dole *et al.*, 2002)
pKETS24	PlacUV5 MCS ori-pSC cmR	MCS in pKESK10
pKETS25	PlacUV5 *leuO* ori-pSC cmR	(Breddermann & Schnetz, 2016)
pKES111	*PyjjQ lacZ* in pACYC	(Stratmann *et al.*, 2008)
pKESL9	*PchiA lacZ* in pKES268	(Salscheider *et al.*, 2014)
pKESL31	*PuspG lacZ* in pKES268	cloned with T734/T735
pKESL32	*PompN lacZ* in pKES268	cloned with T740/T741
pKESL33	*PmicC lacZ* in pKES268	cloned withT738/T739
pKESL34	*Pcas lacZ* in pKES268	cloned with T736/T737
pKESL35	*PnmpC lacZ* in pKES268	cloned with T806/T807
pKESL36	*PsdiA lacZ* in pKES268	cloned with TT808/T809
pKESL38	*PgspC lacZ* in pKES268	cloned with T810/T811
pKESL39	*leuO* in pKETS24	directed *leuO* mutagenesis
pKESL40	*leuO*-S120D(AGC>GAT)+ E111D(GAA>GAC) in pKETS24	directed *leuO* mutagenesis
pKESL61	*leuO*-M243E(ATG>GAA) in pKETS24	directed *leuO* mutagenesis
pKESL62	*leuO*-L263E(CTG>GAA) in pKETS24	directed *leuO* mutagenesis
pKESL64	*leuO*-I125E(ATT>GAA) in pKETS24	directed *leuO* mutagenesis
pKESL65	*leuO*-I125E(ATT>GAA)+ V118I(GTT>ATT) in pKETS24	directed *leuO* mutagenesis
pKESL67	*leuO*-F219E(TTC>GAA) in pKETS24	directed *leuO* mutagenesis
pKESL68	*leuO*-L122E(TTA>GGA)+ H115Y(CAT>TAT) in pKETS24	directed *leuO* mutagenesis
pKESL69	*leuO*-P121D(CCG>GAT) in pKETS24	directed *leuO* mutagenesis
pKESL70	*leuO*-Y168E(TAT>GAA) in pKETS24	directed *leuO* mutagenesis
pKESL71	*leuO*-S120D(AGC>GAT)+ E111D(GAA>GAC) in pKESK22	directed *leuO* mutagenesis
pKESL72	*leuO*-S120D(AGC>GAT)+ E111D(GAA>GAC) +D205N(GAT>AAT) in pKESK22	screen (T568)
pKESL73	*leuO*-M244T(ATG>ACG) in pKESK22	screen (T1281, *)
pKESL74	*leuO*-T127I(ACC>ATC) in pKESK22	screen (T1610)
pKESL75	*leuO*-S128P(TCG>CCG) in pKESK22	screen (T1610, 3x)
pKESL76	*leuO*-H142R(CAT>CGT) in pKESK22	screen (T1610, 2x)
pKESL77	*leuO*-Q210R(CAA>CGA) in pKESK22	screen (T1610, 2x)
pKESL78	*leuO*-R218C(CGT>TGT) in pKESK22	screen (T1610)
pKESL79	*leuO*-V230I(GTA>ATA)+ M244T(ATG>ACG) in pKESK22	screen (T1610, *)
pKESL80	*leuO*-A237V(GCG>GTG) in pKESK22	screen (T1610, 2x)
pKESL101	*leuO*-H254R(CAT>CGT)+ R74R(CGT>CGC) in pKESK22	screen (T1610, 2x)
pKESL102	*leuO*-S120D(AGC>GAT) in pKETS24	directed *leuO* mutagenesis

Table 3: Plasmids		
plasmid	**features**[a]	**reference/ construction**[b]
pKESL103	*leuO*-L122E(TTA>GGA) in pKETS24	directed *leuO* mutagenesis
pKESL104	*leuO*-S120D(AGC>GAT) in pKESK22	directed *leuO* mutagenesis
pKESL105	*leuO*-S120D(AGC>GAT)+ C119S(TGC>AGC) in pKESK22	directed *leuO* mutagenesis
pKESL106	*leuO*-L122E(TTA>GGA) in pKESK22	directed *leuO* mutagenesis
pKESL107	*leuO*-S120D(AGC>GAT)+ M244T(ATG>ACG) in pKESK22	directed *leuO* mutagenesis
pKESL108	*leuO*-S120A(AGC>GCG) in pKESK22	directed *leuO* mutagenesis
pKESL109	*leuO*-M244A(ATG>GCG)+ G240V(GGC>GTC) in pKESK22	directed *leuO* mutagenesis
pKESL110	*leuO*-M244A(ATG>GCG)+ M34V(ATG>GTG) in pKESK22	directed *leuO* mutagenesis
pKESL221	*leuO*-N150A(AAT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL222	*leuO*-N152A(AAC>GCC) in pKESK22	directed *leuO* mutagenesis
pKESL223	*leuO*-H155A(CAT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL224	*leuO*-Q156A(CAG>GCG) in pKESK22	directed *leuO* mutagenesis
pKESL225	*leuO*-H172A(CAT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL226	*leuO*-R173A(CGT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL227	*leuO*-K232A(AAG>GCG) in pKESK22	directed *leuO* mutagenesis
pKESL228	*leuO*-R158A(CGT>GCT)+ E175E(GAA>GAG)in pKESK22	directed *leuO* mutagenesis
pKESL229	*leuO*-Q151A(CAG>GCG) in pKESK22	directed *leuO* mutagenesis
pKESL230	*leuO*-H204A(CAT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL231	*leuO*-N150E(AAT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL232	*leuO*-Q151E(CAG>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL233	*leuO*-N152E(AAC>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL234	*leuO*-H155E(CAT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL235	*leuO*-R158E(CGT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL236	*leuO*-H172E(CAT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL237	*leuO*-H204E(CAT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL238	*leuO*-R218A(CGT>GCT) in pKESK22	directed *leuO* mutagenesis
pKESL239	*leuO*-W226A(TGG>GCG) in pKESK22	directed *leuO* mutagenesis
pKESL240	*leuO*-W226E(TGG>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL241	*leuO*-Q156E(CAG>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL242	*leuO*-R173E(CGT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL243	*leuO*-R218E(CGT>GAG) in pKESK22	directed *leuO* mutagenesis
pKESL244	*leuO*-K232E(AAG>GAG) in pKESK22	directed *leuO* mutagenesis
pKESMS1	*leuO*-M244A(ATG>GCG) in pKESK22	directed *leuO* mutagenesis
pKESMS3	*leuO*-S120D-EBD-His6 in pET-22b(+)	protein purification
pKESMS16	*leuO*-His6 in pET-22b(+)	protein purification
pKESMS17	*leuO*-S120D-His6 in pET-22b(+)	protein purification
pKESMS21	*leuO*-EBD-His6 in pET-22b(+)	protein purification
pKESMS38	*leuO*-His6 in pFDY157	cloned with S326/S400
pKESMS63	*leuO*-DBD in pKESK22	cloned with S326/OA440
pKESMS64	*leuO*-C117S(TGT>AGC)+ L263L(CTG>CTA) in pKESK22	directed *leuO* mutagenesis
pKESMS65	*leuO*-C119S(TGC>AGC) in pKESK22	directed *leuO* mutagenesis
pKESMS66	*leuO*-C119D(TGC>GAT) in pKESK22	directed *leuO* mutagenesis
pKESMS67	*leuO*-C117D(TGT>GAT) in pKESK22	directed *leuO* mutagenesis
pKESMS68	*leuO*-DBD-His6 in pET-22b(+)	protein purification
pKESMS70	*PyafT lacZ* in pKES268	cloned with OA514/OA515
pKESMS71	*PybdO lacZ* in pKES268	cloned with OA516/OA517
pKESMS72	*PybeR lacZ* in pKES268	cloned with OA520/OA521
pKESMS73	*PyghS lacZ* in pKES268	cloned with OA522/OA523
pKESMS74	*PyghT lacZ* in pKES268	cloned with OA524/OA525
pKESMS75	*PenvC lacZ* in pKES268	cloned with OA528/OA529
pKESMS76	*PyjfI lacZ* in pKES268	cloned with OA530/OA531
pKESMS77	*PybeQ lacZ* in pKES268	cloned with OA518/OA519
pKESMS78	*PenvR lacZ* in pKES268	cloned with OA526/OA527
pKESMS79	*leuO*-S120D-His6 in pFDY157	cloned with S326/S400

[a)] The following abbreviations and genetic designations are used: MCS, multiple cloning site, rep[ts], temperature sensitive origin of replication, *aadA* refers to spectinomycin resistance, cmR for chloramphenicol resistance, *neo* for kanamycin resistance (kanR), ampR for ampicillin resistance. DBD refers to the N-terminal DNA-binding domain of LeuO, from amino acid 1 to amino acid 101 of the protein. EBD refers to the C-terminal effector-binding domain of LeuO, from amino acid 108 to amino acid 314 (end of protein).
[b)] For cloning, PCR fragments were amplified with the indicated oligonucleotides. Vectors and PCR fragments were digested with restriction enzymes, gel purified and cloned. All cloned plasmids were confirmed by PCR, restriction digestion, and sequencing of the cloned fragment.
The LeuO mutant screen using either *Pcas-lacZ* (T1281, T1610) or *Pbgl-lacZ* (T568) as reporter, LeuO mutants marked with (2x) and (3x) were isolated two or three times independently (from independent PCR runs), the asterisk (*) marks mutations that were isolated independently but with different secondary mutation or synonymous substitution.

Table 4: Oligonucleotides

oligo	sequence[a)]	application/target
S93	CCGGGCCGACAACAAAGTCA	analysis of *attB* integration
S95	CATATGGGGATTGGTGGCGA	analysis of *attB* integration
S118	TGCGGGCCTCTTCGCTATTA	analysis of *attB* integration
S164	GAGCAGGGGAATTGATCCGGTGGA	analysis of *attB* integration
S326	aagaattcggatccGTGTGACAGTGGAGTTAAGTATGCCAG	*leuO* cloning
S328	ccctgcagctagcTGACCTATTCTGCAATCAGTTAGCG	*leuO* cloning
S400	tactgcagctagcttaatgatgatgatgatgatgGCGTTTGCAAATTGAGACTAATT GC	*leuO* cloning
T334	TGGCGAAGTAATCGCAACATCC	analysis of *attB* integration
T645	aggtggatccTGACCTATTCTGCAATCAGTTAGCG	*leuO* cloning in pKETS24
T734	gcaggtcgacCCAGAGACAGTTCAAAAAAGAGTCAGTC	P*uspG*
T735	gcagtctagaATACATAACCCTTTCTCCCTGTTAATCA	P*uspG*
T736	gcaggtcgacGTCATCCCTGCAAATCCCAAATAAC	P*cas*
T737	gcagtctagaATATGCTCCGACATTTCTCCTGC	P*cas*
T738	gcaggtcgacCAGTACTTTGCTTTTCATTGAATAAATCCT	P*micC*
T739	gcagtctagaGTGACAATAAAGGCATATAACCCGC	P*micC*
T740	gcaggtcgacGTGACAATAAAGGCATATAACCCGC	P*ompN*
T741	gcagtctagaCAGTACTTTGCTTTTCATTGAATAAATCCT	P*ompN*
T792	CGTGTATTTCATCTTTGTGTTTGCgatCCGTTAGACAGCATTCTGAC	*leuO*-S120D mutagenesis
T793	GTATTTCATCTTTGTGTTTGCAGCgatTTAGACAGCATTCTGACCTCG	*leuO*-P121D mutagenesis
T794	ATCTTTGTGTTTGCAGCCCGgaaGACAGCATTCTGACCTCGC	*leuO*-L122E mutagenesis
T795	TTTGCAGCCCGTTAGACAGCgaaCTGACCTCGCAGATTTATAATCAC	*leuO*-I125E mutagenesis
T796	TCAGGAAACGGAGTTTGTGATTAGTgaaGAAGACTTCCATCGTCCTG	*leuO*-Y168E mutagenesis
T800	CGGCGGTTTCGCTCGATCGTgaaGCGTCATTTAGTCAACCTTG	*leuO*-F219E mutagenesis
T801	GCGTATCAGGGCATGGCAgaaATGAGCGTACTTAGCGTGG	*leuO*-M243E mutagenesis
T802	ATTGCGCCGCGTTGGgaaGCTGAAGAGTTCGCTGA	*leuO*-L263E mutagenesis
T803	gaccgctagcGTGTGACAGTGGAGTTAAGTATGCCAG	*leuO* cloning in pKETS24
T806	gcaggtcgacCGACACCCGTTGTTAACTTATCCAT	P*nmpC*
T807	gtcatctagaGCCACTGTTAATTTTTTCATCGTGAG	P*nmpC*
T808	gtcagtcgac CGAGAAGTTTCTGCTGCAATAATAAGA	P*sdiA*
T809	gcagtctaga GCTGAAAAAATCCTTATCCTGCATAGTAAA	P*sdiA*
T810	gcaggtcgac CTTCTCTTCTCGTAGACATAGAACTTCCTG	P*gspC*
T811	gcagtctaga AACGTAGTGTGGGCACGATGTATGT	P*gspC*
T912	GCTGGTGGCACTGGGTAGTTGTTA	analysis of *attB* integration
T945	AGAATGCTGTCttcCGGGCTGCAAACACAAAGAT	*leuO*-L122E mutagenesis
T967	AAATCCGCGCCTGAGCGCAAAAGGAATATAAAAATGACCACGAAGgtgt aggctggagctgcttcg	*ilvE* deletion
T968	TTGTATTTATTGATTAACTTGATCTAACCAGCCCCATTTATCTTCcatatgaa tatcctccttagttcctattcc	*ilvE* deletion
T978	TTGTGTTTGCgcgCCGTTAGACAGCATTCTGACCTC	*leuO*-S120A mutagenesis
T979	GTCTAACGGcgcGCAAACACAAAGATGAAATACACGT	*leuO*-S120A mutagenesis

Table 4: Oligonucleotides

oligo	sequence[a)]	application/target
T980	CATGGCAATGgcgAGCGTACTTAGCGTGGTGTCG	*leuO*-M244A mutagenesis
T981	AAGTACGCTcgcCATTGCCATGCCCTGATACG	*leuO*-M244A mutagenesis
T984	TCAGGGCATGGCAATGgcgAGCGTACTTAGCGTGGTGTCG	*leuO*-M244A mutagenesis
T994	AGTTAAGTATGCCAGAGGTACAAACAGATCATCCAGAGACGGCGGcgaaccccagagtcccgc	*leuO* deletion
T995	GACCTATTCTGCAATCAGTTAGCGTTTGCAAATTGAGACTAATTGCTCcgtcatcgccattaattcactgat	*leuO* deletion
T996	AGTTAAGTATGCCAGAGGTACAAACAGATC	*leuO* cloning
T997	GACCTATTCTGCAATCAGTTAGCGTT	*leuO* cloning
OA001	tcacCATATGCAGTTTAAGGTTTACACCTATAAAAGAG	*ccdB* cloning
OA002	tcacGGATCCACTGGCTGTGTATAAGGGAGCCTG	*ccdB* cloning
OA005	tcaccatATGCCAGAGGTACAAACAGATCATCC	*leuO* cloning in pET-22b(+)
OA006	tcaccatatgGCGAGCAGTGAACGTGTATTTCAT	*leuO*-EBD cloning in pET-22b(+)
OA008	tcacctcgagGCGTTTGCAAATTGAGACTAATTGC	*leuO* cloning in pET-22b(+)
OA182	TTATATGTTTTGCGATTTTTTTTGATATTG	P*leuO* EMSA fragment
OA183	ATTTAATGCATTAATTCTTAACATTAATTGATCA	P*leuO* EMSA fragment
OA184	GTAATTAATTCAATAATCACATTCACTGCAA	P*cas* EMSA fragment
OA185	CATCTTAATATATGTATAGGTTAATTGTATTAAACCAA	P*cas* EMSA fragment
OA186	CAAGTCTTCATTAgctCAGAACACTGAACATCAGCTGCG	*leuO*-N150A mutagenesis
OA187	GTTCAGTGTTCTGagcTAATGAAGACTTGAACATAACATGTATATTTGG	*leuO*-N150A mutagenesis
OA188	AGTCTTCATTAAATgcgAACACTGAACATCAGCTGCGTTATC	*leuO*-Q151A mutagenesis
OA189	TTCAGTGTTcgcATTTAATGAAGACTTGAACATAACATGTATATTT	*leuO*-Q151A mutagenesis
OA190	TCATTAAATCAGgccACTGAACATCAGCTGCGTTATCAG	*leuO*-N152A mutagenesis
OA191	ATGTTCAGTggcCTGATTTAATGAAGACTTGAACATAACATGTA	*leuO*-N152A mutagenesis
OA192	GAACACTGAAgctCAGCTGCGTTATCAGGAAACG	*leuO*-H155A mutagenesis
OA193	GATAACGCAGCTGagcTTCAGTGTTCTGATTTAATGAAGACTTGA	*leuO*-H155A mutagenesis
OA194	AACACTGAACATgcgCTGCGTTATCAGGAAACGGAGTT	*leuO*-Q156A mutagenesis
OA195	GATAACGCAGcgcATGTTCAGTGTTCTGATTTAATGAAGACTT	*leuO*-Q156A mutagenesis
OA196	AACATCAGCTGgctTATCAGGAAACGGAGTTTGTGATTAGTT	*leuO*-R158A mutagenesis
OA197	TTCCTGATAagcCAGCTGATGTTCAGTGTTCTGATTTAAT	*leuO*-R158A mutagenesis
OA198	TTATGAAGACTTCgctCGTCCTGAATTTACCAGCGTACC	*leuO*-H172A mutagenesis
OA199	ATTCAGGACGagcGAAGTCTTCATAACTAATCACAAACTCCG	*leuO*-H172A mutagenesis
OA200	TGAAGACTTCCATgctCCTGAATTTACCAGCGTACCATTATT	*leuO*-R173A mutagenesis
OA201	GTAAATTCAGGagcATGGAAGTCTTCATAACTAATCACAAACTC	*leuO*-R173A mutagenesis
OA202	GTTACTGAAAgctGATGTTTATAACGAACAACATGCGG	*leuO*-H204A mutagenesis
OA203	TTCGTTATAAACATCagcTTTCAGTAACGGGCCCTTAATTG	*leuO*-H204A mutagenesis
OA204	CGCTCGATgctTTCGCGTCATTTAGTCAACCTTG	*leuO*-R218A mutagenesis
OA205	TAAATGACGCGAAagcATCGAGCGAAACCGCCG	*leuO*-R218A mutagenesis
OA206	AGTCAACCTgcgTATGACACGGTAGATAAGCAAGCCA	*leuO*-W226A mutagenesis
OA207	CTACCGTGTCATAcgcAGGTTGACTAAATGACGCGAAACG	*leuO*-W226A mutagenesis
OA208	CGGTAGATgcgCAAGCCAGTATCGCGTATCAGG	*leuO*-K232A mutagenesis
OA209	GCGATACTGGCTTGcgcATCTACCGTGTCATACCAAGGTTGAC	*leuO*-K232A mutagenesis
OA210	CAAGTCTTCATTAgagCAGAACACTGAACATCAGCTGCG	*leuO*-N150E mutagenesis
OA211	GTTCAGTGTTCTGctcTAATGAAGACTTGAACATAACATGTATATTTGG	*leuO*-N150E mutagenesis
OA212	AGTCTTCATTAAATgagAACACTGAACATCAGCTGCGTTATC	*leuO*-Q151E mutagenesis
OA213	CTGATGTTCAGTGTTctcATTTAATGAAGACTTGAACATAACATGTATATTTG	*leuO*-Q151E mutagenesis
OA214	TTCATTAAATCAGgagACTGAACATCAGCTGCGTTATCAG	*leuO*-N152E mutagenesis
OA215	TGATGTTCAGTctcCTGATTTAATGAAGACTTGAACATAACATGT	*leuO*-N152E mutagenesis
OA216	CAGAACACTGAAgagCAGCTGCGTTATCAGGAAACG	*leuO*-H155E mutagenesis
OA217	AACGCAGCTGctcTTCAGTGTTCTGATTTAATGAAGACTTGA	*leuO*-H155E mutagenesis
OA218	TCAGAACACTGAACATgagCTGCGTTATCAGGAAACGGAGT	*leuO*-Q156E mutagenesis
OA219	ATAACGCAGctcATGTTCAGTGTTCTGATTTAATGAAGACTT	*leuO*-Q156E mutagenesis
OA220	TGAACATCAGCTGgagTATCAGGAAACGGAGTTTGTGATTAGT	*leuO*-R158E mutagenesis

Table 4: Oligonucleotides

oligo	sequence[a)]	application/target
OA221	GTTTCCTGATActcCAGCTGATGTTCAGTGTTCTGATTTAAT	*leuO*-R158E mutagenesis
OA222	TATGAAGACTTCgagCGTCCTGAATTTACCAGCGTACC	*leuO*-H172E mutagenesis
OA223	AATTCAGGACGctcGAAGTCTTCATAACTAATCACAAACTCCG	*leuO*-H172E mutagenesis
OA224	TGAAGACTTCCATgagCCTGAATTTACCAGCGTACCATTATT	*leuO*-R173E mutagenesis
OA225	GTAAATTCAGGctcATGGAAGTCTTCATAACTAATCACAAACTC	*leuO*-R173E mutagenesis
OA226	CCGTTACTGAAAgagGATGTTTATAACGAACAACATGCGG	*leuO*-H204E mutagenesis
OA227	GTTCGTTATAAACATCctcTTTCAGTAACGGGCCCTTAATTG	*leuO*-H204E mutagenesis
OA228	CGCTCGATgagTTCGCGTCATTTAGTCAACCTTG	*leuO*-R218E mutagenesis
OA229	AATGACGCGAActcATCGAGCGAAACCGCCG	*leuO*-R218E mutagenesis
OA230	ATTTAGTCAACCTgagTATGACACGGTAGATAAGCAAGCCA	*leuO*-W226E mutagenesis
OA231	CTACCGTGTCATActcAGGTTGACTAAATGACGCGAAACG	*leuO*-W226E mutagenesis
OA232	GACACGGTAGATgagCAAGCCAGTATCGCGTATCAGG	*leuO*-K232E mutagenesis
OA233	TACTGGCTTGctcATCTACCGTGTCATACCAAGGTTGAC	*leuO*-K232E mutagenesis
OA282	ATTATTAAATAAGCACATTTAATCCATTTTGTAG	P*leuO* EMSA fragment
OA283	TGTTGCGAAAACAATCTAATCATAACTAC	P*leuO* EMSA fragment
OA284	ACTTTTAGTTATAATAATTACCATGAATTTTATTACATAA	P*cas* EMSA fragment
OA285	ATTTCCCGGTATGAGATTTTATATTCAC	P*cas* EMSA fragment
OA286	AACTATCACATGAATATTATCATCATAATGAATTT	P*yjjQ* EMSA fragment
OA287	ACGTGCCCGGTTATCTATTATCCT	P*yjjQ* EMSA fragment
OA288	GCGGGAAGGGATATCATTTAATTATA	P*yjjQ* EMSA fragment
OA289	ACATTAAGTTGTTAGGAATACTTAATGAAAAACAG	P*yjjQ* EMSA fragment
OA290	AGTGAATGCTAAGGATAATTTATTCGCT	P*yjjQ* EMSA fragment
OA291	GGTTTATTCAGGATATTCATTCCTTAAATG	P*yjjQ* EMSA fragment
OA432	GTATTTCATCTTgatGTTTGCAGCCCGTTAGACAGCA	*leuO*-C117D mutagenesis
OA433	GGCTGCAAACatcAAGATGAAATACACGTTCACTGCTCG	*leuO*-C117D mutagenesis
OA434	GTATTTCATCTTagcGTTTGCAGCCCGTTAGACAGCA	*leuO*-C117S mutagenesis
OA435	GGCTGCAAACgctAAGATGAAATACACGTTCACTGCTCG	*leuO*-C117S mutagenesis
OA436	CATCTTTGTGTTgatAGCCCGTTAGACAGCATTCTGACC	*leuO*-C119D mutagenesis
OA437	TGTCTAACGGGCTatcAACACAAAGATGAAATACACGTTCACTGC	*leuO*-C119D mutagenesis
OA438	CATCTTTGTGTTagcAGCCCGTTAGACAGCATTCTGACC	*leuO*-C119S mutagenesis
OA439	TGTCTAACGGGCTgctAACACAAAGATGAAATACACGTTCACTGC	*leuO*-C119S mutagenesis
OA440	tacggctagcttaAGGCAATTCATTTTGTACTAGTTGCA	*leuO*-DBD cloning
OA473	TACGTCGGCACACTTCCGTTAT	DNase I footprint Fc, Fd
OA474	CTGCATTTCTATACGTCGGCACACTT	DNase I footprint Fa
OA475	ATGCATTTCAAATCTGCAAGTTATTCGTT	DNase I footprint Fb
OA476	TTGAAATGCATGCATTATTGTCTTTAAAC	DNase I footprint Fc
OA477	CCTACTTAAGTAGGGAAGGTGCACAATG	DNase I footprint Fa, Fb, Fd
OA514	gcaggtcgacAATGAGTTCAGAGAGCCGCAAGA	P*yafT*
OA515	gcaggctagcACAAAGCTTTTTTGAATTCATAATTGGACACTCCCTCGCCT	P*yafT*
OA516	gcaggtcgacAAGGTAATTCATGGATGGTTGAACTATATC	P*ybdO*
OA517	gcagtctagaTAAGTCGAACTTTTTCAAGTCGTAGAGATT	P*ybdO*
OA518	gcaggtcgacTTGCGACTCCATGTCCATATAATTTAT	P*ybeQ*
OA519	gcagtctagaGCAGCAACTTGACGTGAAAATCATTATC	P*ybeQ*
OA520	gcaggtcgacGCAGCAACTTGACGTGAAAATCATTATC	P*ybeR*
OA521	gcagtctagaTTGCGACTCCATGTCCATATAATTTAT	P*ybeR*
OA522	gcaggtcgacTAATGGAGGTGTTATTGACTGCATGA	P*yghS*
OA523	gcagtctagaCATATTACGTTGATTCACGAAAAACCCGGCA	P*yghS*
OA524	gcaggtcgacCATATTACGTTGATTCACGAAAAACCCGGCA	P*yghT*
OA525	gcagtctagaTAATGGAGGTGTTATTGACTGCATGA	P*yghT*
OA526	gcaggtcgacCCTGGCATGTTTCGTCATTACTATTCCTCAA	P*envR*
OA527	gcagtctagaAGCTTCGGCTTTGGTTCTTTTTGCCATGA	P*envR*
OA528	gcaggtcgacAGCTTCGGCTTTGGTTCTTTTTGCCATGA	P*envC*
OA529	gcagtctagaCCTGGCATGTTTCGTCATTACTATTCCTCAA	P*envC*
OA530	gcaggtcgacGTGCGCCAGCGCAGCTAATTTCTCAGAATTATG	P*yjfI*

Table 4: Oligonucleotides

oligo	sequence[a)]	application/target
OA531	gcagtctagaTAGCGCCAACGGATTCCATGTCATA	P*yjfI*
OA576	Phos-agcttGCAGCAG	adaptor pKES334
OA577	Phos-ctagcTGCTGCA	adaptor pKES334

[a)] Nucleotides homologous to the indicated template are shown in capital letters, non-matching parts are shown in lower case letters, and sites for restriction endonucleases are underlined.

4.4 $CaCl_2$-competent cells and transformation

For preparation of $CaCl_2$-competent cells, 25 ml LB medium was inoculated with a fresh single colony and grown to OD_{600} of 0.3. Cultures were harvested on ice and pelleted by centrifugation at 4°C for 10 min at 3000 rpm. The cell pellet was resuspended in 12.5 ml ice-cold 0.1 M $CaCl_2$, incubated for 20 min on ice and pelleted again by centrifugation. The cell pellet was resuspended in 1 ml 0.1 M $CaCl_2$. Cells were used directly for transformation, or for long-term storage at -80°C, glycerol was added (15% final concentration) and incubated for 1 hour. For transformation 10 µl of a ligation or 1 to 100 ng of plasmid DNA was prepared in 50 µl TEN-buffer (20 mM Tris-HCl pH 7.5, 1 mM EDTA, 50 mM NaCl), mixed with 100 µl of competent cells and incubated for 10 min on ice. The cells were incubated for exactly 2 min at 42°C (heat shock) and again for 10 min on ice. The cells were transferred to culture tubes with 1 ml LB medium added and incubated for 1 h at 37°C or appropriate temperature. 100 µl of the culture was plated on LB plates with appropriate antibiotics.

4.5 Electrocompetent cells and electroporation

For preparation of electrocompetent cells 50 ml SOB medium was inoculated from an overnight culture and grown to OD_{600} 0.6. The culture was incubated on ice for 1 h and pelleted by centrifugation at 4°C for 15 min at 3000 rpm. The cells were resuspended in 50 ml ice-cold sterile H_2O, pelleted by centrifugation and resuspended in 25 ml ice-cold sterile H_2O. The cells were again pelleted and resuspended in 2 ml ice-cold 10% glycerol and finally the cell pellet was resuspended in 200 µl ice-cold 10% glycerol. The cells were used directly for electroporation, or frozen at -80°C for long-term storage. For electroporation 1 µl of DNA (e.g. 100 ng PCR fragment for λ-Red recombineering) was mixed with 40 µl of electrocompetent cells. The samples were incubated for 10 min on ice and transferred to a precooled electroporation cuvette (Bio-Rad). Electroporation was carried out with a Bio-Rad Gene Pulser with 1.8 kV for 3 ms. The cells were transferred to 1 ml SOC medium and incubated at 37°C for 1 hr. 100 µl of the culture was plated on LB plates with appropriate antibiotics.

SOB medium (for 1 l): 20 g Bacto Tryptone, 5 g Bacto Yeast Extract, 5 g NaCl, 15 g Bacto Agar

SOC medium: 19.8 ml of 20% glucose added to 1 l of SOB

4.6 Random mutagenesis screen

The random mutagenesis screen was performed to identify LeuO mutants that are constitutively active. For mutagenesis the *leuO* gene fragment was amplified in 25 parallel reactions using a non-proofreading Tag polymerase (Promega). In brief, *leuO* was amplified with oligos S326 and S328 and pKETS25 as a template. The *leuO* gene fragments were purified and digested with restriction endonucleases NheI and EcoRI, and ligated into pKESK22. Transformants of the *Pcas-lacZ* reporter strain (T1281 or T1610) with these ligations were plated on Tryptone X-Gal kanamycin indicator plates and screened for a Lac-positive phenotype. Clones with a Lac-positive phenotype were analyzed by sequencing.

4.7 Site-directed mutagenesis by overlap PCR

Site-directed mutagenesis was used for mutation of the *leuO* gene and carried out by a two step overlap PCR. In the first step, two PCRs were carried out from pKETS25 as a template. For amplification of PCR A the forward primer S326 and the reverse primer carrying the desired mutation was used (see Table 4). In PCR B the forward primer, carrying the desired mutation (corresponding to mutation in PCR A) and the reverse primer S328 was used. The DNA fragments from PCR A and B were gel purified and used as the template in the second step for PCR C with primer S326 and S328. The fragments were then digested with restriction endonucleases NheI and EcoRI and ligated into the vector pKESK22. DH5α cells were transformed with the ligation and the correct mutation was verified by PCR and sequencing.

4.8 Chromosomal integration into *attB* sites

Promoter *lacZ* fusions were integrated into the chromosomal λ attachment site *attB,* of derivates of the Δ*lacZ* strain (S4197) as described (Diederich *et al.*, 1992, Dole *et al.*, 2002). In brief, the target strain was transformed with the integrase-expressing helper plasmid pLDR8 and transformants were selected on LB kanamycin plates at 28°C. Overnight cultures of transformants were grown at 28°C and diluted twenty-fold and grown for 90 min at 37°C to induce integrase expression. The cultures were used to prepare chemically competent cells, as described above. Plasmids carrying the promoter *lacZ* fusion, the *attP* site and a spectinomycin resistance cassette were digested with BamHI. Fragments were purified by gel electrophoresis and 10 ng of the fragments self-ligated. Compentent cells were transformed with 10 µl of this ligation and selected on prewarmed LB spectinomycin plates at 42°C. At 42°C, the integrase promotes the recombination between *attP* and *attB* sites resulting in the integration of the *lacZ* promoter fusion into the chromosome, while simultanously the pLDR8 plasmid is cured at high temperature. The colonies were verified by PCR

analysis using primer pairs S93/S164 (analysis of attB/P' site), S95/T912 (analysis of attP/B' site), S95/S164 (exclusion of dimers), and T334/S118 (analysis of reporter construct) and were tested for kanamycin sensitivity (loss of pLDR8).

4.9 Gene deletion and insertion using λ-Red mediated recombination

Gene deletion of chromosomal genes was performed as described (Datsenko & Wanner, 2000). This method is based on homologous recombination between linear DNA fragments and the chromosomal locus mediated by the λ-Red recombination system. In brief, the chromosomal gene sequence is replaced with a selectable antibiotic resistance gene (chloramphenicol or kanamycin) flanked by FRT sites (Flp Recombination Target). The linear DNA fragments are generated by PCR, using oligonucleotides with a 35 to 50 nucleotide homology to the target chromosomal locus and pKD3 (cmR) or pKD4 (kanR) as a template. Gel-purified PCR fragments (100 ng/µl in H_2O) were electroporated into cells harbouring the helper plasmid pKD46 expressing λ-Red recombinase. Electrocompetent cells were prepared from cultures grown at 28°C in SOB medium supplemented with 10 mM L-arabinose for induction of λ-Red recombinase.

The recombinants were selected at 37°C on LB plates, supplemented with chloramphenicol or kanamycin. The insertion of the resistance cassette at the target gene location was confirmed by PCR. The loss of the helper plasmid was confirmed by ampicillin sensitivity. The FRT sites flanking the resistance cassette allow the deletion of the resistance gene by expressing the Flp recombinase from the helper plasmid pCP20. The respective strain was transformed with pCP20 and selected on LB amp plates at 28°C. Transformants were restreaked on LB plates and incubated at 42°C to induce expression of the Flp recombinase and loss of the plasmid. The loss of the antibiotic resistance casette was confirmed by PCR and antibiotic sensitivity of the clones. The loss of the plasmid pCP20 was confirmed by antibiotic sensitivity of the clones to ampicillin.

4.10 Chromosomal two-step mutagenesis using λ-Red mediated recombination

For introducing point-mutations of *leuO* into the wild-type chromosomal background (BW30270) a two-step method based on the λ-Red recombinase system, was carried out according to (Datsenko & Wanner, 2000) and (Kolmsee & Hengge, 2011). In a first step, the target gene is replaced by a DNA fragment encoding for kanamycin resistance and a *ccdB* toxin with a rhamnose-inducible promoter. In a second step the *leuO*::*kanR-ccdB* cassette is replaced by the mutated *leuO* (*leuO*-S120D, *leuO*-M244T, *leuO*-S120D/M244T) using again the λ-Red system and counter-selected on LB plates supplemented with 12 mM rhamnose. In brief, a fragment of plasmid pKILL45 (carrying the kanamycin resistance cassette and the *ccdB* gene under the control of rhaP3 was amplified by PCR

using oligonucleotides (T994/T995) that carry extensions homologous to *leuO*. The target strain was transformed with the helper plasmid pKD46 expressing λ-Red recombinase and cells were made electrocompetent according to the protocol described above. The gel-purified PCR fragment was electroporated into the cells, and recombinants were selected on LB plates with kanamycin at 37°C. The insertion of the *kanR-ccdB* gene was confirmed by PCR and sequencing. For the second step, this intermediate strain was transformed with pKD46 and electrocompetent cells were prepared. PCR fragments of *leuO*, carrying the mutations (*leuO*-S120D, *leuO*-M244T, *leuO*-S120D/M244T) were amplified from plasmids (pKESL104, pKESL73 and pKESL107, respectively) using oligonucleotides T996 and T997. The intermediate strain *leuO::kanR-ccdB* was transformed with pKD46 and cells were made electrocompetent (chapter 4.5). The gel-purified PCR fragments carrying the *leuO* mutations were used for electroporation into the cells and recombinants were selected on LB plates with rhamnose at 37°C. The induction of toxin *ccdB* with rhamnose ensures the survival of recombinants carrying the *leuO* mutation and kills cells carrying the *ccdB-kanR* casette. The loss of the *ccdB-kanR* casette was analyzed by PCR and sequencing. In addition, sensitivity of cells to kanamycin was confirmed.

4.11 Phage transduction

The phage transduction method is based on the ability of phage T4GT7 to transfer DNA between bacteria (Wilson *et al.*, 1979). In brief, for T4GT7-lysate 100 µl of an overnight culture of the donor strain was incubated with different concentrations of a wild-type T4GT7-lysate for 15 min. Then 1 ml of LB medium was added and the mixture was transferred to a culture tube with 3 ml T4-Topagar. The lysate-mix was poured on a fresh LB plate and incubated for 8 to 14 hours. Plates that showed a confluent lysis were used to isolate the phages by chloroform extraction. For transduction, 100 µl of an overnight culture of the recipient strain was incubated with different concentrations of the bacteriophage lysate for 15 min and plated on LB plates with appropriate antibiotics for selection. The transductants were restreaked at least four times to get rid of contaminating T4GT7 phages. The transfer of the allele (and other relevant loci) was verified by PCR with specific primers.

4.12 Expression analysis

For expression analyses of promoter *lacZ* fusions β-galactosidase assays were performed, as described (Miller, 1992). Briefly, exponential cultures were inoculated from a fresh overnight culture to an OD_{600} 0.1 in LB medium that was supplemented with the specific antibiotic in case of transformants. Where indicated, IPTG was added for induction both to the overnight and the exponential culture to a final concentration of 1 mM. Cultures were grown at 37°C to OD_{600} 0.5 and

then harvested on ice. The β-galactosidase assays were repeated at least three times from independent biological replicates.

4.13 Protein purification of LeuO and LeuO mutants

For purification of C-terminally histidine-tagged $LeuO_{His6}$, $LeuO\text{-}S120D_{His6}$ and $LeuO\text{-}DBD_{His6}$, expression strain BL21(DE3)C41 was transformed with plasmids carrying the respective encoding genes under the control of a *T7* promoter (Table 3). Two liter cultures of LB medium with ampicillin were inoculated to an OD_{600} of 0.1 and grown at 37°C to OD_{600} of 0.6. At this point the protein expression was induced by adding IPTG to a final concentration of 200 µM and the cultures were grown for further 5 h at 28°C. The cultures were harvested and the bacteria pelleted by centrifugation. The cells were resuspended in resuspension buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl), and again 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 cells; 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM Imidazole, 20 µg/ml DNase I (5000 U/ml; New England Biolabs, USA)] and lysed by sonication (40% amplitude, 4 minutes with a 2 seconds pulse/ 2 seconds pause; Sonics Vibracell VCX750 High-Volume Ultrasonic Cell Disrupter). The lysate was cleared by centrifugation (35.000 rpm, 45 min, Ti70 rotor, Beckman Coulter XL70) and the supernatant was loaded onto 1 ml HisTrap HP column (GE Healthcare, Germany) using an ÄKTA fast protein liquid chromatography (FPLC) system (GE Healthcare, Germany). The column was washed with the same buffer (20 mM Tris-HCl pH 8.0, 300 mM NaCl) with increasing imidazole concentrations (10 mM, 20 mM, 40 mM, and 60 mM), with each step equal to at least 10 column volumes. The proteins were eluted with 20 mM Tris-HCl pH 8.0, 300 mM NaCl, 250 mM Imidazole, the 1 ml fractions were collected and analyzed by SDS-PAGE. The protein containing fractions were pooled and the buffer was exchanged with a PD-10 desalting column (GE Healthcare, Germany) to storage buffer (for $LeuO\text{-}DBD_{His6}$: 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 50 mM NDSB-256; for $LeuO_{His6}$ and $LeuO\text{-}S120D_{His6}$: 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT 150 mM NDSB-256). Protein concentrations were measured by Qubit Fluorometric Quantitation system (Invitrogen, Germany). Aliquots of the proteins were stored at -80°C.

4.14 SDS-PAGE

Purified proteins were analyzed by SDS polyacrylamide gel electrophoresis (SDS-PAGE), as described before (Ausubel, 2005). Protein samples were resuspended in 1x SDS-loading buffer [50 mM Tris-HCl (pH 6.8), 1.5% SDS, 5% glycerol, 0.05% bromophenol blue, 125 mM DTT] and boiled for 5 minutes at 95°C and 10 µl were loaded on 12% SDS-polyacrylamide gels containing 0.005% TCE (2,2,2-Trichloroethanol). Gels were run at 100 V and 40 mA for 20 min and at 150 V for approximately 1.5

hours. The gels were analyzed in a fast approach by detecting the flourescence of the protein bands, caused by the ultraviolet light-induced reaction of the the tryptophans in the protein with trihalocompounds of the TCE. In addition, the gels were stained with Coomassie Brilliant blue.

4.15 Crystallization of LeuO-EBD and LeuO-S120D-EBD

For crystallization, C-terminally histidine-tagged LeuO effector-binding domains LeuO-EBD_{His6} and LeuO-S120D-EBD_{His6} were purified as described above for the full-length LeuO protein. Immediately after NiNTA affinity chromatography, the protein containing fractions were pooled and concentrated to 2 mg/ml protein using Vivaspin 20 (10,000 MWCO; Sartorius, Germany). Then the proteins were subjected to size exclusion chromatography (Superdex 200 16/600; GE Healthcare, Germany) at 4°C in 20 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM DTT, 150 mM NDSB-256 using an ÄKTA purifier FPLC (GE Healthcare, Germany). The eluted fractions were analyzed by SDS-PAGE, and fractions containing pure protein were pooled and concentrated to 12.5 mg/ml.

Both proteins were crystallized using the sitting-drop vapor diffusion method. LeuO-S120D-EBD_{His6} protein (12.5 mg/ml) was mixed with precipitant in a ratio 2:1 to form 300 nl drops and incubated at 20°C. Crystals grew in a condition containing 0.1 M Tris, 1.2 M Na/K tartrate pH 8.0. Crystals with dimensions of about 80 x 50 x 50 μm^3 were cryoprotected in 0.1 M Tris, 1.2 M Na/K tartrate pH 8.0, 30% sucrose and flash-cooled in liquid nitrogen.

LeuO-EBD_{His6} protein (12.5 mg/ml) was mixed with precipitant in a ratio 1:2 to form 300 nl drops and incubated at 20°C. Crystals grew in a condition containing 0.2 M Lithium sulfate, 0.1 M Bis-Tris pH 5.5, 25% polyethylene glycol 3350. Crystals with dimensions of 90 x 90 x 50 μm^3 were cryoprotected in 0.2 M Lithium sulfate, 0.1 M Bis-Tris pH 5.5, 32% polyethylene glycol 3350 and flash-cooled in liquid nitrogen.

The X-ray data collection, the structure determination and refinement was done by Anna Montada, Dr. Magdalena Schacherl and Prof. Dr. Ulrich Baumann, from the Institute for Biochemistry, Cologne.

4.16 Electrophoretic shift assay (EMSA)

DNA fragments used for EMSA were amplified by PCR. The concentrations were determined by adsorption measurements using Nanodrop (Thermo scientific, Germany) or Qubit Fluorometric Quantitation system (Invitrogen, Germany) instruments and by comparison of band intensities in agarose gels. Binding of LeuO to the DNA fragments was carried out in a final volume of 10 µl using 10 ng of the 80 bp fragments (final concentration 20 nM). All dilutions of proteins and DNA fragments were set up in binding buffer (20 mM Tris HCl pH 7.5, 100 mM KCl, 2 mM DTT, 10%

glycerol). Protein was added in the indicated concentrations and after protein addition the samples were incubated at 30°C for 20 minutes. Samples were loaded on 8% native polyacrylamide gels (acrylamide:bisacrylamide 29:1; in running buffer 44.5 mM Tris, 44.5 mM Boric acid, 1 mM EDTA pH 8.0). Gels were run at 200 V for 1 h at 4°C and then stained with ethidium bromide (0.5 µg/ml, in running buffer) for 20 minutes.

4.17 DNase I footprinting

DNase I footprinting was carried out as described (Gaugué *et al.*, 2013), with minor modifications. Briefly, 75 pmol of the oligonucleotides were labeled with 75 µCi [γ-^{32}P]ATP (3000 Ci/mmol, Hartmann Analytics, Germany) in the presence of 20 units of T4 polynucleotide kinase (Thermo Fischer Scientific, Germany) in 50 µl T4 polynucleotide kinase buffer for 30 minutes at 37°C. Unincorporated nucleotides were removed from the labeled oligonucleotides using illustra ProbeQuant G50 Micro columns (GE Healthcare, Germany) that were pre-buffered with PCR buffer. Then, DNA fragments were generated by standard PCR in 50 µl with GoTaq polymerase (Promega, Germany) using 20 pmol of the [^{32}P]-labeled oligonucleotide and a non-labeled oligonucleotide as reverse primer. The DNA fragment was purified by agarose gel electrophoresis and eluted with a gel purification kit (Machery Nagel, Germany). The concentration of the eluted fragment was measured using a Qubit Fluorometric Quantitation system (Invitrogen, Germany), and the counts per minute (cpm) were determined by Cerenkov counting. For each footprint reaction 120000 cpm (approximately 40 ng DNA) were used. Binding reactions were carried out for 10 minutes at 30 °C in 40 µl of binding buffer (25 mM HEPES pH 8.0, 100 mM potassium glutamate pH 8.0, 0.5 mg/ml BSA) with increasing concentration of the protein, wild-type $LeuO_{His6}$, LeuO DNA-binding domain (LeuO-DBD_{His6}), and LeuO-S120D_{His6}. Then, 0.02 units (in 4 µl) of DNase I were added (Thermo Fischer Scientific, Germany; diluted 200-fold in 10 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 125 mM KCl, 0.1 mM DTT). Samples were incubated for 60 sec at 30°C and DNase I digestion was stopped by adding 100 µl Tris-HCl phenol pH 8.0 and 200 µl stop buffer (0.5 M Na-acetate pH 5.0, 10 µg/ml herring sperm DNA (Ultra Pure, Thermo Fischer Scientific, Germany), 2.5 mM EDTA) followed by a phenol extraction and ethanol precipitation. The DNA was dried *in vacuo* and resuspended in 5 µl H_2O, followed by addition of 6 µl loading dye (95% formamide, 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol FF, 0.025% ethidium bromide, 0.5 mM EDTA). The samples were heated for 2 min at 90°C and loaded on a 6% denaturing sequencing gel (6% long ranger (Lonza by Biozym Scientific, Germany), 7 M urea, 72 mM Tris-HCl, 72 mM boric acid, 1.6 mM EDTA) next to a sanger sequencing ladder, which was generated using the labeled oligonucleotide and the T7 polymerase sequencing kit (USB corporation, USA). The gel was washed with 10% ethanol, 10% acetic acid, transferred to Whatman 3 MM paper and dried in a gel dryer. Imaging was carried out by exposure

to phosphor-imaging plates and scanning by a Typhoon 7000 imaging system (GE Healthcare, Germany).

4.18 Identification of a consensus LeuO DNA-binding site

The DNase I footprint of LeuO DNA-binding site II at the *cas* promoter region was screened by eye and a 19 bp palindromic sequence of two 7 bp half-sites separated by 5 bp was detected. A similar palindromic sequence was found within DNA-binding site LeuO I (Stratmann *et al.*, 2012) of the *leuO* regulatory region. Then, two 33 bp sequences covering the palindromic motifs of LeuO site II at the *cas* locus and LeuO site I at the *leuO* locus (Figure 18A) were submitted to MEME Suite (Bailey & Elkan, 1994) to generate a motif, which was subsequently submitted to FIMO (Grant *et al.*, 2011) to search the *E. coli* K 12 genome (substr MG1655 uid225) for putative LeuO DNA-binding sites. The 13 sequences with top scores (33.3 - 20.6; see 1. score in Table 5) were filtered regarding their intergenic position as well as their identification in a microarray (Stratmann *et al.*, 2012), and a genomic SELEX screening (Shimada *et al.*, 2011). Six of the 13 putative LeuO binding sequences fulfilled all criteria (see Table 5). In a second motif analysis, 33 bp sequences covering the six putative LeuO binding sites were submitted to MEME Suite to generate a less stringent motif (Figure 18B).

When this motif was submitted to FIMO to search the *E. coli* K12 MG1655 genome for putative LeuO sites 37 LeuO DNA-binding sequences with the best motif scores (cut off value 14.0; see 2. score in Table 5) were identified. These were compared with previous microarray and SELEX data (Stratmann *et al.*, 2012, Shimada *et al.*, 2011), which yielded a striking correlation (see Table 5). A similar motif was obtained with less stringent conditions of filtering the initial 13 sequences.

Table 5: LeuO DNA-binding sites[a)]

center[a]	str.[b]	matched sequence	First scr.[c]	2nd scr.[d]	flanking gene	genome position			MA[e]	flanking gene	genome position			MA[e]	SL[f]	function[g]
4410096	+	ATATTCATTTTATTAATAT	20.8	32.3	*rlmB*	4409275	->	4410006		*yjfI*	4410133	->	4410534	38.7	53.8	yjfI: conserved protein
676686	+	ATATTCATTTAATGAATAT	21.7	31.6	*ybeQ*	675570	<-	676547	43.0	*ybeR*	676711	->	677418	20.1	9.5	ybeQ: conserved protein, ybeR: predicted polypeptide
84024	+	ATATTGATTTGGTGAATAT	33.3	31.1	*leuL*	83622	<-	83708	nd	*leuO*	84368	->	85312	nd	14.7	leuO: LysR-type transcription regulator
2884429	+	ATATTCATACTGTGAATAT	33.3	30.4	*casA*	2882630	<-	2884138	65.4	*cas3*	2884553	<-	2887219	16.3	77.2	cas: CRISPR associated, bacterial defense system
3134088	-	ATATTTATTTGGTTAATAT	21.7	30.2	*yghS*	3133244	<-	3133957	28.0	*yghT*	3134131	->	3134823	12.4	54.7	yghS, yghT: putative proteins with nucleoside triphosphate hydrolase domain
3360901	-	ATATTCATATTTTGTATAT	20.7	28.9	*gltD*	3359198	->	3360616		*gltF*	3361176	->	3361940	56.1	66.8	gltF: periplasmic protein
2388299	+	ATATTGATTTGGTCAATAT	27.0	23.8	*yfbN*	2387710	<-	2388426		*yfbO*	2388635	->	2389057	9.3	14.4	yfbO: putative protein
4591516	+	ATATTTATATTGTGAATAA	21.7	22.9	*yjiY*	4589129	<-	4591279		*tsr*	4591657	->	4593312		37.3	yjiY: inner membrane protein - putative transporter, tsr: methyl-accepting chemotaxis protein - serine-sensing
2899081	-	ATATTGATATTGAGAATAT	27.0	22.8	*ygcS*	2896533	<-	2897870		*ygcU*	2897964	<-	2899418			intragenic
2099760	+	ATATTTATTAAGTTAATAT		21.6	*ugd*	2098447	<-	2099613		*gnd*	2099862	<-	2101268			ugd: UDP-glucose 6-dehydrogenase
1546229	+	AAATTTATTTTTTTAATAT		21.4	*yddL*	1545738	<-	1546028		*yddG*	1546288	<-	1547169			yddL: putative lipoprotein
234112	-	ATATTGATTTAATTAATGT		20.8	*mltD*	232597	<-	233955		*gloB*	234027	<-	234782			mltD: membrane-bound lytic murein transglycosylase D
3799266	+	ATCTTGATTTTGTTTATAT		20.6	*waaU*	3798239	<-	3799312		*waaZ*	3799345	<-	3800196			intragenic
1990189	-	AGATTCACTTTGTGAATAT	20.7	17.9	*yecH*	1989251	<-	1989490		*tyrP*	1989681	->	1990892			intragenic
3136570	-	ATATTTATTCATTTCATAT		17.8	*pitB*	3134872	<-	3136371	31.9	*gss*	3136663	<-	3138522		51.8	pitB: H+ symporter
3770144	-	ATATTCATTATGTTAATAA		17.2	*yibV*	3769345	->	3769680		*yibU*	3769948	->	3770146		7.3	intragenic
1349241	+	ATATTCATTTGATGAATCC		16.7	*rnb*	1346978	<-	1348912		*yciW*	1348980	<-	1350107			rnb: exoribonuclease II
3412572	+	TTATTCATTTCGTGTATAT		16.6	*yhdJ*	3411653	->	3412537		*yhdU*	3412621	->	3412800			yhdU: hypothetical protein
237216	-	ACATTCATTTAATCAATAT		16.4	*aspV*	236931	->	237007		*yafT*	237335	->	238120	7.8	5.9	yafT: predicted lipoprotein, putative aminopeptidase
729459	-	ATATATATTTCATGAATAT		16.4	*ybfA*	729134	->	729340		*rhsC*	729583	->	733776		4.0	rhsC: RhsC protein in rhs element
3267484	+	ACATTTATTTTATCAATAT		16.0	*tdcA*	3266127	<-	3267065		*tdcR*	3267380	->	3267598		12.6	tdcA: LysR-type transcription regulator of metabolism during anaerobiosis
578981	+	AATTTCATATTGTTAATAT		15.9	*borD*	578600	<-	578893		*ybcV*	579184	<-	579594			borD (ybcU): lipoprotein bor homolog from lambdoid prophage DLP12
1371792	-	ATATTTACATGCTGAATAT		15.4	*ycjM*	1370216	->	1371895		*ycjN*	1371909	->	1373201			ycjN: putative periplasmic binding protein ABC family
157242	-	TTATTTATATTGCGAATAT		15.0	*yadN*	156299	<-	156883	13.6	*folK*	157253	<-	157732		7.2	yadN: cryptic fimbrial shaft protein
576887	-	ACATTGATTAGATGAATAT		15.0	*nmpC*	575758	<-	576825	-12.2	*essD*	577398	->	577613		8.3	nmpC: general bacterial porin
2884273	+	ATATTCATTGGTTTAATAC		15.0	*casA*	2882630	<-	2884138	65.4	*cas3*	2884553	<-	2887219	16.3	77.2	cas: CRISPR associated
1872765	-	AGATTGTTTTTGTTAATAT		14.9	*dgcJ*	1872041	->	1873531		*yeaK*	1873574	->	1874077			yeaK: mischarged aminoacyl-tRNA deacylase
3413622	-	ATATTGATGTAGTGAATGT		14.9	*envR*	3412803	<-	3413465	48.3	*envC*	3413864	->	3415021	21.0	81.3	envR (acrS): DNA-binding transcriptional repressor, envC (acrE): lipoprotein, drug efflux pump
3772124	+	ATATTTATGTGATTGATAT		14.7	*yibI*	3771382	<-	3771744		*mtlA*	3772281	->	3774194			mtlA: mannitol-specific PTS enzyme II
4004830	-	TTATTGATAAGGTGAATAT	21.6	14.6	*yigI*	4004230	<-	4004697		*pldA*	4004862	->	4005731	-4.8		yigI: putative thioesterase, pldA: outer membrane phospholipase
637660	-	ACATTGATTAAGTGAATAT		14.6	*ybdO*	636716	<-	637618	23.0	*dsbG*	637827	<-	638573		48.3	putative LysR-type transcription regulator, unknown function
2575953	-	ATATTCATGCGTTGCATAT		14.2	*eutS*	2575470	<-	2575805		*maeB*	2576098	<-	2578377			eutS: predicted structural protein, ethanolamine utilization
3793735	-	GTATACATTCTATTAATAT		14.1	*htrL*	3792826	<-	3793683		*rfaD*	3793987	->	3794919		6.7	htrL (yibB): unknown; rfaD (waaD): LPS biosynthesis

Table 5: LeuO DNA-binding sites[a)]

center[a]	str.[b]	matched sequence	First scr.[c]	2nd scr.[d]	flanking gene	genome position			MA[e]	flanking gene	genome position			MA[e]	SL[f]	function[g]
156276	+	ATATTCATTCAATCAATTT		14.0	*yadV*	155461	<-	156201		*yadN*	156299	<-	156883	13.6		yadV: cryptic pilin chaperone
3288596	-	ATATTCAGTTGGGTTATAT		14.0	*yraI*	3288090	->	3288785		*yraJ*	3288814	->	3291330			yraJ: putative fimbrial usher protein
3640765	-	ATATTGATGCGGTGAATAG	21.6	13.5	*uspA*	3640111	->	3640545		*dtpB*	3640862	->	3642331	-12.2		dtpB: dipeptide/tripeptide: H+ symporter
3302287	+	ATATTGCTTCGGTGAAAAT	20.6	12.9	*yhbU*	3301485	->	3302480		*yhbV*	3302489	->	3303367			yhbV: putative peptidase

[a)] The center refers to the genome position of the matching motif sequence.

[b)] refers to the DNA strand.

[c)] The score refers to the motif score identified with FIMO (Grant *et al.*, 2011); the regulation of the target in the microarray (MA) (Stratmann *et al.*, 2012) and the binding by LeuO in SELEX (SL) (Shimada *et al.*, 2011). The first motif score (First score) was obtained by submitting the motif depicted in the top panel of Figure 6 to FIMO to search the *E. coli* K 12 genome for putative LeuO DNA-binding sites. The 13 sequences displayed scores of 33.3 to 20.6 regarding the sequence identity to the given motif. In a second motif analysis (2nd score; cut off value 14.0) the LeuO binding motif depicted in the bottom panel of Figure 6 was submitted to FIMO, identifying the 37 putative LeuO sites shown here.

[d)] The second score refers to the second motif analysis.

[e)] Targets were identified in a microarray transcriptome analysis and the given values refer to the fold change regulation (Stratmann *et al.*, 2012).

[f)] Targets were identified by genome scale SELEX screening and the given values refer to the peak height (Shimada *et al.*, 2011).

[g)] Descriptions of protein functions were taken from the EcoCyc Database (Karp *et al.*, 2014).

```
A) MEME_1
>84024          TTTTTTGATATTGATTTGGTGAATATTATTGAT
>2884429        ACATAAAATATTCATACTGTGAATATAAAATCT

B) MEME_2
>84024          TTTTTTGATATTGATTTGGTGAATATTATTGAT
>2884429        ACATAAAATATTCATACTGTGAATATAAAATCT
>676686         GCGGCTAATATTCATTTAATGAATATTTAAGGA
>3134088        TAAAAACATATTAACCAAATAAATATTTTTAAT
>4410096        ATTGCAAATATTCATTTTATTAATATTTAAACT
>3360901        GATATTAATATACAAAATATGAATATAAAAAAC
```

Figure 18: FASTA files used for motif generation

The motif was generated using MEME Suite (Bailey & Elkan, 1994).

5 References

Ante, V.M., X.R. Bina & J.E. Bina, (2015a) The LysR-type regulator LeuO regulates the acid tolerance response in *Vibrio cholerae*. *Microbiology.* **161**: 2434-2443.

Ante, V.M., X.R. Bina, M.F. Howard, S. Sayeed, D.L. Taylor & J.E. Bina, (2015b) *Vibrio cholerae leuO* transcription is positively regulated by ToxR and contributes to bile resistance. *J. Bacteriol.* **197**: 3499-3510.

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

Ausubel, F.M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith & K. Struhl, (2005) *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc.

Ayala, J.C., H. Wang, J.A. Benitez & A.J. Silva, (2018) Molecular basis for the differential expression of the global regulator VieA in *Vibrio cholerae* biotypes directed by H-NS, LeuO and quorum sensing. *Mol. Microbiol.* **107**: 330-343.

Bailey, T.L. & C. Elkan, (1994) *Fitting a mixture model by expectation maximization to discover motifs in biopolymers,* p. pp. 28-36.

Balcewich, M.D., T.M. Reeve, E.A. Orlikow, L.J. Donald, D.J. Vocadlo & B.L. Mark, (2010) Crystal Structure of the AmpR Effector Binding Domain Provides Insight into the Molecular Regulation of Inducible AmpC β-Lactamase. *J Mol Biol.* **400**: 998-1010.

Barrangou, R., C. Fremaux, H. Deveau, M. Richards, P. Boyaval, S. Moineau, D.A. Romero & P. Horvath, (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709-1712.

Bina, X.R., D.L. Taylor, A. Vikram, V.M. Ante & J.E. Bina, (2013) *Vibrio cholerae* ToxR downregulates virulence factor production in response to Cyclo(Phe-Pro). *mBio* **4**.

Boivin, S., S. Kozak & R. Meijers, (2013) Optimization of protein purification and characterization using Thermofluor screens. *Protein Expression and Purification* **91**: 192-206.

Breddermann, H. & K. Schnetz, (2016) Correlation of antagonistic regulation of *leuO* transcription with the cellular levels of BglJ-RcsB and LeuO in *Escherichia coli*. *Front. Cell. Infect. Microbiol.* **6**.

Breddermann, H. & K. Schnetz, (2017) Activation of *leuO* by LrhA in *Escherichia coli*. *Mol. Microbiol.* **104**: 664-676.

Browning, D.F. & S.J.W. Busby, (2016) Local and global regulation of transcription initiation in bacteria. *Nat Rev Microbiol* **14**: 638.

Bundy, B.M., L.S. Collier, T.R. Hoover & E.L. Neidle, (2002) Synergistic transcriptional activation by one regulatory protein in response to two metabolites. *Proc. Natl. Acad. Sci. USA.* **99**: 7693-7698.

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

Chen, S., A. Zhang, L.B. Blyn & G. Storz, (2004) MicC, a second small-RNA regulator of Omp Protein expression in *Escherichia coli*. *J. Bacteriol.* **186**: 6689-6697.

Cherepanov, P.P. & W. Wackernagel, (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.

Colyer, T.E. & N.M. Kredich, (1994) Residue threonine-149 of the *Salmonella typhimurium* CysB transcription activator: mutations causing constitutive expression of positively regulated genes of the cysteine regulon. *Mol. Microbiol.* **13**: 797-805.

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

De la Cruz, M.Á., M. Fernández-Mora, C. Guadarrama, M.A. Flores-Valdez, V.H. Bustamante, A. Vázquez & E. Calva, (2007) LeuO antagonizes H-NS and StpA-dependent repression in *Salmonella enterica ompS1*. *Mol. Microbiol.* **66**: 727-743.

Devesse, L., I. Smirnova, R. Lonneborg, U. Kapp, P. Brzezinski, G.A. Leonard & C. Dian, (2011) Crystal structures of DntR inducer binding domains in complex with salicylate offer insights into the activation of LysR-type transcriptional regulators. *Mol. Microbiol.* **81**: 354-367.

Diederich, L., L.J. Rasmussen & W. Messer, (1992) New cloning vectors for integration in the lambda attachment site attB of the *Escherichia coli* chromosome. *Plasmid* **28**: 14-24.

Dillon, S.C., E. Espinosa, K. Hokamp, D.W. Ussery, J. Casadesús & C.J. Dorman, (2012) LeuO is a global regulator of gene expression in *Salmonella enterica* serovar Typhimurium. *Mol. Microbiol.*: no-no.

Dole, S., S. Kühn & K. Schnetz, (2002) Post-transcriptional enhancement of *Escherichia coli bgl* operon silencing by limitation of BglG-mediated antitermination at low transcription rates. *Mol. Microbiol.* **43**: 217-226.

Dover, N. & E. Padan, (2001) Transcription of *nhaA*, the main Na+/H+ antiporter of *Escherichia coli*, is regulated by Na+ and growth phase. *J. Bacteriol.* **183**: 644-653.

Espinosa, E. & J. Casadesús, (2014) Regulation of *Salmonella enterica* pathogenicity island 1 (SPI-1) by the LysR-type regulator LeuO. *Mol. Microbiol.* **91**: 1057-1069.

Ezezika, O.C., S. Haddad, T.J. Clark, E.L. Neidle & C. Momany, (2007) Distinct effector-binding sites enable synergistic transcriptional activation by BenM, a LysR-type regulator. *J Mol Biol.* **367**: 616-629.

Fàbrega, A., J.L. Rosner, R.G. Martin, M. Solé & J. Vila, (2012) SoxS-dependent coregulation of *ompN* and *ydbK* in a multidrug-resistant *Escherichia coli* strain. *FEMS. Microbiol. Lett.* **332**: 61-67.

Fang, M., A. Majumder, K.-J. Tsai & H.-Y. Wu, (2000) ppGpp-dependent *leuO* expression in bacteria under stress. *Biochem. Biophys. Res. Commun.* **276**: 64-70.

Francetic, O., C. Badaut, S. Rimsky & A.P. Pugsley, (2000a) The ChiA (YheB) protein of *Escherichia col*i K-12 is an endochitinase whose gene is negatively controlled by the nucleoid-structuring protein H-NS. *Mol. Microbiol.* **35**: 1506-1517.

Francetic, O., D. Belin, C. Badaut & A.P. Pugsley, (2000b) Expression of the endogenous type II secretion pathway in *Escherichia coli* leads to chitinase secretion. *EMBO. J.* **19**: 6697-6703.

García-Lara, J., L.H. Shang & L.I. Rothfield, (1996) An extracellular factor regulates expression of *sdiA*, a transcriptional activator of cell division genes in *Escherichia coli*. *J. Bacteriol.* **178**: 2742-2748.

Gaugué, I., D. Bréchemier-Baey & J. Plumbridge, (2013) DNase I footprinting to identify protein binding sites. *Bio-protocol* **3**.

Gemmill, R.M., J.W. Jones, G.W. Haughn & J.M. Calvo, (1983) Transcription initiation sites of the leucine operons of *Salmonella typhimurium* and *Escherichia coli*. *J Mol Biol.* **170**: 39-59.

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

Guadarrama, C., A. Medrano-López, R. Oropeza, I. Hernández-Lucas & E. Calva, (2014a) The *Salmonella enterica* Serovar Typhi LeuO global regulator forms tetramers: Residues involved in oligomerization, DNA binding, and transcriptional regulation. *J. Bacteriol.* **196**: 2143-2154.

Guadarrama, C., T. Villasenor & E. Calva, (2014b) The subtleties and contrasts of the LeuO Regulator in *Salmonella* Typhi: implications in the immune response. *Front. Immunol.* **5**: 581.

Henikoff, S., G.W. Haughn, J.M. Calvo & J.C. Wallace, (1988) A large family of bacterial activator proteins. *Proc. Natl. Acad. Sci. USA.* **85**: 6602-6606.

Hernández-Lucas, I. & E. Calva, (2012) The coming of age of the LeuO regulator. *Mol. Microbiol.* **85**: 1026-1028.

Higashi, K., T. Tobe, A. Kanai, E. Uyar, S. Ishikawa, Y. Suzuki, N. Ogasawara, K. Kurokawa & T. Oshima, (2016) H-NS facilitates sequence diversification of horizontally transferred DNAs during their integration in host chromosomes. *PLoS Genet* **12**: e1005796.

Hirakawa, H., A. Takumi-Kobayashi, U. Theisen, T. Hirata, K. Nishino & A. Yamaguchi, (2008) AcrS/EnvR represses expression of the *acrAB* multidrug efflux genes in *Escherichia coli*. *J. Bacteriol.* **190**: 6276-6279.

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

Jiang, Y.-L., X.-P. Wang, H. Sun, S.-J. Han, W.-F. Li, N. Cui, G.-M. Lin, J.-Y. Zhang, W. Cheng, D.-D. Cao, Z.-Y. Zhang, C.-C. Zhang, Y. Chen & C.-Z. Zhou, (2018) Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR. *Proc. Natl. Acad. Sci. USA.* **115**: 403-408.

Jo, I., I.-Y. Chung, H.-W. Bae, J.-S. Kim, S. Song, Y.-H. Cho & N.-C. Ha, (2015) Structural details of the OxyR peroxide-sensing mechanism. *Proc. Natl. Acad. Sci. USA.* **112**: 6443-6448.

Jones, R.M., D.L. Popham, A.L. Schmidt, E.L. Neidle & E.V. Stabb, (2018) *Vibrio fischeri* DarR directs responses to d-aspartate and represents a group of similar LysR-type transcriptional regulators. *J. Bacteriol.*

Jørgensen, C. & G. Dandanell, (1999) Isolation and characterization of mutations in the *Escherichia coli* regulatory protein XapR. *J. Bacteriol.* **181**: 4397-4403.

Juncker, A.S., H. Willenbrock, G. von Heijne, S. Brunak, H. Nielsen & A. Krogh, (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. *Protein Sci.* **12**: 1652-1662.

Karp, P., D. Weaver, S. Paley, C. Fulcher, A. Kubo, A. Kothari, M. Krummenacker, P. Subhraveti, D. Weerasinghe, S. Gama-Castro, A. Huerta, L. Muñiz-Rascado, C. Bonavides-Martinez, V. Weiss, M. Peralta-Gil, A. Santos-Zavaleta, I. Schröder, A. Mackie, R. Gunsalus, J. Collado-Vides, I. Keseler & I. Paulsen, (2014) The EcoCyc Database. *EcoSal Plus*.

Klauck, E., J. Böhringer & R. Hengge-Aronis, (1997) The LysR-like regulator LeuO in *Escherichia coli* is involved in the translational regulation of *rpoS* by affecting the expression of the small regulatory DsrA-RNA. *Mol. Microbiol.* **25**: 559-569.

Kobayashi, K., N. Tsukagoshi & R. Aono, (2001) Suppression of hypersensitivity of *Escherichia coli acrB* mutant to organic solvents by integrational activation of the *acrEF* operon with the IS1 or IS2 Element. *J. Bacteriol.* **183**: 2646-2653.

Kolmsee, T. & R. Hengge, (2011) Rare codons play a positive role in the expression of the stationary phase sigma factor RpoS (sigmaS) in *Escherichia coli*. *RNA Biol* **8**: 913-921.

Krisko, A., T. Copic, T. Gabaldón, B. Lehner & F. Supek, (2014) Inferring gene function from evolutionary change in signatures of translation efficiency. *Genome Biology* **15**: R44.

Krissinel, E. & K. Henrick, (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol.* **372**: 774-797.

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

Kvint, K., L. Nachin, A. Diez & T. Nyström, (2003) The bacterial universal stress protein: function and regulation. *Curr. Opin. Microbiol.* **6**: 140-145.

Laishram, R.S. & J. Gowrishankar, (2007) Environmental regulation operating at the promoter clearance step of bacterial transcription. *Genes Dev.* **21**: 1258-1272.

Lee, D.J., S.D. Minchin & S.J.W. Busby, (2012) Activating transcription in bacteria. *Annu. Rev. Microbiol.* **66**: 125-152.

Lerche, M., C. Dian, A. Round, R. Lönneborg, P. Brzezinski & G.A. Leonard, (2016) The solution configurations of inactive and activated DntR have implications for the sliding dimer mechanism of LysR transcription factors. *Scientific Reports* **6**: 19988.

Liu, Z., M. Yang, G.L. Peterfreund, A.M. Tsou, N. Selamoglu, F. Daldal, Z. Zhong, B. Kan & J. Zhu, (2011) *Vibrio cholerae* anaerobic induction of virulence gene expression is controlled by thiol-based switches of virulence regulator AphB. *Proc. Natl. Acad. Sci. USA.* **108**: 810-815.

Lomize, A.L., M.A. Lomize, S.R. Krolicki & I.D. Pogozheva, (2017) Membranome: a database for proteome-wide analysis of single-pass membrane proteins. *Nucleic. Acids. Res.* **45**: D250-D255.

Lunin, V.V., C. Chang, T. Skarina, E. Gorodischenskaya, A.M. Edwards, A. Joachimiak & A. Savchenko, (2005) The crystal structure of putative transcriptional regulator Pa0477 from *Pseudomonas aeruginosa*. In. https://www.rcsb.org/structure/2esn, pp.

Maddocks, S.E. & P.C.F. Oyston, (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology.* **154**: 3609-3623.

Madhusudan, S., A. Paukner, Y. Klingen & K. Schnetz, (2005) Independent regulation of H-NS-mediated silencing of the *bgl* operon at two levels: upstream by BglJ and LeuO and downstream by DnaKJ. *Microbiology.* **151**: 3349-3359.

Majumder, A., M. Fang, K.-J. Tsai, C. Ueguchi, T. Mizuno & H.-Y. Wu, (2001) LeuO expression in response to starvation for branched-chain amino acids. *J. Biol. Chem.* **276**: 19046-19051.

Medina-Aparicio, L., J.E. Rebollar-Flores, A.L. Gallego-Hernández, A. Vázquez, L. Olvera, R.M. Gutiérrez-Ríos, E. Calva & I. Hernández-Lucas, (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.

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.

Miroux, B. & J.E. Walker, (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.

Momany, C. & E.L. Neidle, (2012) Defying stereotypes: the elusive search for a universal model of LysR-type regulation. *Mol. Microbiol.* **83**: 453-456.

Monferrer, D., T. Tralau, M.A. Kertesz, I. Dix, M. Sola & I. Uson, (2010) Structural studies on the full-length LysR-type regulator TsaR from *Comamonas testosteroni* T-2 reveal a novel open conformation of the tetrameric LTTR fold. *Mol. Microbiol.* **75**: 1199-1214.

Mosca, R. & T.R. Schneider, (2008) RAPIDO: a web server for the alignment of protein structures in the presence of conformational changes. *Nucleic. Acids. Res.* **36**: W42-W46.

Nguyen Le Minh, P., C. Velázquez Ruiz, S. Vandermeeren, P. Abwoyo, I. Bervoets & D. Charlier, (2018) Differential protein-DNA contacts for activation and repression by ArgP, a LysR-type (LTTR) transcriptional regulator in *Escherichia coli*. *Microbiological Research* **206**: 141-158.

Pannen, D., M. Fabisch, L. Gausling & K. Schnetz, (2016) Interaction of the RcsB Response Regulator with Auxiliary Transcription Regulators in *Escherichia coli*. *J. Biol. Chem.* **291**: 2357-2370.

Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng & T.E. Ferrin, (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605-1612.

Picossi, S., B.R. Belitsky & A.L. Sonenshein, (2007) Molecular mechanism of the regulation of *Bacillus subtilis gltAB* expression by GltC. *J Mol Biol.* **365**: 1298-1313.

Pul, Ü., R. Wurm, Z. Arslan, R. Geißen, N. Hofmann & R. Wagner, (2010) Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **75**: 1495-1512.

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

Ruangprasert, A., S.H. Craven, E.L. Neidle & C. Momany, (2010) Full-length structures of BenM and two variants reveal different oligomerization schemes for LysR-type transcriptional regulators. *J Mol Biol.* **404**: 568-586.

Salmon, K., C. Yang & G. Hatfield, (2006) Biosynthesis and regulation of the branched-chain amino acids. *EcoSal Plus*.

Salscheider, S.L., A. Jahn & K. Schnetz, (2014) Transcriptional regulation by BglJ–RcsB, a pleiotropic heteromeric activator in *Escherichia coli*. *Nucleic. Acids. Res.* **42**: 2999-3008.

Schell, M.A., (1993) Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**: 597-626.

Schell, M.A., P.H. Brown & S. Raju, (1990) Use of saturation mutagenesis to localize probable functional domains in the NahR protein, a LysR-type transcription activator. *J. Biol. Chem.* **265**: 3844-3850.

Schubert, S.M., (2013) Modulation of the activity of LeuO, a transcriptional regulator in *Escherichia coli*. In: Institute for Genetics, Cologne, Germany. University of Cologne, pp.

Shimada, T., A. Bridier, R. Briandet & A. Ishihama, (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.

Stoebel, D.M., A. Free & C.J. Dorman, (2008) Anti-silencing: overcoming H-NS-mediated repression of transcription in Gram-negative enteric bacteria. *Microbiology.* **154**: 2533-2545.

Stratmann, T., S. Madhusudan & K. Schnetz, (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.

Stratmann, T., Ü. Pul, R. Wurm, R. Wagner & K. Schnetz, (2012) RcsB-BglJ activates the *Escherichia coli leuO* gene, encoding an H-NS antagonist and pleiotropic regulator of virulence determinants. *Mol. Microbiol.* **83**: 1109-1123.

Taylor, J.L., R.S. De Silva, G. Kovacikova, W. Lin, R.K. Taylor, K. Skorupski & F.J. Kull, (2012) The crystal structure of AphB, a virulence gene activator from *Vibrio cholerae*, reveals residues that influence its response to oxygen and pH. *Mol. Microbiol.* **83**: 457-470.

The PyMOL Molecular Graphics System, V.S., LLC. . In., pp.

The UniProt Consortium, (2017) UniProt: the universal protein knowledgebase. *Nucleic. Acids. Res.* **45**: D158-D169.

Toledano, M.B., I. Kullik, F. Trinh, P.T. Baird, T.D. Schneider & G. Storz, (1994) Redox-dependent shift of OxyR-DNA contacts along an extended DNA-binding site: A mechanism for differential promoter selection. *Cell* **78**: 897-909.

Tropel, D. & J.R. van der Meer, (2004) Bacterial transcriptional regulators for degradation pathways of aromatic compounds. *Microbiology and Molecular Biology Reviews* **68**: 474-500.

Ueguchi, C., T. Ohta, C. Seto, T. Suzuki & T. Mizuno, (1998) The *leuO* gene product has a latent ability to relieve *bgl* silencing in *Escherichia coli*. *J. Bacteriol.* **180**: 190-193.

van Keulen, G., L. Girbal, E.R.E. van den Bergh, L. Dijkhuizen & W.G. Meijer, (1998) The LysR-type transcriptional regulator CbbR controlling autotrophic CO2 fixation by *Xanthobacter flavus* is an NADPH sensor. *J. Bacteriol.* **180**: 1411-1417.

Venkatesh, G.R., F.C. Kembou Koungni, A. Paukner, T. Stratmann, B. Blissenbach & K. Schnetz, (2010) BglJ-RcsB heterodimers relieve repression of the *Escherichia coli bgl* operon by H-NS. *J. Bacteriol.* **192**: 6456-6464.

Wang, Q. & J.M. Calvo, (1993) Lrp, a global regulatory protein of *Escherichia coli*, binds co-operatively to multiple sites and activates transcription of *ilvIH*. *J Mol Biol.* **229**: 306-318.

Wek, R.C. & G.W. Hatfield, (1988) Transcriptional activation at adjacent operators in the divergent-overlapping *ilvY* and *ilvC* promoters of *Escherichia coli*. *J Mol Biol.* **203**: 643-663.

Westra, E.R., Ü. Pul, N. Heidrich, M.M. Jore, M. Lundgren, T. Stratmann, R. Wurm, A. Raine, M. Mescher, L. Van Heereveld, M. Mastop, E.G.H. Wagner, K. Schnetz, J. Van Der Oost, R. Wagner & S.J.J. Brouns, (2010) H-NS-mediated repression of CRISPR-based immunity in *Escherichia coli* K12 can be relieved by the transcription activator LeuO. *Mol. Microbiol.* **77**: 1380-1393.

Westra, E.R., D.C. Swarts, R.H.J. Staals, M.M. Jore, S.J.J. Brouns & J. van der Oost, (2012) The CRISPRs, they are A-Changin': How prokaryotes generate adaptive immunity. *Annu. Rev. Genet.* **46**: 311-339.

Wiebe, H., D. Gurlebeck, J. Gross, K. Dreck, D. Pannen, C. Ewers, L.H. Wieler & K. Schnetz, (2015) YjjQ represses transcription of *flhDC* and additional loci in *Escherichia coli*. *J. Bacteriol.* **197**: 2713-2720.

Wilson, G.G., K.Y. Young, G.J. Edlin & W. Konigsberg, (1979) High-frequency generalised transduction by bacteriophage T4. *Nature* **280**: 80-82.

Zhai, Y. & M.H. Saier, (2002) The β-barrel finder (BBF) program, allowing identification of outer membrane β-barrel proteins encoded within prokaryotic genomes. *Protein Sci.* **11**: 2196-2207.

Danksagung

An erster Stelle möchte ich mich herzlich bei Prof. Dr. Karin Schnetz bedanken, dass ich meine Doktorarbeit in ihrer Arbeitsgruppe schreiben konnte, für das interessante Projekt und die hervorragende Betreuung in der ganzen Zeit.

Prof. Dr. Niels Gehring danke ich für die Übernahme des Zweitgutachtens und die Mitwirkung in meinem *Thesis Committee*.

Besonders danken möchte ich Prof. Dr. Ulrich Baumann für die fachliche Unterstützung bei der Lösung der Proteinstruktur von LeuO und für die Mitwirkung im *Thesis Committee*. Ich bedanke mich außerdem bei der gesamten AG Baumann, insbesondere bei Anna Montada und Dr. Magdalena Schacherl für die nette Zusammenarbeit und für die Unterstützung.

Ich danke den Kollegen aus der AG Gehring, AG Leptin, AG Riemer und AG Zuccaro für die tollen Karnevals- und Weihnachtsfeiern, die netten Gespräche beim Montagskuchen, und das gute Klima im Laboralltag.

Ich bedanke mich bei allen aktuellen und ehemaligen Kollegen der Arbeitsgruppe: Aathmaja, Cihan, Derk, Hannes, Katrin und allen Studenten. Vielen Dank für die gute Zeit, die Unterstützung im Labor und das ein oder andere Mauerbier.

Ein besonderer Dank gilt meiner Mutter, die mich immer motiviert hat mein Bestes zu geben. Ich danke meinem Bruder und seiner Familie für die willkommenen Ablenkungen und die Erinnerung daran, dass Arbeit nicht alles ist. Ich danke Janine und Boris für ihren emotionalen Rückhalt in der stressigen Schlussphase. Zum Schluss aber nicht zuletzt, möchte ich mich bei meinem Mann Tom für seine unermüdliche Unterstützung in der ganzen Zeit bedanken.

Erklärung

Ich versichere, dass ich die von mir abgegebene Dissertation selbstständig angefertigt habe, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von der unten angegebenen Teilpublikation - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.

Die Bestimmungen der Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Karin Schnetz, Institut für Genetik, betreut worden.

Wesentliche Teile dieser Dissertation werden zur Publikation in einer wissenschaftlichen Zeitung eingereicht.

Teilpublikation im Rahmen dieser Arbeit:

Susann M Fragel, Anna Montada, Ralf Heermann, Ulrich Baumann, Magdalena Schacherl, and Karin Schnetz

Functional characterization of the pleiotropic LysR-type transcription regulator LeuO of *Escherichia coli*

Köln, den 07. September 2018 _____

Susann Marlies Fragel