Insights into the molecular signal perception mechanism of the membrane-integrated transcriptional activator CadC of *Escherichia coli*

Dissertation

der Fakultät für Biologie

der Ludwig-Maximilians-Universität München

vorgelegt von

Ina Maria Suntka Haneburger

aus München

München

27. Oktober 2011

GUTACHTER

1. Prof. Dr. Kirsten Jung
2. Prof. Dr. Dirk Schüler
3. Prof. Dr. Barbara Conradt
4. Prof. Dr. Martin Parniske
5. Prof. Dr. Ute Vothknecht
6. Prof. Dr. Thorsten Mascher

TAG DER MÜNDLICHEN PRÜFUNG: 19. DEZEMBER 2011

EHRENWÖRTLICHE VERSICHERUNG

Hiermit versichere ich ehrenwörtlich, dass die vorliegende Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt wurde. Des Weiteren erkläre ich, dass ich nicht anderweitig ohne Erfolg versucht habe, eine Dissertation einzureichen oder mich der Doktorprüfung zu unterziehen. Die vorliegende Dissertation liegt weder ganz, noch in wesentlichen Teilen einer anderen Prüfungskommission vor.

München, Ina Haneburger

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources. As well I declare, that I have not submitted a dissertation without success and not passed the oral exam. The present dissertation (neither the entire dissertation nor parts) has not been presented to another examination board.

Munich, Ina Haneburger

Contents

Nomenclature

Gene products of *cadC* are numbered in a way that the first methionine of the wild type protein is designated "1" in the amino acid sequence. N-terminal affinity tags are left unconsidered.

All full-length CadC derivatives produced in this study exhibit an N-terminal deca-histidine (His) tag attached to the N-terminus of the protein via a nine amino acid spacer (SSGHIEGRH). The complete notation His_{10}-CadC was reduced to the designation CadC by means of shortness. The same elision is done on the level of genes.

The periplasmic domain of CadC was produced as a hybrid protein with thioredoxin. After thrombin cleavage, the resulting protein encompasses 32 amino acids (GSGMKETAAAKFERQHMDSPDLGTDDDDKAMA) in front of a hexa-His tag N-terminally fused to amino acids 188 - 512 of the periplasmic domain of CadC. This protein construct is designated CadC188-512 or $CadC_{pd}$.

Amino acid substitutions in CadC are termed as follows: The native amino acid is designated in one-letter code, followed by the respective amino acid position in the protein. The amino acid introduced by (site-directed) mutagenesis is terminally added in one-letter code (Example: CadC_D471N).

Unless otherwise noted, nucleotide positions indicate the distance from the transcriptional start site (+1).

Abbreviations

σ^s	stationary phase-specific sigma factor (*rpoS*)
σ^{70}	primary ("housekeeping") sigma factor (*rpoD*)
AAA^+	ATPases associated with various cellular activities
AAT	amino acid transporter family of transport proteins
APC	amino acid/polyamine/organocation superfamily of transport proteins
AR	acid resistance
ASP	acid shock protein
ATP	adenosine-5'-triphosphate
$CadC_{pd}$	periplasmic (sensor) domain of CadC (encompassing residues 188 - 512)
cAMP	cyclic adenosine-5'-monophosphate
CM	cytoplasmic membrane
CP	cytoplasm
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
EMSA	electrophoretic mobility shift assay
EPR	electron paramagnetic resonance
GABA	γ-amino butyric acid
His_n tag	affinity tag composed of n histidine residues
LB	Luria Bertani medium
MAD	multiple wavelength anomalous dispersion
Ni-NTA	nickel-charged resin (Ni^{2+}-nitrilotriacetic acid)
PAGE	polyacrylamide gel electrophoresis
PDB	Protein Data Bank (www.pdb.org)
PEG	polyethylene glycol
pmf	proton motive force
PP	periplasm
ppGpp	guanosine tetraphosphate
TCS	two-component signal transduction system
TM	transmembrane domain
$T\text{-}N_{11}\text{-}A$	conserved promoter motif, where T and A are spaced by 11 nucleotides

Publications and Manuscripts Originating from this Thesis

CHAPTER 2
Fritz, G., Koller, C., Burdack, K., Tetsch, L., Haneburger, I., Jung, K. and Gerland, U. (2009), Induction kinetics of a conditional pH stress response system in *Escherichia coli*. *J. Mol. Biol.* 2: 272-286.

CHAPTER 3
Eichinger, A., Haneburger, I., Koller, C., Jung, K. and Skerra, A. (2011), Crystal structure of the sensory domain of *Escherichia coli* CadC, a member of the ToxR-like protein family. *Prot. Sci.* 20: 656-669.

CHAPTER 4
Haneburger, I., Eichinger, A., Skerra, A., and Jung, K. (2011), New Insights into the signaling mechanism of the pH-responsive, membrane-integrated transcriptional activator CadC of *Escherichia coli*. *J. Biol. Chem.* 286: 10681-10689.

CHAPTER 5
Haneburger, I., Fritz, G., Jurkschat, N., Tetsch, L., Eichinger, A., Skerra, A., Gerland, U. and Jung, K. (2011), The feedback-inhibitor cadaverine suppresses pH response by binding to the pH susceptible site of CadC in *Escherichia coli*. *Manuscript.*

CHAPTER 6
Ruiz, J., Haneburger, I. and Jung, K. (2011), Identification of ArgP and Lrp as transcriptional regulators of *lysP*, the gene encoding the specific lysine permease of *Escherichia coli*. *J. Bacteriol.* 193: 2536-2548.

Contributions to Publications and Manuscripts Presented in this Thesis

CHAPTER 2

G. Fritz, C. Koller, K. Burdack, L. Tetsch, I. Haneburger, K. Jung and U. Gerland designed the concept of the study. G. Fritz performed mathematical modeling. C. Koller and K. Burdack performed the transient expression dynamic experiments. L. Tetsch measured lysine-dependent dose-reponse and I. Haneburger pH-dependent dose-response. G. Fritz, C. Koller, K. Burdack, L. Tetsch, I. Haneburger, K. Jung and U. Gerland wrote the manuscript and discussed the results with the other co-authors.

CHAPTER 3

A. Eichinger, I. Haneburger, C. Koller, K. Jung and A. Skerra developed the concept of the study. C. Koller established the purification protocol. I. Haneburger overproduced and purified the wild type protein and two $CadC_{pd}$ variants for crystallization. A. Eichinger set up crystallization conditions and performed x-ray and data analyses. I. Haneburger performed the initial crystallization screen for the two variants. A. Eichinger, I. Haneburger, C. Koller, A. Skerra and K. Jung discussed the results and wrote the manuscript.

CHAPTER 4

I. Haneburger, A. Skerra and K. Jung developed the concept of the study. I. Haneburger constructed the variants and analyzed the pH dependent *cadBA* expression and membrane integration. I. Haneburger, A. Eichinger, A. Skerra and K. Jung wrote the manuscript.

CHAPTER 5

I. Haneburger, G. Fritz and K. Jung designed the study. I. Haneburger and G. Fritz contributed equally to this work. I. Haneburger conducted the experiments performed by N. Jurkschat, performed affinity measurements of $CadC_{pd}$_T229A_E447Q and constructed and analyzed variants with exchanges in the binding site close to the acidic patch. G. Fritz performed mathematical modeling. K. Burdack and I. Haneburger performed transient expression analysis. N. Jurkschat constructed and analyzed CadC variants with exchanges in the internal cavity and CadC_Y453I and CadC_T475A. I. Haneburger, G. Fritz, N. Jurkschat, L. Tetsch, A. Eichinger, A. Skerra, U. Gerland and K. Jung wrote the manuscript.

CHAPTER 6

J. Ruiz, I. Haneburger and K. Jung developed the concept of the study. J. Ruiz analyzed the lysine dependent *lysP*-expression, identified ArgP and Lrp and determined DNA-binding affinity and binding area via EMSA and DNase footprinting analyses. I. Haneburger analyzed the influence of basic amino acids on *lysP* expression and the importance of the T-N_{11}-A motif *in vivo*. J. Ruiz, I. Haneburger and K. Jung wrote the manuscript.

I hereby confirm the above statements:

Ina Haneburger Prof. Dr. Kirsten Jung

Summary

Escherichia coli is a neutrophilic bacterium that can survive drastic changes in the external pH (pH 5 - 9) while sustaining the intracellular pH at pH 7.4 - 7.8. Signal transduction systems that detect external signals are vital to the maintenance of the intracellular pH as they induce appropriate cellular responses. Among the different response systems, manifold prototypes can be found. The simplest conceivable way of transducing information from the exterior to the interior of a cell is to sense information outside (in the periplasmic domain) and to transduce the information over the membrane (transmembrane domain) into the cytoplasm to the cytoplasmic effector domain in one single polypeptide. The ToxR-like regulator CadC of *E. coli* is such a one-component system and activates transcription of an acid stress response system (*cadBA* operon) at acidic pH when lysine is available. CadC senses and integrates two stimulating and one inhibiting signal and regulates *cadBA* transcription in response to these signals (lysine, low pH, cadaverine). During the course of this thesis the so-far-unknown molecular mechanisms of signal perception by CadC have been investigated.

First, transcriptional analyses of the *cadBA* operon were performed at high temporal resolution. Together with the determination of the amount of cadaverine and the CadA activity of the same samples, these data allowed the compilation of a mathematical model. With the help of the model the cellular response for a *lysP* mutant was predicted. Furthermore, the model helped to ascertain that the transient *cadBA* expression can be ascribed to cadaverine-mediated feedback inhibition.

In line with the goals of this thesis, the first three dimensional structure out of the group of ToxR-like transcriptional activators, the crystal structure of the periplasmic (sensor) domain of CadC, was solved at 1.8 Å resolution. In addition to the wild type domain, the structures of two CadC variants were determined. The periplasmic domain possesses a so far unknown protein fold; only the C-terminal subdomain shows similarities to helical bundles of known three dimensional structures. For the first time, dimerization of CadC was proven within the crystals.

With the help of the crystal structure and an elaborate mutagenesis study, a patch of acidic amino acids (Asp_198, Asp_200, Glu_461, Glu_468 and Asp_471) was identified that is essential for detection of acidification of the surrounding milieu. Remarkably, these acidic residues are located at the CadC dimer interface. By virtue of localization of these amino acids, it is likely that upon acidification the residues get protonated (neutralized) and hence repelling surfaces between the monomers are omitted. In consequence, CadC becomes active and induces transcription of the *cadBA* operon. Additionally, the importance of the first periplasmic subdomain of CadC for

signal transduction was demonstrated.

The mechanism of feedback inhibition was further elucidated with the finding that cadaverine likely binds two distinct binding sites, leading to inactivation of CadC. With the help of the mathematical model the number of binding sites (four within a dimer) and cooperative binding were established.

Finally, the transcriptional fine-tuning of *lysP* (the gene encoding the lysine permease, LysP, associated with the Cad system) and *cadBA* was analyzed. This analysis resulted in the identification of ArgP as *lysP* activator. Furthermore, the global regulator, Lrp, was found to exert a stimulating effect on *lysP* and *cadBA* expression.

Altogether, the data, gathered in the course of this thesis, complement the knowledge about signal-dependent regulation of the Cad system with molecular insights and allowed compilation of an updated model for this *E. coli* acid tolerance system. In addition, these data help to understand how multiple signals can be externally sensed and integrated to generate an appropriate intracellular response by a single polypeptide.

Zusammenfassung

Das neutrophile Bakterium *Escherichia coli* kann drastische Veränderungen des externen pH-Werts (pH 5 - 9) unter Beibehaltung eines intrazellulären pH-Werts von 7,4 - 7,8 überleben. Signaltransduktionssysteme, die externe Signale wahrnehmen, sind essentiell für die Aufrechterhaltung des intrazellulären pH-Werts, da sie die entsprechenden zellulären Antworten induzieren. Unter den verschiedenen Antwortsystemen existieren viele Varianten. Der einfachste vorstellbare Weg, Information von Außen ins Zellinnere zu übertragen, ist die Information mittels *eines* Proteins außen wahrzunehmen (in der periplasmatischen Domäne) und über die Membran (Transmembrandomäne) in das Zytoplasma zur Effektordomäne weiterzuleiten. Der ToxR-ähnliche Regulator CadC von *E. coli* ist ein solches Ein-Komponenten System und aktiviert die Transkription eines Säurestress-Systems (*cadBA* Operon) bei saurem pH und in Gegenwart von Lysin. CadC sensiert und integriert zwei stimulierende und ein inhibierendes Signal und reguliert die *cadBA* Transkription als Antwort auf diese Signale (Lysin, niedriger pH-Wert, Cadaverin). Im Rahmen dieser Doktorarbeit wurden die - bis dahin unbekannten - molekularen Mechanismen der Signalwahrnehmung durch CadC untersucht.
Zunächst wurden zeitlich hochaufgelöste Transkriptionsuntersuchungen des *cadBA* Operons durchgeführt. Gemeinsam mit der Bestimmung der Cadaverin-Konzentration und der Enzymaktivität in den gleichen Proben, ermöglichten diese Daten die Erstellung eines mathematischen Modells. Mit Hilfe diesen Modells konnte die zelluläre Antwort einer *lysP* Mutante vorhergesagt werden. Außerdem konnte mit diesem Modell die transiente *cadBA* Expression auf die Cadaverin-vermittelte "Feedback"-Inhibierung zurückgeführt werden.
Im Rahmen dieser Arbeit wurde die Kristallstruktur der periplasmatischen (Sensor-) Domäne von CadC, die erste dreidimensionale Struktur aus der Gruppe der ToxR-ähnlichen Regulatoren mit einer Auflösung von 1.8 Å aufgelöst. Zusätzlich zur periplasmatischen Domäne des Wildtyp CadCs wurden die Strukturen der periplasmatischen Domänen von zwei CadC Varianten aufgeklärt. Die periplasmatische Domäne besitzt einen bisher unbekannten Faltungstyp, von dem nur die C-terminale Subdomäne Ähnlichkeiten zu helikalen Bündeln von bekannten dreidimensionalen Strukturen besitzt. Zum ersten Mal konnte eine Dimerisierung von CadC in den Kristallen nachgewiesen werden.
Mit Hilfe der Kristallstruktur und einer ausführlichen Mutagenesestudie, wurde ein Bereich mit sauren Aminosäuren (Asp_198, Asp_200, Glu_461, Glu_468 and Asp_471) identifiziert, der essentiell für die Wahrnehmung der Ansäuerung des umgebenden Milieus ist. Bemerkenswerterweise liegen diese Aminosäuren im Bereich der Grenzfläche

zwischen den Monomeren im CadC Dimer. Aufgrund der Lage der Aminosäuren kommt es bei einer Ansäuerung wahrscheinlich zur Protonierung der sauren Seitenketten (Neutralisierung) und dadurch zur Aufhebung von abstoßenden Kräften zwischen den CadC Monomeren. In der Konsequenz wird CadC aktiv und induziert die Transkription des *cadBA* Operons. Zusätzlich konnte gezeigt werden, daß für die Reizweiterleitung die erste periplasmatische Subdomäne von CadC essentiell ist.
Der Mechanismus der *Feedback*-Inhibierung konnte weiter aufgeklärt werden: Cadaverin bindet dabei wahrscheinlich an zwei unterschiedliche Bindestellen, wodurch es zur Inaktivierung von CadC kommt. Mit Hilfe eines mathematischen Modells konnte die Zahl der Bindestellen (vier in einem Dimer) sowie eine kooperative Bindung nachgewiesen werden.
Abschließend wurde noch die transkriptionelle Fein-Regulation von *lysP* (kodierend für eine mit dem Cad System assoziierte Lysin Permease) und von *cadBA* untersucht. Dabei wurde ArgP als Aktivator für *lysP* identifiziert. Für den globalen Regulator Lrp wurde eine stimulierende Wirkung auf die *cadBA* und *lysP* Expression nachgewiesen.
Alles in allem ergänzen die Ergebnisse dieser Arbeit das Wissen zur signalabhängigen Regulation des Cad Systems mit molekularen Details und ermöglichen das Erstellen eines aktualisierten Aktivierungsmodells für dieses Säuretoleranz-System von *E. coli*. Ferner helfen die Daten dieser Arbeit zu verstehen, wie in einem Polypeptid mehrere externe Signale aufgenommen und zusammengefasst werden und eine angemessene intrazelluläre Antwort generiert wird.

1 Introduction

1.1 Bacteria and Acid Stress

All terrestrials face varying environmental conditions such as day and night, the rise and drop in temperature, and periods of rain and drought. As a result, all organisms have developed some sort of mechanism in order to deal with their environments. Especially bacteria that belong to the smallest living organisms have to adapt to the slightest changes. As bacteria are single cell organisms, alterations in the environment have direct effects on the cell and life-sustaining processes such as metabolism, transcription and translation. In order to survive life-threatening stresses, bacteria depend on the ability to sense and respond to changes in their surroundings [Boor, 2006].
Within their natural habitats, bacteria encounter a broad range of environmental stresses. Enterobacteria, like *Escherichia coli* and *Salmonella typhimurium* frequently cope with hostile conditions such as acidic stress during passage into and out of the host [Bearson *et al.*, 1997]. Within a host, acid stress is comprised of low pH and the presence of weak (organic) acids. At low pH these weak acids, including the volatile fatty acid products of fermentation, exist in an uncharged, protonated state and are thus able to diffuse through the membrane into the cytoplasm of the bacteria. In the neutral cytoplasm, the fatty acids dissociate and thereby lower the internal pH [Bearson *et al.*, 1997]. As neutrophilic enterobacterium, *E. coli* depends on a neutral environment to survive and grow over the long term. Therefore, maintaining a neutral pH within the bacterium's cytoplasm is crucial for the stability and function of macromolecules and for the perpetuation of the proton motive force (pmf), essential to ATP generation, locomotion and nutrient transport [Zilberstein *et al.*, 1984]. For example, *E. coli* maintains its cytoplasmic pH within pH 7.4 - 7.8 over an external pH range of 5.0 - 9.0 [Slonczewski *et al.*, 2009]. Beside passive protective mechanisms such as the impermeability of the cytoplasmic membrane for charged compounds as well as the buffering capacity of the cytoplasm, enterobacteria possess different mechanisms to survive these unfavorable conditions (see figure 1.1).

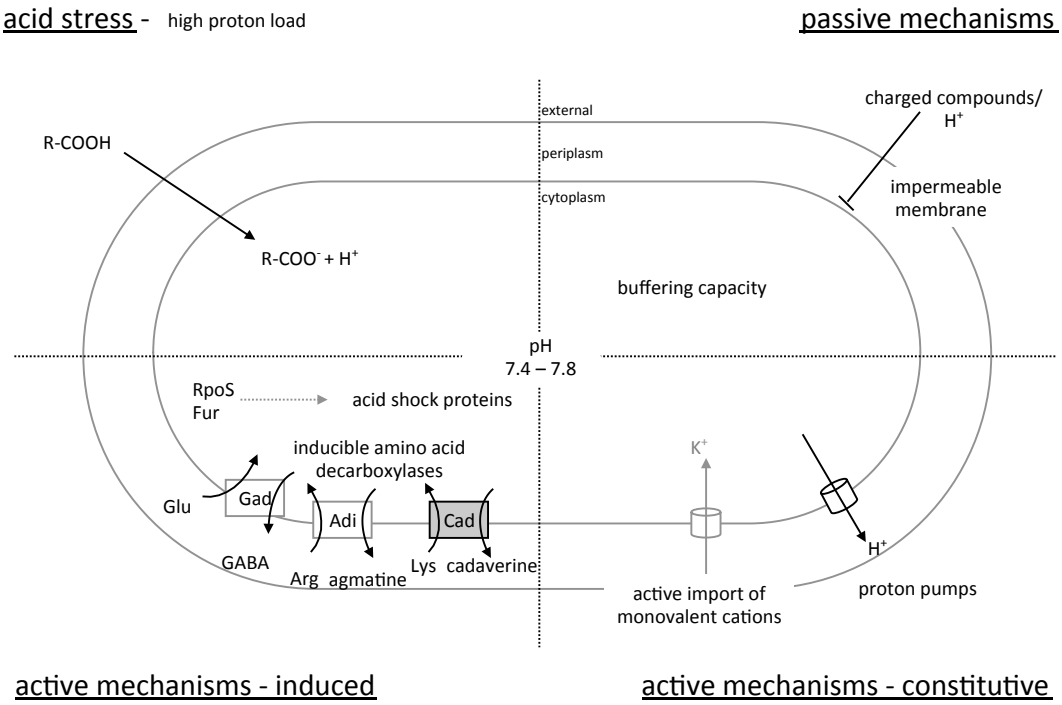

Figure 1.1: Acid stress and resistance in *E. coli*. When the surrounding pH decreases, *E. coli* has to sustain an immense increase in external proton concentration. Along with this, protonated organic acids may diffuse into the cell and dissociate in the more alkaline cytoplasm thereby decreasing the internal pH (upper left corner). Besides passive mechanisms, as the proton impermeable cytoplasmic membrane and the buffering capacity of the cytoplasm (upper right corner), *E. coli* possesses several active systems (lower part). Constitutively expressed proton pumps and ion coupled transporters counterbalance the internal proton concentration and the net charge of the cytoplasm. In addition, acid resistance of *E. coli* is conferred by the action of inducible acid resistance systems. In response to acid stress, multiple regulators (i.e. RpoS (σ^S), Fur) induce expression of acid shock proteins (ASPs) that function in the protection, repair and degradation of macromolecules. As well degradative amino acid decarboxylase systems are induced in response to acid stress (Gad, Adi, Cad). Conversion of amino acids to amines and consumption of cytoplasmic protons lead to an increase in the cytoplasmic pH. The expression of the Gad system is dependent on acidic pH and glutamate (Glu) availability. The Adi system is induced at acidic pH if arginine (Arg) is present and leads to the production of agmatine that is excreted in exchange with arginine. The lysine decarboxylase (CadA) and the lysine/cadaverine antiporter CadB form the Cad system that helps to survive mild acidic stress.

Active proton pumps antagonize the proton influx by actively exporting protons to the periplasmic space. To prevent a further influx of protons, positively charged monovalent ions, such as potassium, are enriched in the cytoplasm by the action of transporters [Bakker and Mangerich, 1981; Booth, 1985; Dosch *et al.*, 1991; Kroll and Booth, 1983]. In addition to these constitutively expressed systems, *E. coli* possesses inducible acid resistance (AR) systems (see figure 1.1). As a result of acid stress, expression of acid

shock proteins (ASPs) is induced. Different subsets of acid shock proteins are induced by the action of diverse regulators such as the stationary phase sigma factor (σ^S or RpoS), PhoP, OmpR or Fur [Bearson *et al.*, 1997; Heyde and Portalier, 1990]. All acid shock proteins function in the protection, repair and degradation of macromolecules. In the presence of the amino acids glutamate, arginine or lysine induction of the inducible amino acid decarboxylase systems (Gad, Adi or Cad, respectively) occurs at low pH. These systems consist of pairs of amino acid decarboxylases and antiporters. The decarboxylases are pyridoxal phosphate-containing enzymes (glutamate decarboxylases GadA and GadB, arginine decarboxylase AdiA and the lysine decarboxylase CadA) that catalyze the conversion of the α-carboxy groups of their amino acid substrate to amine groups by consumption of a cytoplasmic proton. In this reaction, CO_2, as well as the corresponding amine (γ-amino butyric acid (GABA)/ agmatine or cadaverine), are produced. Subsequently, the cognate antiporters export the decarboxylation end product in exchange with the amino acid substrate [Foster, 2004]. Some *E. coli* strains possess an additional amino acid decarboxylase system, the ornithine (Spe) decarboxylase system, that seems to have developed recently and appears to be of minor importance [Applebaum *et al.*, 1977].

1.2 The Cad System

The lysine-dependent amino acid decarboxylase system of *E. coli* consists of the inducible, cytoplasmic lysine decarboxylase CadA and the transmembrane proteins CadB (lysine-cadaverine antiporter) and CadC (transcriptional activator; figure 1.2; Auger *et al.* [1989]; Sabo *et al.* [1974]; Meng and Bennett [1992a]; Watson *et al.* [1992]; Meng and Bennett [1992b]).

CadA catalyzes the decarboxylation of lysine to cadaverine. In this reaction an intracellular proton is consumed and carbon dioxide and cadaverine are produced. CadA is a 715 amino acid protein (81 kDa) that forms decameric oligomers out of dimer subunits in order to be active [Kanjee *et al.*, 2011; Sabo *et al.*, 1974; Meng and Bennett, 1992a]. As other amino acid decarboxylases, CadA is dependent on the co-factor pyridoxal 5-phosphate [Sabo *et al.*, 1974; Sher and Mallette, 1954]. CadA was shown to work optimally at a pH of 5.7 [Sabo *et al.*, 1974] and to interact with the AAA^+ (ATPases associated with various cellular activities)-ATPase RavA and the small alarmone ppGpp [Kanjee *et al.*, 2011; Snider *et al.*, 2006]. The interaction between CadA and RavA results in a cage-like complex where two CadA decamers associate with at

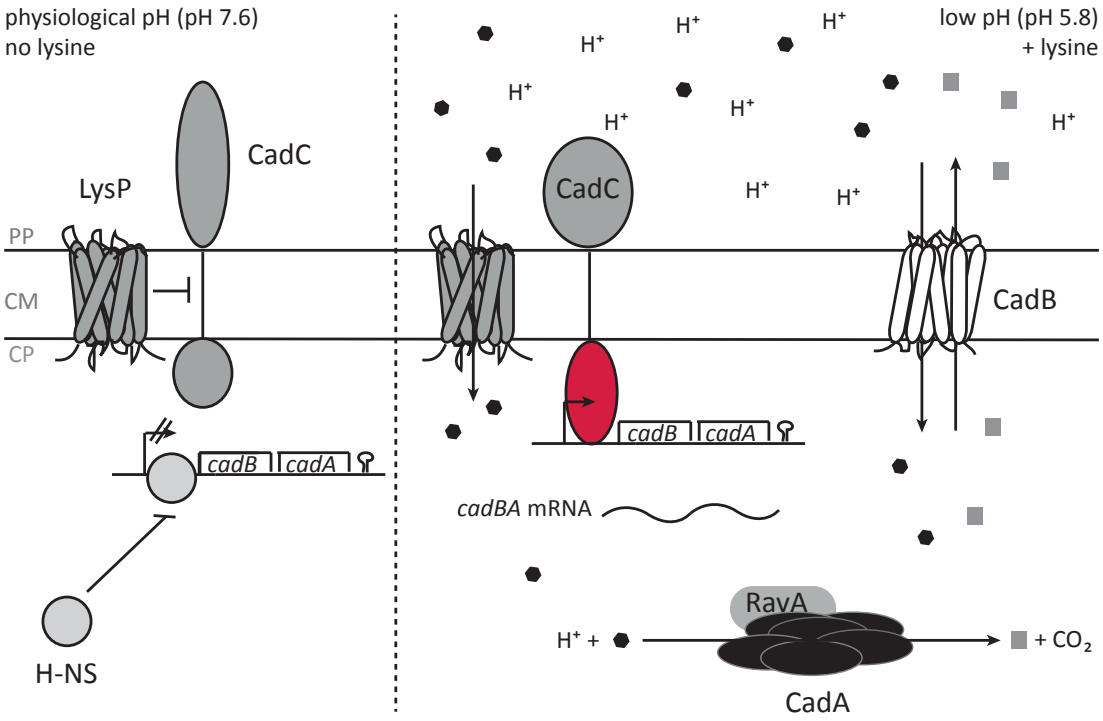

Figure 1.2: The Cad system of *E. coli*. Left: At physiological pH and in the absence of lysine, LysP prevents activation of *cadBA* transcription. In addition the global regulator H-NS inhibits initiation. **Right:** When the external pH decreases and lysine becomes available inhibition by LysP gets abrogated. Thus CadC is able to bind to P_{Cad} and activates transcription. As a consequence CadA and CadB are produced. CadA decarboxylates lysine (black hexagons) under consumption of a cytoplasmic proton and releases cadaverine (grey squares) as well as carbon dioxide. Subsequently, cadaverine is exported by CadB to the periplasmic space. After reaching a certain threshold concentration cadaverine and CO_2 act as feedback inhibitors (not depicted here). PP: periplasm; CM: cytoplasma membrane; CP: cytoplasm.

most five RavA hexamers. While association with RavA does not interfere with CadA activity [Snider *et al.*, 2006], ppGpp strongly inhibits CadA enzymatic activity. As ppGpp is known to mediate stringent response (stress response of bacteria in response to amino acid starvation) it was assumed that interaction with ppGpp links CadA activity to the nutritional state of the cell and diminishes lysine degradation under amino acid limiting conditions [Kanjee *et al.*, 2011].

CadB is the inner membrane bound transporter of the Cad system which expels cadaverine in exchange with lysine (K_m for cadaverine 390 μM). CadB consists of 444 amino acids that form 12 transmembrane-spanning helices [Meng and Bennett, 1992a; Soksawatmaekhin *et al.*, 2006]. It shares high similarity with PotE (30.7% overall identity),

which is the putrescine-ornithine antiporter of the Spe system [Soksawatmaekhin *et al.*, 2004]. The antiport catalyzed by CadB promotes the membrane potential. Together with the consumption of a cytoplasmic proton during decarboxylation, this proton motive metabolic cycle causes neutralization of the cytoplasm and an increase in the ATP level of the cell [Soksawatmaekhin *et al.*, 2004]. In addition, expelled cadaverine binds to and thus inhibits the cation-selective porins, OmpC and OmpF, aiding in acid survival [Samartzidou and Delcour, 1999; Samartzidou *et al.*, 2003].

1.2.1 Regulatory Mechanisms and Associated Proteins

Overall, expression of the afore-mentioned *cad* genes, *cadA* and *cadB*, is mainly regulated by the transcriptional activator CadC. *cadA* and *cadB* are encoded in an operon that is located at min 93.7 on the *E. coli* chromosome and are transcribed bicistronically [Auger *et al.*, 1989; Meng and Bennett, 1992a]. Transcription is induced at acidic pH and concomitantly available external lysine [Neely and Olson, 1996; Kannan *et al.*, 2008; Meng and Bennett, 1992a] and is σ^{70}-dependent. If the cells are shifted back to neutral pH, expression of *cadBA* is shut off within minutes [Neely and Olson, 1996]. In addition, the *cadBA* operon is only transiently expressed, i.e. the number of *cadBA* transcripts decreases approximately 20 minutes after the shift (no transcript detectable 60 min post-induction, Neely and Olson [1996]) although inducing conditions persist. Under anaerobic conditions transcription is increased up to ten fold [Sabo *et al.*, 1974]. Transcription is controlled by the P_{Cad} promoter and is activated through direct binding of the inner-membrane protein CadC (see section 1.2.2, Küper and Jung [2005]). Küper and Jung [2005] identified two binding sites for CadC in the *cadBA* promoter region that are necessary for activation: Cad1 (-144 to -112 bp) and Cad2 (-89 to -59 bp).

LysP - mediates lysine-dependent inhibition

Inhibition of *cadBA* transcription in the absence of lysine was shown to be mediated by the lysine-specific permease LysP. LysP consists of 489 amino acids with a predicted mass of 53.7 kDa [Steffes *et al.*, 1992]. According to Ellis *et al.* [1995] LysP consists of 12 transmembrane helices with five cytoplasmic and six periplasmic loops, with both termini in the cytoplasm. LysP belongs to the APC (amino acid/polyamine/organocation) superfamily, and, more precisely, to the AAT (amino acid transporter) family [Jack *et al.*, 2000] of proteins, due to its topology and the use of lysine as its substrate. Be-

cause of its inhibitory effect on *cadBA* expression LysP was also termed CadR [Neely *et al.*, 1994; Tabor *et al.*, 1980]. Mutants lacking *lysP* displayed a lysine-independent expression, while overexpression prevented activation of *cadBA* transcription even in the presence of lysine [Tetsch *et al.*, 2008; Neely *et al.*, 1994]. Replacement of the periplasmic sensor domain of CadC with the comparable domain of *Vibrio harveyi*'s ToxR, which is not known to sense lysine, did not affect lysine-dependent regulation. Instead, lysine-dependency was lost after substitution of the CadC transmembrane helix with helices from PutP (a 12 TM - helix proline transporter of *E. coli*). Moreover, the periplasmic domain of CadC did not bind lysine *in vitro*. Therefore, the role of the transmembrane helix of CadC in LysP-mediated inhibition was further analyzed, resulting in identification of an aromatic patch crucial to inhibition. The aromatic patch spanned from Phe159 to 165, with Phe165 playing the most significant role. In accordance, it was concluded that CadC is not a direct sensor of external lysine. It was suggested that inhibition in the absence of lysine is realized by a physical interaction of LysP with CadC via the transmembrane domains that constrain CadC and render it inactive [Tetsch *et al.*, 2008]. Due to its regulatory function, LysP belongs to the trigger transporters [Tetsch and Jung, 2009], which transport substrates and interact with regulatory proteins, thus influencing gene expression.

H-NS - repression at non-inducing conditions

In addition to LysP, the global regulator H-NS (histone-like nucleoid structuring protein) was shown to be essential for repression of *cadBA* expression under non-inducing conditions, i.e. neutral pH and aerobic conditions [Shi *et al.*, 1993]. H-NS is one of the most abundant proteins in the bacterial cell (approx. 20,000 copies) and thus a major component of the nucleoid. It is known to preferentially bind to curved and AT-rich DNA sequences commonly found at promoters [Schröder and Wagner, 2002]. As a nucleoid associated protein H-NS affects (represses) transcription of genes coding for versatile proteins (see for a review Fang and Rimsky [2008]). Küper and Jung [2005] confirmed the repressive effect of H-NS on *cadBA* expression *in vivo* and proposed a model where transcription is repressed via H-NS mediated transcriptional silencing. Accordingly, H-NS binds to binding sites within the P_{Cad} promoter and more distant sites thus leading to curvature of the DNA and consequently preventing transcription. At inducing conditions, CadC is thought to displace H-NS by binding to the promoter region and to activate transcription afterwards.

Feedback inhibition by cadaverine and CO_2

Along with LysP and H-NS, the end-products of lysine decarboxylation (cadaverine and CO_2) inhibit *cadBA* expression [Neely *et al.*, 1994; Takayama *et al.*, 1994]. Takayama *et al.* [1994] demonstrated that carbon dioxide and sodium carbonate had a repressive effect on *cadB* expression. So far, the molecular mechanism of this inhibition and whether it is a direct effect of CO_2 or its derivatives is not understood. In contrast it is known that the second feedback inhibitor, cadaverine, binds to the periplasmic domain of CadC *in vitro* with an apparent affinity of 96 ± 18 μM [Tetsch *et al.*, 2008]. In addition, inhibition by cadaverine was not linked to the transmembrane domain [Tetsch *et al.*, 2008]. Also, several amino acid substitutions in the periplasmic domain of CadC are known to abrogate inhibition of *cadBA* expression by cadaverine (R265C; Q266P; G284D; N263K; P499L; T475A and L497S - Dell *et al.* [1994]; Neely *et al.* [1994]). Nonetheless, the molecular details of cadaverine-mediated inhibition are not well understood either.

1.2.2 CadC

As outlined above, CadC is the transcriptional activator that activates transcription of the *cadBA* operon. CadC integrates multiple signals [lysine - mediated by LysP; cadaverine and CO_2 (see section 1.2.1) as well as the pH (see below)].

CadC - a ToxR-like transcriptional activator

The *cadC* gene is located directly upstream of the *cadBA* promoter (P_{Cad}) and is constitutively expressed [Neely and Olson, 1996; Watson *et al.*, 1992]. CadC is a 512 amino acid inner membrane protein (58 kDa) that belongs to the ToxR-like transcriptional activators [Miller *et al.*, 1987]. ToxR of *Vibrio cholerae* and other *Vibrio*-species, the master protein of ToxR-like proteins, regulates virulence-associated genes. Beside the ToxR proteins of the *Vibrionaceae* the following proteins also belong to the ToxR-like regulators: CadC of *E. coli*, *S. typhimurium*, *V. cholerae* [Merrell and Camilli, 2000] and *Vibrio vulnificus* [Rhee *et al.*, 2002, 2005]; TcpP of *V. cholerae* (co-regulator of virulence genes; Häse and Mekalanos [1998]); PsaE of *Yersinia pseudotuberculosis* (necessary for expression of fimbriae genes; Yang and Isberg [1997]) and WmpR of *Pseudoalteromonas tunicata* (involved in regulation of type IV pili and biofilm; Egan *et al.* [2002]). These proteins are characterized by a common topology consisting of an N-terminal cytoplasmic DNA-binding domain, one single transmembrane helix and a C-terminal, periplasmic sensor domain. Similar to CadC, ToxR is believed to directly

sense and transduce external stimuli into the cytoplasm.

DNA binding and transcriptional activation

DNA-binding of ToxR-like regulators is accomplished by a winged helix-turn-helix motif that can also be found in RO_{II}-type response regulators such as PhoB, OmpR of *E. coli* and VirG of *Agrobacterium tumefaciens* [Martinez-Hackert and Stock, 1997; Watson *et al.*, 1992; Parkinson and Kofoid, 1992]. In contrast to these response regulators, ToxR-like regulators are not chemically modified during signal-transduction. In addition to this highly conserved DNA-binding domain, ToxR-like regulators contain a large unstructured loop preceding the transmembrane domain. Currently the role of this extended loop region is not understood but seems to be important for proper transcriptional activation [Buchner, 2011].

CadC - an atypical one-component system

As mentioned above, lysine-availability is detected via an interaction with the lysine permease LysP. Alongside that mechanism, CadC is thought to sense external cadaverine and acidification of the periplasm via its periplasmic sensor domain [Dell *et al.*, 1994; Tetsch *et al.*, 2008]. Thus CadC integrates multiple signals and transmits the external state into the cytoplasm where it alters gene expression (of the *cadBA* operon). CadC, as well as all other ToxR-like regulators, combine sensory, transducing and effector function in the same polypeptide, and thereby represent the simplest model for signal transduction. In addition, signal transduction takes place without chemical modification dissimilar to the classical two component signal transduction systems that use phosphotransfer (see figure 1.3 and Stock *et al.* [2000] for a review on two-component signal transduction). In contrast to CadC, signal transduction by two-component systems employs a histidine kinase (that mostly becomes autophosphorylated upon stimulus perception) and a response regulator protein. These two proteins interact and change their activation state upon transfer of phosphoryl groups between each other. For a long time, two-component systems were viewed as the dominating system to link environmental signals to cellular responses. However, one component systems represent a more rapid and direct way of signal transduction, as the input and output domains are combined into a single polypeptide. The majority of one-component systems are cytoplasmic, soluble proteins while only 3% are membrane-bound proteins as CadC [Ulrich *et al.*, 2005].

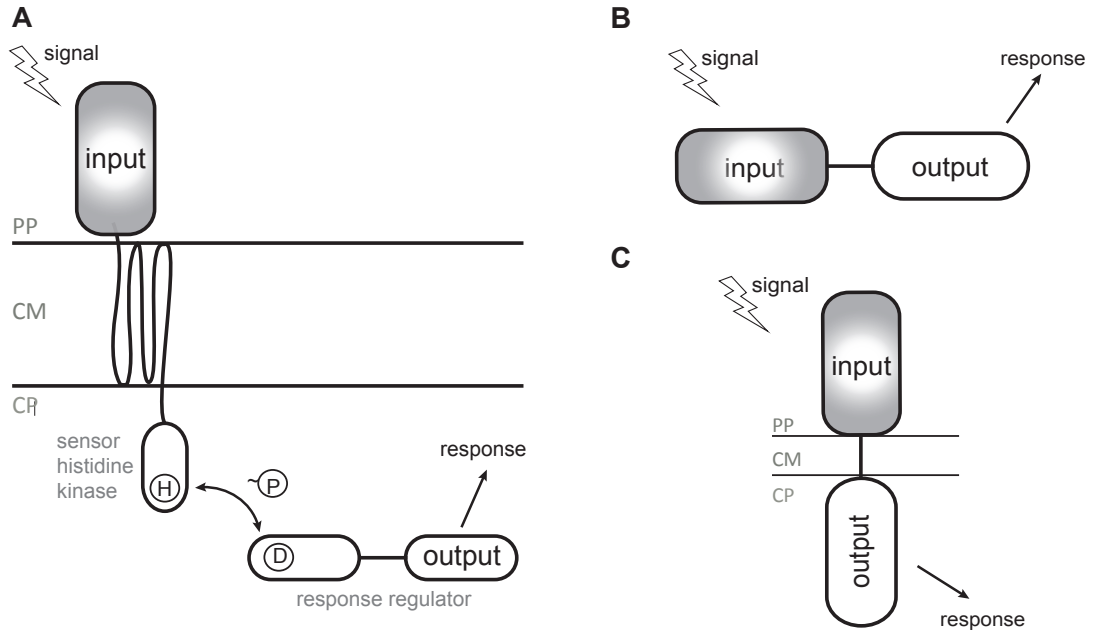

Figure 1.3: Modules of prokaryotic signal transduction. A. Prototypic two-component signal transduction system (TCS). In classical TCS an external stimuli (detected in the input domain) leads to autophosphorylation of the histidine (H) residue. The phosphoryl group is transferred to a conserved aspartate residue (D) in the receiver domain of the cognate response regulator. Phosphorylation induces cellular responses such as DNA-binding and transcriptional activation (output domain/ response). **B. and C.** However, the simplest way to convert a certain signal into a cellular output is to combine output and input domain in a single polypeptide as it is realized in one-component systems. Most of these one-component systems are soluble, intracellular proteins (**B**) but a small proportion (3%) contains transmembrane domains (**C**) and thus directly transduces an extracellular signal into the cytoplasm [Ulrich *et al.*, 2005]. PP: periplasm; CM: cytoplasma membrane; CP: cytoplasm.

1.3 Scope of this Thesis

The main objective of this thesis is to elucidate the molecular signal perception mechanism(s) of CadC. Understanding how signals are processed and transferred to another compartment by a single polypeptide will greatly improve our understanding of transmembrane signaling.

The first study sheds light on how the different modules of the Cad system collaborate in creating a certain response. To achieve this, a combination of wet experimentation and mathematical modeling is employed.

Since molecular analyses are much easier if one can interpret data with the help of a three dimensional structure, the crystal structure of the sensor domain (periplasmic do-

main) of CadC is solved in the second study. In addition, the tertiary structure of two variants that are perturbed in pH-dependent regulation is solved in order to identify signal-dependent structural changes.
With the structural information in hand, a direct mutagenesis screen is carried out in the third analysis. Thus, information about the molecular mechanism behind detecting acidification of the surrounding environment is gathered.
Since the third analysis focuses on pH detection, the fourth study attempts to understand cadaverine-dependent shut-off of *cadBA* expression. Using the current literature available regarding CadC, variants are constructed and analyzed with respect to cadaverine dependent regulation. The data are fed into the mathematical model formulated in the first branch of this study, allowing conclusions to be drawn about cadaverine-dependent regulation on a qualitative level.
Finally, the transcriptional regulation of *lysP* is analyzed in order to obtain a complete picture of all the components interacting within the Cad system and to gain a better understanding of this simple signal transduction system.
In conclusion, the data that is compiled will be incorporated into a more sophisticated and detailed model of Cad regulation.

1.4 References for Introduction

Applebaum, D. M., Dunlap, J. C., and Morris, D. R., 1977. Comparison of the biosynthetic and biodegradative ornithine decarboxylases of *Escherichia coli*. *Biochemistry*, 16(8):1580–1584.

Auger, E.A., Redding, K.E., Plumb, T., Childs, L.C., Meng, S.Y., and Bennett, G.N., 1989. Construction of *lac* fusions to the inducible arginine- and lysine decarboxylase genes of *Escherichia coli* K12. *Mol Microbiol*, 3(5):609–620.

Bakker, E. P. and Mangerich, W. E., 1981. Interconversion of components of the bacterial proton motive force by electrogenic potassium transport. *Journal of Bacteriology*, 147(3):820–826.

Bearson, S., Bearson, B., and Foster, J. W., 1997. Acid stress responses in enterobacteria. *FEMS microbiology letters*, 147(2):173–180.

Boor, K. J., 2006. Bacterial stress responses: What doesn't kill them can make them stronger. *PLoS Biology*, 4(1):18–20.

Booth, I. R., 1985. Regulation of cytoplasmic pH in bacteria. *Microbiology and Molecular Biology Reviews*, 49(4):359–378.

Buchner, S., 2011. *personal communication.*

Dell, C. L., Neely, M. N., and Olson, E. R., 1994. Altered pH lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. *Molecular microbiology*, 14(1):7–16.

Dosch, D. C., Helmer, G. L., Sutton, S. H., Salvacion, F. F., and Epstein, W., 1991. Genetic analysis of potassium transport loci in *Escherichia coli*: evidence for three constitutive systems mediating uptake potassium. *Journal of bacteriology*, 173(2): 687–696.

Egan, S., James, S., and Kjelleberg, S., 2002. Identification and characterization of a putative transcriptional regulator controlling the expression of fouling inhibitors in *Pseudoalteromonas tunicata*. *Applied and environmental microbiology*, 68(1):372–378.

Ellis, J., Carlin, A., Steffes, C., Wu, J., Liu, J., and Rosen, B. P., 1995. Topological analysis of the lysine-specific permease of *Escherichia coli*. *Microbiology*, 141(8): 1927–1935.

Fang, F. C. and Rimsky, S., 2008. New insights into transcriptional regulation by H-NS. *Current opinion in microbiology*, 11(2):113–120.

Foster, J. W., 2004. *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nature Reviews Microbiology*, 2(11):898–907.

Häse, C. C. and Mekalanos, J. J., 1998. TcpP protein is a positive regulator of virulence gene expression in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences*, 95(2):730–734.

Heyde, M. and Portalier, R., 1990. Acid shock proteins of *Escherichia coli*. *FEMS microbiology letters*, 69(1-2):19–26.

Jack, D. L., Paulsen, I. T., and Saier, M. H., 2000. The amino acid/ polyamine/organocation (APC) superfamily of transporters specific for amino acids, polyamines and organocations. *Microbiology*, 146(8):1797–1814.

Kanjee, U., Gutsche, I., Alexopoulos, E., Zhao, B., El Bakkouri, M., Thibault, G., Liu, K., Ramachandran, S., Snider, J., Pai, E. F., and Houry, W. A., 2011. Linkage between the bacterial acid stress and stringent responses: the structure of the inducible lysine decarboxylase. *The EMBO Journal*, 30:931–944.

Kannan, G., Wilks, J., Fitzgerald, D., Jones, B., BonDurant, S., and Slonczewski, J., 2008. Rapid acid treatment of *Escherichia coli*: transcriptomic response and recovery. *BMC microbiology*, 8(1):37.

Kroll, R. G. and Booth, I. R., 1983. The relationship between intracellular pH, the pH gradient and potassium transport in *Escherichia coli*. *Biochemical Journal*, 216(3): 709–716.

Küper, C. and Jung, K., 2005. CadC-mediated activation of the *cadBA* promoter in *Escherichia coli*. *J Mol Microbiol Biotechnol*, 10(1):26–39.

Martinez-Hackert, E. and Stock, A. M., 1997. Structural relationships in the OmpR family of winged-helix transcription factors. *J Mol Biol*, 269(3):301–312.

Meng, S. Y. and Bennett, G. N., 1992a. Nucleotide sequence of the *Escherichia coli cad* operon: a system for neutralization of low extracellular pH. *Journal of bacteriology*, 174(8):2659–2669.

Meng, S. Y. and Bennett, G. N., 1992b. Regulation of the *Escherichia coli cad* operon: location of a site required for acid induction. *Journal of bacteriology*, 174(8):2670.

Merrell, D. S. and Camilli, A., 2000. Regulation of *Vibrio cholerae* genes required for acid tolerance by a member of the "ToxR-like" family of transcriptional regulators. *Journal of Bacteriology*, 182(19):5342–5350.

Miller, V.L., Taylor, R.K., and Mekalanos, J.J., 1987. Cholera toxin transcriptional activator ToxR is a transmembrane DNA binding protein. *Cell*, 48(2):271–279.

Neely, M. N. and Olson, E. R., 1996. Kinetics of expression of the *Escherichia coli cad* operon as a function of pH and lysine. *J Bacteriol*, 178(18):5522–5528.

Neely, M. N., Dell, C. L., and Olson, E. R., 1994. Roles of LysP and CadC in mediating the lysine requirement for acid induction of the *Escherichia coli cad* operon. *J Bacteriol*, 176(11):3278–3285.

Parkinson, J. S. and Kofoid, E. C., 1992. Communication modules in bacterial signaling proteins. *Annual review of genetics*, 26(1):71–112.

Rhee, J. E., Rhee, J. H., Ryu, P. Y., and Choi, S. H., 2002. Identification of the *cadBA* operon from *Vibrio vulnificus* and its influence on survival to acid stress. *FEMS microbiology letters*, 208(2):245–251.

Rhee, J. E., Kim, K. S., and Choi, S. H., 2005. CadC activates pH-dependent expression of the *Vibrio vulnificus cadBA* operon at a distance through direct binding to an upstream region. *Journal of bacteriology*, 187(22):7870–7875.

Sabo, D. L., Boeker, E. A., Byers, B., Waron, H., and Fischer, E. H., 1974. Purification and physical properties of inducible *Escherichia coli* lysine decarboxylase. *Biochemistry*, 13(4):662–670.

Samartzidou, H. and Delcour, A. H., 1999. Excretion of endogenous cadaverine leads to a decrease in porin-mediated outer membrane permeability. *Journal of bacteriology*, 181(3):791–798.

Samartzidou, H., Mehrazin, M., Xu, Z., Benedik, M. J., and Delcour, A. H., 2003. Cadaverine inhibition of porin plays a role in cell survival at acidic pH. *Journal of bacteriology*, 185(1):13–19.

Schröder, O. and Wagner, R., 2002. The bacterial regulatory protein H-NS - A versatile modulator of nucleic acid structures. *Biological chemistry*, 383(6):945–960.

Sher, I. H. and Mallette, M. F., 1954. Purification and study of L-lysine decarboxylase from *Escherichia coli* B. *Archives of Biochemistry and Biophysics*, 53(2):354–369.

Shi, X., Waasdorp, B. C., and Bennett, G. N., 1993. Modulation of acid-induced amino acid decarboxylase gene expression by *hns* in *Escherichia coli*. *Journal of bacteriology*, 175(4):1182–1186.

Slonczewski, J. L., Fujisawa, M., Dopson, M., and Krulwich, T. A., 2009. Cytoplasmic pH measurement and homeostasis in bacteria and archaea. *Advances in Microbial Physiology*, 55:1–79.

Snider, J., Gutsche, I., Lin, M., Baby, S., Cox, B., Butland, G., Greenblatt, J., Emili, A., and Houry, W. A., 2006. Formation of a distinctive complex between the inducible bacterial lysine decarboxylase and a novel AAA+ ATPase. *Journal of Biological Chemistry*, 281(3):1532–1546.

Soksawatmaekhin, W., Kuraishi, A., Sakata, K., Kashiwagi, K., and Igarashi, K., 2004. Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. *Mol Microbiol*, 51(5):1401–1412.

Soksawatmaekhin, W., Uemura, T., Fukiwake, N., Kashiwagi, K., and Igarashi, K., 2006. Identification of the cadaverine recognition site on the cadaverine-lysine antiporter CadB. *J Biol Chem*, 281(39):29213–29220.

Steffes, C., Ellis, J., Wu, J., and Rosen, B. P., 1992. The *lysP* gene encodes the lysine-specific permease. *Journal of bacteriology*, 174(10):3242–3249.

Stock, A. M., Robinson, V. L., and Goudreau, P. N., 2000. Two-component signal transduction. *Annual review of biochemistry*, 69(1):183–215.

Tabor, H., Hafner, E. W., and Tabor, C. W., 1980. Construction of an *Escherichia coli* strain unable to synthesize putrescine, spermidine, or cadaverine: characterization of two genes controlling lysine decarboxylase. *Journal of Bacteriology*, 144(3):952–956.

Takayama, M., Ohyama, T., Igarashi, K., and Kobayashi, H., 1994. *Escherichia coli* *cad* operon functions as a supplier of carbon dioxide. *Molecular microbiology*, 11(5): 913–918.

Tetsch, L. and Jung, K., 2009. How are signals transduced across the cytoplasmic membrane? Transport proteins as transmitter of information. *Amino acids*, 37(3): 467–477.

Tetsch, L., Koller, C., Haneburger, I., and Jung, K., 2008. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol Microbiol*, 67(3):570–583.

Ulrich, L. E., Koonin, E. V., and Zhulin, I. B., 2005. One-component systems dominate signal transduction in prokaryotes. *Trends in microbiology*, 13(2):52–56.

Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L., and Olson, E. R., 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli* *cad* operon by external pH. *J Bacteriol*, 174(2):530–540.

Yang, Y. and Isberg, R. R., 1997. Transcriptional regulation of the *Yersinia pseudotuberculosis* pH 6 antigen adhesin by two envelope-associated components. *Molecular microbiology*, 24(3):499–510.

Zilberstein, D., Agmon, V., Schuldiner, S., and Padan, E., 1984. *Escherichia coli* intracellular pH, membrane potential, and cell growth. *Journal of bacteriology*, 158 (1):246–252.

2 Induction kinetics of a conditional pH stress response system in *Escherichia coli*

Reprinted from The Journal of Molecular Biology, 393 (2), 2009, G. Fritz, C. Koller, K. Burdack, L. Tetsch, I. Haneburger, K. Jung and U. Gerland; Induction Kinetics of a Conditional pH Stress Response System in *Escherichia coli*, 272-286, Copyright (2009), with permission from Elsevier.

doi:10.1016/j.jmb.2009.08.037 J. Mol. Biol. (2009) 393, 272–286

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Induction Kinetics of a Conditional pH Stress Response System in *Escherichia coli*

Georg Fritz[1,2], Christiane Koller[3], Korinna Burdack[3], Larissa Tetsch[3], Ina Haneburger[3], Kirsten Jung[3] and Ulrich Gerland[1,2]*

[1]*Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany*

[2]*Center for NanoScience, Ludwig-Maximilians-Universität München, Theresienstraße 37, 80333 München, Germany*

[3]*Munich Center of Integrated Protein Science at the Department of Biology I, Microbiology, Ludwig-Maximilians-Universität München, Großhaderner Straße 2-4, 82152 Martinsried, Germany*

Received 18 May 2009; received in revised form 6 August 2009; accepted 13 August 2009
Available online 21 August 2009

Edited by M. Gottesman

The analysis of stress response systems in microorganisms can reveal molecular strategies for regulatory control and adaptation. In this study, we focused on the Cad module, a subsystem of *Escherichia coli's* response to acidic stress that is conditionally activated at low pH only when lysine is available. When expressed, the Cad system counteracts the elevated H^+ concentration by converting lysine to cadaverine under the consumption of H^+ and exporting cadaverine in exchange for external lysine. Surprisingly, the *cad* operon displays a transient response, even when the conditions for its induction persist. To quantitatively characterize the regulation of the Cad module, we experimentally recorded and theoretically modeled the dynamics of important system variables. We established a quantitative model that adequately describes and predicts the transient expression behavior for various initial conditions. Our quantitative analysis of the Cad system supports negative feedback by external cadaverine as the origin of the transient response. Furthermore, the analysis puts causal constraints on the precise mechanism of signal transduction via the regulatory protein CadC.

Keywords: CadB; CadA; lysine decarboxylase; acid stress response; quantitative modeling

Introduction

During their natural life cycle, gastrointestinal bacteria are faced with acid stress while passing the extreme low pH of the stomach and being exposed to volatile fatty acids in the intestine. *Escherichia coli*'s remarkable ability to sustain growth over multiple decades of H^+ concentrations[1] and its potential to survive extremely low pH are implemented by a battery of pH homeostasis[2–4] and acid tolerance[5–8] systems. In recent years, it was increasingly recognized that each of these subsystems is specifically activated under certain environmental conditions,[6,8] while the orchestration of the different responses is just beginning to be explored. However, a system-level study of the acid stress response requires detailed quantitative analysis of the individual modules.

One of the conditional stress response modules is the Cad system,[9–13] which is induced only when acidic stress occurs in a lysine-rich environment. The Cad module protects members of the *Enterobacteriaceae* against anorganic and organic acids in the intestinal tract[14] and against fermentation acids under phosphate-limiting conditions.[15] The three principal components of the Cad system are the enzyme CadA, the transport protein CadB, and the regulatory protein CadC (see Fig. 1). The decarboxylase CadA converts the amino acid lysine into cadaverine, a reaction that effectively consumes H^+.[16] The antiporter CadB imports the substrate, lysine, and exports the product, cadaverine. Together, CadA and CadB

*Corresponding author. E-mail address: gerland@lmu.de.

0022-2836/$ - see front matter

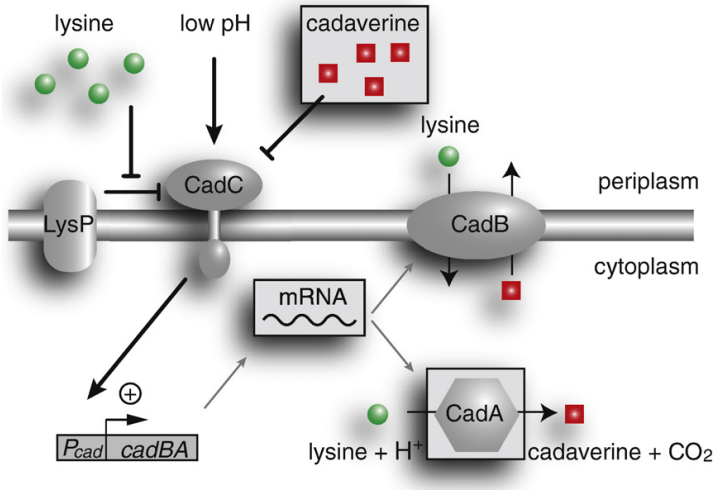

Fig. 1. Qualitative model of the Cad system in *E. coli* (simplified). The Cad system is conditionally activated by low pH and high lysine levels. Lysine inhibits the repressive effects of the lysine permease LysP on the receptor CadC, while low pH activates CadC directly. The active form of CadC activates transcription of *cadBA*, encoding the lysine decarboxylase CadA and the lysine/cadaverine antiporter CadB. The Cad system imports lysine, decarboxylates it under consumption of a cytoplasmic proton, and exports the product cadaverine in exchange for another lysine molecule. The net effects of these reactions are the expulsion of a proton from the cytoplasm and the excretion of the basic polyamine cadaverine. Finally, it is believed that external cadaverine deactivates CadC. In our experiments, we recorded the time evolution of the variables in gray boxes.

reduce the intracellular H^+ concentration and thereby contribute to pH homeostasis.[10,17] The cytoplasmic membrane protein CadC not only senses the external conditions[12,18,19] but also regulates the response by binding directly to the DNA and activating the transcription of *cadBA*.[20] Similar to other members of the ToxR family,[21] CadC thereby performs signal transduction in a single component, without the phosphorylation step employed by two-component systems.[22] Figure 1 also depicts the lysine permease LysP, which is not part of the *cad* operon but essential for its function since CadC senses lysine indirectly via interaction with LysP.[12,18,19,23] In contrast, the external (periplasmic) pH is believed to be sensed directly by CadC through a pH-dependent conformational transition and/or proteolytic cleavage.[19,24] The signal integration performed by CadC then ensures that CadA and CadB are produced only under the appropriate external conditions of low pH and lysine abundance.

However, CadC also senses a third input, which seems surprising from a physiological point of view: Dell *et al.* identified several CadC derivatives with single amino replacements that responded differently to cadaverine in comparison with wild-type CadC.[19] These data provided initial evidence that cadaverine influences CadC-mediated *cadBA* expression. Subsequently, it was found that cadaverine represses the long-term expression of the *cad* operon,[12] and we could show that the periplasmic domain of CadC in fact binds cadaverine *in vitro*.[23] As cadaverine is the end product of the decarboxylase reaction, it was suggested that it accumulates in the medium and causes a delayed transcriptional down-regulation of *cadBA* expression.[13] Although many stress response systems display a similar transient response,[25,26] their regulation strategy appears to be fundamentally different: for instance, the osmo-stress response of yeast directly follows its stimulus (low osmolarity) and remains active until the osmolarity returns back to physiological levels, while *cadBA* expression is down-regulated although the stress persists.

In this study, we explored whether negative feedback via external cadaverine can account for the transient response of the Cad system on a quantitative level. In particular, we were interested in the factors that determine the *duration* and the *amplitude* of the transient response and asked how the addition of external cadaverine affects these characteristic quantities. It is known, for instance, that external cadaverine reduces the long-term activity of the Cad system,[12] but one would like to know whether it shortens the duration of the transient expression pulse or reduces the amplitude of the pulse. Or does it affect both of these properties?

To address these questions, we quantitatively measured the dynamics of the Cad system in three important variables at high time resolution: the *cadBA* transcript, the activity of the lysine decarboxylase CadA, and the concentration of excreted cadaverine. Based on the existing qualitative model (cf., Fig. 1), we formulated a quantitative model for the Cad system and tested its agreement with the experimental response dynamics. We found that our quantitative model coherently describes the dynamical response of the wild-type Cad system within a physiological parameter regime. The available data constrain the key biochemical parameters to a narrow regime. For instance, we inferred the effective *in vivo* deactivation threshold for the Cad system and compared it with a previously measured *in vitro* binding threshold.[23] Using the quantitative model, we formulated predictions for the response dynamics of the Cad system under conditions with initially added cadaverine and in a mutant strain with a defective lysine permease, LysP. The suc-

cessful experimental validation of these predictions strongly supports the existence of the postulated feedback inhibition mechanism via cadaverine in the Cad system. Finally, we discuss the causal constraints of our results on the signal transduction mechanism by CadC, helping discriminate between two contradicting models.

Results

Transient expression dynamics

To probe the transient response of the Cad module, we first grew *E. coli* (strain MG1655) to exponential phase at pH 7.6 in minimal medium. We then induced the Cad module by transferring cells into fresh minimal medium with 10 mM lysine and buffered at pH 5.8. After shifting *E. coli* to inducing conditions, we monitored cell growth by determining the number of colony-forming units. Due to the acid stress, cell growth was significantly impaired, but cells remained viable over the 4-h period of the experiment (see Fig. S3). In the habitat of the intestines, fast and/or uniform growth of the cells is also not expected. Hence, the analysis of the Cad module under the present conditions is reasonable, as it mimics an *E. coli* population in its natural environment.

The induction defined the starting point, $t = 0$ min, for our measurements of the response, which we performed initially at intervals of 5 min until $t = 30$ min and then at longer intervals of 30 min. To quantify the response, we assayed the *cadBA* mRNA level, the specific CadA activity, and the external cadaverine concentration (see Materials and Methods for all experimental details). We found that the external pH, shown in Fig. 2a, remained low, even slightly decreasing from the induction level, over the entire 4-h period of the experiment. Transcription of the *cad* operon began immediately after induction, and mRNA rapidly accumulated, as shown in Fig. 2b. At $t \approx 25$ min, the mRNA level peaked and then rapidly decreased, reaching its low pre-induction level at about $t = 90$ min. The response on the protein level, as quantified by the specific activity of CadA shown in Fig. 2c, was slower, exhibiting a slight delay after induction and reaching a plateau level at $t \approx 25$ min, which was sustained over the time of the experiment. The activity of the Cad module led to the production and secretion of cadaverine, which accumulated in the medium, as shown in Fig. 2d.

The transient expression of the Cad module shown in Fig. 2 is in qualitative agreement with previous induction experiments that studied the system in a less quantitative manner.[13] The biochemical mechanism for the transient behavior remained unclear, however. It was suggested that the external cadaverine level exerts negative feedback on the activity of the regulator CadC.[12,13] Alternatively, the Cad module might, for instance, directly affect and control the level of its input stimuli. In other words, the activity of the Cad module might reduce the external lysine concentration below its induction threshold or shift the external pH level outside its range for induction.

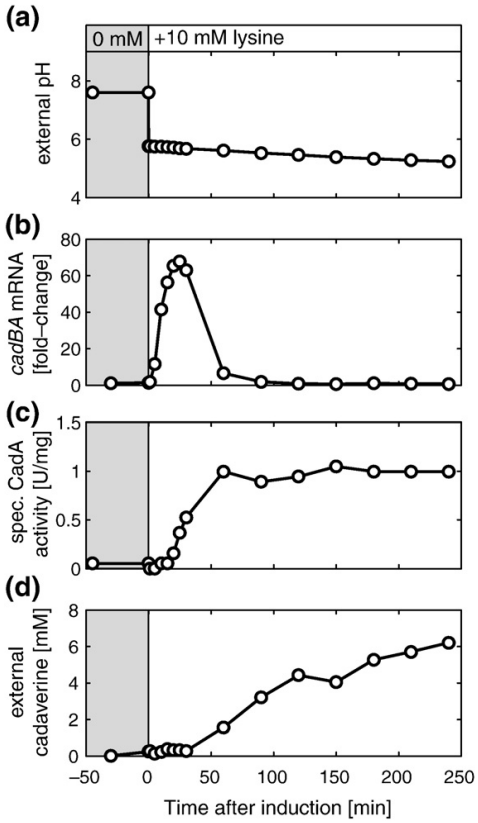

Fig. 2. Induction kinetics of the *cad* operon in *E. coli* MG1655. (a) The Cad system was induced at $t = 0$ min by a shift from pH 7.6 to pH 5.8 and simultaneous addition of 10 mM lysine. (b–d) The time evolution of the *cadBA* mRNA (b), the specific CadA activity (c), and the extracellular cadaverine concentration (d) was determined as described in Materials and Methods. All values are average values from duplicate repetitions. Specific CadA activity is given as U/(mg protein), with 1 U = 1 μmol cadaverine/min. For each sample, the cell density was adjusted to an optical density at 600 nm of 1, corresponding to 150 μg protein/mL. Concentration of total RNA was determined, and for each time point, equal amounts of RNA were tested with a radiolabeled probe directed against *cadBA* mRNA. Moreover, the signal intensity of *cadBA* transcript was compared with the signal intensity of *rpoD* mRNA that is constitutively transcribed. From these data, fold changes of *cadBA* transcription relative to the pre-induction value were calculated. For a detailed discussion, see the main text.

Dose–response curves

To address these possible alternative explanations and to characterize the ranges and the intensity of

the total response under our experimental conditions, we next determined the "dose-response" behavior of the Cad module. We have seen above (see Fig. 2c) that the CadA activity reaches a steady-state plateau about 60 min after induction. We took this plateau value as a proxy for the total (cumulative) response of the Cad system and studied its dependence on the input signals. To this end, we induced the Cad module with different external pH levels and initial lysine concentrations and assayed samples 90 min after induction for their CadA activity (see Materials and Methods for details). Panels (a) and (b) of Fig. 3 show the lysine dependence and the pH dependence of the response, respectively. The data in Fig. 3a indicate that when induced with pH 5.8, the Cad module is barely active at lysine concentrations below 0.5 mM, whereas it is fully active for lysine levels exceeding 5 mM. In between these values, the activity increases sigmoidally with the inducing lysine concentration. Similarly, at a given lysine induction level of 10 mM, the activity depends sigmoidally on the inducing pH (see Fig. 3b), with no significant activity above pH 6.8 and full activity at pH 5.8 and lower.

Taken together, the pH dependence of the total response shown in Fig. 3b and the time series in Fig. 2a show that the pH level did not leave the range for induction during the course of our experiment in Fig. 2. Hence, the transient behavior of the *cadBA* expression is clearly not mediated by a decrease of the external pH stimulus.

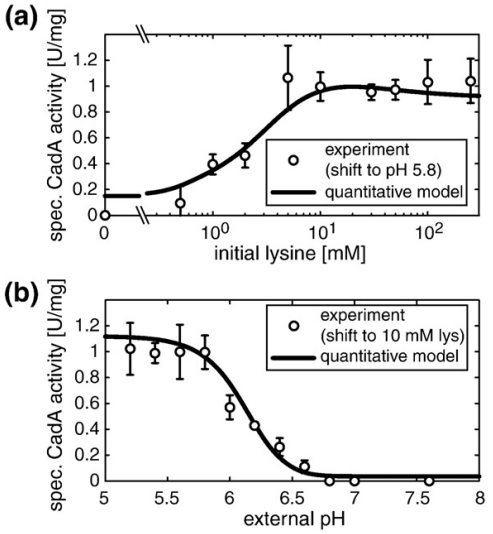

Fig. 3. Dose–response curves of the wild-type Cad system. (a) Dependence of the CadA response on the inducing lysine concentration. The specific CadA activity was determined 90 min after induction with pH 5.8 and the indicated lysine concentration (see Materials and Methods). Similarly, the dependence on the inducing pH in (b) was obtained by induction with 10 mM lysine and the indicated pH. Again, specific CadA activity was determined 90 min after induction. The continuous lines show the fit result of our quantitative model. Values are mean values from at least three independent experiments. Error bars represent standard deviation from these experiments.

The dose–response curves in Fig. 3 characterize the input–output behavior of the system when the Cad module is regarded as a "black box" signal processing unit. In particular, we can read off the apparent activation thresholds of the Cad module (i.e., the pH and lysine levels at which the module displays half-maximal activity). Such apparent thresholds constitute the first level of description in a top–down system analysis. Conversely, in a bottom–up analysis, the first level of description is via biochemical interaction parameters, while effective parameters, such as apparent thresholds, emerge from the interplay of molecular interactions. Quantitative modeling and analysis of this interplay are the only ways to make a connection between the two levels of description. In the following, we want to make such a connection and then leverage it to estimate the molecular activation thresholds of CadC from the apparent behavior of the Cad module.

Construction of a quantitative model

It is clear from the abovementioned data that a minimal quantitative model of the Cad module must describe the integration of the input signals pH and lysine, as well as the effect of cadaverine on the activity of CadC. Furthermore, it must describe the regulation and expression of the *cadBA* operon and the functioning of the CadA and CadB proteins, such that we may relate the quantitative model to the observed dynamics of the three system variables monitored in our experiments. Our construction of such a quantitative model is guided by the qualitative model in Fig. 1 and the known biochemistry of the Cad module. In that, our approach takes advantage of the wealth of prior knowledge about molecular interactions, which determine the topology (i.e., the "wiring") of the signaling network. In general, when the molecular players and their interactions are unknown, network reconstruction techniques can be used to gain qualitative and quantitative insights into the regulation of the system (see, for example, Refs. 27 and 28). However, here we are faced with a simpler problem where a predefined network topology has to be tested for agreement with the experimental data.

Signal integration

The three external signals known to affect the activity of the Cad module are the time-dependent lysine concentration, $l(t)$, the cadaverine concentration, $c(t)$, and the pH. The membrane protein CadC, which receives and combines these signals into a single response,[12,18,19] is constitutively expressed,[19] and hence we take the total amount of CadC per cell, C_0, to be constant. The signals then modulate only the fraction of active CadC molecules per cell, $C(t)/$

C_0. We assume that the different signals regulate CadC independently such that the CadC activity is described by the product form,

$$C(t)/C_0 = f(\mathrm{pH}(t)) \times g(l(t)) \times h(c(t)) \qquad (1)$$

Indeed, all experimental data available show no indication for a coupled effect of the input signals on CadC.[13,23] Another assumption implicitly made by Eq. (1) is that the fraction of active receptors is always equilibrated to the current levels of the input signals—that is, it does not depend on the signal levels prior to time t. This assumption is also plausible since the typical timescale for conformational transitions in receptors (see, for example, Ref. 29) is much shorter than the timescale of our experiments. The functions f, g, and h in Eq. (1) take on values between 0 and 1 and are assumed to be of the Hill form typical for cooperative binding reactions. The pH dependence is parameterized as follows:

$$f(\mathrm{pH}) = \frac{1}{1 + 10^{\frac{\mathrm{pH}-\mathrm{pH}_0}{\Delta\mathrm{pH}}}} \qquad (2)$$

with pH_0 denoting the pH value at which f reaches half-maximal activity and ΔpH determining the width of the response curve. Similarly, lysine dependence and cadaverine dependence respectively take the following forms:

$$g(l) = \frac{(l/K_l)^{n_l}}{1 + (l/K_l)^{n_l}} \quad h(c) = \frac{1}{1 + (c/K_c)^{n_c}} \qquad (3)$$

where K_l and K_c are the effective *in vivo* activation thresholds for the direct and indirect regulatory interactions of lysine and cadaverine with CadC, respectively. As usual, the Hill coefficients n_l and n_c parameterize the cooperativity of the binding reactions and determine the maximal sensitivity for signal detection. The difference in form between $g(l)$ and $h(c)$ stems from the fact that lysine activates the Cad module, whereas cadaverine represses it. On the other hand, the difference from Eq. (2) is due to pH being logarithmically related to the H^+ concentration.

Transcriptional regulation

In its activated conformation, CadC directly binds to the *cadBA* promoter P_{cad} and activates *cadBA* expression.[20] Generally, transcriptional regulation in bacteria can be described by quantitative "thermodynamic" models.[30,31] For the present case, an appropriate form for the transcriptional activity from P_{cad} as a function of the abundance of active CadC, $C(t)$, is derived in Materials and Methods. The resulting rate equation for the time evolution of the mRNA level m then takes the following form:

$$\frac{\mathrm{d}}{\mathrm{d}t} m(t) = \nu_\mathrm{m} \left(\frac{1 + (C(t)/K_C)^2 f}{1 + (C(t)/K_C)^2} \right)^2 - \lambda_\mathrm{m} m(t) \qquad (4)$$

with the basal transcription rate ν_m, the degradation rate λ_m, the fold change f between basal and maximal transcription rates, and K_C denoting the binding threshold for CadC-DNA binding. The particular choice of the exponents in the first term is motivated by the observation that the *cadBA* promoter appears to be regulated by two binding sites for dimeric forms of CadC.[20]

Kinetics of enzyme expression and catalysis

On the protein level, we have a similar interplay of synthesis and decay as in Eq. (4):

$$\frac{\mathrm{d}}{\mathrm{d}t} A(t) = \nu_\mathrm{p} m(t) - \lambda_\mathrm{p} A(t) \qquad (5)$$

where A is the abundance of CadA per cell and ν_p and λ_p are the translation rate and the degradation rate, respectively. The level of the transporter CadB is taken to be proportional to that of CadA since they are translated from the same mRNA. We neglect possible post-transcriptional regulation, for which there seems to be no experimental indication.[11] We also assume that we can subsume the transport and turnover of lysine to cadaverine through CadB and CadA by a single effective reaction since little is known about the microscopic rates and affinities of the coupled transport and decarboxylase reactions. As detailed in Materials and Methods, this assumption leads us to:

$$\frac{\mathrm{d}}{\mathrm{d}t} l(t) = -\nu_\mathrm{max} A(t) \frac{l(t)}{K_\mathrm{m} + l(t)} \qquad (6)$$

This simplified reaction corresponds to an effective Michaelis–Menten process with external lysine as the substrate, an effective maximal lysine turnover rate ν_max, and an effective Michaelis constant K_m. Also implicit in Eq. (6) is the assumption that growth of the bacterial population over the period of the experiment is negligible. For the external cadaverine level, we assume the flux balance,

$$\frac{\mathrm{d}}{\mathrm{d}t} c(t) = -\frac{\mathrm{d}}{\mathrm{d}t} l(t) \qquad (7)$$

implying that the sum of external lysine and cadaverine is conserved at all times, $l(t) + c(t) = \mathrm{const}$. This flux balance appears justified, given experimental results with a LysP-deficient mutant strain that we report and discuss further below. We take the pH to be a constant over the duration of our kinetic experiments since the pH changes only very little in our buffered medium and the Cad module is not very sensitive to the pH over this regime (see above).

Data interpretation with the quantitative model

We now demonstrate that the simple quantitative model constructed above is indeed a powerful tool. First, we tested the extent to which this model is compatible with the data sets reported above. For this test, we also included another data set from

Table 1. Repression of the long-term Cad response by cadaverine

Initial cadaverine	Relative *cadBA* expression	Model value
0 μM	1.00	0.97
20 μM	0.89	0.92
80 μM	0.60	0.74
320 μM	0.12	0.12
1300 μM	0.00	0.05

The data were taken from the work of Neely *et al.*[12] *cadBA* expression (central column), as determined from the β-galactosidase activity of a *cadA-lacZ* fusion, was measured in cells that were grown in medium, pH 5.8, with 10 mM lysine and the indicated cadaverine concentrations for 3 h.[12] The model values (rightmost column) show the fit result of our quantitative model.

Neely it *et al.*,[12] who determined the cadaverine-dependent dose response of the Cad module (see Table 1). This additional data set further constrains our model and probes consistency with the existing literature.

In total, the quantitative model has 14 parameters. We constrained each of these to a range inferred from typical physiological values and other information in the literature (see Table 2). We then fitted our model to all data sets simultaneously using standard least-squares minimization of the residual χ^2, as described in Materials and Methods. The curves corresponding to the best-fit parameters are shown in Figs. 3 and 4 (blue lines). The overall agreement with the experimental data is good, both for the response dynamics in Fig. 4a–c and for the dose–response curves in Fig. 3 and Table 1. Note, however, that not all model parameters are individually well constrained by the data. This becomes apparent by plotting the correlations between the quality of the fit, characterized by the residual χ^2, and the fit parameters, as shown in Fig. 5. As the fit becomes better (lower χ^2), most parameter values are confined to a narrow interval, indicating that the information contained in the experimental data accurately determines their values. For instance, the *cadBA* mRNA lifetime τ_m determines the decay time of the transient expression peak in Fig. 4a and is therefore strongly constrained in our model. In contrast, some of the parameters display a wide variation even at the lowest χ^2 values (e.g., the transcription and translation rates ν_m and ν_p). For these cases, where individual parameters are "sloppy,"[35] certain combinations of these parameters are well

Table 2. Parameters of the quantitative model

	Parameter	LB	UB	Estimated value	Comment
Sensory module					
K_l	Threshold for CadC activation by lysine	1	20	$3.6\pm^{5.8}_{0.6}$ mM	Bounds suggested by Fig. 3a
K_c	Threshold for CadC inactivation by cadaverine	50	1000	$235\pm^{32}_{49}$ μM	Bounds suggested by Fig. 3a
n_l	Hill exponent for CadC regulation by lysine	1	5	$1.1\pm^{0.2}_{0.1}$	
n_c	Hill exponent for CadC regulation by cadaverine	1	5	$2.8\pm^{0.9}_{0.3}$	
pH_0	pH threshold for CadC activation	—	—	6.2	Estimated from data in Fig. 3b
ΔpH	Width of the transition from active to inactive CadC	—	—	0.5	Estimated from data in Fig. 3b
Expression module					
C_0/K_C	Total CadC per cell in relation to the threshold for CadC-promoter binding	0.1	10	$1.1\pm^{2.6}_{0.1}$	The level of CadC is just sufficient to activate the pathway,[18] suggesting that *in vivo* $C_0/K_C \approx 1$
ν_m	Basal transcription rate	0.001	0.1	$4.3\pm^{14.1}_{2.2} \times 10^{-3}$ min^{-1}	
f	Fold change between basal and maximal transcription rates	10	1000	$698\pm^{170}_{452}$	Typical range[32,33]
ν_p	Effective translation rate	10^{-4}	10^{-1}	$4.2\pm^{8.6}_{2.3} \times 10^{-3}$ U/mg/min	Effective parameter with broad range
τ_m	mRNA half-life ($\ln 2/\lambda_m$)	1	50	$13.8\pm^{0.4}_{1.2}$ min	Typical range[34]
τ_p	Protein half-life ($\ln 2/\lambda_p$)	1	10^4	$29\pm^{2137}_{4}$ h	CadA is expected to be stable[16]
ν_{max}	Maximal rate for lysine turnover via CadA and CadB	10^{-4}	10	$1.3\pm^{1.4}_{0.5} \times 10^{-3}$ mM/min/(U/mg)	Effective parameter with broad range
K_m	Effective Michaelis constant for lysine turnover via CadA and CadB	1	100	$26\pm^{37}_{12}$ mM	Effective parameter with broad range

LB indicates lower bound; UB, upper bound. The estimated parameter values are shown as (best-fit value)$\pm^{\sigma^+}_{\sigma^-}$, where σ^+ and σ^- indicate the asymmetric standard errors in the positive direction and in the negative direction, respectively, see Eq. (14) in Materials and Methods.

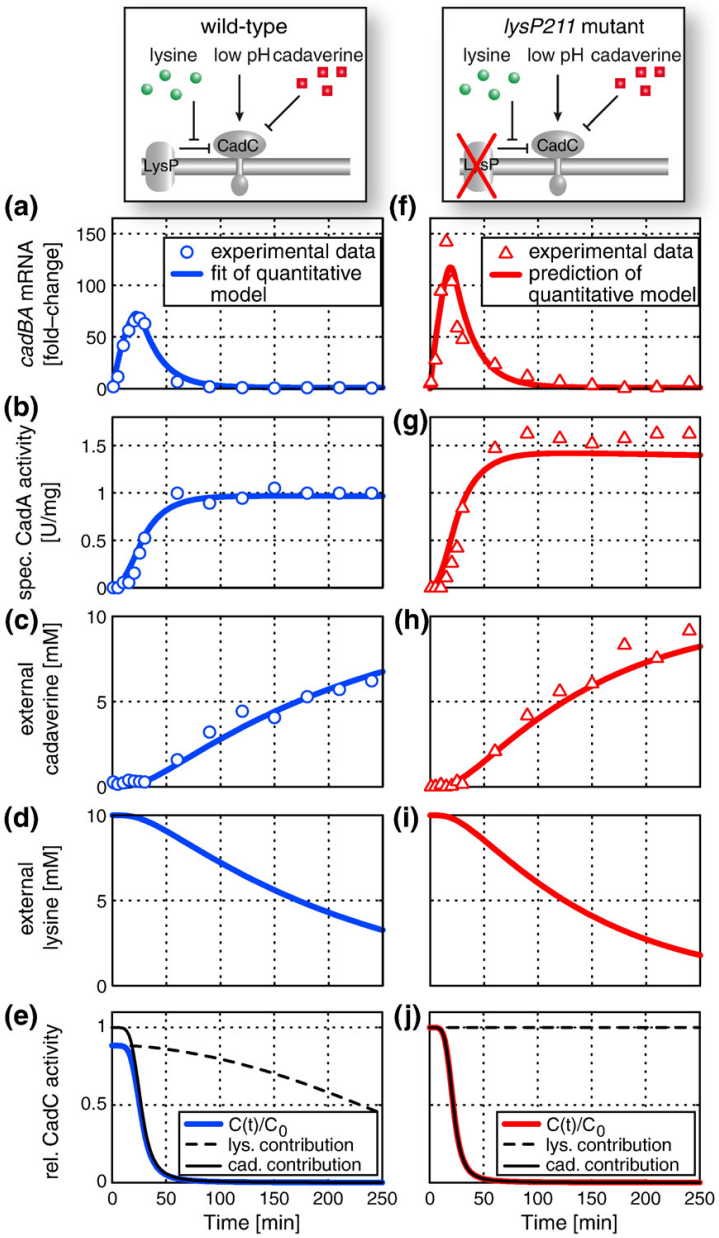

Fig. 4. Fit of our quantitative model (blue lines, left column) to the experimental data of the wild-type induction kinetics (blue circles) and the parameter-free prediction of our model for a LysP-deficient mutant (red lines, right column), together with the corresponding experimental data (red triangles), are shown. The external lysine concentration, shown in (d) and (i), and the relative CadC activity, shown in (e) and (j), have not been measured but were inferred from the quantitative model by a global fit to the wild-type data, as described in the main text and Materials and Methods. The dashed and continuous black lines in (e) and (j) show the lysine and cadaverine contributions, $g(l(t))$ and $h(c(t))$, in Eq. (3), to the signal integration function in Eq. (1).

constrained by the data sets. Pairwise scatter plots, as shown in Fig. S1, identify correlations and anti-correlations between the parameters and help to reveal the appropriate combinations. For instance, the product of the transcription and translation rates is much better determined by the data than the individual rates.

Please refer to Supplementary Material for the full discussion of our estimated parameter values in the light of previously published literature values. It is noteworthy that the best-fit value for the effective Hill coefficient n_c for the regulation of CadC by cadaverine is close to 3 and is relatively well constrained by the data. This suggests that a molecular mechanism for cooperativity is at work, possibly a multimerization of CadC proteins in the membrane. Another interesting observation from Table 2 concerns the half-life of *cadBA* mRNA, which was well constrained by the data to a value of almost 14 min. A global analysis of RNA half-lives in *E. coli*[36] found an extremely short half-life of less than 2 min for the *cadBA* mRNA, suggesting that an active degradation mechanism is involved. Our Northern blot data for the *lysP211* mutant, shown in Fig. 4f, do indeed

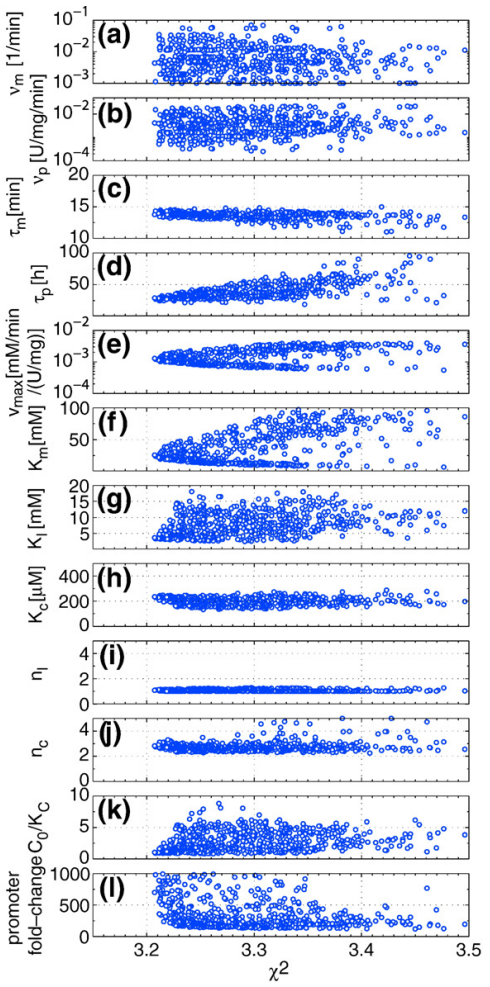

Fig. 5. Correlations between the goodness of fit and the estimated parameters. The points correspond to local optima in the parameter space, for which the difference between the quantitative model and the experimental data in Fig. 2, Fig. 3, and Table 1 is minimized (see Materials and Methods for all details). As the fit quality increases (lower χ^2), most parameters are confined to narrow intervals, indicating that their values are well constrained by the experimental data. However, some parameters display significant variation even for the lowest χ^2 values and one finds, from parameter–parameter correlation analysis in Fig. S1, that only combinations of those are well confined by our data.

suggest a rapid decay of the mRNA at high levels directly after the peak, followed by a slower decay at lower levels. Our quantitative model only allows for a single degradation rate, which leads to the intermediate half-life of 14 min as a best-fit value. The changing degradation rate could be rationalized under the assumption that the *cadBA* mRNA has a relatively weak binding affinity to the degrading enzyme such that active degradation only contributes significantly at high mRNA levels, whereas a slower passive decay is at work at low mRNA levels. Figure S2 illustrates that this mechanism could indeed account for the observed shape of the mRNA curve in Fig. 4f. However, this explanation would raise the question of why the rapid active decay is not observed in the data of the wild type in Fig. 4a as such. Possibly, the kinetics of LysP unbinding from CadC is slow (contrary to our model assumption of a rapid binding equilibrium) such that it interferes with the reception of the cadaverine signal and thereby broadens the onset of transcriptional down-regulation. Further experiments would be required to test these hypotheses.

Given the compatibility of the model and data, we next used the model to infer quantitative characteristics of the module that are not directly assayed. For instance, the activity of the central regulator CadC as a function of the external signals lysine, pH, and cadaverine is a biochemical characteristic that is pivotal to the function of the module but difficult to measure directly. In our model, this quantitative characteristic is represented by the signal integration function, Eq. (1). On the other hand, the experimental data shown in Fig. 3 probe the final output of the Cad module on the protein level (specific CadA activity), which is the integrated result of a nonlinear dynamical system with feedback. Figure 6 compares this final system output (black curves, as in Fig. 3, but normalized to 1) with the inferred activity of the CadC regulator (blue curves). The latter represents the three sigmoidal functions that make up the signal integration function, Eq. (1), with the parameters determined from the global fitting procedure described above. We observe from Fig. 6 that the final system output behaves qualitatively similar to the inferred biochemical activity of CadC. However, in each case, the apparent activation threshold for the system response (point of half-maximal CadA activity, black curve) is shifted with respect to the inferred biochemical activation threshold (point of half-maximal CadC activity, blue curve). These shifts are due to the fact that the total CadA activity depends not only on the characteristics of the regulator but also on the biochemical properties and timescales of the negative feedback loop. In principle, the feedback can even lead to nonmonotonic behavior in the dose response, despite the underlying monotonic dependence of the signal integration function on the levels of the external signals†. The inferred *in vivo* biochemical activation thresholds, K_I=3.6 mM and K_c=235 µM, and the values of the Hill coefficients describing the sensitivities to the signals can be read off directly from the blue curves, as indicated in Fig. 6.

It is useful to compare these biochemical activation thresholds with the actual concentrations

† The model predicts indeed a weak nonmonotonic effect in the lysine dependence; however, this is not a robust prediction.

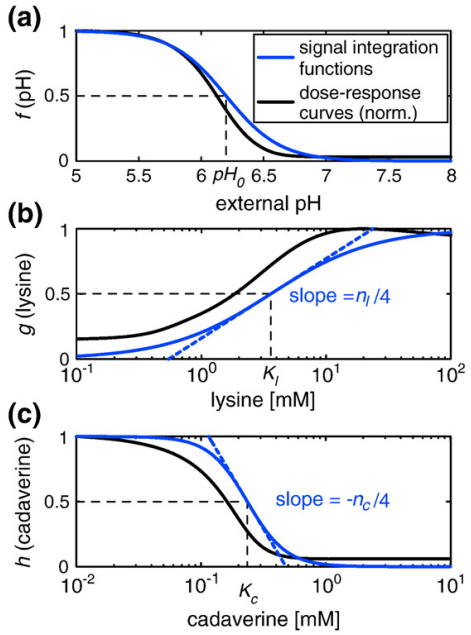

Fig. 6. Comparison of the extracted signal integration functions (blue curves) with the normalized dose–response curves (black curves). The signal integration functions describe the dependence of CadC activity on pH level (a), on lysine (b), and on cadaverine (c). They correspond to functions *f*, *g*, and *h* in Eqs. (2) and (3) and are plotted for the best-fit parameters listed in Table 2. The apparent activation thresholds of the dose–response curves (black curves) are shifted with respect to the inferred biochemical thresholds of CadC (blue curves) since the dose–response curves also depend on the biochemical properties of the nonlinear feedback in the Cad module. From the extracted signal integration functions, we can also read off the sensitivity of CadC on its input signals since their maximal slopes are determined by the Hill coefficients n_l and n_c.

encountered in our kinetic induction experiments, in particular at the time of transcriptional down-regulation ($t = 30$ min). From the plots in Fig. 4c and d, we see that the lysine concentration at this time point is still about a factor of 3 higher than the activation threshold, whereas the cadaverine concentration of 300 μM exceeds the deactivation threshold. The individual regulatory contributions from lysine and cadaverine to the CadC activity [i.e., $g(l(t))$ and $h(c(t))$] are plotted in Fig. 4e. The lysine curve (dashed line) displays only a very weak impact on the CadC activity, whereas the increase of cadaverine is the primary effect causing the down-regulation of the CadC activity (continuous line). Hence, the analysis with our quantitative model strongly suggests that the negative feedback via external cadaverine can quantitatively explain the timing of the transient response in the wild-type Cad system and that the decreasing lysine stimulus is not involved in this behavior.

Prediction and experimental analysis under altered conditions

We have so far analyzed the Cad module only in the wild-type strain and only with a single induction protocol. To obtain a more complete picture of its quantitative behavior, we constructed a mutant strain, MG1655-*lysP211*, with a truncated and inactive form of the lysine permease LysP (see Materials and Methods for details). Qualitatively, we expected that this mutation would completely abolish the lysine requirement for the activation of CadC since the truncated form of LysP would be unable to repress CadC[23] (an early study also indicated a derepressed activation of the Cad system by a spontaneous mutation in *lysP*[9]). Within our quantitative model, the *lysP211* mutation was mimicked by setting the lysine-dependent activity function in Eq. (1) equal to its maximal value [i.e., $g(l) = 1$], thereby rendering it independent of the inducing lysine concentration. Apart from this "*in silico* mutation," we left the model and the parameter values unchanged.

We verified that the LysP-deficient mutant was indeed inducible by a shift from pH 7.6 to pH 5.8 alone and did not require lysine for its induction (data not shown). We then performed the same kinetic induction experiments as with the wild-type strain [see the results in Fig. 4f–j for the *lysP211* mutant (red triangles) and the quantitative model (red curves)]. We clearly saw that the expression in the mutant remains transient, again supporting the conclusion that a signal different from lysine mediates the transcriptional down-regulation. The shape of the response is altered, however, with a stronger initial expression and a peak in the mRNA level that has a larger amplitude and reaches its maximum earlier than for the wild-type strain. Exactly these features are expected also on the basis of the quantitative model: The stronger initial expression is due to the full relief of the LysP repression, and the negative feedback via cadaverine sets in at an earlier time since the cadaverine threshold is more rapidly reached (see Fig. 4h and j). Also, the CadA activity, shown in Fig. 4g, is expected to reach a higher steady-state plateau, as observed in the experiment. Interestingly, at the end of the experiment, the cadaverine level in Fig. 4h reaches almost the 10 mM level of initially added lysine, which is in line with our flux balance assumption in the model (see previous section).

Next, we considered altered environmental conditions for the induction of the Cad module. Neely *et al.*[12] had already shown that induction with initially added cadaverine causes a significantly reduced long-term activity of the *cad* operon (cf., Table 1). Yet, it is not clear whether this diminished long-term activity is caused by a *cadBA* expression pulse of similar strength but reduced duration, by a pulse of reduced strength with similar duration, or by a combination of both. To resolve this question, we again performed kinetic induction experiments with the wild-type strain under identical conditions

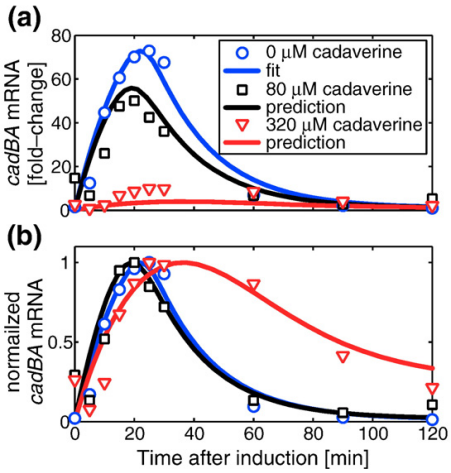

Fig. 7. Experimental test of the kinetic model prediction. The lines show the parameter-free model prediction for induction of the Cad system by a shift to pH 5.8 and 10 mM lysine, together with the cadaverine concentration indicated in the legend. The experimental data in (a) (symbols) were recorded and scaled as described in Materials and Methods. In (b), all data were normalized to their maximal values.

as in Fig. 2 but additionally with 80 or 320 μM cadaverine supplied at the time of induction. In the latter case, the initially supplied cadaverine already slightly exceeded the inferred inactivation threshold of $K_c = 235$ μM such that a strong effect on the response could be expected. The resulting data are shown in Fig. 7a (squares and triangles), together with the original data (no added cadaverine, circles) for comparison. Figure 7b shows the same data but with all curves normalized to peak height 1 in order to emphasize the shape of the response. We observed that the primary effect of the addition of initial cadaverine is to reduce the strength of the response. This is also predicted by the quantitative model (continuous lines), rather accurately for the 80 μM cadaverine data set, while the reduction for 320 μM cadaverine is predicted to be stronger than that observed experimentally.

It should be noted that transcription was detected by Northern blot analysis, so that the shape of the response could be more accurately determined than the absolute amplitude. It is then interesting to observe that the quantitative model predicts a widening of the expression peak and a shift of the maximum to a later time for the highest cadaverine concentration (320 μM, red curve). Again, this behavior results from the nonlinear feedback in the Cad module—that is, the addition of initial cadaverine decreases the initial rate of mRNA production such that the CadA level increases more slowly and the negative feedback also sets in more slowly. This predicted change in the shape of the expression peak for 320 μM agrees remarkably well with the experimental observation (see Fig. 7b). This finding provides strong evidence that the quantitative characteristics of the Cad module are well described and understood with the help of our mathematical model.

Discussion

Conditional stress response with feedback inhibition

In this work, we analyzed the kinetics of a conditional pH stress response system, the lysine decarboxylase system of *E. coli*, which exhibits only transient induction, even when pH stress persists. Our results strongly suggest that the additional stimulus for the conditional response (i.e., a lysine-rich environment) is also not responsible for the transient behavior. Rather, our kinetic and dose-response experiments in combination with our quantitative model clearly indicate that negative feedback via the product of the decarboxylation reaction, cadaverine, leads to the down-regulation of the response.

Cadaverine has previously been linked to the transient behavior in pivotal work on this stress response system.[12,13] This link was based on observations that external addition of cadaverine significantly reduces the long-term Cad activity and that a CadA$^-$ mutant displayed persistent *cadBA* expression. However, if the transcriptional shutoff was mediated by a decrease of the lysine stimulus, the persistent *cadBA* expression could alternatively be explained by the lack of lysine consumption in the CadA$^-$ mutant. Also, while a reduction in the steady-state activity implies that the Cad system is generally repressed by cadaverine, it is not clear that the timing of its down-regulation is set via this negative regulatory interaction. None of the previous studies directly measured the system-induced dynamics of the cadaverine concentration or studied the kinetics of the Cad system with externally added cadaverine. By performing these quantitative experiments with a high time resolution and by interpreting them with a quantitative theoretical model, we obtained evidence for a causal relation between the time of transcriptional down-regulation and the increase of the external cadaverine concentration above its deactivation threshold. For instance, the decreasing amplitude in the dynamical response after adding external cadaverine and a more subtle delayed down-regulation, both predicted by the quantitative model, were strikingly confirmed by our kinetic measurements.

Signal transduction mechanism

Our system-level study of the Cad module also permits some conclusions about the involved molecular interactions and the signal transduction mechanism. Using our quantitative model, we were able to estimate the relevant *in vivo* activation and

deactivation thresholds of the regulatory protein CadC. For the inactivation of CadC by external cadaverine, we found a threshold of $K_c = 235$ μM. This value is surprisingly close to the *in vitro* binding constant of 96 μM for the interaction of cadaverine with the periplasmic domain of CadC.[23] In contrast, CadC has almost no affinity for lysine. There is recent evidence that CadC is inhibited at low lysine concentrations via a transmembrane domain interaction with the lysine permease LysP, whereas the interaction is released at high lysine levels.[23] In the present work, we determined the effective *in vivo* lysine activation threshold to be $K_l = 3.6$ mM. This result is somewhat surprising since the Michaelis constant K_m for lysine transport by LysP is much lower at ~10 μM.[37] However, the LysP–CadC interaction and the K_m of LysP do not necessarily need to have a direct correspondence.

An interesting open question concerns the signal transduction mechanism of CadC. Two alternative models have been proposed: (i) a reversible conformational transition of CadC activates its cytoplasmic N-terminal domain and allows it to bind to the promoter while remaining integrated in the membrane and (ii) a shift to acidic pH induces cleavage of the cytoplasmic domain, allowing it to diffuse freely to the promoter. The existence of the negative feedback by external cadaverine puts causal constraints on these microscopic mechanisms. If the cleavage mechanism were realized, it is not clear how external cadaverine could down-regulate *cadBA* expression after it has been induced since the freely diffusing cytoplasmic domains would no longer be able to recognize this signal. A high turnover of cleaved CadC could solve this problem by rapidly eliminating unresponsive activators. However, the increased degradation of CadC under inducing conditions would have to be balanced by an elevated *cadC* expression. This, however, is in disagreement with previous observations where it was found that *cadC* expression is constitutive.[18,19] Hence, the existence of the negative feedback by external cadaverine leads us to favor the reversible model.

Top–down system analysis of a functional module

The Cad system of *E. coli* is a functional module with few closely connected molecular components. How such modules integrate, process, and respond to external signals is a central question—but generally also a difficult one. The approach taken in the current study is akin to a "top–down" system analysis, where input signals are controlled and the output(s) and key internal system variables are measured. This is in contrast to a biochemical "bottom–up" approach, where each component would first be characterized separately, followed by their pairwise interactions, gradually moving upward in complexity. In our study, we were able to bring these two complementary approaches into a first contact for the Cad module, with the help of our quantitative model that provided the means to estimate relevant *in vivo* values for biochemical interaction parameters and the quantitative form of the signal integration function displayed in Fig. 6. Without the quantitative model, we would have been unable to extract this "hidden information" from the experimental data. Direct *in vivo* measurements of signal integration functions have been performed for the regulatory circuit controlling chemotaxis in *E. coli* using sophisticated single-molecule techniques. Since these powerful techniques are not easily transferred to the large class of functional modules of interest, the indirect approach taken in the present work may often be a welcome alternative.

Conclusions and outlook

Our quantitative analysis of the Cad system provides a first step toward a system-level understanding of the complex acid stress response network in microbes. Analogous quantitative studies of the two other major amino acid decarboxylase systems, glutamate and arginine decarboxylase,[8] could reveal important insights into how these modules are orchestrated in the complex environment of their host. The presence of multiple amino acids in the natural environment of *E. coli* suggests that these conditional stress response systems are often induced in parallel. It will be interesting to study how these systems are coordinated to provide an effective and robust way of pH homeostasis and acid tolerance response.

Materials and Methods

Bacterial strains and growth conditions

E. coli MG1655[38] was used as wild-type strain. The *lysP211* mutant was obtained via undirected mutagenesis.[9,39] For this purpose, cells of *E. coli* MG1655 were grown on minimal agar plates with 0.2% (w/v) glucose as sole carbon source[40] containing 100 μg/mL of thiosine (*S*-aminoethyl cysteine), a toxic lysine analog that leads to spontaneous mutations in *lysP*. One mutant, designated MG1655-*lysP211*, had a nucleotide exchange at position 211 in *lysP*, resulting in a stop codon and hence a truncated and inactive form of LysP (70 amino acids). *E. coli* strains MG1655 and MG1655-*lysP211* were grown aerobically in shaking flasks under noninducing conditions at pH 7.6 in 5 L of phosphate-buffered minimal medium[41] containing 0.4% (w/v) glucose as sole carbon source at 37 °C to an optical density at 600 nm of 0.5 (noninducing conditions). Subsequently, cells were collected by centrifugation (10 min, 4000*g* at 37 °C) and transferred into fresh prewarmed minimal medium, pH 5.8, containing 10 mM L-lysine (L-lysine-hydrochloride, Roth) and 0.4% (w/v) glucose in a 5-L fermenter (Biostat B, Satorius BBI Systems GmbH) (inducing conditions). Cultivation of cells was continued anaerobically at 37 °C. At the indicated times, samples were taken and then centrifuged at 4000*g* (4 °C) for 5 min, and cell pellets as well as supernatants were separately stored at −80 °C until further use. The number of colony-forming units was determined after incubation of 100 μL of various dilutions on LB agar plates overnight at 37 °C.[42]

Lysine decarboxylase assay

Specific activity of the lysine decarboxylase CadA was measured by resuspending cells corresponding to 10 mL of culture in 1 mL of Ldc buffer [100 mM Na-acetate, pH 6.0, 1 mM ethylenediaminetetraacetic acid, 0.1 mM pyridoxal phosphate, 10 mM β-mercaptoethanol, 10% (w/v) glycerol]. Lysozyme (0.1 mg/mL) was added, and the mixture was incubated on ice for 30 min. The cell suspension was sonified (10%–20% amplitude; 0.5-s output pulse; digital Branson Sonifier II 250), and the lysate was centrifuged at 15,000*g* at 4 °C for 15 min. Activity of lysine decarboxylase in cell-free extracts was measured as described previously[43] using 5 μg of protein per assay. Specific activity is defined as 1 U/mg = 1 μmol cadaverine/(min × mg protein).

Measurement of extracellular cadaverine

The extracellular cadaverine concentration was determined according to the spectrophotometric method described by Phan *et al.*[44] Briefly, 10 μL of culture supernatant was diluted 5-fold with H_2O_{dest}, 120 μL of Na_2CO_3 (1 M) and 120 μL of TNBS (2,4,6-trinitrobenzene-sulfonic acid; 10 mM) were then added, and the mixture was incubated for 4 min at 40 °C. After extraction with 1 mL of toluene, the absorption of the organic phase (containing *N*,*N*-bistrinitrophenylcadaverine) at 340 nm was measured. The cadaverine concentration was calculated based on a standard curve using cadaverine dihydrochloride (Sigma) between 0 and 500 nmol.

Preparation of RNA

Total RNA was isolated according to the method of Aiba *et al.*[45] Briefly, cells were resuspended in cold 20 mM Tris–HCl, pH 8.0, and subsequently lysed by addition of 20 mM sodium acetate, pH 5.5, 0.5% (w/v) SDS, and 1 mM ethylenediaminetetraacetic acid, pH 8.0. Then, RNA was extracted with prewarmed (60 °C) acid phenol, and the mixture was centrifuged at 12,000*g*. After an additional extraction of RNA using phenol/chloroform/isoamyl alcohol (25:24:1), RNA was precipitated with 100% ethanol at −20 °C overnight. The precipitate was washed with 70% (v/v) ethanol, and the dry RNA pellet was dissolved in 35 μL of H_2O_{dest}. RNA concentration was determined by measuring the absorption at 260 nm. All solutions were prepared with 0.1% (v/v) DEPC (diethylpyrocarbonate).

Northern blot analysis

cadBA mRNA was quantitatively determined by Northern blot analysis. This technique does not allow determination of the absolute number of *cadBA* transcripts, but it determines the relative induction level (ratio between induced cells and pre-induced cells) by using a radioactive labeled probe that specifically detects *cadBA* mRNA. Nevertheless, the readout scales linearly with *cadBA* mRNA concentration. Northern blot analysis was performed following a protocol described earlier.[46] Briefly, 20 μg of RNA was separated by electrophoresis in 1.2% (w/v) agarose–1.1% (v/v) formaldehyde gels in Mops (4-morpholinepropanesulfonic acid) buffer. RNA was transferred to a Hybond nylon membrane (GE Healthcare) by capillary blotting. Hybridization was performed following a standard protocol[42] using an [α-^{32}P]dCTP-labeled PCR fragment of the first 400 base pairs of *cadBA*. Radioactive labeling was quantified with a PhosphoImager. As control, expression of *rpoD*, a housekeeping gene of *E. coli*, was analyzed. Signal intensity of *cadBA* mRNA was normalized to the signal intensity of *rpoD* mRNA. If not indicated otherwise, the data are given as fold change of *cadBA* transcription relative to the pre-induction value. Additionally, in Figs. 4 and 7, the absolute magnitude of the mRNA fold change was rescaled such that the integral over the expression curve was proportional to the long-term CadA activity.

Model details

From thermodynamic models of transcriptional regulation reviewed by Bintu *et al.*,[30,31] the effective transcription rate, ν_m^{eff}, as a function of two activators, A and B, with independent binding sites (binding constants K_A and K_B) is given by

$$\nu_m^{eff} = \nu_m \left(\frac{1 + (A/K_A)f_A}{1 + (A/K_A)}\right)\left(\frac{1 + (B/K_B)f_B}{1 + (B/K_B)}\right) \qquad (8)$$

cf., Ref. 30, Table 1, case 10. We make the simplifying assumption that the binding sites are identical and that

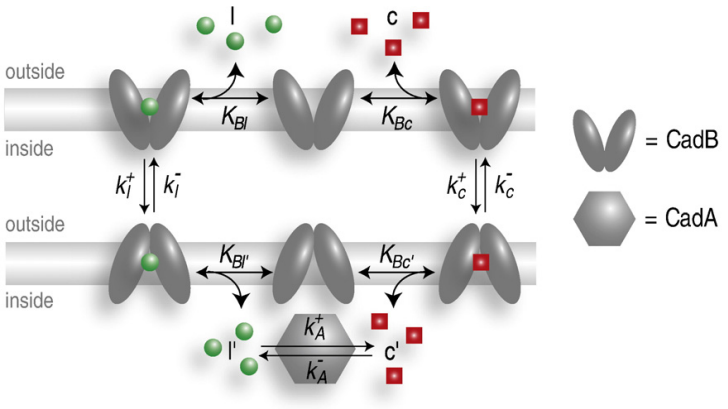

Fig. 8. Kinetic scheme of lysine and cadaverine transport and turnover. Antiport of lysine (*l*) and cadaverine (*c*) by CadB is modeled in analogy to the homologous arginine–ornithine antiporter ArcD in *L. lactis* by a single-site Ping Pong Bi–Bi mechanism.[47] Interconversion of lysine and cadaverine by CadA is modeled by a reversible first-order reaction.[48] The prime (′) indicates internal quantities. The rates (k_l^+, k_l^-, k_c^+, k_c^-, k_A^+, k_A^-) and equilibrium constants (K_{Bl}, $K_{Bl'}$, K_{Bc}, $K_{Bc'}$) are indicated next to the reaction steps.

each site can only be bound by a dimer of CadC (C_2).[20] Setting $A=B=C_2$, exploiting mass action $K=C^2/C_2$, and introducing the effective binding constant $K_C=\sqrt{KK_{C_2}}$ lead to the first term in Eq. (4).

The effective Michaelis–Menten form of the lysine turnover rate in Eq. (6) was derived as follows: The transport of lysine and cadaverine via CadB was modeled in analogy to the homologous arginine–ornithine antiporter ArcD in *Lactococcus lactis*[47] (see Fig. 8). For low external cadaverine and internal lysine concentrations, a general form of the inwardly directed flux of lysine is as follows:[47]

$$v_{\text{eff}} = k_l^+ B \frac{[\text{l}]}{K_{Bl} + \left(1 + \frac{k_l^+}{k_c^-} + \frac{k_l^+ K_{Bc'}}{k_c^- [\text{c}']}\right)[\text{l}]} \tag{9}$$

where B is the number of CadB molecules per cell and [l] and [c] are the concentrations of lysine and cadaverine on the outer surface of the membrane, respectively. Internal solute concentrations are marked by a prime (′). The parameters k_l^+, k_c^-, K_{Bl}, and $K_{Bc'}$ are defined in the kinetic scheme of Fig. 8. Here, it is assumed that (i) the conformational transition in CadB, which mediates the transport, does not occur without bound lysine or cadaverine, (ii) the antiporter and its substrates are in binding equilibrium at each surface of the membrane, and (iii) the membrane translocation reaction of the carrier is slow and rate limiting. The internal cadaverine concentration [c′] is determined by an interplay of the (reversible) decarboxylation through CadA[48] and by the export via CadB. In steady state, we find

$$[\text{c}'] = \frac{k_A^+}{k_A^-}[\text{l}'] - \frac{v_{\text{eff}}}{k_A^- A} \tag{10}$$

where A is the number of CadA molecules per cell. If the equilibration between lysine and cadaverine through CadA is fast compared with the transport through CadB, the second term is negligible and the internal cadaverine level is solely determined by the internal lysine level [l′]. In steady state, it turns out that [l′] is not affected by the Cad module, since the 1:1 stoichiometry of the antiporter ensures that lysine decarboxylation is balanced by lysine import. If we further take advantage of the fact that CadB and CadA are transcribed polycystronically and that there seems to be no post-transcriptional regulation,[11] we can set CadB proportional to CadA (i.e., $B=\alpha A$). Taken together, in the limit of rapid CadA kinetics and low external cadaverine and internal lysine concentrations, Eq. (9) reduces to the simple effective Michaelis–Menten form in Eq. (6) with CadA as the enzyme, $v_{\max}=(\alpha\times k_l^+)/\eta$ as the effective maximal turnover rate, and $K=K_{Bl}/\eta$ as the effective Michaelis constant, where $\eta = 1 + \frac{k_l^+}{k_c^-} + \frac{k_l^+ k_A^- K_{Bc'}}{k_c^- k_A^+ [\text{l}']}$.

Parameter estimation

The parameters of our quantitative model were estimated by using a trust-region reflective Newton method (MATLAB, The MathWorks, Inc.) to minimize the total χ^2, defined by

$$\chi^2(\vec{\theta}) = \chi^2_{\text{kin}}(\vec{\theta}) + \chi^2_{\text{pH}}(\vec{\theta}) + \chi^2_{\text{lys}}(\vec{\theta}) + \chi^2_{\text{cad}}(\vec{\theta}) \tag{11}$$

with respect to the parameter vector $\vec{\theta}=(\theta_1,\ldots,\theta_M)$, where $M=14$ is the total number of model parameters (cf., Table 2). The contribution of the kinetic data is calculated from

$$\chi^2_{\text{kin}}(\vec{\theta}) = \sum_{i=1}^{3} \frac{1}{N_i \overline{\Delta t_i}} \sum_{j=1}^{N_i} \Delta t_{ij} \left(\frac{y_{ij} - \tilde{y}_{ij}(\vec{\theta})}{\sigma_{ij}} \right)^2 \tag{12}$$

where y_{1j}, y_{2j}, and y_{3j} are the experimental data of the *cadBA* mRNA, the CadA activity, and the cadaverine concentration at time t_j, respectively. Similarly, $\tilde{y}_{ij}(\vec{\theta})$ denotes the corresponding values of the quantitative model for a given parameter set $\vec{\theta}$ and σ_{ij} is the standard error of each measurement (estimates from our experiments: $\sigma_{1j}=5$, $\sigma_{2j}=0.1$ U/(mg protein), and $\sigma_{3j}=0.5$ mM $\forall j$). N_i is the number of data points of a given time series, $\Delta t_{ij} \equiv (t_{i(j+1)} - t_{i(j-1)})/2$ is the time between subsequent data points, and $\overline{\Delta t_i}$ is the mean time between the data points in data set i. The contributions of the dose–response curves are similarly defined and are exemplarily shown for the pH-dependent response:

$$\chi^2_{\text{pH}}(\vec{\theta}) = \frac{1}{N_{\text{pH}}} \sum_{i=1}^{N_{\text{pH}}} \left(\frac{A_\tau(\text{pH}_i) - \tilde{A}_\tau(\text{pH}_i, \vec{\theta})}{\sigma_i} \right)^2 \tag{13}$$

where $A_\tau(\text{pH}_i)$ and $\tilde{A}_\tau(\text{pH}_i, \vec{\theta})$ are the experimental and theoretical CadA activities after time τ at a given pH level pH_i ($i=1,..N_{\text{pH}}$), respectively. σ_i denotes the standard error of the measurement, and N_{pH} is the total number of data points. The other contributions, χ^2_{lys} and χ^2_{cad}, are defined equivalently to Eq. (13). However, they differ in the time at which the CadA activity was determined experimentally—that is, $\tau_{\text{pH}}=1.5$ h, $\tau_{\text{lys}}=1.5$ h, and $\tau_{\text{cad}}=3$ h.

To account for the presence of local optima and to quantify the uncertainty in the estimated parameters, we performed 1000 independent fits with randomly chosen initial parameter sets (within their physiological ranges). In Fig. 5, the final χ^2 values are plotted against the final parameters. We followed Wall *et al.*[49] to compute the asymmetric errors σ_+ and σ_- with respect to the optimal parameter values $\vec{\theta}^{\text{opt}}$ listed in Table 2. The squared errors for parameter θ_k were calculated using the following equations:

$$\sigma^2_{k,+} = \frac{\sum_{i:\theta_{k,i}>\theta_k^{\text{opt}}} \left(\theta_{k,i} - \theta_k^{\text{opt}}\right)^2 e^{-\chi_i^2/2}}{\sum_{i:\theta_{k,i}>\theta_k^{\text{opt}}} e^{-\chi_i^2/2}} \quad \text{and}$$
$$\sigma^2_{k,-} = \frac{\sum_{i:\theta_{k,i}<\theta_k^{\text{opt}}} \left(\theta_{k,i} - \theta_k^{\text{opt}}\right)^2 e^{-\chi_i^2/2}}{\sum_{i:\theta_{k,i}<\theta_k^{\text{opt}}} e^{-\chi_i^2/2}} \tag{14}$$

where $\sigma_{k,i}$ is the value of parameter θ_k in the ith fit, θ_k^{opt} is the value of θ_k in the fit with the lowest value of χ^2, and χ_i^2 is the value of χ^2 for the ith fit. In using the likelihood function $e^{-\chi^2/2}$, we assume that the errors in the measurements are independent and normally distributed with widths equal to the standard error of the mean.

Acknowledgements

It is a pleasure to thank T. Hwa and M. Saier for helpful discussions. This work was supported by the Excellence Cluster "Nanosystems Initiative Munich"

and by Deutsche Forschungsgemeinschaft through grants JU270/5-3 and Exc114/1. I. H. was supported by a fellowship from Elite Netzwerk Bayern.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.08.037

References

1. Gale, E. F. & Epps, H. M. R. (1942). The effect of the pH of the medium during growth on the enzymic activities of bacteria (*Escherichia coli* and *Micrococcus lysodeikticus*) and the biological significance of the changes produced. *Biochem. J.* **36**, 600–618.
2. Booth, I. R. (1985). Regulation of cytoplasmic pH in bacteria. *Microbiol. Rev.* **49**, 359–378.
3. Park, Y.-K., Bearson, B., Bang, S. H., Bang, I. S. & Foster, J. W. (1996). Internal pH crisis, lysine decarboxylase and the acid tolerance response of *Salmonella typhimurium*. *Mol. Microbiol.* **20**, 605–611.
4. Richard, H. & Foster, J. W. (2004). *Escherichia coli* glutamate- and arginine-dependent acid resistance systems increase internal pH and reverse transmembrane potential. *J. Bacteriol.* **186**, 6032–6041.
5. Castanie-Cornet, M.-P., Penfound, T. A., Smith, D., Elliott, J. F. & Foster, J. W. (1999). Control of acid resistance in *Escherichia coli*. *J. Bacteriol.* **181**, 3525–3535.
6. Booth, I. R., Cash, P. & O'Byrne, C. (2002). Sensing and adapting to acid stress. *Antonie van Leeuwenhoek*, **81**, 33–42.
7. Merrell, D. S. & Camilli, A. (2002). Acid tolerance of gastrointestinal pathogens. *Curr. Opin. Microbiol.* **5**, 51–55.
8. Foster, J. W. (2004). *Escherichia coli* acid resistance: tales of an amateur acidophile. *Nat. Rev. Microbiol.* **2**, 898–907.
9. Popkin, P. S. & Maas, W. K. (1980). *Escherichia coli* regulatory mutation affecting lysine transport and lysine decarboxylase. *J. Bacteriol.* **141**, 485–492.
10. Meng, S. Y. & Bennett, G. N. (1992). Nucleotide sequence of the *Escherichia coli cad* operon: a system for neutralization of low extracellular pH. *J. Bacteriol.* **174**, 2659–2669.
11. Meng, S. Y. & Bennett, G. N. (1992). Regulation of the *Escherichia coli cad* operon: location of a site required for acid induction. *J. Bacteriol.* **174**, 2670–2678.
12. Neely, M. N., Dell, C. L. & Olson, E. R. (1994). Roles of LysP and CadC in mediating the lysine requirement for acid induction of the *Escherichia coli cad* operon. *J. Bacteriol.* **176**, 3278–3285.
13. Neely, M. N. & Olson, E. R. (1996). Kinetics of expression of the *Escherichia coli cad* operon as a function of pH and lysine. *J. Bacteriol.* **178**, 5522–5528.
14. Torres, A. G. (2009). The *cad* locus of *Enterobacteriaceae*: more than just lysine decarboxylation. *Anaerobe*, **15**, 1–6.
15. Moreau, P. L. (2007). The lysine decarboxylase CadA protects *Escherichia coli* starved of phosphate against fermentation acids. *J. Bacteriol.* **189**, 2249–2261.
16. Sabo, D. L., Boeker, E. A., Byers, B., Waron, H. & Fischer, E. H. (1974). Purification and physical properties of inducible *Escherichia coli* lysine decarboxylase. *Biochemistry*, **13**, 662–670.
17. Soksawatmaekhin, W., Kuraishi, A., Sakata, K., Kashiwagi, K. & Igarashi, K. (2004). Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. *Mol. Microbiol.* **51**, 1401–1412.
18. Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L. & Olson, E. R. (1992). Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J. Bacteriol.* **174**, 530–540.
19. Dell, C. L., Neely, M. N. & Olson, E. R. (1994). Altered pH and lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. *Mol. Microbiol.* **14**, 7–16.
20. Küper, C. & Jung, K. (2005). CadC-mediated activation of the *cadBA* promoter in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **10**, 26–39.
21. Pfau, J. D. & Taylor, R. K. (1998). Mutations in *toxR* and *toxS* that separate transcriptional activation from DNA binding at the cholera toxin gene promoters. *J. Bacteriol.* **180**, 4724–4733.
22. Laub, M. T. & Goulian, M. (2007). Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.* **41**, 121–145.
23. Tetsch, L., Koller, C., Haneburger, I. & Jung, K. (2008). The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol. Microbiol.* **67**, 570–583.
24. Lee, Y. H., Kim, J. H., Bang, I. S. & Park, Y. K. (2008). The membrane-bound transcriptional regulator CadC is activated by proteolytic cleavage in response to acid stress. *J. Bacteriol.* **190**, 5120–5126.
25. Jung, K. & Altendorf, K. (2002). Towards an understanding of the molecular mechanisms of stimulus perception and signal transduction by the KdpD/KdpE system of *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **4**, 223–228.
26. Klipp, E., Nordlander, B., Kruger, R., Gennemark, P. & Hohmann, S. (2005). Integrative model of the response of yeast to osmotic shock. *Nat. Biotechnol.* **23**, 975–982.
27. Wahde, M. & Hertz, J. (2000). Coarse-grained reverse engineering of genetic regulatory networks. *BioSystems*, **55**, 129–136.
28. Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L. & Palsson, B.Ø. (2009). Reconstruction of biochemical networks in microorganisms. *Nat. Rev. Microbiol.* **7**, 129–143.
29. Sourjik, V. & Berg, H. C. (2002). Receptor sensitivity in bacterial chemotaxis. *Proc. Natl Acad. Sci. USA*, **99**, 123–127.
30. Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J. & Phillips, R. (2005). Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* **15**, 116–124.
31. Bintu, L., Buchler, N. E., Garcia, H. G., Gerland, U., Hwa, T., Kondev, J. *et al.* (2005). Transcriptional regulation by the numbers: applications. *Curr. Opin. Genet. Dev.* **15**, 125–135.
32. Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1–I2 regulatory elements. *Nucleic Acids Res.* **25**, 1203–1210.
33. Kuhlman, T., Zhang, Z., Saier, M. H., Jr & Hwa, T. (2007). Combinatorial transcriptional control of the lactose operon of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **104**, 6043–6048.

34. Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S. & Cohen, S. N. (2002). Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. *Proc. Natl Acad. Sci. USA*, **99**, 9697–9702.
35. Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R. & Sethna, J. P. (2007). Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput. Biol.* **3**, 1871–1878.
36. Selinger, D. W. & Rosenow, C. (2003). Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.* **13**, 216–223.
37. Rosen, B. P. (1971). Basic amino acid transport in *Escherichia coli*. *J. Biol. Chem.* **246**, 3653–3662.
38. Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M. *et al.* (1997). The complete genome sequence of *Escherichia coli* K-12. *Science*, **277**, 1453–1474.
39. Tabor, H., Hafner, E. W. & Tabor, C. W. (1980). Construction of an *Escherichia coli* strain unable to synthesize putrescine, spermidine, or cadaverine: characterization of two genes controlling lysine decarboxylase. *J. Bacteriol.* **144**, 952–956.
40. 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, Cold Spring Harbor, NY.
41. Epstein, W. & Kim, B. S. (1971). Potassium transport loci in *Escherichia coli* K-12. *J. Bacteriol.* **108**, 639–644.
42. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
43. Lemonnier, M. & Lane, D. (1998). Expression of the second lysine decarboxylase gene of *Escherichia coli*. *Microbiology*, **144**, 751–760.
44. Phan, A. P. H., Ngo, T. T. & Lenhoff, H. M. (1982). Spectrophotometric assay for lysine decarboxylase. *Anal. Biochem.* **20**, 193–197.
45. Aiba, H., Adhya, S. & de Crombrugghe, B. (1981). Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J. Bacteriol.* **256**, 11905–11910.
46. Jung, K., Krabusch, M. & Altendorf, K. (2001). Cs^+ induces the *kdp* operon by lowering the intracellular K^+ concentration. *J. Bacteriol.* **183**, 3800–3803.
47. Driessen, A. J., Molenaar, D. & Konings, W. N. (1989). Kinetic mechanism and specificity of the arginine–ornithine antiporter of *Lactococcus lactis*. *J. Biol. Chem.* **264**, 10361–10370.
48. Koppelman, R., Mandeles, S. & Hanke, M. E. (1958). Use of enzymes and radiocarbon in estimation of the equilibrium constants for the decarboxylation of lysine and glutamate. *J. Biol. Chem.* **230**, 73–80.
49. Wall, M. E., Markowitz, D. A., Rosner, J. L., & Martin, R. G. (2009). Model of transcriptional activation by MarA in *Escherichia coli*. arXiv:0902.0959v1.

3 Crystal structure of the sensory domain of *Escherichia coli* CadC, a member of the ToxR-like protein family

Crystal structure of the sensory domain of *Escherichia coli* CadC, a member of the ToxR-like protein family

Andreas Eichinger,[1] Ina Haneburger,[2] Christiane Koller,[2] Kirsten Jung,[2] and Arne Skerra[1]*

[1]Munich Center for Integrated Protein Science and Lehrstuhl für Biologische Chemie, Technische Universität München, Freising-Weihenstephan, Germany

[2]Munich Center for Integrated Protein Science and Department of Biology I, Microbiology, Ludwig-Maximilians-Universität, München, Germany

Received 23 September 2010; Revised 27 December 2010; Accepted 9 January 2011
DOI: 10.1002/pro.594
Published online 24 January 2011 proteinscience.org

Abstract: The membrane-integral transcriptional activator CadC comprises sensory and transcriptional regulatory functions within one polypeptide chain. Its C-terminal periplasmic domain, $CadC_{pd}$, is responsible for sensing of environmental pH as well as for binding of the feedback inhibitor cadaverine. Here we describe the crystal structure of $CadC_{pd}$ (residues 188–512) solved at a resolution of 1.8 Å via multiple wavelength anomalous dispersion (MAD) using a $ReCl_6^{2-}$ derivative. $CadC_{pd}$ reveals a novel fold comprising two subdomains: an N-terminal subdomain dominated by a β-sheet in contact with three α-helices and a C-terminal subdomain formed by an eleven-membered α-helical bundle, which is oriented almost perpendicular to the helices in the first subdomain. Further to the native protein, crystal structures were also solved for its variants D471N and D471E, which show functionally different behavior in pH sensing. Interestingly, in the heavy metal derivative of $CadC_{pd}$ used for MAD phasing a $ReCl_6^{2-}$ ion was found in a cavity located between the two subdomains. Amino acid side chains that coordinate this complex ion are conserved in CadC homologues from various bacterial species, suggesting a function of the cavity in the binding of cadaverine, which was supported by docking studies. Notably, $CadC_{pd}$ forms a homodimer in solution, which can be explained by an extended, albeit rather polar interface between two symmetry-related monomers in the crystal structure. The occurrence of several acidic residues in this region suggests protonation-dependent changes in the mode of dimerization, which could eventually trigger transcriptional activation by CadC in the bacterial cytoplasm.

Keywords: cadaverine; pH-sensing; periplasmic domain; ToxR-like protein; X-ray structure

Introduction

Bacteria adapt to changes of external pH by various physiological processes and alterations in gene expression. Thus, several genes are induced in the neutrophilic bacterium *Escherichia coli* upon exposure to acidic conditions to maintain the cytoplasmic pH in the physiological range between 7.6 and 7.8.[1] In particular, the degradative amino acid decarboxylase systems *adi*, *gad*, and *cad* are among those genes.[2] Under conditions of an acidic environment and in the presence of external lysine the membrane-integral transcriptional activator CadC is triggered and induces transcription of the *cadBA* operon, which encodes the cadaverine/lysine-antiporter CadB and the lysine decarboxylase CadA. CadA intracellularly decarboxylates lysine to cadaverine

Christiane Koller's current address is Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-Universität, Munich, Germany.

Grant sponsor: the Deutsche Forschungsgemeinschaft; Grant numbers: Exc114/1 and JU270/5-3.

*Correspondence to: Prof. Dr. Arne Skerra, Lehrstuhl für Biologische Chemie, Technische Universität München, Emil-Erlenmeyer-Forum 5, 85350 Freising-Weihenstephan, Germany. E-mail: skerra@tum.de

under consumption of one proton. The resulting cadaverine is exported by CadB, which in turn takes up lysine form the external medium.[3]

CadC is an integral membrane protein of 512 amino acids comprising an N-terminal cytoplasmic DNA-binding domain (residues 1–159), a transmembrane helix (residues 160–187), and a C-terminal periplasmic domain (residues 188–512).[4] CadC belongs to the family of ToxR-like proteins, which includes CadC of *Salmonella enterica* serovar Typhimurium and *Vibrio cholerae* (regulating the *cadBA* operon),[5] ToxR from *V. cholerae* (regulating cholera toxin, pilus, and outer-membrane protein expression),[6] TcpP from *V. cholerae* (ToxR-coregulator of virulence gene expression),[7] PsaE from *Yersinia pestis* (regulating fimbriae genes),[8] and WmpR from *Pseudoalteromonas tunicata* (regulating iron uptake).[9] Despite low sequence homology (outside the CadC orthologs) proteins of this family are characterized by a common three-domain topology and all combine sensory and transcriptional regulatory functions within one polypeptide chain. Therefore, these proteins represent the simplest known transmembrane signaling system that transduces information across the lipid bilayer without involving chemical modification.

The DNA-binding function of CadC is mediated by its cytoplasmic domain that contains a helix-turn-helix motif similar to the RO_{II} subgroup of DNA-binding domains found in response regulators such as PhoP from *Bacillus subtilis*,[10] VirG from *Agrobacterium tumefaciens*[11] or OmpR from *Escherichia coli*.[12] The DNA-binding sites were identified and the interaction between purified CadC and the promoter region was demonstrated *in vitro*.[13] Expression of the *cadBA* operon is activated at low external pH and concomitantly available lysine. Recently, it was shown that CadC is not a direct sensor of the external lysine concentration. Instead, lysine is co-sensed via interplay with the lysine-specific permease LysP.[14,15] Still, the pH-sensory function is assigned to the periplasmic domain of CadC.[16,17] Several CadC variants with single amino acid replacements in the periplasmic domain affecting the pH-dependent *cadBA* expression were previously identified.[16,17] For example, replacement of Asp471 against Asn resulted in a pH-independent activation of expression. In contrast, CadC with Glu at the same position was no longer able to respond to acidification of the environment. Thus, single amino acid replacements can fix CadC either in an ON or an OFF state, respectively.

A combination of biochemical experiments and computational simulations revealed that the transcriptional stimulating activity of CadC is further regulated via feedback inhibition by the product of the CadA decarboxylation reaction, namely cadaverine.[18] Consistently, cadaverine binds to the periplasmic domain of CadC with moderate affinity (K_D = 96 μM).[14] Nevertheless, it is hardly understood how CadC adopts its active state after receiving the pH signal and, also, how cadaverine inactivates CadC. To gain insight into the structural mechanisms, we have carried out a crystallographic analysis of the periplasmic domain $CadC_{pd}$ and its two variants D471E and D471N.

Results

Three-dimensional structure of the periplasmic domain of CadC

The X-ray structure of $CadC_{pd}$ (residues 188–512) was solved via multiple wavelength anomalous dispersion (MAD) at a resolution of 2.3 Å by using a crystal soaked with K_2ReCl_6. Further crystal structures were solved for the native apo-protein at a resolution of 1.8 Å as well as for the $CadC_{pd}$ variants D471E and D471N at resolutions of 1.9 and 2.2 Å, respectively (Table I). $CadC_{pd}$ appears as an irregularly shaped protein with a maximal diameter of about 60 Å and a diameter of approximately 30 Å at its narrowest site. The entire periplasmic protein domain is essentially composed of two subdomains: the first subdomain comprises a mixed, parallel/antiparallel β-sheet packed on one side against two α-helices as part of a three-helix bundle, whereas the second subdomain is a pure α-helical bundle encompassing 11 helices (Fig. 1). Taken together, α-helices and β-strands account for 70% of the 314 ordered residues in the periplasmic domain of the wild type CadC apo-protein.

Following to an N-terminal loop at positions 191–194, the polypeptide chain first enters a five-stranded β-sheet at strand B. The chain then adopts a flexible loop not fully defined in the electron density, runs back through an α-helix (helix 1) and enters the β-sheet again as strand A, which is arranged parallel to strand B. After a long loop the chain proceeds back to the β-sheet as strand C, again parallel to B. Two more strands, D and E, connected by shorter loops, follow and are aligned antiparallel to strand C and to each other. The chain then enters an element of two antiparallel α-helices (helices 2 and 3) which, together with helix 1, form the three-helix bundle in the N-terminal subdomain. Notably, all secondary structure elements within this subdomain are oriented in the same direction. For residue Cys208, which is part of the flexible loop connecting strand A and helix 1, no electron density was observed with any dataset, indicating a presumably free (reduced) state in these crystal structures. Because of its proximity to the only other thiol side chain of the well resolved residue Cys272, however, a disulfide bond could be formed between these two amino acids and, for the purpose of illustration, has been modelled here [cf. Fig. 1(A)]. Indeed, the presence of such a disulfide bridge was detected for

Table I. *Crystallographic and Refinement Statistics*

Protein	$CadC_{pd} \bullet ReCl_6^{2-}$			$CadC_{pd}$		
Dataset	Peak	Inflection	Remote	Wild Type	D471E	D471N
Space group		$P6_122$		$P6_122$	$P6_122$	$P3_221$
Unit cell dimensions, a, b, c [Å], $\alpha = \beta = 90°$, $\gamma = 120°$	80.56 80.56 199.78	80.88 80.88 200.30	81.00 81.00 200.65	83.90 83.90 199.09	83.63 83.63 199.49	79.73 79.73 126.01
Wavelength [Å]	1.17652	1.17705	1.07813	0.91841	0.91841	0.91841
Resolution range [Å][a]	69.77–2.30 (2.36–2.30)	70.04–2.60 (2.74–2.60)	70.15–2.90 (3.06–2.90)	72.66–1.80 (1.85–1.80)	72.43–1.90 (1.95–1.90)	69.05–2.20 (2.26–2.20)
$I/\sigma I$[a]	6.2 (2.1)	7.4 (2.2)	9.0 (2.5)	5.6 (2.7)	5.2 (2.6)	20.2 (4.6)
R_{merge}[a,b][%]	9.5 (35.5)	8.0 (34.1)	6.7 (30.6)	8.5 (27.6)	8.6 (27.6)	5.2 (36.7)
Unique reflections	16,693	12,472	9,169	39,276	33,530	24,043
Multiplicity[a]	14.1 (14.5)	13.9 (14.4)	13.8 (14.3)	14.3 (15.1)	13.5 (13.7)	5.6 (5.1)
Completeness[a]	98.5 (97.9)	99.2 (98.8)	99.4 (99.2)	99.9 (100.0)	99.9 (100.0)	99.6 (99.9)
Anomalous multiplicity[a]	7.7 (7.7)	7.7 (7.7)	7.7 (7.7)	–	–	–
Anomalous completeness[a]	99.5 (98.9)	99.8 (99.5)	99.9 (99.7)	–	–	–
Phasing and refinement:						
Figure of merit centric, acentric reflections		0.697, 0.419		–	–	–
R_{cryst}, R_{free}[c][%]		23.0, 27.1		19.1, 21.8	20.2, 23.1	22.8, 27.9
Protein atoms		2,268		2,497	2,507	2,362
Ligand atoms		112		–	–	–
Water molecules		34		237	234	114
Average B factor [Å^2]		24.34		17.04	22.38	33.60
Geometry:						
R.m.s.d. bond lengths, angles [Å, °]		0.018, 1.752		0.012, 1.217	0.014, 1.300	0.018, 1.513
Ramachandran analysis:						
Core, allowed, generously allowed, disallowed [%]		91.3, 8.3, 0.0, 0.4		93.0, 6.7, 0.0, 0.4	93.4, 5.9, 0.3, 0.3	93.2, 6.1, 0.4, 0.4

[a] Values in parentheses are for the highest resolution shell.
[b] R_{free} is the R_{cryst} with 5% of the reflections that were randomly selected and excluded from refinement: $R_{cryst} = \sum_h ||F_o(h)| - |F_C(h)|| / \sum_h |F_o(h)|$
[c] $R_{merge} = \sum_h \sum |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$

$CadC_{pd}$ in biochemical experiments (Tetsch *et al.*, in preparation).

Separated by a loop that forms a sharp turn between residues Leu330 and Leu336, the polypeptide chain enters the C-terminal subdomain of $CadC_{pd}$. Its altogether 11 α-helices (nos. 4–14) are organized as a bundle of five antiparallel α-helical pairs, whereby helix 13 comprises just one turn. These pairs are twisted in a clockwise direction, resulting in a spiral bundle such that the C-terminus of helix 4 is situated close to the N-terminal end of helix 12. The very short helix 13 together with helix 14 form the C-terminus of the structure. Notably, the bundle of α-helices in the second subdomain is roughly oriented perpendicular to the three-helix motif of the first subdomain.

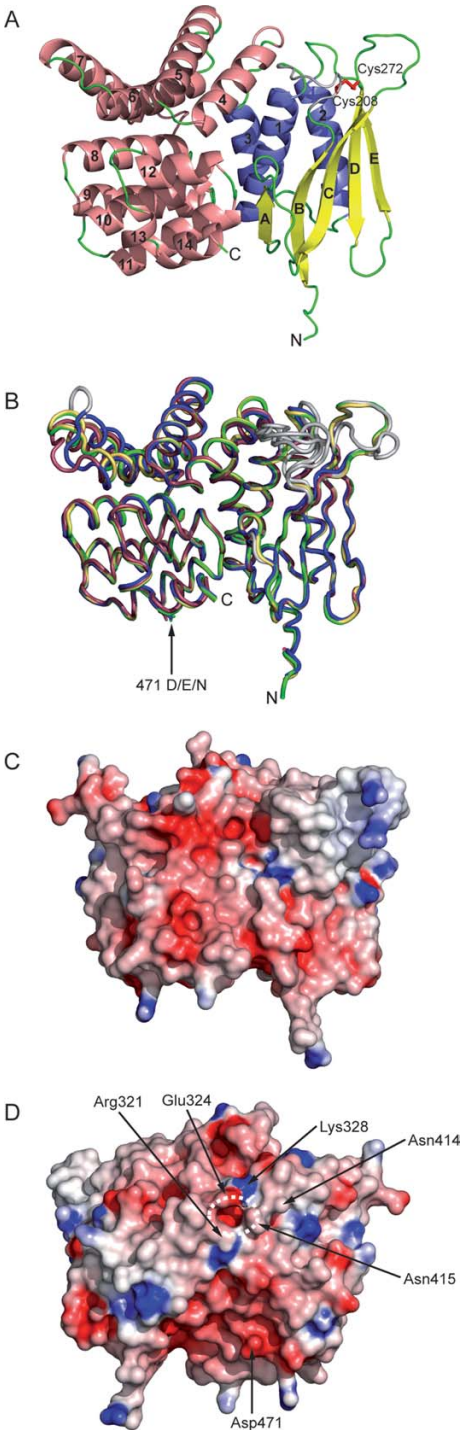

Figure 1. Overview of the $CadC_{pd}$ crystal structure. (A) Cartoon representation of the apo-protein solved at 1.8 Å. The protein consists of two subdomains. Secondary structure elements are depicted in yellow (β-strands), blue (α-helices of the N-terminal subdomain), and salmon (α-helices of the C-terminal subdomain). The two short 3_{10}-helices 334–336 and 411–413 are also shown in salmon. The disordered loop 207–214 has been modelled and is depicted in light gray. The disulfide bond connecting residues Cys208 and Cys272 is shown as red sticks and labeled. (B) Superposition of the $CadC_{pd}$ apo-protein (yellow-orange) with the $ReCl_6^{2-}$ complex (raspberry) and the variants D471E (green) and D471N (blue). Disordered loops were modelled with plausible geometry and are depicted in light gray. (C) Electrostatic surface representation of the $CadC_{pd}$ apo-structure in an orientation identical to (B). Negatively charged areas are colored red ($-10\ k_BT/e$), positively charged areas are colored blue ($10\ k_BT/e$). (D) The molecule in the same representation rotated vertically by around 180° with respect to (C), illustrating a narrow entrance to the central cavity indicated by a dashed ellipse. Residues mentioned in the text are labeled. An interactive view is available in the electronic version of the article.

Comparison between the $CadC_{pd}$ apo-structure and its rhenate complex

The crystal structures of the apo-form of $CadC_{pd}$ and its rhenate complex are very similar even though a rather large number of altogether 16 bound $ReCl_6^{2-}$ complex ions was identified: the root mean square deviation (r.m.s.d.) of the polypeptide chains was 0.79 Å upon superposition of the 282 Cα atoms of common ordered residues in all four models (192–204, 217–245, 251–261, 273–293, 298–394, and 399–509). However, three exposed regions showed significant structural differences [cf. Fig. 1(B)]. One deviation occurs at a stretch formed by residues 294–300, which are part of a loop at the N-terminus of α-helix 2. Second, the stretch 331–339 forms part of a loop at the N-terminus of α-helix 4 and is located close to a bound $ReCl_6^{2-}$ ion at the protein surface. Finally, the segment 388–399 forms a loop connecting α-helices 6 and 7, whose Cα positions are shifted up to 6.7 Å. Two $ReCl_6^{2-}$ ions were bound in the vicinity of this site.

Structures of the $CadC_{pd}$ variants D471E and D471N

Residue Asp471 was previously identified in a random mutagenesis study to be involved in pH-sensing.[16] Recent site-directed mutagenesis experiments

furthermore revealed that replacement of Asp471 with Glu (D471E) resulted in a CadC variant preventing *cadBA* expression, also under acidic conditions (OFF state), whereas the variant with an Asn side chain at this position (D471N) led to a pH-insensitive active protein (ON state).[17] The X-ray structures of wild type $CadC_{pd}$ and its variant D471E were superimposable with an r.m.s.d. value of 0.11 Å using 282 Cα atoms, indicating a virtually identical conformation [Fig. 1(B)]. However, the deviations in the variant D471N, which was crystallized under acidic conditions in a different space group (see Table I), were more significant and reflected by a higher r.m.s.d. value of 0.95 Å.

Considerable differences in the structures of wild type $CadC_{pd}$ and its variant D471N occur at residues 296–297 in a loop between β-strand E and α-helix 2 in the N-terminal subdomain. There is also discrepancy in a long stretch comprising residues 320–340, which connects the N- and the C-terminal subdomains. It includes α-helix 3, a loop followed by a 3_{10}-helix, and the N-terminal part of α-helix 4 [cf. Fig. 1(B)]. Another deviation is evident at the amino acid stretch 388–415 comprising the C-terminus of α-helix 6, a loop, α-helix 7, another loop, a 3_{10}-helix, and the connecting loop to α-helix 8. In contrast, α-helices 9–14 do not show significant deviations, in particular α-helix 11, which carries the mutation at position 471. Notably, among all those differences only two loops are involved in crystal packing contacts: stretch 293–298 with contacting residues Val293, Thr295, Asn296, and Asn298 and stretch 390–393 with contacting residues His390 and Pro391. Overall, it seems that wild type $CadC_{pd}$ was crystallized in an inactive state also mimicked by $CadC_{pd}$(D471E), whereas $CadC_{pd}$(D471N) should better represent the active state, which may also be reflected by the fact that the D471N mutation led to a different crystallization behavior (and altered space group) of $CadC_{pd}$.

Crystal contacts of the $CadC_{pd}$ wild type protein and its variants D471E and D471N

The wild type $CadC_{pd}$ (both apo-protein and Re-complex) and the variant D471E isomorphously crystallized in the same space group $P6_122$ and, hence, both proteins form the same crystal packing contacts with altogether five neighboring symmetry-equivalent molecules. Interestingly, a detailed analysis of these packing interactions using the PISA server[19] indicated the formation of a functional dimer with one of the neighboring molecules, which was similarly detectable also in the $P3_221$ crystal form of the D471N variant (Fig. 2). The putative dimers in all three crystal structures yielded a complexation significance score (CSS) of 1.000. Furthermore, the shape complementarity of the wild type CadC dimer was calculated as $S_C = 0.76$ (using a 1.7 Å probe sphere radius), which is in the range of physiologically functional protein dimers.[20] These structural findings were backed by biochemical results of size exclusion chromatography experiments, revealing dimerization of the periplasmic domain also in solution [cf. Fig. 2(D)].

The dimer formation can be described as an extended contact between two more or less planar and surprisingly polar surfaces on both monomers with local C2 symmetry. In addition, each subunit forms a short arm with its N-terminal loop residues 190–196, which nestles in a crevice of the opposite subunit. Loop residues 248–252 form one side of this niche and residues 504–511 from the C-terminal α-helix 14 form the other side [cf. Fig. 2(A)]. The total dimer interface of the wild type apo-protein buries a solvent accessible area of 1,761 $Å^2$ per monomer, which corresponds to 11.2% of the total molecular surface. The dimer interface comprises 53 residues from each monomer with an extensive network of 11 hydrogen bonds and 6 salt bridges, including 50 buried water molecules, but exhibits only few (20) hydrophobic side chains. In case of $CadC_{pd}$(D471E) the interface area is slightly larger with 1,994 $Å^2$ (12.6% of the total molecular surface) comprising 58 residues with 11 hydrogen bonds and 2 salt bridges. However, the interface area of $CadC_{pd}$(D471N) is just marginally smaller, with 1,732 $Å^2$ (10.8%), than the one of the wild type protein, comprising 51 residues with 9 hydrogen bonds and 2 salt bridges.

Figure 2. Analysis of $CadC_{pd}$ dimerization. (A) Stereo view of one $CadC_{pd}$ dimer related by a crystallographic symmetry axis. One monomer is colored as in Figure 1(A), the other one is uniformly colored cyan. The N- and C-termini are both directed towards the lipid bilayer indicated at the bottom. (B) Hydrophobic surface representation of the two $CadC_{pd}$ monomers rotated by 90° about a horizontal axis (with the N-termini directed toward the front) and then rotated by ±90° apart to reveal the dimer interface. The interface forming residues are colored according to increasing hydrophobicity from green (hydrophilic) over white (neutral, including the polypeptide backbone) to brown (hydrophobic).[47] Residues Asp471 and Lys255, which form a salt bridge, are labeled. (C) The same view as in (B) with electrostatic surface representation as in Figure 1(C and D). (D) Analytical size exclusion chromatography of the recombinant $CadC_{pd}$ at pH 7.5. The protein eluted with a retention volume of 14.4 mL, which corresponds to a molecular size of 64.8 kDa (left) as determined from a plot of log MW of calibration proteins against their elution volumes (right): cytochrome c (12.4 kDa), carbonic anhydrase (29 kDa), ovalbumin (43 kDa), bovine serum albumin (66.3 kDa), alcohol dehydrogenase (150 kDa), β-amylase (200 kDa), apoferritin (443 kDa). The apparent size of $CadC_{pd}$ is considerably larger than the calculated mass of the proteolytic product after thrombin cleavage (41.8 kDa), indicating stable dimer formation in solution.

Interestingly, all the different side chains at position 471, that is Asp, Glu, or Asn, participate in a (pairwise symmetric) dimer contact. The wild type residue Asp471 forms several hydrogen bonds with its carboxylate group, both to two water molecules and to the NZ atom of Lys255 in the neighboring

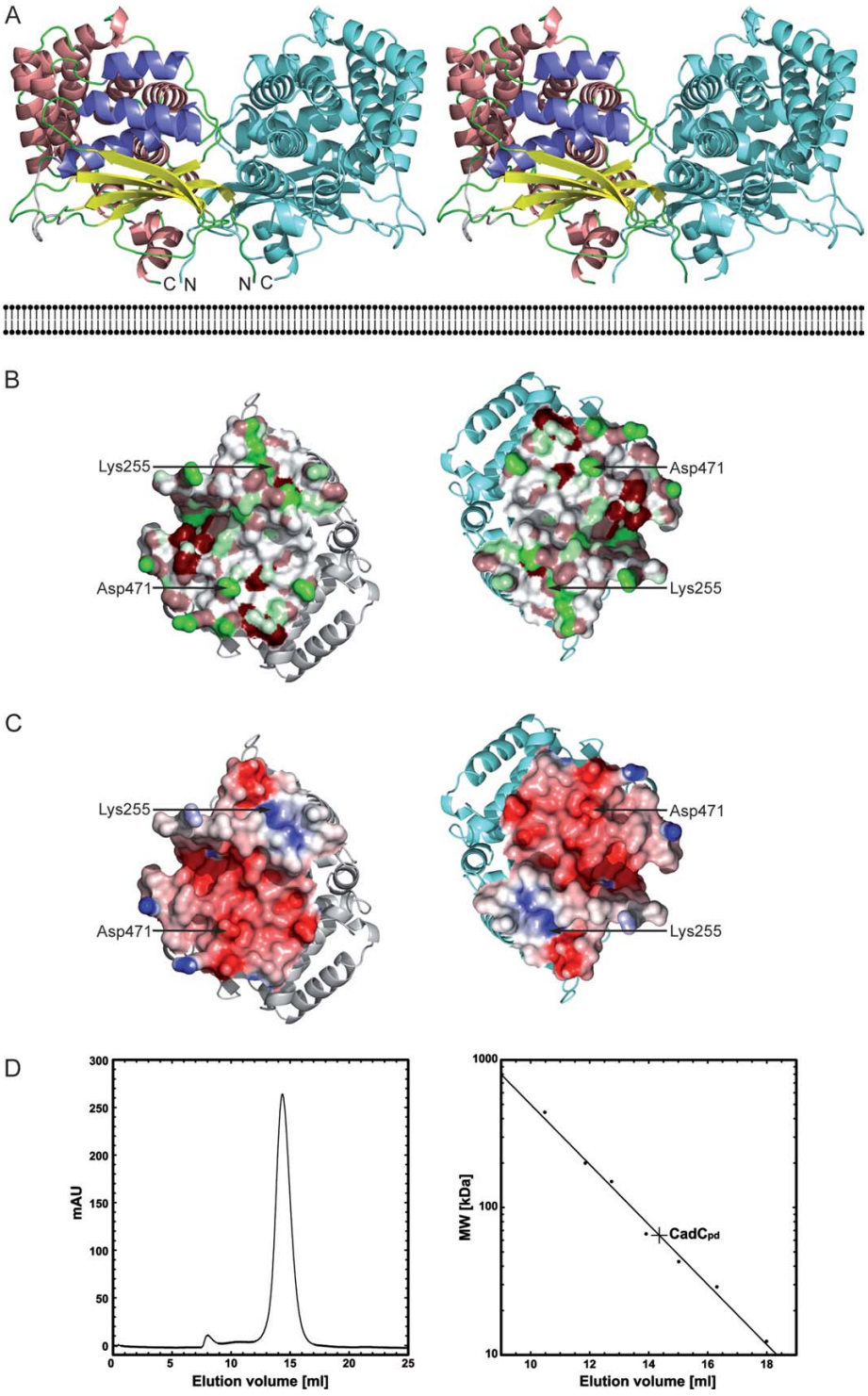

Figure 2.

molecule. In the isomorphously crystallized variant D471E both carboxylate oxygens of the longer Glu471 side chain form direct hydrogen bonds to the NZ atom of Lys255. In the variant D471N the mutated residue Asn471 forms one direct hydrogen bond to Lys255 NZ and a second one to a water molecule. Apart from residue Asp471, there are up to seven other acidic amino acids involved in the dimer interface: Asp198, Asp248, Glu249, Asp280, Asp283, Asp445, and Glu468. Of these, Glu249 and Asp283 form salt bridges with residues Arg191 and Arg467, respectively. Such a polar and wet, loosely packed, and both-sided negatively charged interface is unusual for a stable protein dimer. In fact, these peculiar biophysical properties suggest a charge-driven sensing mechanism of the CadC transcriptional activator as will be further discussed below.

The putative ligand cavity and the $ReCl_6^{2-}$ complex of $CadC_{pd}$

The two subdomains of $CadC_{pd}$ enclose a cylinder-shaped cavity in between, which is about 20 Å long and has a diameter of 13 Å. A narrow hole, lateral to the dimer interface described above, gives entrance to this cavity [Fig. 1(D)] from the solvent. Its diameter is about 5 Å and it is lined by residues Arg321, Glu324, Lys328, Asn414, and Asn415. Side chains of altogether 46 residues form the central cavity in $CadC_{pd}$. Its internal surface is made by 20 aliphatic residues, of which 11 side chains (Leu347, Leu348, Leu355, Leu381, Val385, Val425, Leu428, Leu451, Val454, Leu455, and Ile493) form a hydrophobic pocket within the α-helical subdomain. Ten residues contribute polar side chains, of those Arg321 and Arg480 are positively charged and Asp225, Asp241, Glu378, and Glu447 are negatively charged [Fig. 3(A)]. Notably, in the heavy atom derivative that had been prepared for phase determination the cavity harbored one of the $ReCl_6^{2-}$ complex ions [Fig. 3(C)] while in the $CadC_{pd}$ apo-structure the cavity was filled with a cluster of 20 ordered solvent molecules arranged in typical hydrogen bonding distances.

Considering the in total doubly negative charge of the hexachlororhenate ion it is surprising that this ligand became accommodated in the mostly negatively charged cavity. Yet, the metal-coordinating chloride ions might be polarized by the electron-deficient Re^{IV} central ion to such an extent that positive partial charges appear at the surface of this complex ion. Apart from the electrostatic aspect, the $ReCl_6^{2-}$ sterically fits well into the cavity and forms close contacts with five residues: Asp225, Thr229, His344, Glu447, and Trp450. In fact, their polar groups are in hydrogen bonding distance to the six chloride ions. In addition, there are three water molecules located in the cavity that contribute hydrogen bonds: one water molecule contacts Thr229, Asn232, and Met448, the second one contacts Ser417 and Ile418, and the third one forms hydrogen bonds to residues Gln421 and Glu447.

Docking studies with cadaverine

In principle, the cavity between the two subdomains of the periplasmic region would be nicely suited to accommodate the positively charged cadaverine ligand of CadC, both with regard to sterical fit and to electrostatic complementarity. Unfortunately, soaking experiments with this ligand followed by crystallographic analysis did not give rise to well defined electron density. To identify side chains that could possibly interact with cadaverine within the presumed binding site, we performed a computer simulation using the program AutoDock[21] with cadaverine initially provided in a random orientation. All the 10 best scoring results yielded cadaverine in a mostly extended conformation and bound with similar orientation in the cavity [Fig. 3(D)]. Cadaverine was surrounded by residues Glu447, Gln421, and Tyr374 on one side and Asp225 as well as Thr229 on the other. The polar or negatively charged side chains were appropriately positioned to form hydrogen bonds to both positively charged amino groups of the ligand. AutoDock reported calculated dissociation constants for these 10 best $CadC_{pd}$-cadaverine complexes in the range between 12 and 28 μ*M*, which is in the same order of magnitude as the K_D values from 96±18 μ*M* previously measured for the wild type protein via fluorescence titration.[14]

Comparison between $CadC_{pd}$ and proteins with related folds

A search for structurally related proteins in the protein data bank (PDB) with DALI[22] revealed several hits with *Z*-scores ranging from 11.8 to 9.5. Among the best 20 candidates no protein with a recognizable structural homology was found. Only parts of two structures (hit nos. 8 and 15) showed some local similarity with the α-helical subdomain of $CadC_{pd}$ (Fig. 4). The peroxisomal targeting signal 1 binding domain of *Trypanosoma brucei* peroxin 5[23] (PDB entry 3CVN; *Z*-score = 10.6) exhibits a twisted 18 α-helix bundle. A fragment of this structure formed by α-helices 3 to 12 shows some resemblance with $CadC_{pd}$ [Fig. 4(B)]. However, it is less compact and the mutual arrangement of the α-helices is more entwined in this section while one of the α-helices clearly has a different orientation. The transcription factor MalT domain III of *Escherichia coli*[24] (PDB entry 1HZ4; *Z*-score = 9.9) likewise forms a twisted 18 α-helix structure. A fragment comprising α-helices 6–15 again shows similarity with the α-helical subdomain of $CadC_{pd}$. Interestingly, this protein is also a transcriptional activator. The other 18 hits were predominantly all-α-helical proteins but did not exhibit significant similarity with $CadC_{pd}$. Another search for structurally related proteins using the

HHPred server[25] revealed one additional protein with partial similarity to $CadC_{pd}$: the tetratricopeptide repeat protein NlpI of *Escherichia coli* (PDB entry 1XNF).[26] A fragment of this lipoprotein comprising α-helices 4 to 13 grossly resembles the α-helical architecture of $CadC_{pd}$, yet the helix arrangement is clearly more open. Consequently, $CadC_{pd}$ exhibits a novel fold, both with the 11-α-helix bundle and the mixed β-sheet/α-helix assembly in its two subdomains.

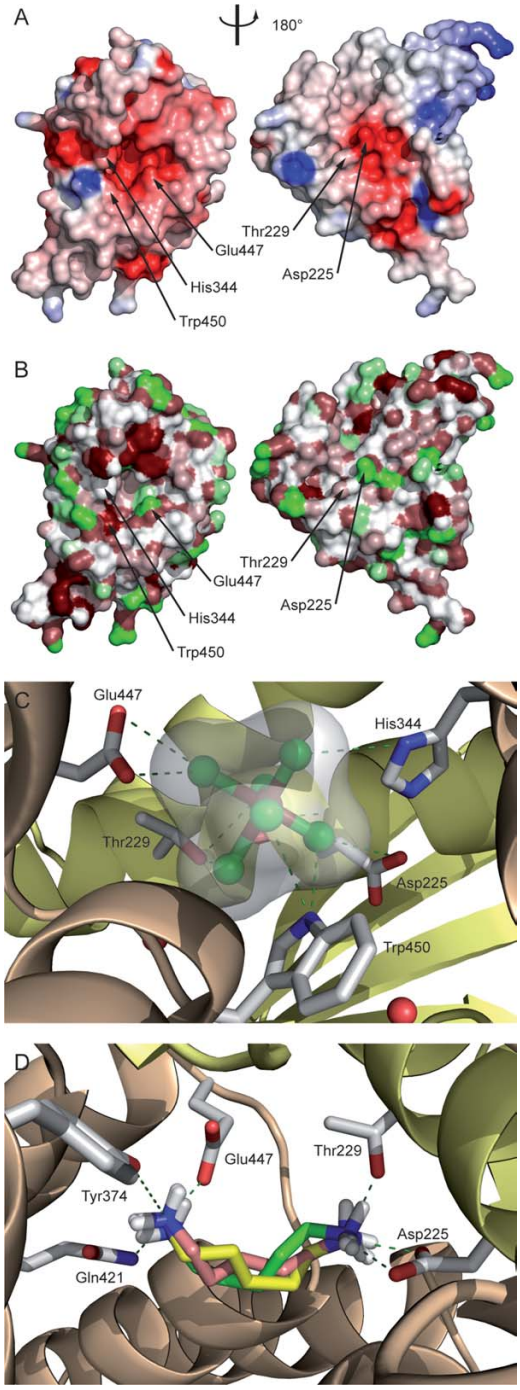

Figure 3. Analysis of the cavity and the interface between the two subdomains of $CadC_{pd}$. (A) Electrostatic surface properties (cf. Fig. 1) of the C-terminal α-helical subdomain (residues 333–509; left) and the N-terminal predominantly β-sheet subdomain (residues 190–332; right). Both moieties were rotated by ±90° about a vertical axis to visualize the internal cavity surfaces. Residues that contact the bound $ReCl_6^{2-}$ ion are labeled. (B) Hydrophobic surface representation colored as in Figure 2B. (C) Complexation of $ReCl_6^{2-}$ within the central cavity. Interacting side chains of CadC within a distance of 4 Å to the rhenate central ion are indicated. (D) Results from a docking simulation with cadaverine. The three cadaverine conformations with highest scores as well as the putative interacting residues in the cavity are depicted as sticks. Cadaverine carbon atoms are colored differently according to scoring rank order: 1, yellow; 2, salmon; 3, green. Oxygen atoms are shown in red, nitrogen atoms are shown in blue, hydrogen atoms are shown in white, and hydrogen bonds are indicated as dark green dashed lines.

Homologous proteins with similar sequence motifs

A BLAST[27] search with the amino acid sequence of $CadC_{pd}$ in the NCBI nonredundant protein sequences database[28] revealed homologous CadC proteins from 24 different bacterial species with sequence identities ranging from 91 to 21%, namely (with GI numbers of the NCBI database given in parentheses) *Escherichia albertii* (170766700), *Escherichia fergusonii* (218549242), *Klebsiella variicola* (288936734), *Klebsiella pneumoniae* (206578436), *Salmonella enterica* (161502288), *Yersinia ruckeri* (238756479), *Salmonella typhimurium* (16765877), *Serratia proteamaculans* (15737200), *Serratia odorifera* (293394724), *Edwardsiella tarda* (294635142), *Edwardsiella ictaluri* (238918793), *Vibrio cholerae* (147674140), *Vibrio orientalis* (261250472), *Vibrio harveyi* (156972567), *Vibrio shilonii* (149190651), *Vibrio metschnikovii* (260774390), *Vibrio vulnificus* (27365392), *Vibrio mimicus* (262163818), *Vibrio parahaemolyticus* (28899667), *Vibrio fischeri* (59712667), *Vibrio alginolyticus* (91228716), *Aeromonas hydrophila* (117620838), *Aliivibrio salmonicida* LFI1238 (209695912), and *Acidobacteria bacterium* Ellin345 (94968198). Apart from these hits there were only two shorter sequences with similarity to $CadC_{pd}$, yet corresponding to proteins of other classes: the putative serine/threonine kinase of *Psychroflexus torquis* (91215656) with 24% identity to the CadC sequence 266–381 and the TPR repeat-containing serine/threonine protein kinase of

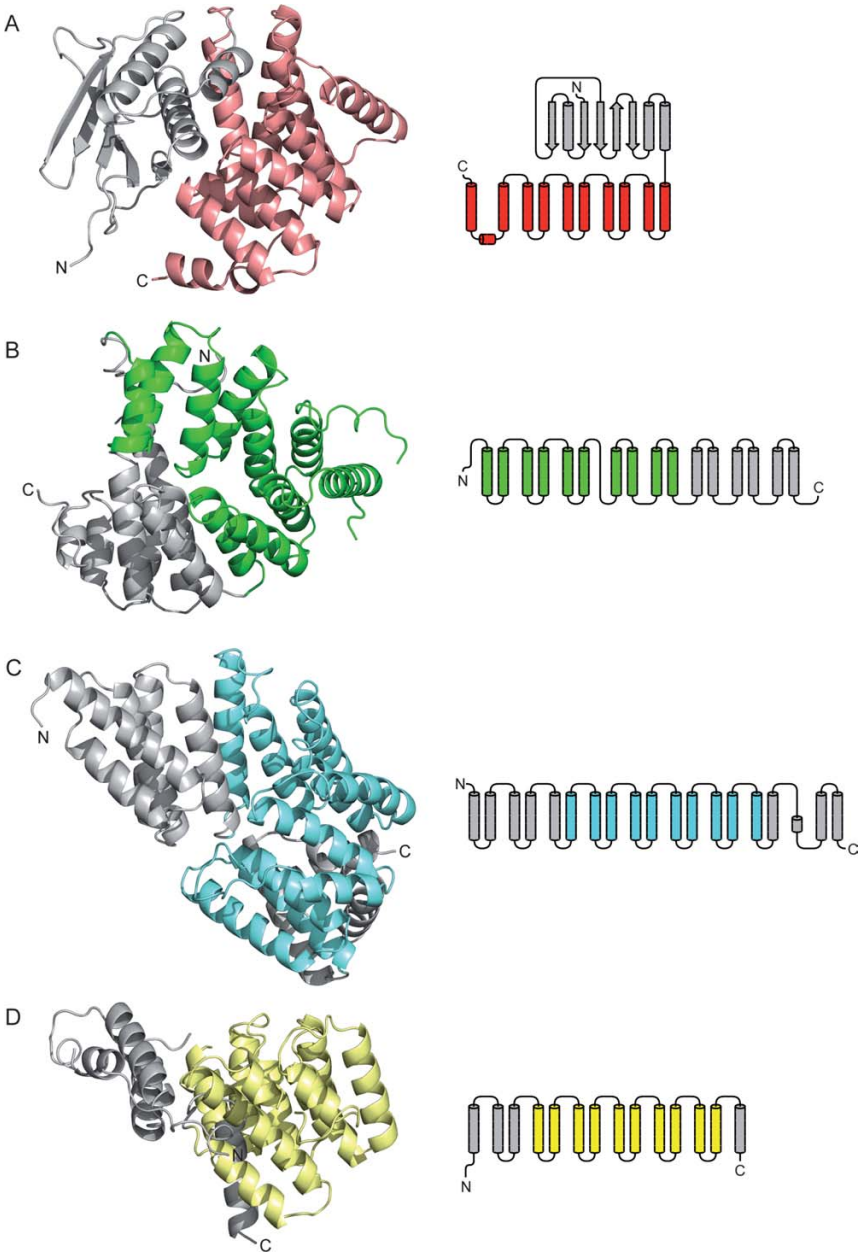

Figure 4. Comparison of the crystal structure and secondary structure topology of $CadC_{pd}$ with its most closely related α-helical bundle proteins. Subdomains to be considered for the comparison are colored whereas other structural elements are shown in gray. (A) $CadC_{pd}$ with its C-terminal subdomain (residues 333–512) comprising 11 α-helices (salmon). (B) The peroxisomal targeting signal 1 binding domain of *Trypanosoma brucei* peroxin 5 (PDB entry 3CVN) forms a twisted 16 α-helix bundle; helices 3–12 are colored green. (C) The transcription factor MalT domain III protein of *Escherichia coli* (1HZ4) forms another twisted 18 α-helix bundle; helices 6–15 are colored blue. (D) The tetratricopeptide repeat protein NlpI of *Escherichia coli* (1XNF) with α-helices 4–13 colored yellow.

Acidobacteria bacterium (94969422) with 27% identity to the CadC sequence 248–387. Interestingly, several CadC orthologs revealed conserved residues at the positions that line the putative cadaverine cavity described above (Fig. 5).

Discussion

CadC belongs to the ToxR-like regulators that encompass biochemically non-modified one-component systems with similar gross topology, including several low pH-induced transcription regulators.

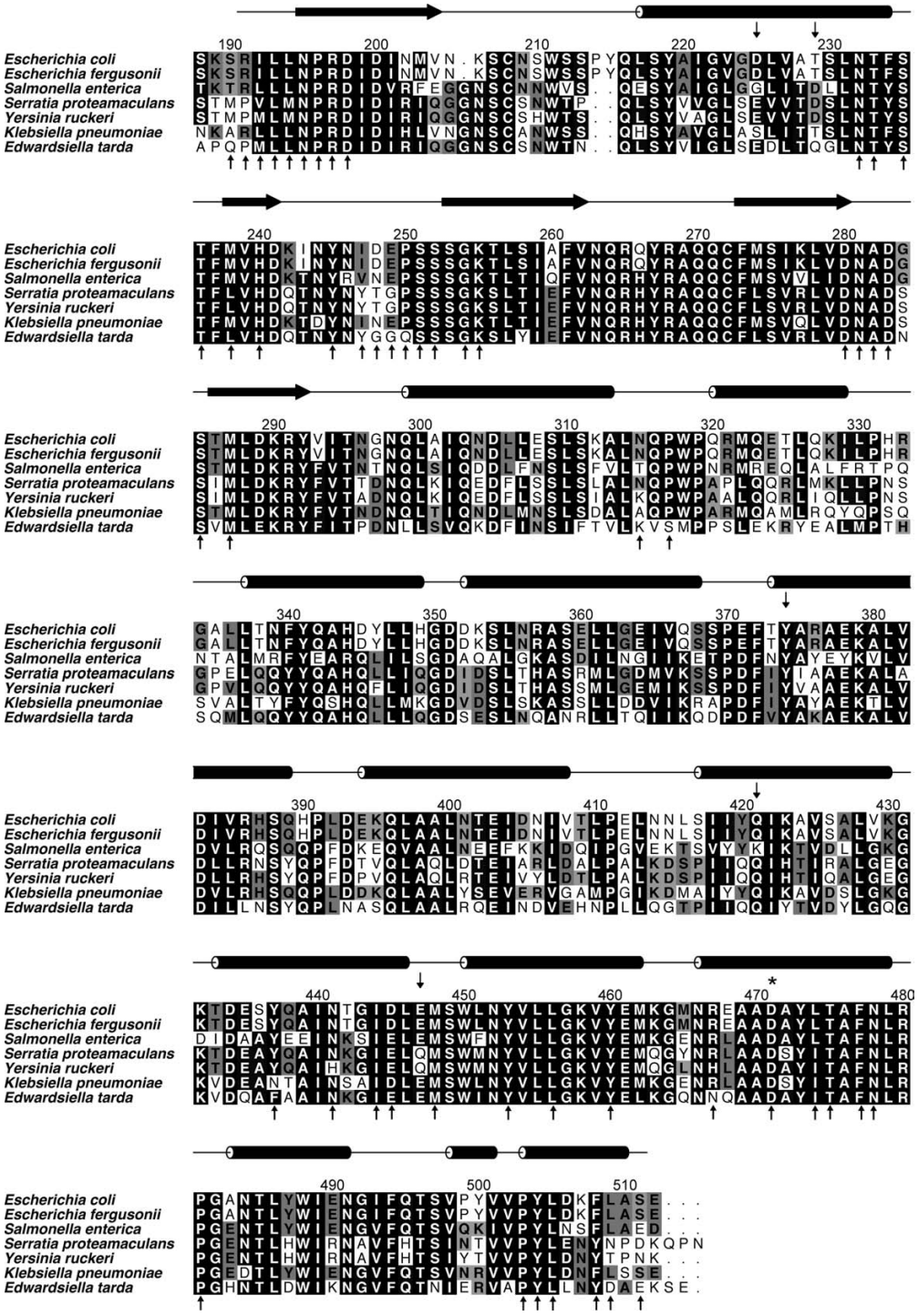

Figure 5. Sequence alignment of $CadC_{pd}$ orthologs (residues 188–512) from a representative set of γ-proteobacteria. Amino acids lining the characteristic cavity of $CadC_{pd}$ are labeled as small vertical arrows above the sequences while residues involved in the dimer interface are indicated by arrows underneath, many of which involve fully conserved positions. Secondary structure elements are shown as horizontal arrows (β-strands) and rods (α-helices), respectively, above the alignment according to the crystal structure of the *Escherichia coli* $CadC_{pd}$ protein determined in this study. Amino acid Asp471, which is important for pH-sensing, is marked by an asterisk. GI numbers (www.ncbi.nlm.nih.gov/Sitemap/sequenceIDs.html) are given in parentheses: *Escherichia coli* MG1655 (16131959), *Escherichia fergusonii* ATCC 35469 (218549242), *Salmonella enterica* serovar typhimurium LT2 (16765877), *Serratia proteamaculans* 568 (157323768), *Yersinia ruckeri* ATCC 29473 (238756479), *Klebsiella pneumoniae* NTUH-K2044 (218117876), *Edwardsiella tarda* EIB202 (267983771).

Because of the important role of these proteins for bacterial virulence, for example ToxR of *Vibrio cholerae*, there is elevated interest in their biomolecular mechanism. Albeit there is no clear correlation between the primary structures of the periplasmic domains of CadC and ToxR, even in terms of size ($CadC_{pd}$: 325 amino acids; $ToxR_{pd}$: 100 amino acids), it is conceivable that the general mechanism of signal perception and transduction across the inner bacterial membrane should be related.

The X-ray analysis of $CadC_{pd}$ reported here provides the first known tertiary structure for a sensory domain of a (pH-activated) ToxR-like regulator. A characteristic feature of the $CadC_{pd}$ structure is its composition of two subdomains with a cavity at their interface that is suited to accommodate cadaverine, the feedback inhibitor of the Cad system. The docking results demonstrate that the side chains of the residues Glu447, Asp225, Thr229, Tyr374, and Gln421 within this pocket are appropriately oriented to bind the positively charged ligand. This finding is in agreement with the binding of the hexachlororhenate ion in almost the same place. In this case, the side chains of residues Asp225, Thr229, His344, Glu447, and Trp450 are involved. Notably, the predicted dissociation constants for the $CadC_{pd}$-cadaverine complex from the docking experiment of 12–28 μM are in the same order of magnitude as the experimentally determined value of 96 ± 18 μM.[14] The lower estimated value may be caused by incorrect assumptions for the electrostatic contribution to the binding interaction. However, the measured moderate affinity explains the lack of success in our soaking experiments with this ligand. Generally, the flexibility of the cadaverine ligand and the preponderance of electrostatic interactions should facilitate adaptation of the mode of binding to partly different charged and/or polar residues that line the ligand pocket in other species (cf. Fig. 5). For example, it seems that in bacteria that carry Gly instead of Asp at position 225 the role of the charged side chain could be taken over by the adjacent Asp229 (instead of Thr).

Another remarkable feature of the $CadC_{pd}$ crystal structure is the formation of a biological dimer as revealed by PISA and CS analysis and experimentally backed by size exclusion chromatography. The quaternary structure of this dimer would allow an arrangement with respect to the lipid bilayer that is compatible with a preceding transmembrane segment leading into the N-terminal subdomain dominated by the five-stranded β-sheet. Thus, it is conceivable that pH-dependent protonation of negatively charged side chains, including residue Asp471, should affect the interface of the dimer, resulting in an altered mutual orientation of the periplasmic domains and, thus, to a general conformational effect that becomes transduced via the transmembrane segment and ultimately triggers transcriptional activity on the cytoplasmic side. The observation that the mutation D471N (representing the ON state) both has a pronounced effect on the solubility of the protein (see Materials and Methods) and leads to a different crystal packing supports the notion of minute conformational changes at the dimer interface that should functionally influence transcriptional activity. This would also be consistent with an additional role of cadaverine binding at the cavity within each periplasmic domain on the mode of dimer association, apart from the pH-sensing effect. Notably, the complex between $CadC_{pd}$ and the rhenate ligand showed a significantly smaller interface area of 1368 $Å^2$ per monomer with fewer hydrogen bonds (4) and salt bridges (2).

In conclusion, the three-dimensional structure of $CadC_{pd}$ elucidated in this study establishes a new protein superfamily as there is no significant structural similarity to any other protein in the PDB. Only sub-structures of three large α-helical bundle proteins show certain resemblance to the C-terminal α-helical subdomain of $CadC_{pd}$. In particular, there is no structurally related protein with the N-terminal β-sheet subdomain of $CadC_{pd}$. The results from a BLAST search underline these findings because there is no independent protein in the database that shows significant sequence similarity to $CadC_{pd}$, including members of the ToxR-like protein family itself. However, CadC homologues with high sequence similarity are found in several γ-proteobacterial genera like the *Salmonella*, *Serratia*, and *Klebsiella*. Thus, it can be expected that these CadC-related proteins form a novel structural family with similar functions in bacterial cell physiology. Mutagenesis experiments are currently under way to further study the structural mechanisms that underly the pH- and ligand-dependent transcriptional activation of this interesting bacterial membrane sensor protein.

Materials and Methods

Overexpression and purification of wild type $CadC_{pd}$ and its variants D471E and D471N

$CadC_{pd}$ was produced as a fusion protein consisting of N-terminal thioredoxin, a double His_6-tag interspersed with thrombin and enterokinase cleavage sites, and the CadC fragment comprising residues 188–512 (Swiss-Prot entry P23890). This hybrid protein was overproduced in *Escherichia coli* Origami B (DE3)/pLysS using the plasmid pET32a-CadC188-512 as described before.[14] The structural genes for the $CadC_{pd}$ variants D471E and D471N were obtained by site-directed mutagenesis and expressed like the wild type protein.[17] Briefly, the fusion proteins were purified from the bacterial whole cell extract by Ni-NTA affinity chromatography. After that, thioredoxin was cleaved off using the Thrombin Cleavage Capture Kit

(Merck, Darmstadt, Germany). The reaction product was dialyzed against 20 m*M* Tris/HCl pH 8.0, 130 m*M* NaCl and applied to a Q Sepharose Fast Flow anion exchange column (GE Healthcare, Munich, Germany), followed by elution with a gradient ranging from 130 m*M* to 1 *M* NaCl in 20 m*M* Tris/HCl pH 8.0. Fractions containing $CadC_{pd}$ were pooled and concentrated to 11 mg/mL by ultrafiltration. The wild type protein was finally purified by size exclusion chromatography using an S75 analytic HR 10/30 column (GE Healthcare) in the presence of 10 m*M* Tris/HCl pH 7.5, 150 m*M* NaCl. $CadC_{pd}$ was then concentrated with 10 kDa cut-off Vivaspin tubes (Vivascience, Hannover, Germany) yielding two batches with concentrations of 8.8 and 13 mg/mL, respectively, and filtrated with 0.45 μm Spin-X centrifuge tubes (Corning Inc., Corning, NY) before crystallization. To measure the apparent size of the wild type $CadC_{pd}$ protein, a size exclusion chromatography was performed on a calibrated S200 analytic HR 10/30 column (GE Healthcare) in the presence of 150 m*M* NaCl, 10 m*M* Tris/HCl pH 7.5. The $CadC_{pd}$ variant D471E was purified in the same way and finally concentrated to 14 mg/mL. The variant D471N, however, formed a gel-like precipitate after the first dialysis against 20 m*M* Tris/HCl pH 8.0, 130 m*M* NaCl, resulting in an initial protein concentration of merely 2.8 mg/mL. Following centrifugation and sterile filtration, the protein was purified via anion exchange chromatography as described above, and, after adjusting the NaCl concentration to 500 m*M*, subjected to ultrafiltration. The concentrated sample, which turned opaque after resting on ice for an hour, was cleared by filtration through a Spin-X tube, resulting in a final protein concentration of 6.9 mg/mL, and used for crystallization trials without size exclusion chromatography.

Protein crystallization and preparation of heavy atom derivatives

$CadC_{pd}$ was crystallized using the hanging drop vapor diffusion technique, yielding crystals within 1 week at 20°C. The precipitants were 1.4 to 1.7 *M* $(NH_4)_2SO_4$ in combination with 4.5–11% (v/v) 2-propanol (without additional buffer except for the 10 m*M* Tris/HCl pH 7.5 from the original protein solution; cf. above). Crystals of native $CadC_{pd}$ were harvested from their drops, dipped into the reservoir solution supplemented with 30% (v/v) glycerol, and flash-frozen. Crystals of the variant D471E were grown under the same conditions, whereas crystals of the variant D471N were obtained in an acidic milieu with 20% (w/v) PEG 6000, 100 m*M* Na-citrate pH 5. In the latter case, the droplets were set up with a 2:1 protein to precipitant volume ratio to compensate for the lower concentration of the original protein solution. A search for heavy atom compounds (Hampton Research, Aliso Viejo, CA) that bind to wild type $CadC_{pd}$ was performed by means of native polyacrylamide gel electrophoresis (PAGE).[29] To this end, 5 μL of each heavy atom compound with a concentration of 10 m*M* in aqueous solution was added to 5 μL of a 2 mg/mL $CadC_{pd}$ solution in 10 m*M* Tris/HCl pH 7.5, 150 m*M* NaCl and the mixture was applied to 6.5% (w/v) PAGE in the presence of 750 m*M* Tris/HCl pH 7.5 using 190 m*M* glycine, 10 m*M* Tris/HCl pH 8.3 as running buffer. A detectable band shift of the protein-heavy atom complex led to the identification of K_2ReCl_6 as a promising compound. To prepare the heavy atom derivative for X-ray data collection, crystals of $CadC_{pd}$ were first transferred from the mother liquor described above to 2.5 *M* Li_2SO_4 to avoid the formation of metal-ammonia complexes. Then, a solution of 30 m*M* K_2ReCl_6 and 1.7 *M* Li_2SO_4 was cautiously added to the drop such that the final $ReCl_6^{2-}$ concentration reached 5 m*M*. Within about 10 min the initially pale yellow color of the droplet changed to dark violet. On the next day, the deeply colored crystals were washed in 2.5 *M* Li_2SO_4 to remove precipitated rhenium salt. Then, the soaking was repeated for 4 h, again followed by washing of the crystals with 2.5 *M* Li_2SO_4. Finally, crystals were transferred to 2.5 *M* Li_2SO_4, 30% (v/v) glycerol, using MicroMount loops (MiTeGen, Ithaca, NY) and flash-frozen in liquid nitrogen.

Data collection and processing, model building, and refinement

For the structure solution of $CadC_{pd}$ via MAD, three datasets from a single Re-soaked crystal of space group $P6_122$ were collected at BESSY beam line 14.2. Diffraction up to a resolution of 2.3 Å (peak), 2.6 Å (inflection), and 2.9 Å (high energy remote) was measured at 100 K (Table I). Data were processed with MOSFLM, scaled with SCALA,[30] and heavy atom sites were identified with SHELXD.[31] Refinement of the Re sites, MAD phasing, solvent flattening, and calculation of a native electron density map were carried out with SHARP.[32] The initial atomic model of the Re-complexed $CadC_{pd}$ was built using ARP/wARP[30] and manually corrected with Coot.[33] Altogether 16 $ReCl_6^{2-}$ sites were verified by inspecting the anomalous density map, and their occupancy was adjusted to yield B-factors similar to surrounding protein atoms. The geometry of the $ReCl_6^{2-}$ complex ion itself was obtained from the Inorganic Crystal Structure Database,[34] revealing Re-CI distances in the range from 2.353 to 2.367 Å, consistent with literature data.[35] Hence, restraints during refinement were set for Re-Cl distances to 2.36 Å and for Cl-Re-Cl angles to 90°. After initial refinement of the protein-rhenate complex the crystal structures of the native protein and its variants were solved by molecular replacement with Phaser[36] using corresponding data sets collected at BESSY beam lines 14.1 and 14.2 (Table I). In the case of the

D471N crystal, diffraction data were processed with XDS.[37] Water molecules were added to all models using ARP/wARP, rotamers of Asn and Gln residues were adjusted with NQ-Flipper[38] and, subsequently, all protein models were refined with Refmac5[30] in iterative cycles interspersed with manual correction.

The electron density map of native $CadC_{pd}$ at a resolution of 1.8 Å (Table I) allowed model building for 314 of the 362 residues present in the recombinant protein, namely Ser190–Lys206 and Tyr215–Ser511. For the $ReCl_6^{2-}$ derivative of $CadC_{pd}$, there was electron density for altogether 286 residues: Ile192–Val204, Leu217–Tyr245, Ser251–Phe261, Phe273–Val293, and Asn298–Leu509. For the $CadC_{pd}$ variant D471E the electron density map at 1.9 Å resolution covered residues Lys189–Lys206 and Tyr215–Ser511, while in the case of D471N there was electron density at 2.2 Å resolution for residues Ser190–Val204, Gln216–Asn263, Cys272–Glu394, and Ala399–Ser511. The C-terminal residue Glu512 as well as the N-terminal His_6-tag, including the remaining upstream linker of the processed fusion protein (47 residues), were invisible in all structures. The models were validated with PROCHECK[39] and WHAT_CHECK.[40] Secondary structure elements were assigned using DSSP[41] and secondary structure topologies were depicted with TopDraw.[42] Crystal packing contacts and protein interfaces were analyzed with PISA[19] and with SC[20] from CCP4.[30] Graphics were prepared with PyMOL[43] and GRASP.[44] Structural superposition was carried out with PyMOL while docking studies were performed with AutoDock 4.[21] Evolutionary related $CadC_{pd}$ sequences were searched with BLAST[27] in the NCBI non-redundant protein sequences database.[28] Sequence alignments were subsequently made with ClustalW2[45] and edited using Aline[46] with a low similarity cut off value of 0.241 (similarity groups DE FWY HKR ILMV NQ ST). Structurally related proteins were searched with DALI[22] and HHPred.[25] The atomic coordinates and structure factors of the $CadC_{pd}$ models have been deposited at the RCSB Protein Data Bank under the accession codes 3LY7 (wild type protein), 3LYA (Re-derivative), 3LY8 (variant D471E), and 3LY9 (variant D471N).

Acknowledgment

The authors wish to thank Korinna Burdack and Josef Danzer for technical assistance and Eberhardt Herdtweck and Martin U. Schmidt for database searches. The synchrotron measurements at BESSY and Free University Berlin at BESSY beamlines 14.1 and 14.2 were supported by the BMBF (project funding reference number 05 ES3XBA/5); we thank Jörg Schulze, Georg Zocher, and Sandra Pühringer for assistance. I.H. received a research scholarship from the Elite Network of Bavaria.

References

1. Padan E, Zilberstein D, Rottenberg H (1976) The proton electrochemical gradient in *Escherichia coli* cells. Eur J Biochem 63:533–541.
2. Gale EF, Epps HM (1942) The effect of the pH of the medium during growth on the enzymic activities of bacteria (*Escherichia coli* and *Micrococcus lysodeikticus*) and the biological significance of the changes produced. Biochem J 36:600–618.
3. Soksawatmaekhin W, Kuraishi A, Sakata K, Kashiwagi K, Igarashi K (2004) Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. Mol Microbiol 51:1401–1412.
4. Watson N, Dunyak DS, Rosey EL, Slonczewski JL, Olson ER (1992) Identification of elements involved in transcriptional regulation of the Escherichia coli cad operon by external pH. J Bacteriol 174:530–540.
5. Merrell DS, Camilli A (2000) Regulation of *Vibrio cholerae* genes required for acid tolerance by a member of the "ToxR-like" family of transcriptional regulators. J Bacteriol 182:5342–5350.
6. Miller VL, Taylor RK, Mekalanos JJ (1987) Cholera toxin transcriptional activator *toxR* is a transmembrane DNA binding protein. Cell 48:271–279.
7. Häse CC, Mekalanos JJ (1998) TcpP protein is a positive regulator of virulence gene expression in *Vibrio cholerae*. Proc Natl Acad Sci USA 95:730–734.
8. Yang Y, Isberg RR (1997) Transcriptional regulation of the *Yersinia pseudotuberculosis* pH6 antigen adhesin by two envelope-associated components. Mol Microbiol 24:499–510.
9. Egan S, James S, Kjelleberg S (2002) Identification and characterization of a putative transcriptional regulator controlling the expression of fouling inhibitors in *Pseudoalteromonas tunicata*. Appl Environ Microbiol 68:372–378.
10. Seki T, Yoshikawa H, Takahashi H, Saito H (1987) Cloning and nucleotide sequence of *phoP*, the regulatory gene for alkaline phosphatase and phosphodiesterase in *Bacillus subtilis*. J Bacteriol 169:2913–2916.
11. Melchers LS, Thompson DV, Idler KB, Schilperoort RA, Hooykaas PJ (1986) Nucleotide sequence of the virulence gene *virG* of the *Agrobacterium tumefaciens* octopine Ti plasmid: significant homology between *virG* and the regulatory genes *ompR*, *phoB* and *dye* of *E. coli*. Nucleic Acids Res 14:9933–9942.
12. Comeau DE, Ikenaka K, Tsung KL, Inouye M (1985) Primary characterization of the protein products of the *Escherichia coli ompB* locus: structure and regulation of synthesis of the OmpR and EnvZ proteins. J Bacteriol 164:578–584.
13. Kuper C, Jung K (2005) CadC-mediated activation of the *cadBA* promoter in *Escherichia coli*. J Mol Microbiol Biotechnol 10:26–39.
14. Tetsch L, Koller C, Haneburger I, Jung K (2008) The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. Mol Microbiol 67: 570–583.
15. Tetsch L, Jung K (2009) The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. Mol Microbiol 73:982–991.
16. Dell CL, Neely MN, Olson ER (1994) Altered pH and lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. Mol Microbiol 14:7–16.
17. Haneburger I, Eichinger A, Skerra A, Jung K (2011) New insights into the signaling mechanism of the pH-

responsive, membrane-integrated transcriptional activator CadC of *Escherichia coli* J Biol Chem, in press. [doi:10.1074/jbc.M110.196923].
18. Fritz G, Koller C, Burdack K, Tetsch L, Haneburger I, Jung K, Gerland U (2009) Induction kinetics of a conditional pH stress response system in *Escherichia coli*. J Mol Biol 393:272–286.
19. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372: 774–797.
20. Lawrence MC, Colman PM (1993) Shape complementarity at protein/protein interfaces. J Mol Biol 234: 946–950.
21. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and and empirical binding free energy function. J Comp Chem 19: 1639–1662.
22. Holm L, Kaariainen S, Rosenstrom P, Schenkel A (2008) Searching protein structure databases with DaliLite v.3. Bioinformatics 24:2780–2781.
23. Sampathkumar P, Roach C, Michels PA, Hol WG (2008) Structural insights into the recognition of peroxisomal targeting signal 1 by *Trypanosoma brucei* peroxin 5. J Mol Biol 381:867–880.
24. Steegborn C, Danot O, Huber R, Clausen T (2001) Crystal structure of transcription factor MalT domain III: a novel helix repeat fold implicated in regulated oligomerization. Structure 9:1051–1060.
25. Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res 33:W244–248.
26. Wilson CG, Kajander T, Regan L (2005) The crystal structure of NlpI. A prokaryotic tetratricopeptide repeat protein with a globular fold. FEBS J 272: 166–179.
27. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410.
28. Bleasby AJ, Akrigg D, Attwood TK (1994) OWL–a nonredundant composite protein sequence database. Nucleic Acids Res 22:3574–3577.
29. Boggon TJ, Shapiro L (2000) Screening for phasing atoms in protein crystallography. Structure 8: R143–149.
30. CCP4 (1994) The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50: 760–763.
31. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. Acta Crystallogr D Biol Crystallogr 58:1772–1779.
32. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr D Biol Crystallogr 59:2023–2030.
33. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132.
34. Bergerhoff G, Brown ID, Inorganic Crystal Structure Database. In Allen FH, Bergerhoff G, Sievers R, Eds. (1987) Crystallographic Databases. Chester: IUCr.
35. Barbour LJ, MacGillivray LR, Atwood JL (1996) Crystal and molecular structure of $[H_3O\bullet 18\text{-crown-}6]_2$-$[ReCl_6]$ isolated from a liquid clathrate medium. J Chem Cryst 26:59–61.
36. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Cryst 40:658–674.
37. Kabsch W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J Appl Cryst 26:795–800.
38. Weichenberger CX, Sippl MJ (2007) NQ-Flipper: recognition and correction of erroneous asparagine and glutamine side-chain rotamers in protein structures. Nucleic Acids Res 35:W403–406.
39. Laskowski RA, MacArthur MW, Mos DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26: 283–291.
40. Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. Nature 381:272.
41. Kabsch W, Sander C (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22: 2577–2637.
42. Bond CS (2003) TopDraw: a sketchpad for protein structure topology cartoons. Bioinformatics 19:311–312.
43. DeLano WL (2002) The PyMOL Molecular Graphics System. California: DeLano Scientific.
44. Nicholls A, Sharp KA, Honig B (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 11: 281–296.
45. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948.
46. Bond CS, Schüttelkopf AW (2009) ALINE: a WYSIWYG protein-sequence alignment editor for publication-quality alignments. Acta Crystallogr D Biol Crystallogr 65: 510–512.
47. Eichinger A, Nasreen A, Kim HJ, Skerra A (2007) Structural insight into the dual ligand specificity and mode of high density lipoprotein association of apolipoprotein D. J Biol Chem 282:31068–31075.

4 New insights into the signaling mechanism of the pH-responsive, membrane-integrated transcriptional activator CadC of *Escherichia coli*

This research was originally published in THE JOURNAL OF BIOLOGICAL CHEMISTRY. I. Haneburger, A. Eichinger, A. Skerra, and K. Jung. New Insights into the Signaling Mechanism of the pH-responsive, Membrane-integrated Transcriptional Activator CadC of *Escherichia coli*. 286: 10681-10689.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 12, pp. 10681–10689, March 25, 2011

New Insights into the Signaling Mechanism of the pH-responsive, Membrane-integrated Transcriptional Activator CadC of *Escherichia coli*[*] [S]

Received for publication, October 25, 2010, and in revised form, December 22, 2010 Published, JBC Papers in Press, January 7, 2011, DOI 10.1074/jbc.M110.196923

Ina Haneburger[‡§], Andreas Eichinger[‡¶], Arne Skerra[‡¶], and Kirsten Jung[‡§1]

From the [‡]Center of Integrated Protein Science Munich and [§]Department of Microbiology, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany and the [¶]Lehrstuhl für Biologische Chemie, Technische Universität München, 85350 Freising-Weihenstephan, Germany

The membrane-integrated transcriptional regulator CadC of *Escherichia coli* activates expression of the *cadBA* operon at low external pH with concomitantly available lysine, providing adaptation to mild acidic stress. CadC is a representative of the ToxR-like proteins that combine sensory, signal transduction, and DNA-binding activities within a single polypeptide. Although several ToxR-like regulators such as CadC, as well as the main regulator of *Vibrio cholerae* virulence, ToxR itself, which activate gene expression at acidic pH, have been intensively investigated, their molecular activation mechanism is still unclear. In this study, a structure-guided mutational analysis was performed to elucidate the mechanism by which CadC detects acidification of the external milieu. Thus, a cluster of negatively charged amino acids (Asp-198, Asp-200, Glu-461, Glu-468, and Asp-471) was found to be crucial for pH detection. These amino acids form a negatively charged patch on the surface of the periplasmic domain of CadC that stretches across its two subdomains. The results of different combinations of amino acid replacements within this patch indicated that the N-terminal subdomain integrates and transduces the signals coming from both subdomains to the transmembrane domain. Alterations in the phospholipid composition did not influence pH-dependent *cadBA* expression, and therefore, interplay of the acidic surface patch with the negatively charged headgroups is unlikely. Models are discussed according to which protonation of these acidic amino acid side chains reduces repulsive forces between the two subdomains and/or between two monomers within a CadC dimer and thereby enables receptor activation upon lowering of the environmental pH.

Bacteria that prefer the human intestine as a natural habitat encounter a broad range of environmental stresses, including acidic pH (1). To survive a decrease in the external pH, neutrophilic bacteria such as *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio cholerae* possess different stress response systems to maintain the cytoplasmic pH in the physiological range. One system that is important for the response to mild acidic stress in *E. coli* is the Cad system (1). The Cad system of *E. coli* consists of the cytoplasmic protein CadA and the transmembrane proteins CadB and CadC. The lysine decarboxylase CadA converts lysine upon consumption of a cytoplasmic proton to cadaverine and carbon dioxide. Subsequently, CadB transports the basic reaction product cadaverine in exchange for lysine out of the cell (2, 3). Transcription of the *cadBA* operon in *E. coli* is activated when the external pH drops below pH 6.6 if external lysine is available (4, 5). The main activator of the Cad system is the inner membrane protein CadC (6).

CadC is a member of the family of ToxR-like transcriptional activators that comprise several low pH-regulated activators, including the main regulator for virulence of *V. cholerae*, ToxR. ToxR-like regulators are characterized by a conserved modular composition. They exhibit a cytoplasmic DNA-binding domain with a winged helix-turn-helix motif, a single transmembrane domain, and a periplasmic sensor domain. In contrast to histidine kinase/response regulator systems, in which signal transduction is accompanied by phosphorylation reactions, signal transduction in ToxR-like regulators is mediated without chemical modification (7). Thus, ToxR-like regulators are one-component systems that exhibit sensory, signal transduction, and effector function within one polypeptide (8). Although transcriptional activation by ToxR as well as CadC has been intensively investigated using genetic and biochemical methods, the precise molecular mechanism is hardly understood. Previously, we showed that CadC is not the direct sensor for lysine but senses lysine via interplay with the lysine-permease LysP as co-sensor (9). Using random mutagenesis, Dell *et al.* (10) identified several CadC variants with mutations in the periplasmic domain that were impaired in the pH-dependent regulation of the Cad system. Therefore, the pH sensory property was assigned to the periplasmic domain of CadC. Recently, the three-dimensional structure of the periplasmic domain of *E. coli* CadC was solved (41). In this study, we employed a structure-guided mutational analysis that led to the identification of a negatively charged patch comprising acidic amino acids on the CadC periplasmic domain surface as crucial for sensing low pH.

* This work was supported by Deutsche Forschungsgemeinschaft Grants JU270/5-1 and Exc114/1 and by an Elitenetzwerk Bayern research scholarship (BayEFG) (to I. H.).

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and Table 1.

The atomic coordinates and structure factors (code 3LY7) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] To whom correspondence should be addressed: Dept. Biologie I, Mikrobiologie, Biozentrum, Ludwig-Maximilians-Universität München, Grosshadernerstr. 2–4, 82152 Martinsried, Germany. Tel.: +49-89-2180-74500; Fax: +49-89-2180-74520; E-mail: jung@lmu.de.

EXPERIMENTAL PROCEDURES

Bacterial Strains and Growth Conditions—*E. coli* JM109 or DH5α (11, 12) was used as a carrier for the plasmids described. All plasmids used in this study are listed in supplemental Table 1. *E. coli* BL21(DE3)pLysS (13) was used for expression of *cadC* and its variants from the T7 promoter. *E. coli* EP314 (14), which carries a *cadA-lacZ* fusion gene and a Tn10 transposon in *cadC*, was complemented with plasmids (pET16b-based) encoding *cadC* and its variants and used for *cadBA* transcriptional analysis. For strain maintenance, plasmid preparation, and protein overproduction, strains were grown in LB medium (15). To monitor signal transduction *in vivo*, *E. coli* EP314 transformed with the indicated plasmids was grown in minimal medium (16). The pH was adjusted to either pH 5.8 or pH 7.6 using the corresponding phosphate-buffered medium. *E. coli* strains with modified phospholipid content (*E. coli* AD93 with or without plasmid pDD72 (17) and HDL1001 (18)) were grown in LB medium that was pH-adjusted using 100 mM MOPS (pH 7.6 or 7.0) or MES (pH 5.8). To enable growth of AD93, 50 mM $MgCl_2$ was added to the medium (17). Where indicated, lysine was added to a final concentration of 10 mM. Antibiotics were added for selection purpose at concentrations of 100 μg/ml (ampicillin), 50 μg/ml (kanamycin), and 34 μg/ml (chloramphenicol).

Construction of cadC Variants—All *cadC* variants were constructed by either one- or two-step PCR using mismatched primers (19). To facilitate construction, four unique restriction sites (XhoI, 1338T→C and 1341A→G; XmaI, 1443A→C; SacI, 1002T→A, 1005G→T, and 1006T→C; and SacII, 588A→G and 591C→G; numbers indicate the nucleotide positions in the *cadC* sequence) were introduced into plasmid pET16b-cadC2 (9) by silent mutation, resulting in plasmids pET16b-cadC3 (XhoI), pET16b-cadC4 (XhoI and XmaI), pET16b-cadC5 (XhoI and SacI), pET16b-cadC6 (XhoI, XmaI, and SacI), and pET16b-cadC7 (XhoI, XmaI, SacI, and SacII) (supplemental Table 1).

Detection of CadC in the Membrane Fraction—Verification of production and membrane integration was carried out for all CadC variants as described previously (9).

Measurement of CadC Signal Transduction Activity in Vivo—Cultivation of cells and measurements of the signal transduction activity of different CadC variants by β-galactosidase assays were performed as described previously (9). To analyze the influence of changes in the phospholipid composition, cells of an overnight culture (grown in LB medium at pH 7.6) were inoculated into fresh LB medium (pH 7.6) with the A_{600} adjusted to 0.05. Cells were grown under aerobic conditions at 37 °C to mid-logarithmic growth phase ($A_{600} \sim 0.7$) and harvested by centrifugation (37 °C). Pellets were resuspended and transferred to prewarmed LB medium (pH 5.8–7.6) by adjusting the A_{600} to 0.7. After growth for another 1.5 h, cells were harvested by centrifugation. To test CadA activity, a modified spectrophotometric assay for lysine decarboxylase as described by Lemonnier and Lane (20) was used. For this purpose, the cell density was adjusted to $A_{600} = 1$ with 20 mM potassium P_i buffer (pH 5.8). Subsequently, 200 μl of this cell suspension was mixed with 20 μl of chloroform to disrupt cells. After settling of the chloroform, 10 μl of the chloroform-free supernatant was added to 120 μl of either a lysine-free (control) or a lysine-containing reaction mixture consisting of 16 mM potassium P_i buffer (pH 5.8) and 0.1 mM pyridoxal phosphate (Sigma) with and without 5 mM lysine (Sigma) 37 °C. The enzymatic reaction was performed at 37 °C for 15 min and thereafter stopped by the addition of 120 μl of 1 M Na_2CO_3. Subsequently, lysine and cadaverine were derivatized by adding 120 μl of 10 mM picrylsulfonic acid (2,4,6-trinitrobenzenesulfonic acid; Sigma), followed by incubation at 40 °C for 4 min. To each sample toluene (1 ml) was added, followed by vigorously mixing for 20 s. The toluene and aqueous layers were separated by centrifugation. *N*,*N*′-Bistrinitrophenyllysine is soluble in water but not in toluene, whereas the reverse is true for *N*,*N*′-bistrinitrophenyl cadaverine (21). The absorption of *N*,*N*′-bistrinitrophenyl cadaverine (toluene phase) was determined at 340 nm, and the specific activity of CadA (μmol/min*mg) was calculated.

Three-dimensional Structure Analysis—Analysis of the three-dimensional structure of the periplasmic domain (Protein Data Bank code 3LY7) was carried out using the PyMOL molecular viewer, version 0.99rc6 (22). Calculation of electrostatic surface was carried out with the program GRASP (23).

RESULTS

Histidine Residues Are Not Involved in pH Detection by CadC—Moderate acidification (pH 6.6) of the surrounding medium in the presence of lysine is sufficient to activate transcription of the *cadBA* operon in *E. coli*. Full induction of this operon is accomplished at or below pH 5.8 (4). These values are close to the pH value at which histidine side chains become protonated (pK_a around 6). In addition, histidines have already been shown to mediate the pH response in several proteins such as the sensor kinase PhoQ of *Salmonella typhimurium* (24) and the histidine kinase ArsS of *Helicobacter pylori* (25). The periplasmic domain of CadC contains six histidine residues (His-240, His-332, His-345, His-349, His-387, and His-390), which were all replaced with either leucine or glutamine to substitute the imidazole ring with different side chains of similar size. Plasmids encoding the corresponding CadC variants were introduced into the reporter strain *E. coli* EP314, which lacks a functional chromosomal *cadC* gene and carries a *cadA-lacZ* fusion gene (9, 14). With the exception of CadC-H240Q, all variants of CadC with histidine replacements induced *cadBA* expression at low pH but not at pH 7.6, similar to wild-type CadC (Fig. 1*A*). Because all CadC variants were inserted into the membrane (Fig. 1*B*), the lack of response of CadC-H240Q seemed to be related to defects in sensing and/or signal transduction (off-state). To test whether the off-state of CadC-H240Q was due to the loss of the positive charge of the protonated imidazole ring or other structural effects, arginine and leucine were introduced at this position, respectively. Indeed, replacement of His-240 with leucine and also with arginine resulted in pH-responsive CadC variants, albeit the signaling capacity of CadC-H240R was somewhat reduced (Fig. 1*A*). These results indicate that the imidazole side chain of His-240 is not important for pH sensing but that this amino acid seems to play a more intricate role in signaling.

Glu-461, Glu-468, and Asp-471 Are Involved in pH Sensing—Full induction of *cadBA* transcription occurs at pH 5.8, a pH value at which initial protonation of aspartate (side chain pK_a of

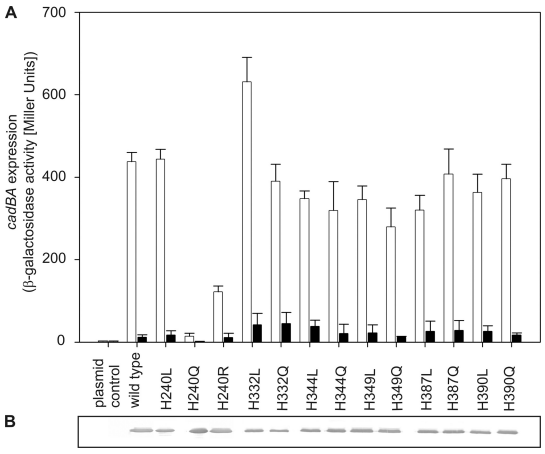

FIGURE 1. *A*, influence of histidine replacements in the periplasmic domain of CadC on *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA*′::*lacZ*) that was complemented with plasmid-encoded *cadC* or the mutant indicated. Cells were cultivated under microaerobic conditions in minimal medium in the presence of 10 mM lysine at pH 7.6 (*black bars*) or at pH 5.8 (*white bars*). The activity of the reporter enzyme β-galactosidase was determined according to Miller (40) and served as a measure for *cadBA* expression. The experiment was performed in triplicate, and *error bars* indicate S.D. *B*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli*. *E. coli* BL21(DE3)pLysS was transformed with plasmids encoding either wild-type or mutant *cadC*. Each lane contained 25 μg of membrane protein. CadC was detected by a monoclonal mouse antibody against the His tag and an alkaline phosphatase-coupled secondary antibody.

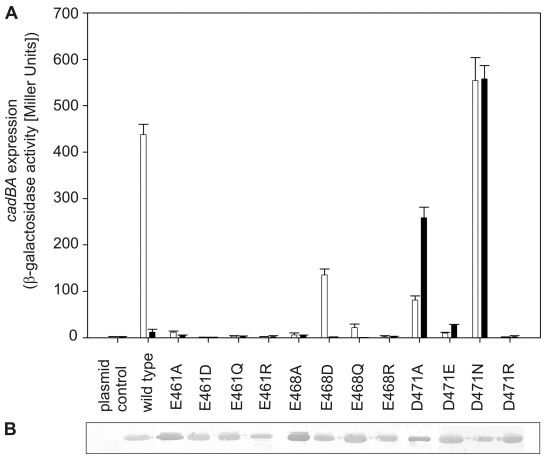

FIGURE 2. *A*, influence of replacements of acidic amino acids in the C-terminal region of the CadC periplasmic domain on *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA*′::*lacZ* fusion) that was complemented with plasmid-encoded *cadC* or the indicated *cadC* mutant. Cells were cultivated under microaerobic conditions in minimal medium with 10 mM lysine at pH 7.6 (*black bars*) or at pH 5.8 (*white bars*). The reporter enzyme β-galactosidase was determined as described in the legend to Fig. 1. *B*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli* (see Fig. 1).

TABLE 1

Influence of replacements of aspartate and glutamate residues in the C-terminal subdomain of the periplasmic domain of CadC on *cadBA* expression

Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA*′::*lacZ*) that was complemented with plasmid-encoded *cadC* or the indicated *cadC* mutant. See also Fig. 2.

CadC variant	*cadBA* expression (β-galactosidase activity)	
	pH 5.8 + lysine	pH 7.6 + lysine
	Miller units	
Plasmid control	1.9	2.0
Wild-type CadC	437.7	11.8
D434N	417.8	25.1
D434R	250.4	12.5
E435Q	465.2	12.7
E435R	270.4	10.3
D445N	384.3	10.2
D445R	263.6	9.3
E447Q	789.4	31.5
E447R	699.6	8.8
E490Q	281.0	53.8
E490R	338.7	14.5
D506N	358.8	29.6
D506R	243.9	6.6
E512Q	359.7	35.5
E512R	272.7	35.4

3.9) or glutamate (side chain pK_a of ~4.3 (26)) residues, depending on their electrostatic environment, could be responsible for pH sensing. The periplasmic domain contains 19 aspartates and 15 glutamates. Because almost all pH-sensitive amino acids previously identified by Dell *et al.* (10) were located close to the C terminus, we initially focused on 10 C-terminal aspartate and glutamate residues (Asp-434, Glu-435, Asp-445, Glu-447, Glu-461, Glu-468, Asp-471, Glu-490, Asp-506, and Glu-512). Each amino acid was substituted with the related carboxamide amino acid devoid of charge (*i.e.* asparagine instead of aspartate and glutamine instead of glutamate) to mimic the uncharged protonated state and also with arginine to further introduce a positive net charge. The effect of each amino acid replacement was tested as described for the histidine variants. Most of the resulting CadC variants activated *cadBA* expression like wild-type CadC (Fig. 2 and Table 1). However, replacement of Glu-461, Glu-468, and Asp-471 caused either nonresponsive CadC variants (off-state) or pH-insensitive, *i.e.* constitutively activated, CadC variants (on-state) (Fig. 2*A*). Therefore, the role of these residues was further analyzed by additional amino acid replacements. For Glu-461, shortening the side chain (CadC-E461A), removing the negative charge (CadC-E461Q), and introducing a positive charge (CadC-E461R) resulted in nonresponsive proteins. Even the introduction of aspartate at this position (CadC-E461D) abolished the response to the external low pH signal (off-state). Hence, any alteration of Glu-461 led to CadC variants that were arrested in the off-state. More or less the same results were observed when Glu-468 was replaced. In this case, only the conservative replacement with aspartate (CadC-E468D) resulted in a pH-responsive sensor, albeit with lower activity (Fig. 2*A*). Notably, the replacement of aspartate at position 471 evoked different effects. Small uncharged amino acids such as alanine and glycine (CadC-D471A/G) (Fig. 2*A*) (10), as well as the larger uncharged asparagine (CadC-D471N) converted CadC to the on-state. On the other hand, the charged amino acids glutamate (CadC-D471E) and arginine (CadC-D471R) at this position led to nonresponsive CadC proteins, which no longer activated *cadBA* expression when exposed to low pH (Fig. 2*A*). As the three-dimensional structure of the periplasmic domain of

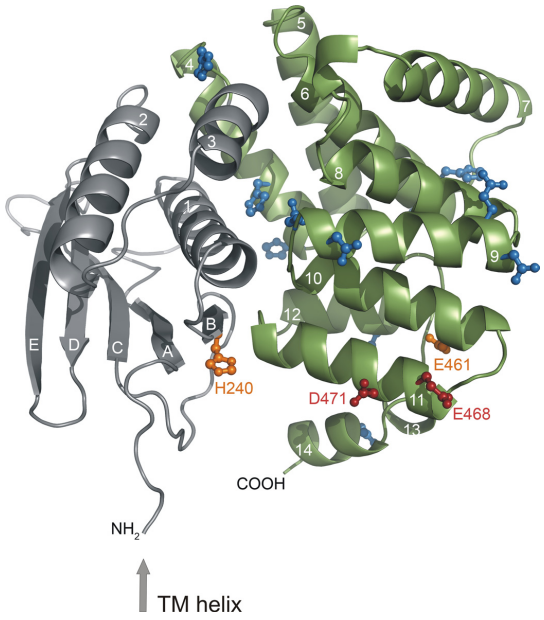

FIGURE 3. **Localization of histidine, glutamate, and aspartate residues investigated in the first mutagenesis screen within the three-dimensional structure of the periplasmic domain of CadC (41).** The preceding transmembrane (*TM*) helix is indicated by an *arrow*. The N-terminal subdomain of the periplasmic domain is colored *gray*, and the C-terminal subdomain is colored *green*. The side chains of residues that were mutated but did not alter the pH response of the corresponding CadC variants are colored *blue*. Amino acids whose mutagenesis suggested a role in pH sensing are indicated in *orange* (off-state) or *red* (on-state). Secondary structure elements are labeled with *white letters* (41).

CadC was recently solved (41), the amino acids involved in pH sensing identified thus far could be located in the tertiary structure. The CadC periplasmic domain consists of two subdomains (Fig. 3). Only His-240 is located in the N-terminal subdomain, whereas Glu-461, Glu-468, and Asp-471 are located in the C-terminal subdomain, which forms a bundle of α-helices (Fig. 3). Strikingly, Glu-468 and Asp-471 are located within the same α-helix (helix 11), like Leu-479 and Thr-475, which were previously identified by Dell *et al.* (10). Glu-461 is located in helix 10, and its side chain is oriented toward helices 11 and 12. Interestingly, His-240 is found in close proximity to α-helices 10 and 11 and displayed on the same protein face as Glu-468 and Asp-471.

Helix 11 Harbors pH-sensing Residues—Beside Glu-468 and Asp-471 within helix 11, two CadC variants with altered pH responses were already known from the random mutagenesis experiments of Dell *et al.* (10), namely CadC-T475A and CadC-L479S. To further clarify the role of helix 11 in pH sensing, we systematically exchanged all remaining surface-exposed residues (Arg-467, Leu-474, and Asn-478) (Fig. 4). In addition, we mutated Arg-480 and Phe-477 to analyze the influence of the charged residue (Arg-480) and the hydrophobic side chain (Phe-477) (Fig. 4). Beyond these residues, helix 11 contains four alanines. Because these alanines are part of the contact area with the neighboring helices 10 and 14, they were not replaced. Substitution of arginine at position 467 (CadC-R467A/Q/E/K) clearly abolished the pH response (off-state) of the corresponding CadC variants (Fig. 4*B*), although the mutations did not affect biosynthesis or membrane integration (Fig. 4*C*). Substitution of the other surface-exposed residues (CadC-L474A and CadC-N478A) led to CadC variants that remained in the on-state (Fig. 4*B*). Because changes of Arg-467, Leu-474, and Asn-

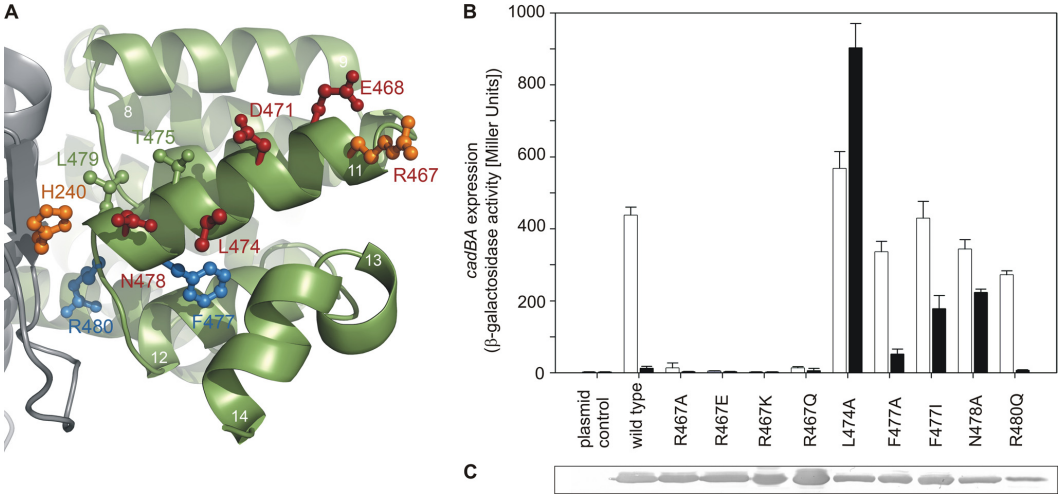

FIGURE 4. **Helix 11 of the periplasmic domain of CadC.** *A*, amino acids of helix 11 that were investigated are colored *blue* (wild type-like phenotype), *orange* (off-state), or *red* (on-state). Side chains that are depicted in *green* were previously identified by random mutagenesis (Thr-475 and Leu-479) (10). The color code of the subdomains and the labeling of secondary structure elements is the same as described in the legend to Fig. 3. *B*, influence of amino acid replacements in helix 11 in the periplasmic domain of CadC on *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA'*::*lacZ* fusion) that was complemented with plasmid-encoded *cadC* or the indicated *cadC* mutant as described in the legend to Fig. 1. *C*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli* BL21(DE3)pLysS.

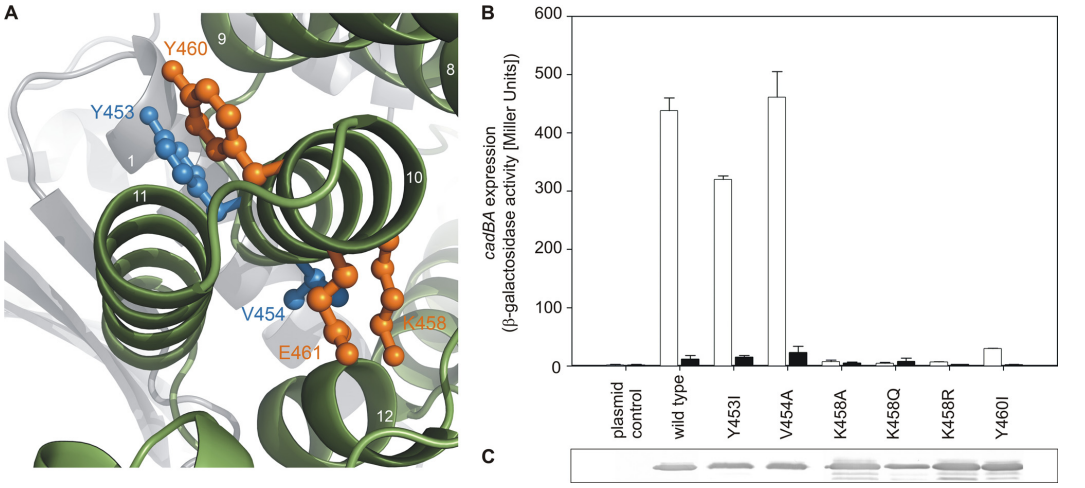

FIGURE 5. **Helix 10 of the periplasmic domain of CadC.** *A*, amino acids of helix 10 of the periplasmic domain of CadC embrace helix 11. Amino acids that were investigated are colored *blue* (wild type-like phenotype) or *orange* (off-state). The color code of the subdomains and the labeling of secondary structure elements is the same as described in the legend to Fig. 3. *B*, influence of amino acid replacements in helix 10 in the periplasmic domain of CadC on *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA'*::*lacZ* fusion) that was complemented with plasmid-encoded *cadC* or the indicated *cadC* mutant as described in the legend to Fig. 1. *Black bars*, pH 7.6; *white bars*, pH 5.8. *C*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli* (see Fig. 1).

478 (all unprotonable or positively charged amino acids) caused drastic effects on the pH-dependent *cadBA* expression, it is conceivable that these amino acids, together with Thr-475 and Leu-479, are responsible either for the correct positioning of the protonable residues within this helix, namely Glu-468 and Asp-471, or for their electrostatic force field. Exchange of Phe-477 with different residues resulted in opposing phenotypes: isoleucine at this position led to an almost pH-insensitive CadC variant (on-state), whereas the smaller alanine (CadC-F477A) did not alter pH sensing (Fig. 4*B*). Because removal of the aromatic side chain (CadC-F447A) did not affect pH-dependent activation of CadC, it is likely that the effect seen for CadC-F477I was due to steric alterations evoked through insertion of the β-branched isoleucine, which is less well accommodated in α-helices. Finally, exchange of the second inward-facing residue, Arg-480 (CadC-R480Q), did not significantly affect the pH-sensing capability of CadC.

Helix 10 Exerts a Stabilizing Effect—As Glu-461 is located in helix 10, the role of this helix in detecting low external pH was investigated in greater detail. Tyr-453, Val-454, Lys-458, and Tyr-460, together with Glu-461, decorate this helix on opposite sides (Fig. 5*A*). When Tyr-453 and Tyr-460 were replaced with isoleucine, the resulting CadC variants exhibited different phenotypes. CadC-Y453I regulated *cadBA* expression similarly to wild-type CadC, whereas CadC-Y460I blocked *cadBA* transcription (off-state) (Fig. 5*B*). Tyr-453 and Val-454, which were also inert to replacement (Fig. 5*B*), are located more at the N-terminal end of helix 10, whereas Tyr-460, as well as Lys-458 and Glu-461, *i.e.* sensitive residues, are situated at the C-terminal end. As in the case of Glu-461, each amino acid replacement of Lys-458 (CadC-K458A/Q/R) led to a pH-nonresponsive protein in the off-state (Fig. 5*B*), although these variants were synthesized and integrated into the membrane like wild-type CadC (Fig. 5*C*). The helix 10 residues analyzed are oriented toward helix 11 and seem to embrace it. For this reason and also because they are barely exposed to the surface of the protein, it is assumed that the nontitratable residues provide a kind of structural scaffold for the correct positioning of the pH-responsive residues Glu-468 and Asp-471 in helix 11.

Identification of a Negatively Charged Patch That Extends across Both Subdomains—Comparison of the location of amino acids involved in pH detection and of the electrostatic charge at the protein surface (Fig. 6*A*) revealed that the residues in helix 11 are part of a broader negatively charged patch. Importantly, one-half of this patch is formed by the N-terminal subdomain, whereas the second half is located on the C-terminal subdomain (encompassing residues of helix 11) of the periplasmic domain of CadC. Based on this notion, all other charged amino acids contributing to this patch were analyzed by mutagenesis: Asp-198, Asp-200, Glu-249, and also the positively charged Lys-242. The role of the aspartates and the glutamate was probed by removal of the negative charge (substitution with asparagine or glutamine, respectively) or by shortening the side chain (alanine). Furthermore, CadC variants that have a conservative exchange at these positions (mutual exchange between glutamate and aspartate, respectively) were tested. Mutations of Asp-198 and Asp-200 indicated the necessity of the aspartate side chain at these two positions. Each replacement (D198A/E/N and D200A/E/N) transformed the resulting CadC variant to a pH-nonresponsive off-state (Fig. 6*B*), although these proteins were properly integrated into the membrane (Fig. 6*C*). In contrast, substitution of Glu-249 did not significantly influence pH detection (Fig. 6*B*); all corresponding CadC variants (CadC-E249A/D/Q) activated *cadBA* expression in a wild type-like manner. In addition, we analyzed Lys-242 to explore the role of this positively charged amino acid within the

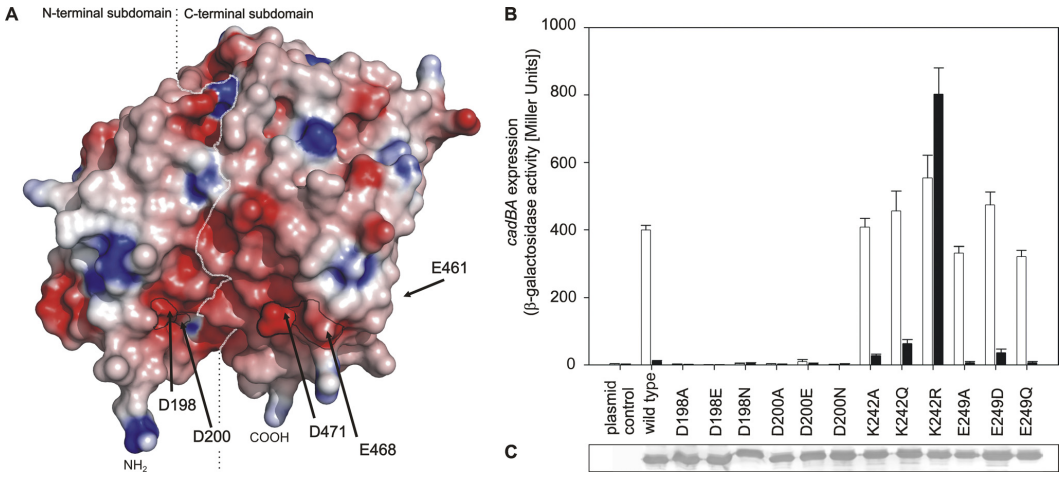

FIGURE 6. **Acidic residues from both subdomains form a contiguous negatively charged surface patch.** *A*, electrostatic surface coloring of the periplasmic domain of CadC. Negatively charged surfaces are shown in *red*, whereas positively charged areas are shown in *blue*. The positions of Asp-198, Asp-200, Glu-461, Glu-468, and Asp-471 at the protein surface are indicated by *arrows*. The boundary between the two subdomains is indicated by a *white dashed line*. The N (NH_2) and C (*COOH*) termini are indicated. *B*, influence of replacements of amino acids in the N-terminal subdomain on *cadBA* expression. Reporter gene assays were performed as described in the legend to Fig. 1. *Black bars*, pH 7.6; *white bars*, pH 5.8. *C*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli*.

cluster. Unexpectedly, only the mutation to arginine (CadC-K242R), which retained the positive charge, altered the CadC-mediated pH-dependent activation, whereas CadC-K242A and CadC-K242Q showed a response like the wild-type protein (Fig. 6*B*). At first glance, this result would suggest that the Lys side chain naturally occurs in CadC in the unprotonated state such that its replacement with the more basic arginine would lead to an electrostatic effect. However, the salt bridge between the Lys Nϵ group and the negatively charged carboxylate side chain of Asp-200 provides a more plausible explanation, as this electrostatic interaction is probably different and possibly even stronger with the shorter arginine side chain. Notably, Asp-200 also plays a crucial role in this context, as described above.

Role of Phospholipids in Sensing—Ca^{2+} sensing mediated by the histidine kinase PhoQ of *S. typhimurium*, for example, supposedly involves metal ion bridging between negatively charged side chains of the protein and the phospholipid headgroups of the cytoplasmic membrane (27). To test whether CadC might sense a higher H^+ concentration via changes in a similar kind of interaction, *cadBA* expression was analyzed in different *E. coli* phospholipid mutant strains. These experiments were based on the hypothesis that, at neutral pH, the repulsive force between CadC and the phospholipids is stronger than after protonation of residues in the negatively charged surface patch. Hence, the *cadBA* expression profile was determined in two phospholipid mutants, whereby major changes in the phospholipid composition are expected to result in significant effects on CadC-mediated output. *E. coli* AD93 (17) has an inactivated allele of the gene encoding the phosphatidylserine synthase (*pss93*::*kan*), which catalyzes the committed step in the synthesis of phosphatidylethanolamine. Due to this defect, the mutant has a much higher content of negatively charged phospholipids (phosphatidylglycerol and cardiolipin). This property can be reverted by expressing *pss* from a plasmid (pDD72) (17). The second mutant, *E. coli* HDL1001 (18), is genetically modified such that expression of *pgsA*, which is crucial for the production of phosphatidylglycerol and cardiolipin, is under the control of the *lac* promoter. This mutant is characterized by an extremely low content of these two phospholipids (<2%). Both mutants were grown in medium at different pH values (pH 7.6, 7.0, and 5.8), and *cadBA* expression was recorded by measuring CadA activity (supplemental Fig. S1). In both mutants, *cadBA* expression was induced in a pH-dependent manner, and no shift either in the onset of expression or in the induction level was observed in comparison with the corresponding control strains (supplemental Fig. S1). Thus, alteration of the phospholipid composition was without effect on *cadBA* expression, and it can be concluded that the activation mechanism of CadC does not involve specific interactions with the phospholipid headgroups.

The N-terminal Subdomain Transduces Signals to the Transmembrane Domain—To gain insight into the signaling mechanism between the two subdomains within the periplasmic region of CadC, different amino acid replacements leading to the on- or off-state were combined within one polypeptide. In this way, a putative intramolecular complementation of different phenotypes was tested. For these experiments, the amino acid replacements D198E (off-phenotype) and K242R (on-phenotype) in the N-terminal subdomain and the replacements D471E (off-phenotype) and D471N (on-phenotype) in the C-terminal subdomain were assembled in all pairwise combinations. Interestingly, an intramolecular complementation was not achieved with any combination (Fig. 7). However, amino acid exchanges in the N-terminal subdomain were always dominant over exchanges in the C-terminal subdomain. Thus, the replacements D198E and K242R transformed CadC into either the off- or on-state, respectively, regardless of the second site

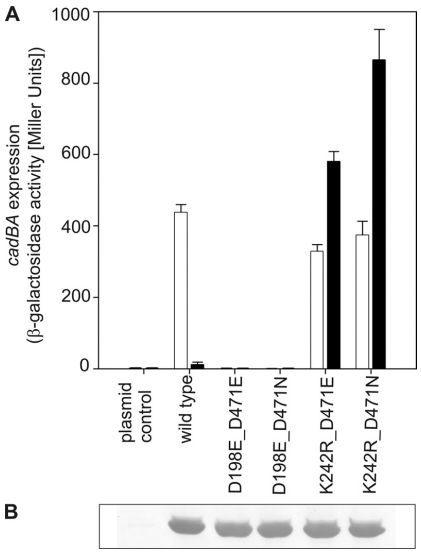

FIGURE 7. **Intramolecular complementation of signaling off- and on-state mutants.** *A*, influence of different combinations of amino acid replacements in the two subdomains on *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA'*::*lacZ* fusion) as described in the legend to Fig. 1. *Black bars*, pH 7.6; *white bars*, pH 5.8. *B*, verification of production and integration of CadC variants into the cytoplasmic membrane of *E. coli*.

mutation. These results suggest that the N-terminal subdomain integrates the signals coming from both subdomains and transduces them to the transmembrane and thus to the cytoplasmic domain (see Fig. 3).

DISCUSSION

The capability to adapt to changes in the external pH is vital throughout all three domains of life. Changes in the intracellular pH can have severe effects such as induction of apoptosis in human tissues (28–30) and cell death in bacteria. To cope with an acidification of the environment, different mechanisms have evolved. Besides complete adaptation of the mode of life (*i.e.* for acidophilic bacteria), cells often respond by changes in gene expression. Therefore, pH sensors are crucial to monitor the pH of the environment. Different strategies to detect acidic stress are conceivable, *e.g.* pH-dependent protein folding and unfolding (31), but direct measurement of protons would provide the most straightforward and rapid response.

The ToxR-like regulator CadC of *E. coli* senses acidification of the environment through its periplasmic domain, thereby enabling the bacterium to respond to unfavorable conditions before experiencing challenges to the cellular interior. Until now, the molecular mechanism by which CadC senses an increase in the external proton concentration was not clear.

Histidine residues often mediate pH detection through protonation as shown for the pH-dependent folding switch of the *Pseudomonas syringae* effector protein AvrPto (32), for interaction of the diphtheria toxin T domain (33), and the pH-dependent dimer dissociation of the dynein light chain LC8 (34). Therefore and due to the fact that the activation range of *cadBA* expression (4) essentially matches the pK_a value of the free histidine side chain, we analyzed the role of all histidines present in the periplasmic domain of CadC by site-directed mutagenesis. Our mutational analyses revealed, however, that protonation of histidines is not essentially involved in sensing low pH.

Further analyses of protonable residues in the periplasmic domain of CadC led to the identification of a negatively charged patch at the protein surface (Fig. 6). This patch is formed by Asp-198, Asp-200, Glu-461, Glu-468, and Asp-471 and extends across both subdomains. Strikingly, almost all residues are located on the same side of the protein surface, with Asp-198 and Asp-200 in the N-terminal subdomain and Glu-468 and Asp-471 in the C-terminal subdomain (Fig. 6). Glu-461 is almost buried within the protein but close to Glu-468. Substitution of the titratable residues within this patch had drastic effects (on-state, CadC-D471A/N; and off-state, CadC-D198A/E/N, CadC-D200A/E/N, CadC-E461A/D/Q, CadC-E468A/D/Q, and CadC-D471E/R). This might be due to a concerted effect of the entire group of negatively charged amino acids, also considering that the pK_a of titratable side chains depends on the electrostatic environment and thus on the protonation state of groups in its surroundings. One could speculate that sensing the external H^+ concentration with such an elaborate network could protect the cell from investing in an energy-consuming process (protein expression) under the wrong conditions and allows a graduated response of CadC to alterations of the external pH as it was already demonstrated (4).

Interestingly, a similar network of acidic residues was found to be involved in the pH-dependent gating of the acid-sensing ion channel 1a (35). In contrast, neither the crystal structure nor the site-directed mutagenesis studies presented here provided any evidence for the formation of specific salt bridges at neutral pH that may be broken upon proton binding within the CadC monomer, as was discussed for the pH-sensing mechanism of the bacterial potassium channel KcsA (36). Instead, it seems likely that the overall electrostatic potential at the protein surface (and associated variations in the Coulomb force field) plays a role in pH-dependent conformational changes.

Perturbation of the negatively charged patch, even via replacement of unprotonable or positively charged residues, led to CadC variants with either pH-insensitive (on-state, CadC-L479A and CadC-N478A) or nonresponsive (off-state, CadC-H240Q, CadC-L479A, CadC-N478A, CadC-T475A, CadC-L474A, CadC-R467A/E/K/Q, and CadC-K458A/Q/R) phenotypes. Because of the severe effects caused by these amino acid exchanges, one might speculate that these residues are vital for the precise structural or dynamic integrity of the protein, thus indirectly affecting pH sensing.

For the membrane-integrated kinase PhoQ, it was proposed that the positively charged calcium ions form bridges between the carboxylate side chains of the protein and the negatively charged headgroups of phospholipids (27). A similar bridging mechanism was described for the C2 membrane-binding domain of protein kinase Cα and its interaction with phosphatidylserine (37). Alternatively, protonation of the negatively charged surface patch of CadC might reduce repulsive forces between CadC and phospholipid headgroups. Thus, protonation instead of metal ion binding might lead to alterations in the

protein orientation with respect to the lipid bilayer and hence affect activity. However, based on our results obtained with two different phospholipid mutants of *E. coli*, the protonation-mediated involvement of phospholipids seems unlikely.

Three putative mechanisms by which protonation of the negatively charged residues within the surface patch affects the conformation or dynamics of CadC are conceivable: (1) reduction of intramolecular Coulomb repulsion, (2) reduction of intermolecular Coulomb repulsion, and (3) repositioning of monomer subunits within a functional dimer relative to each other. Mechanism 1 is based on the assumption that repelling surfaces exist between the two subdomains within one CadC monomer, more precisely between the two halves of the negatively charged patch. The two subdomains could alter their mutual arrangement via hinge movement at the connecting loop (Leu-330–Arg-333). Structural changes could be propagated to the transmembrane helix, resulting in binding of the cytoplasmic domain to the DNA and/or RNA polymerase. According to Mechanism 2, protonation would weaken the electrostatic repulsion between two CadC monomers, leading to pH-dependent oligomerization and corresponding signal transduction across the cytoplasmic membrane. Indeed, two CadC-binding sites have been identified in the P_{cad} promoter (38), such that a CadC dimer would be needed for efficient *cadBA* activation. This hypothesis is supported by recent observations from our structural analysis according to which the periplasmic domain of CadC forms a dimer (41). It is important to note that this protein-protein interaction is based mainly on polar interactions. Notably, the pH-responsive residues identified in this study are located mostly at the dimer interface (Fig. 8). Thus, it is conceivable that protonation of the corresponding residues would reduce intermolecular repulsion and allow approximation or, at least, spatial rearrangement (corresponding to Mechanism 3) of the two monomers. Unfortunately, such a signal-dependent change in the oligomerization state is difficult to prove experimentally due to the naturally extremely low copy number of CadC. Notably, none of the crystal structures solved so far indicates major pH-dependent structural changes. The crystal structure of the periplasmic wild-type domain, as well as the domain of the off-state variant CadC-D471E and the on-state variant CadC-D471N, showed nearly the same structure (41). This could be related to the fact that only a truncated version was crystallized so far that lacks the ability to undergo the changes of the full-length protein. Mechanism 3 also takes into account that a CadC dimer is needed to activate *cadBA* expression. In this case, the dimer is persistent, and the monomers would change their position relative to each other only as a consequence of a change in pH. Such a movement could be achieved, for example, by a rotation similar to that described for signaling in HAMP domains (39).

Regardless of the precise mechanism, the pH signal arising from protonation of different side chains within the acidic patch seems to be integrated and further transduced by the N-terminal subdomain. This notion is based on our study of combined amino acid replacements within the N- and C-terminal subdomains. All tested amino acid replacements in the N-terminal subdomain were found to dominate the phenotype of the resulting CadC variants (CadC-D198E, off-state; and CadC-K242R, on-state), regardless of the amino acid replacement in the N-terminal subdomain.

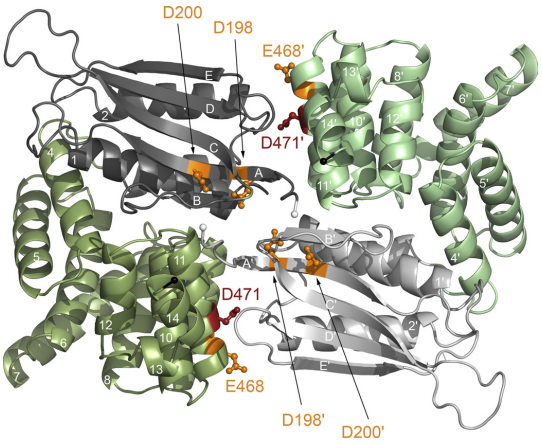

FIGURE 8. **Residues involved in pH sensing are located at the CadC periplasmic domain dimer interface.** To distinguish the monomers, they are colored in *dark* and *pale colors*, respectively. Residues that are crucial for pH sensing are indicated in *orange* (off-state) or *red* (on-state). The N-terminal subdomains of the periplasmic domain are colored *gray*, and the C-terminal subdomains are colored *green*. The N termini, which are preceded by the transmembrane helices, are indicated by *white spheres*, whereas the C termini are indicated by *black spheres*.

In conclusion, it is proposed that CadC senses a decrease in the external pH through direct binding of protons and associated conformational and/or oligomerization effects. Because of the dose-dependent activation of *cadBA* expression between pH 6.6 and 5.8 (4), it is conceivable that multiple residues are involved in pH detection, giving rise to a network of pH-responsive side chains that may sense the external proton concentration in a cooperative manner. Depending on the external pH, the charge of the identified surface patch varies due to different protonation propensities of each titratable residue within the network. Once a certain degree of charge neutralization is achieved, cooperative conformational changes and/or oligomerization occurs, resulting in the activation of *cadBA* expression in the cytoplasm.

Acknowledgments—We thank Sophie Buchner and Ulla Kyrylenko for help with construction and analysis of some CadC variants.

REFERENCES

1. Bearson, S., Bearson, B., and Foster, J. W. (1997) *FEMS Microbiol. Lett.* **147,** 173–180
2. Auger, E. A., Redding, K. E., Plumb, T., Childs, L. C., Meng, S. Y., and Bennett, G. N. (1989) *Mol. Microbiol.* **3,** 609–620
3. Soksawatmaekhin, W., Kuraishi, A., Sakata, K., Kashiwagi, K., and Igarashi, K. (2004) *Mol. Microbiol.* **51,** 1401–1412
4. Fritz, G., Koller, C., Burdack, K., Tetsch, L., Haneburger, I., Jung, K., and Gerland, U. (2009) *J. Mol. Biol.* **393,** 272–286
5. Meng, S. Y., and Bennett, G. N. (1992) *J. Bacteriol.* **174,** 2670–2678
6. Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L., and Olson, E. R. (1992) *J. Bacteriol.* **174,** 530–540
7. Miller, V. L., Taylor, R. K., and Mekalanos, J. J. (1987) *Cell* **48,** 271–279
8. Ulrich, L. E., Koonin, E. V., and Zhulin, I. B. (2005) *Trends Microbiol.* **13,** 52–56

9. Tetsch, L., Koller, C., Haneburger, I., and Jung, K. (2008) *Mol. Microbiol.* **67,** 570–583
10. Dell, C. L., Neely, M. N., and Olson, E. R. (1994) *Mol. Microbiol.* **14,** 7–16
11. Taylor, R. G., Walker, D. C., and McInnes, R. R. (1993) *Nucleic Acids Res.* **21,** 1677–1678
12. Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene* **33,** 103–119
13. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol.* **189,** 113–130
14. Neely, M. N., Dell, C. L., and Olson, E. R. (1994) *J. Bacteriol.* **176,** 3278–3285
15. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
16. Epstein, W., and Kim, B. S. (1971) *J. Bacteriol.* **108,** 639–644
17. DeChavigny, A., Heacock, P. N., and Dowhan, W. (1991) *J. Biol. Chem.* **266,** 5323–5332
18. Heacock, P. N., and Dowhan, W. (1989) *J. Biol. Chem.* **264,** 14972–14977
19. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77,** 51–59
20. Lemonnier, M., and Lane, D. (1998) *Microbiology* **144,** 751–760
21. Phan, A. P., Ngo, T. T., and Lenhoff, H. M. (1982) *Anal. Biochem.* **120,** 193–197
22. DeLano, W. L. (2008) *The PyMOL Molecular Graphics System*, DeLano Scientific, Palo Alto, CA
23. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11,** 281–296
24. Prost, L. R., Daley, M. E., Le Sage, V., Bader, M. W., Le Moual, H., Klevit, R. E., and Miller, S. I. (2007) *Mol. Cell* **26,** 165–174
25. Müller, S., Götz, M., and Beier, D. (2009) *PLoS One* **4,** e6930
26. Nozaki, Y., and Tanford, C. (1967) *J. Biol. Chem.* **242,** 4731–4735
27. Cho, U. S., Bader, M. W., Amaya, M. F., Daley, M. E., Klevit, R. E., Miller, S. I., and Xu, W. (2006) *J. Mol. Biol.* **356,** 1193–1206
28. Lagadic-Gossmann, D., Huc, L., and Lecureur, V. (2004) *Cell Death Differ.* **11,** 953–961
29. Matsuyama, S., Llopis, J., Deveraux, Q. L., Tsien, R. Y., and Reed, J. C. (2000) *Nat. Cell Biol.* **2,** 318–325
30. Yu, J., Tian, S., Metheny-Barlow, L., Chew, L. J., Hayes, A. J., Pan, H., Yu, G. L., and Li, L. Y. (2001) *Circ. Res.* **89,** 1161–1167
31. Damaghi, M., Bippes, C., Köster, S., Yildiz, O., Mari, S. A., Kühlbrandt, W., and Muller, D. J. (2010) *J. Mol. Biol.* **397,** 878–882
32. Dawson, J. E., Seckute, J., De, S., Schueler, S. A., Oswald, A. B., and Nicholson, L. K. (2009) *Proc. Natl. Acad. Sci. U.S.A.* **106,** 8543–8548
33. Perier, A., Chassaing, A., Raffestin, S., Pichard, S., Masella, M., Ménez, A., Forge, V., Chenal, A., and Gillet, D. (2007) *J. Biol. Chem.* **282,** 24239–24245
34. Nyarko, A., Cochrun, L., Norwood, S., Pursifull, N., Voth, A., and Barbar, E. (2005) *Biochemistry* **44,** 14248–14255
35. Liechti, L. A., Bernèche, S., Bargeton, B., Iwaszkiewicz, J., Roy, S., Michielin, O., and Kellenberger, S. (2010) *J. Biol. Chem.* **285,** 16315–16329
36. Thompson, A. N., Posson, D. J., Parsa, P. V., and Nimigean, C. M. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 6900–6905
37. Verdaguer, N., Corbalan-Garcia, S., Ochoa, W. F., Fita, I., and Gómez-Fernández, J. C. (1999) *EMBO J.* **18,** 6329–6338
38. Kuper, C., and Jung, K. (2005) *J. Mol. Microbiol. Biotechnol.* **10,** 26–39
39. Hulko, M., Berndt, F., Gruber, M., Linder, J. U., Truffault, V., Schultz, A., Martin, J., Schultz, J. E., Lupas, A. N., and Coles, M. (2006) *Cell* **126,** 929–940
40. Miller, J. H. (1992) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
41. Eichinger, A., Haneburger, I., Koller, C., Jung, K., Skerra, A. (2011) *Protein Sci.*, in press

5 The feedback-inhibitor cadaverine suppresses pH response by binding to the pH susceptible site of CadC in *Escherichia coli*

Ina Haneburger[‡,1,2], Georg Fritz[‡,3], Nicole Jurkschat[1,2], Larissa Tetsch[1,2], Andreas Eichinger[1,4], Arne Skerra[1,4], Ulrich Gerland[3] and Kirsten Jung[1,2]

[1]from the Center of Integrated Protein Science Munich at the [2]Department of Microbiology, Ludwig-Maximilians-Universität München, 82152 Martinsried, Germany, the [3]Arnold Sommerfeld Center for Theoretical Physics and the Center for NanoScience at the Ludwig-Maximilians-Universität München, 80333 München, Germany and the [4]Lehrstuhl für Biologische Chemie, Technische Universität München, 85350 Freising-Weihenstephan, Germany.

To whom correspondence should be addressed: Kirsten Jung, Dept. Biologie I, Mikrobiologie, Biozentrum, Ludwig-Maximilians-Universität München, Großhadernerstr. 2 - 4, 82152 Martinsried, Germany, Tel.: +49-89-2180-74500; Fax: +49-89-2180-74520; E-mail: jung@lmu.de

Keywords: acid stress; mathematical model; one-component system; polyamine; ToxR

Background: CadC is a membrane-integrated sensor and transcriptional activator that integrates various external signals.
Results: CadC is inactivated by binding of the feedback inhibitor cadaverine to the activation site.
Conclusion: It is suggested that cadaverine binding interferes with the active CadC dimer conformation.
Significance: This study provides molecular details on feedback regulation mediated by a periplasmic sensor domain.

5.1 Summary

At acidic pH and in the presence of lysine the pH-sensor CadC activates transcription of the *cadBA* operon encoding the lysine-cadaverine antiporter CadB and the lysine decarboxylase CadA. In effect, these proteins convert external lysine into external cadaverine under the consumption of a cytoplasmic proton, thus contributing to acid stress adaptation in *Escherichia coli*. Earlier, it was shown that *cadBA* expression is feedback inhibited by cadaverine, and a cadaverine binding site was predicted within the central cavity of the periplasmic domain of CadC. Our present study demonstrates that the predicted binding site only partially accounts for the cadaverine sensitivity of CadC *in vivo*. Moreover, a second, pivotal binding site was identified, which is localized at the pH-responsive patch of amino acids at the surface of the periplasmic domain of CadC, directly at the homo-dimerization interface. These data were incorporated into a previously established mathematical model for the Cad module, and the temporal response upon acid shock was simulated for two cadaverine-insensitive CadC variants. Based on these data a model of sequential binding of cadaverine to CadC is suggested. Binding of cadaverine within the central cavity renders the second site accessible for cadaverine, and binding of cadaverine finally leads to the inactivation of CadC. Altogether, the data reveal binding of a feedback inhibitor directly to the activation site of a receptor, and represent a striking example for the deactivation of a pH-sensor.

5.2 Introduction

Rapid adaptation to changing environmental conditions is pivotal for the reproductive success of microbes. For instance, enterobacteria, such as *E. coli*, thrive in- and outside the human host, whereby growth conditions change suddenly and deviate drastically from their optimal ranges [Bearson *et al.*, 1997]. Adaptation to these ever-changing conditions requires accurate monitoring of critical parameters and a precise and specific extra/intracellular information flow in order to respond with appropriate alterations in gene expression and protein activity [Chung *et al.*, 2006]. Therefore, it is still challenging to understand how stress response systems integrate and process multiple input signals and generate conditional responses to environmental stimuli.

One such conditional stress response system is the Cad module of *E. coli*, in which the one-component receptor system CadC detects external acidification and activates transcription of the *cadBA* operon [Watson *et al.*, 1992]. *cadA* encodes the inducible lysine decarboxylase and converts lysine under consumption of a cytoplasmic proton to cadav-

erine and carbon dioxide [Gale and Epps, 1944]. Cadaverine is subsequently excreted via the lysine/cadaverine antiporter CadB. During this cycle, the proton consumption in the cytoplasm and the export of alkaline cadaverine are believed to be the main mechanisms to tackle acid stress [Meng and Bennett, 1992]. Since induction of *cadBA* under stress conditions (acidic pH, low oxygen tension) results in a drastic increase of the corresponding proteins (CadA increases to approx. 2% of cellular protein, see Stim-Herndon *et al.* [1996]), transcription has to be tightly controlled. In fact, it is known that the Cad module is only activated if in addition to the primary stimulus (low pH), the substrate for the decarboxylase is also present (high lysine level) [Sabo *et al.*, 1974] and the product of the reaction is absent (low cadaverine level) [Fritz *et al.*, 2009; Neely *et al.*, 1994]. Hence, CadC unifies the remarkable abilities of integrating and processing three external cues, transmitting this information to the cytoplasm and, accordingly, activating transcription of *cadBA*. Activation in the presence of lysine is most likely accomplished by removal of an inhibitory interaction of CadC with the lysine permease LysP mediated by the transmembrane domain [Tetsch *et al.*, 2008]. Acidification of the external milieu is most probably sensed by protonation of a patch of acidic amino acids at the dimer interface within the periplasmic domain of CadC [Eichinger *et al.*, 2011; Haneburger *et al.*, 2011]. This neutralization enables conformational changes that lead to activation of *cadBA* transcription. While these studies have shed light on the mechanisms of lysine sensing and low pH detection, the molecular details of feedback inhibition of CadC by cadaverine are still unclear.

Crystal structure analysis of the periplasmic sensor domain of CadC ($CadC_{pd}$) revealed a central cavity between the two periplasmic subdomains of each CadC monomer, which is adequate for accommodation of the positively charged cadaverine - both with respect to sterical fit and electrostatic complementarity [Eichinger *et al.*, 2011].

In the present study we functionally analyze the molecular details of CadC inactivation by the feedback inhibitor cadaverine. We identify a second, pivotal cadaverine binding site, that overlaps with the patch of pH-responsive amino acids. Mathematical simulations support a sequential inactivation model for CadC. It is proposed that binding of cadaverine to the first site induces conformational changes which render the second site accessible for cadaverine. Binding of cadaverine then lead to rapid inactivation of CadC.

5.3 Experimental Procedures

Bacterial strains and growth conditions - *E. coli* strains JM109 or DH5α [Taylor *et al.*, 1993; Yanisch-Perron *et al.*, 1985] were used as cloning hosts for the plasmids listed in Table S1. *E. coli* EP314 [Neely *et al.*, 1994], which carries a *cadA'–lacZ* fusion gene and a deletion in *cadC*, was complemented with plasmids (pET16b-based) encoding *cadC* and its variants and was used to determine *cadBA* expression. For strain propagation, plasmid preparation and protein overproduction, strains were grown on Luria Bertani (LB) medium [Miller *et al.*, 1992]. To monitor signal transduction *in vivo*, *E. coli* EP314 transformed with the indicated plasmids was grown in minimal medium as described below and previously published [Epstein and Kim, 1971; Tetsch *et al.*, 2008]. Antibiotics were added for selection at concentrations of 100 μg ml^{-1} (ampicillin), 50 μg ml^{-1} (kanamycin) and 34 μg ml^{-1} (chloramphenicol). For monitoring *cadBA* expression as well as CadA activity and cadaverine accumulation MG1655Δ*cadC* [MG1655 *cadC*::*kan*; MG1655 [Blattner *et al.*, 1997] x P1(W3110-Δ*cadC* [Soksawatmaekhin *et al.*, 2004])] was transformed with plasmids pET16b-cadC2 [Tetsch *et al.*, 2008], pET16b-cadC_Y453I [Haneburger *et al.*, 2011] and pET16b-cadC_T475A, respectively. These three strains were then cultivated as described earlier [Fritz *et al.*, 2009].

Measurement of extracellular cadaverine, intracellular CadA activity and cadBA *expression* - Enzymatic activity, cadaverine content and *cadBA* transcription were determined as described earlier [Fritz *et al.*, 2009].

Construction of cadC *variants* - Mutagenesis was performed by either one- or two-step PCR using the appropriate oligonucleotides (Table S1). To facilitate construction, *cadC* genes harboring singular restriction sites were employed [Haneburger *et al.*, 2011; Tetsch *et al.*, 2008].

In vivo *analysis of cadaverine dependent inhibition* - Cultivation of cells in minimal medium with 10 mM lysine and measurements of signal transduction activity of different CadC variants by ß-galactosidase assays were performed as described before [Tetsch *et al.*, 2008]. In addition, cells were cultivated with addition of cadaverine/HCl at the indicated concentrations to test for the effect on CadC-dependent *cadBA* expression.

Determination of in vitro *affinity for cadaverine* - The periplasmic domains of CadC_T229A_E447Q and wild type CadC were produced and purified as described in [Tetsch *et al.*, 2008]. Purification was complemented by an additional size exclusion chromatography step (Superdex 200 10/300 GL, Amersham Pharmacia), to remove the thioredoxin cleavage product of the initial fusion protein [Tetsch *et al.*, 2008]. The purified protein was used for scintillation proximity assays (with ^{14}C-cadaverine from Hartmann ana-

lytic) according to Quick and Javitch [2007] or tryptophan fluorescence measurements as described in Tetsch *et al.* [2008] with a slightly changed buffer (Trp-buffer: 50 mM sodium phosphate buffer (pH 7.0) and 150 mM NaCl).

Computational Model - To analyze the *in vivo* dynamics of the Cad module at a quantitative level, we previously established a mathematical model for the dynamics of the three system variables *cadBA* mRNA, CadA activity, and external cadaverine monitored in our experiments [Fritz *et al.*, 2009]. Briefly, the model was guided by the known biochemistry of the Cad module, describing the integration of the input signals by CadC, a quasi-static CadC activity function and accounting for the dynamics of transcription, translation and enzyme kinetics of the Cad module via ordinary differential equations. Here we have refined this model by including the perception that CadC is active as a dimer [Eichinger *et al.*, 2011]. The fraction of active CadC dimer per cell ($CadC_2$) at time t ($C_2(t)$), with respect to the maximal number of $CadC_2$ dimer per cell (C_0) is defined as a function of its input signals - pH(t), external lysine, l(t), and external cadaverine, $c(t)$ - and now reads

$$C_2(t)/C_0 = f(pH(t)) \times g(l(t)) \times h(c(t)) \tag{5.1}$$

wherein the functions f, g, and h in Eq. (5.1) take on values between 0 and 1 and are assumed to be of the Hill form typical for cooperative binding reactions. For instance, the cadaverine dependence takes the form

$$h(c(t)) = \frac{1}{1 + (c(t)/K_c)^{n_c}} \tag{5.2}$$

wherein K_c is the apparent cadaverine affinity of $CadC_2$ and n_c the corresponding Hill exponent (see Fritz *et al.* [2009] for all details of the other regulation functions). Transcriptional regulation of P_{Cad} by $CadC_2$ is described by a quantitative thermodynamic model analogous to Fritz *et al.* [2009] and the resulting rate equation for the time evolution of *cadBA* mRNA, $m(t)$, takes the form

$$\frac{d}{dt}m(t) = \nu_m \left(\frac{1 + \omega(C_2/K_{C_2})}{1 + C_2/K_{C_2}} \right)^2 - \lambda_m m(t) \tag{5.3}$$

Here, ν_m is the basal transcription rate of the *cadBA* promoter (P_{Cad}), ω the fold change between basal and maximal transcription rate, K_{C_2} the affinity of a single $CadC_2$ binding site on the *cadBA* promoter (P_{Cad} harbors two binding sites [Küper and Jung, 2005] - presumably for the dimer form of CadC - leading to the square in the first term of the right hand side) and λ_m the *cadBA* mRNA degradation rate. The kinetics of

enzyme expression, conversion of lysine to cadaverine and subsequent antiport by CadB were modeled as before [Fritz *et al.*, 2009]. Again, parameters were estimated by least-squares minimization. Briefly, all previously published data [Fritz *et al.*, 2009] were fitted with the refined model described above (the full list of fit parameters is given in Table S2). To account for the somewhat faster Cad response due to the slightly higher copy number of the plasmid-encoded CadC used here, compared to the previously used strain MG1655 [Fritz *et al.*, 2009], the transcription rate of P_{Cad} was varied as the only fit parameter (Table S2). Subsequently, the dynamical responses of the two mutant strains harboring CadC_T229A_E447Q and CadC_Y453I were fitted by varying only the cadaverine affinity K_c, the Hill exponent n_c and the maximal number of $CadC_2$ dimers C_0 (see figures S2 and S3 for the fit results).

5.4 Results

Cadaverine binding within the internal cavity - The product of lysine decarboxylation, cadaverine, has been shown to inhibit expression of the *cadBA* operon [Neely *et al.*, 1994]. Direct binding of cadaverine to the periplasmic domain of CadC was determined *in vitro* with an apparent K_d of 96 μM [Tetsch *et al.*, 2008]. The crystal structure of $CadC_{pd}$ was recently solved [Eichinger *et al.*, 2011]. As crystallization in the presence of cadaverine did not result in a $CadC_{pd}$ complex structure, docking studies with cadaverine were performed. The ten best scoring results of the docking studies performed earlier [Eichinger *et al.*, 2011] proposed a binding mode of cadaverine with an extended conformation within the central cavity of the periplasmic domain of CadC. According to these findings one amine group of cadaverine was coordinated by Glu447, Gln421 and Tyr374 and the second amine group by Asp225 and Thr229. To verify the results of these docking studies, we constructed CadC variants with single, double and triple amino acid substitutions at these positions. In addition, we analyzed the residues Trp450 and His344, found to coordinate the hexachlororhenate ion in the CadC complex that had been prepared for MAD phasing [Eichinger *et al.*, 2011]. Cadaverine-dependent down-regulation of transcription was determined in reporter assays. *E. coli* strain EP314 lacks the native *cadC* gene and possesses a *cadA'-lacZ* fusion gene. Therefore, this strain is unable to produce cadaverine via CadA (as a large part of *cadA* is replaced by *lacZ*) and, hence, cadaverine-dependent inhibition only relies on the addition of external cadaverine while CadC-dependent *cadBA* promoter activity can be monitored via β-galactosidase measurements. Complementation of this strain with plasmid-encoded CadC or CadC variants (pET16b-based, see Exper-

Table 5.1: Effect of amino acid substitutions within the central cavity and at the dimerization interface on cadaverine-dependent inhibition of *cadBA* expression.

central cavity			dimerization interface		
CadC variant	**IC_{50} (50% inhibition at x[mM] cadaverine)**	**residual activity at 10.4 mM cadaverine [%]**	**CadC variant**	**IC_{50} (50% inhibition at x[mM] cadaverine)**	**residual activity at 10.4 mM cadaverine [%]**
Y374A	0.6	8.6 ± 8.0	E468D	0.4	1.6 ± 0.3
Y374F	0.8	1.7 ± 0.4	D471A	0.4	1.8 ± 0.5
D225A	0.9	1.7 ± 1.3	H240L	0.8	2.2 ± 2.1
W450A	0.9	7.0 ± 10.5	D471N	1.0	1.7 ± 0.4
wild type	**1.0**	**1.1 ± 0.9**	**wild type**	**1.0**	**1.1 ± 0.9**
W450F	1.1	6.9 ± 6.4	F477A	1.4	1.9 ± 1.5
D225N	1.2	7.0 ± 8.9	Y453F	2.5	11.1 ± 1.3
D225N_E447Q	1.8	14.4 ± 12.2	N478A	5.5	26.4 ± 1.9
Q421A	1.8	4.1 ± 0.5	F477I	» 10	53.9 ± 6.6
H344L_E447Q	1.9	1.3 ± 1.3	L474A	» 10	89.7 ± 8.3
T229A_H344L	2.0	9.2 ± 4.0	Y453A	» 10	71.4 ± 9.4
H344L	2.0	7.0 ± 1.6	Y453I	» 10	73 ± 11.9
D225N_T229A_E447Q	2.1	7.0 ± 2.0	T475A	» 10	88.0 ± 14.0
T229A_H344L_E447Q	2.2	11.6 ± 3.8	T475S	» 10	62.7 ± 3.5
T229A	2.8	11.0 ± 4.7			
E447Q	3.4	13.9 ± 6.9			
D225W	3.4	13.2 ± 2.0			
T229A_E447Q	3.6	21.4 ± 1.3			
T229A_Q421A_E447Q	3.8	19.1 ± 4.2			

imental Procedures) allows the survey of CadC-dependent *cadBA* expression depending on externally applied stimuli. A spectrum of cadaverine concentrations (0 mM; 0.65 mM; 1.3 mM; 2.6 mM; 5.2 mM; 10.4 mM) was tested to determine the inhibitory effects on *cadBA* expression. Increasing cadaverine concentrations inhibited wild type CadC in a dose-dependent manner (figure 5.1A). Approximately 1 mM cadaverine was required to inhibit CadC wild type-dependent *cadBA* expression to about 50% (IC_{50}; table 5.1). In the presence of 10.4 mM cadaverine expression was inhibited by nearly 100% (residual activity: 1%; figure 5.1A). Most of the variants (CadC_D225A, CadC_D225N, CadC_Y374A, CadC_Y374F, CadC_W450A, CadC_W450F, CadC_Q421A, CadC_D225N_T229A_E447Q, CadC_H344L, CadC_H344L_E447Q, CadC_T229A_H344L) with substitutions in the central cavity were inhibited by cadaverine like wild type CadC ($IC_{50} < 2$ mM; figure S1). Only the variants CadC_D225W, CadC_-

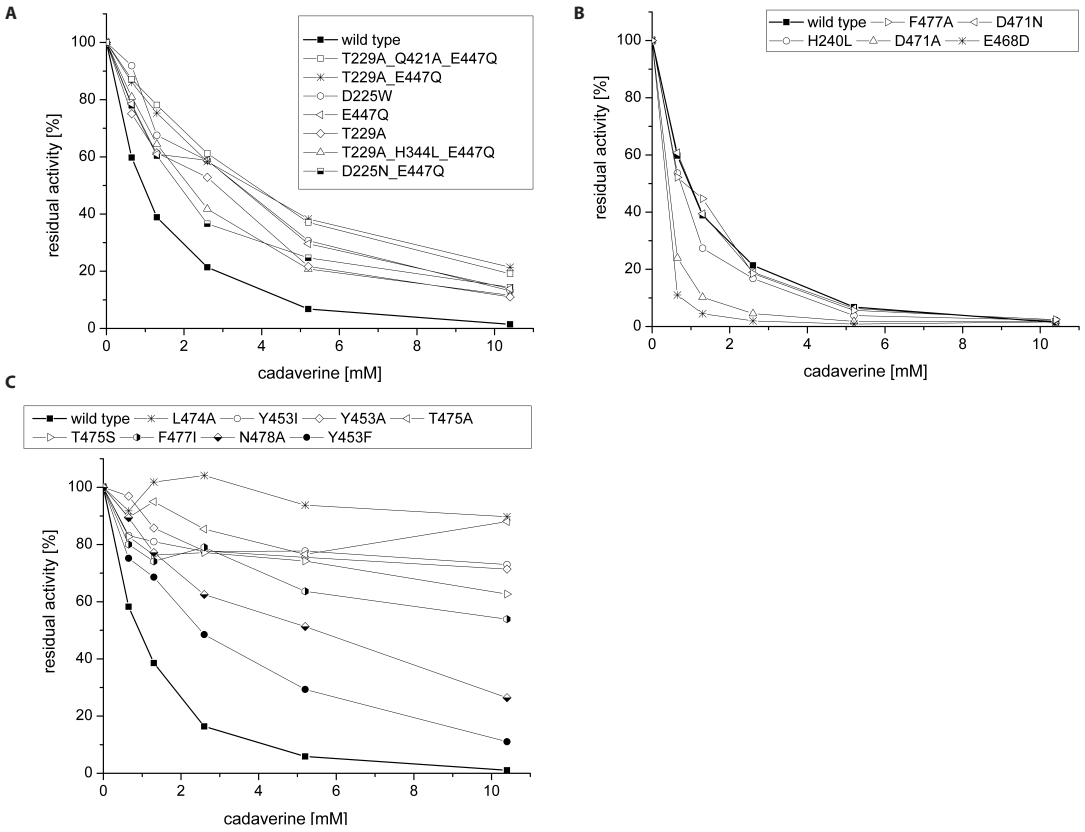

Figure 5.1: Inhibitory effect of cadaverine on CadC-dependent *cadBA* expression. Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10*, *cadA'*::*lacZ*) that was complemented with plasmid-encoded CadC (wild type) or the indicated CadC variants. An overnight culture (pH 7.6) was shifted to pH 5.8 with the indicated cadaverine concentrations (0.65; 1.3; 2.6; 5.2; 10.4 mM) and harvested in the mid-logarithmic growth phase. β-galactosidase activity was measured and the percentage of residual activity was calculated in relation to the same condition without cadaverine. Results are given as mean from at least three independent experiments. For clarity, the corresponding standard deviations were omitted. **A.** CadC variants with substitutions in the central cavity. **B.** CadC variants with substitutions close to the patch of acidic amino acids that do not influence cadaverine dependent inhibition. **C.** CadC variants with substitutions close to the patch of acidic amino acids that influence cadaverine dependent inhibition.

T229A and CadC_E447Q as well as the variants with multiple amino acid substitutions, CadC_D225N_E447Q, CadC_T229A_E447Q, CadC_T229A_H344L_E447Q and CadC_T229A_Q421A_E447Q were inhibited to a lesser extend by cadaverine (figure 5.1A). In addition, the IC_{50} of those variants impaired in cadaverine sensing was significantly increased (table 5.1). Polyamine binding often occurs at acidic or aromatic residues [Kashiwagi *et al.*, 1996, 2000; Soksawatmaekhin *et al.*, 2006; Sugiyama *et al.*, 1996; Vassylyev *et al.*, 1998]. Therefore, several substitutions at positions Asp225,

Tyr374 and Trp450 were analyzed (D225A, D225N, D225W; Y374A, Y374F; W450A, W450F). Replacement by alanine, asparagine or phenylalanine did not affect cadaverine sensing, compromising the idea that Asp225, Tyr374 and Trp450 directly coordinate cadaverine. Nonetheless, replacement by the more bulky tryptophan at position Asp225 reduced the inhibitory effect of cadaverine (table 5.1, IC_{50}: 3.4 mM cadaverine). Altogether, these experimental data indicated that the central cavity is involved in cadaverine sensing, as certain side chain substitutions within the cavity clearly reduced the inhibitory effect of cadaverine on CadC-dependent *cadBA* expression, although the precise nature of cadaverine coordination within the cavity remains elusive.

Identification of additional cadaverine coordinating residues – As no CadC variant with side chain substitutions within the central cavity was completely impaired in cadaverine sensing, we searched for additional cadaverine coordinating residues. In a random mutagenesis attempt Dell *et al.* [1994] identified amino acid substitutions that resulted in pH-independent CadC variants that were not affected by the addition of 2 mM cadaverine. Remarkably, variants CadC_T475A and CadC_L479S, both impaired in cadaverine-dependent inhibition, carry substitutions that co-localize with the patch of acidic amino acids in the periplasmic domain which was previously identified to be crucial for detection of acidification [Haneburger *et al.*, 2011].

Consequently, we examined whether cadaverine sensing and pH detection are linked to each other. To test this hypothesis, we investigated the effect of increasing cadaverine concentration on *cadBA* expression regulated by CadC variants harboring substitutions close to the patch of acidic amino acids. For this purpose, we used CadC variants that were previously constructed to analyze pH-dependent signaling [Haneburger *et al.*, 2011]. CadC_H240L, CadC_D471A, CadC_D471N, CadC_E468D and CadC_F477A were strongly affected by increasing concentration of cadaverine similar to wild type CadC (figure 5.1B, table 5.1). CadC_N478A and CadC_Y453F were less affected by cadaverine (figure 5.1C, table 5.1). In contrast, cadaverine exerted nearly no inhibiting effect on CadC_Y453I, CadC_F477I, CadC_T475A, and CadC_L474A (figure 5.1C, table 5.1). As the activity of these variants in the presence of 10.4 mM cadaverine still exceeded 60%, the IC_{50} can be assumed to be above 10 mM (table 5.1).

To understand whether the inability of these variants to sense cadaverine was evoked just by replacement of the amino acid side chain or by indirect effects such as conformational changes, Tyr453 was exchanged by phenylalanine (to preserve the aromatic character), isoleucine (insertion of an aliphatic non-aromatic side chain) or alanine (to remove nearly the complete side chain). Our data reveal that CadC is inactivated in the presence of cadaverine as long as an aromatic side chain is present at position 453 (wt and Y453F) (figure 5.1C). However, cadaverine can no longer exert its inhibitory

effect if the aromatic side chain is removed either through substitution by a short (Ala) or an aliphatic side chain (Ile) (figure 5.1B and C). For residue Thr475 the influence of a substitution by serine was tested. Since the variant CadC_T475S was not inactivated by cadaverine (cadaverine did not inhibit *cadBA* expression in a strain harboring this variant as well as CadC_T475A), it is likely that for Thr475 the nature of the branched amphipathic side chain is crucial. Since CadC_F477A is inhibited by cadaverine, the effect seen with CadC_F477I seems not to be caused by the lack of the aromatic side chain. It seems likely that the branched aliphatic side chain of isoleucine interferes with the 3D structure of CadC thereby impairing signal sensing. CadC_F477I reportedly [Haneburger *et al.*, 2011] is also impaired in pH detection, as it induces *cadBA* expression irrespective of the external pH. Together with the results for CadC_T475A and those for CadC_L479A from Dell *et al.* [1994] these data suggest a second cadaverine binding site close to the patch of acidic amino acids at the dimer interface encompassing residues Tyr453, Thr475, Leu474 and Leu479, since CadC variants with substitutions at these positions activate expression of *cadBA* irrespective of the presence of cadaverine. Closer inspection of the three-dimensional structure of CadC suggested that cadaverine binding to the second binding site might arise from a small number of residues at the dimer interface (figure 5.2, red residues). It appears that there are two clearly defined and separate binding sites at the dimer interface; one on each monomer. Although cadaverine binding and pH detection seem to be mediated by the same protein region, it is likely that they are independent of each other, as some variants are only affected in one of the two functions: CadC_Y453I induces *cadBA* expression in response to acid stress but is not inhibited by cadaverine; in contrast, CadC_D471A and CadC_D471N activate *cadBA* expression irrespective of the external pH [Haneburger *et al.*, 2011] and are inhibited by cadaverine similarly to wild type.

Probing in vitro *affinity* – We intended to test the *in vitro* affinity of the periplasmic domains of distinct cadaverine insensitive variants (CadC_T229A_E447Q, CadC_T475A and CadC_Y453I [Haneburger *et al.*, 2011]) for cadaverine. However, Trx-$CadC_{pd}$_Y453I and Trx-$CadC_{pd}$_T475A were produced in inclusion bodies and could not be used for the assay. As both variants contain substitutions at the dimer interface, this could indicate that substitutions at the dimer interface lead to alterations in the protein conformation that render the protein more insoluble. For CadC_T475A, both a tryptophan fluorescence study [Tetsch *et al.*, 2008] as well as a scintillation proximity assay failed, since we measured a linear dependency between cadaverine concentrations and the corresponding signals. Thus, it can be assumed that affinity for cadaverine is greatly reduced in this variant.

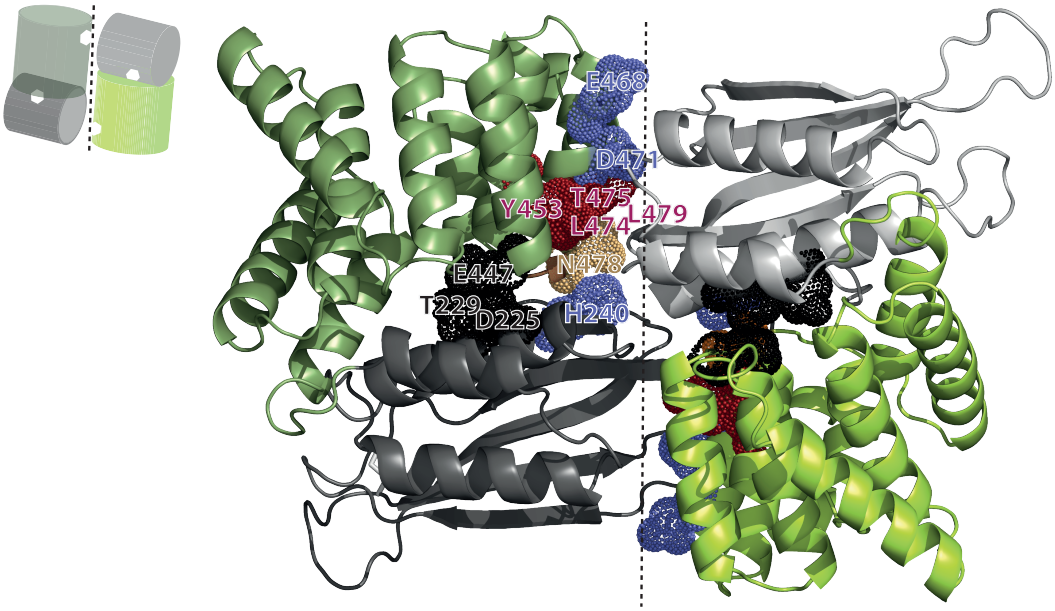

Figure 5.2: Localization of cadaverine-sensing residues in a $CadC_{pd}$ dimer. Left upper corner: simplified view of the $CadC_{pd}$ dimer. Colors and domain organization as in the large picture. White hexagons indicate cadaverine binding sites. Cartoon representation of $(CadC_{pd})_2$: The dimer interface is indicated by a dashed line. Residues involved in cadaverine binding at the interface, which were investigated in the experiments are presented as red spheres (Tyr453, Leu474, Thr475 and Leu479) or orange spheres (Asn478). Blue spheres represent the residues, that when substituted did not lead to an altered cadaverine-dependent inhibition (His240, Glu468, Asp471). Residues important for cadaverine sensing in the central cavity are depicted as black spheres (Asp225, Thr229, Glu447). The two monomers are depicted in light and dark colors, respectively (grey: N-terminal subdomain, green: C-terminal subdomain). The dimer is presented in top view towards the cytoplasmic membrane.

Dynamics of the Cad module of CadC variants impaired in cadaverine sensing - In order to infer key biochemical parameters of the Cad module *in vivo*, we previously established a quantitative mathematical model [Fritz *et al.*, 2009]. The model parameters were calibrated with a comprehensive set of experimental data, and the parameter-free predictions for the expression behavior under various initial conditions and for a LysP mutant deficient in lysine sensing were confirmed experimentally [Fritz *et al.*, 2009]. Likewise, here we reasoned that two CadC variants impaired in cadaverine sensing (CadC_T229A_E447Q and CadC_Y453I) should produce a significantly altered response after a shift to low pH. We used our model to interpret the experimental changes in the dynamics of the Cad module.

First we refined our previously published model [Fritz *et al.*, 2009], by incorporating the result of Eichinger *et al.* [2011] that CadC is active as a dimer ($CadC_2$; see also

Experimental Procedures). Then we re-analyzed all data previously collected for the wild type Cad module [Fritz *et al.*, 2009] and extracted the *in vivo* parameters for the interaction between $CadC_2$ and its input signals H^+, lysine and cadaverine (the full list of fit-parameters is given in Table S2). Specifically, for the interaction of wild type CadC with cadaverine we found an apparent *in vivo* affinity of $K_c = 170\ \mu M$ and an effective Hill exponent $n_c = 3.5$. As the Hill exponent is a measure for the cooperativity involved in the interaction between ligand and protein, the obtained value pointed to (at least) four cadaverine binding sites within one $CadC_2$ dimer [Abeliovich, 2005]. This would be consistent with the idea that each CadC monomer has one binding site in its central cavity, and one at the patch of acidic amino acids (at the dimer interface). Similarly, the deduced affinity parameter $K_c = 170\ \mu M$ of cadaverine for $CadC_2$ was in good agreement with the measured *in vitro* affinity (see above and Tetsch *et al.* [2008]). Next, we experimentally characterized the influence of the proposed binding site in the central cavity of CadC on the dynamics of the Cad module. To this end, we shifted the strain harboring CadC_T229A_E447Q from pH 7.6 to a lysine-rich medium (10 mM) with pH 5.8 and recorded the response of *cadBA* mRNA, specific CadA activity and external cadaverine level as a function of time (figure 5.3; green circles). CadA activity and cadaverine concentration was only determined for selected time points, as it is known from earlier studies that CadA activity reaches a plateau after about 60 min and cadaverine concentration increases almost linearly [Fritz *et al.*, 2009]. As a control, we applied the same protocol to a strain with identical genetic background harboring wild type CadC (figure 5.3A-C; black circles; see Experimental Procedures for all details). In this case (see also [Fritz *et al.*, 2009]) *cadBA* expression was quickly turned on (figure 5.3A; black circles) and as a consequence, both the CadA activity (figure 5.3B, black circles) as well as the cadaverine concentration in the medium increased (figure 5.3C; black circles). Once external cadaverine reached a certain threshold, transcription of *cadBA* was turned off again, thus leading to a transient expression of *cadBA*. Qualitatively, the CadC_T229A_E447Q harboring strain displayed a similar transient induction, but with a significantly widened expression peak of *cadBA* (figure 5.3A; green circles), a stronger increase in CadA activity (figure 5.3B; green circles) and a faster accumulation of cadaverine (figure 5.3C; green circles). These results clearly demonstrated that the substitutions in CadC_T229A_E447Q had impact on the dynamics of the Cad module *in vivo*, indicating that the targeted residues are involved in cadaverine sensing.

In order to extract more quantitative information from the dynamics of the Cad module, we leveraged our mathematical model to infer the key biochemical parameters relevant for cadaverine binding to CadC. Specifically, we used the binding constant of cadaverine

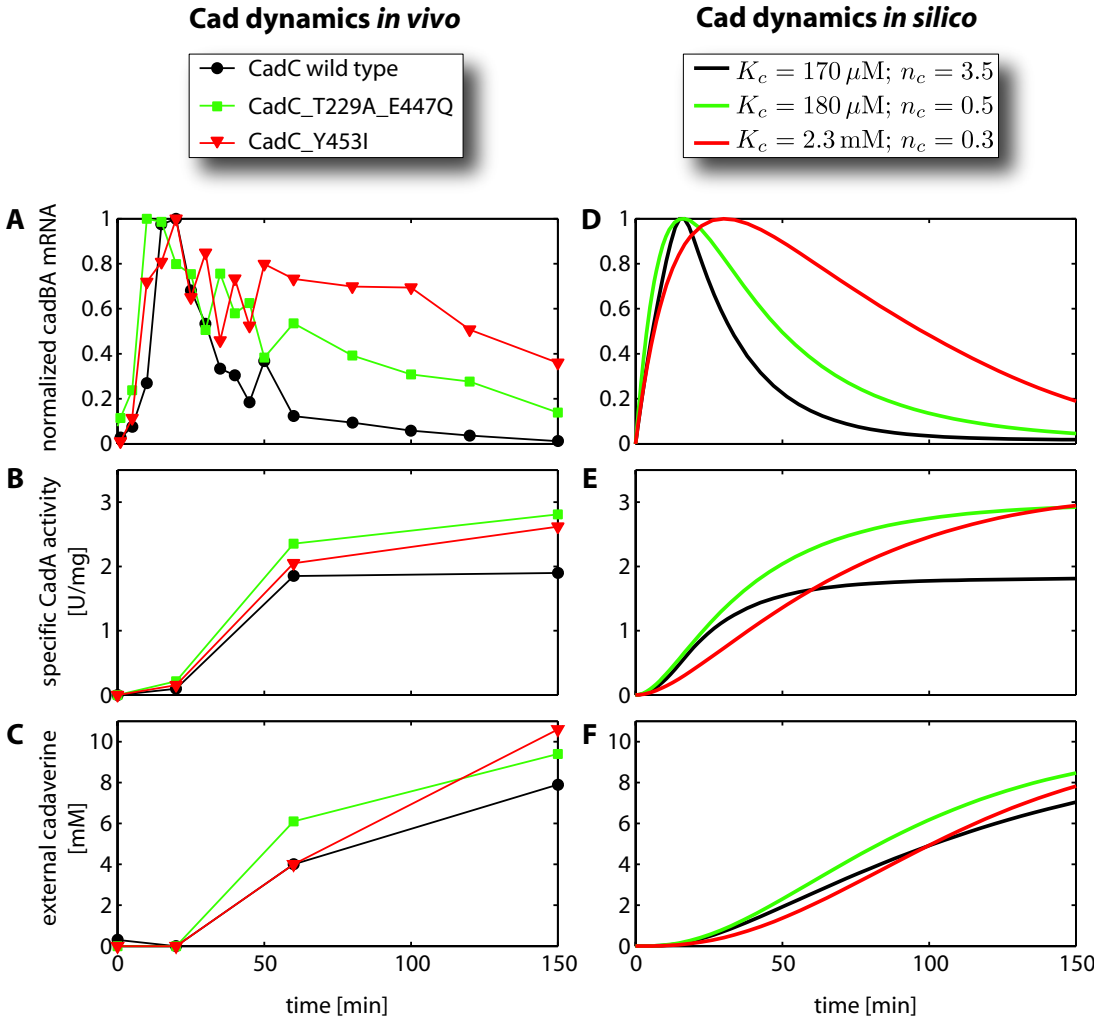

Figure 5.3: Induction kinetics of the Cad module *in vivo* (A-C) and *in silico* (D-E) for *E. coli* strains harboring different CadC variants (see legend). For a detailed list of all model parameters please refer to Table S2.

for CadC, K_c, the Hill exponent, n_c, and the maximal number of active $CadC_2$ dimers, C_0, as fit parameters for the CadC_T229A_E447Q harboring strain and set all other model parameters to those of the wild type Cad module, see Experimental Procedures for details. C_0 was included as fit parameter since the dimerization of CadC might be affected by amino acid substitutions. The fit results (figure 5.3D-F; green lines) displayed good quantitative agreement with the experimental data (figure 5.3A-C; green circles), capturing both the widened expression peak of *cadBA* mRNA as well as the stronger increase in CadA activity and cadaverine level when compared to the dynamics

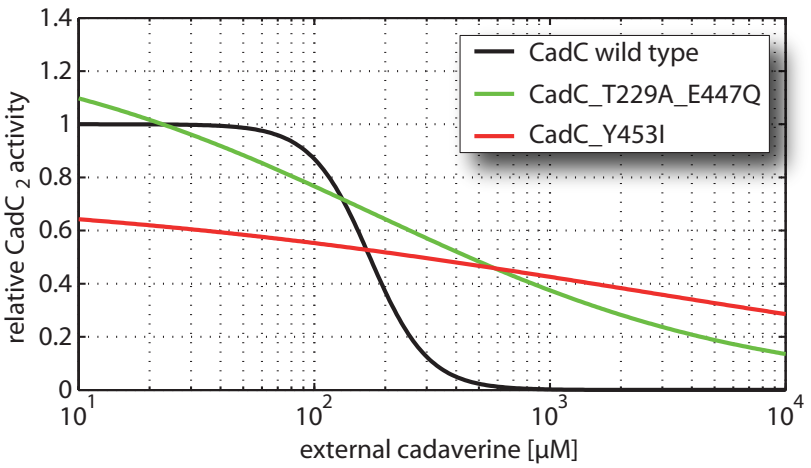

Figure 5.4: Simulated *in vivo* activities of CadC dimers wild type $CadC_2$ dimer (black line), $(CadC\_T229A\_E447Q)_2$ dimer (green line) and $(CadC\_Y453I)_2$ dimer (red line) as a function of external cadaverine. All curves were obtained by equation (5.2) with the corresponding parameters estimated for the *in vivo* data, see Table S2.

of the wild type strain. Surprisingly, the Hill exponent was the only parameter that changed significantly between the wild type (n_c = 3.5) and CadC_T229A_E447Q (n_c = 0.5), indicating that the residues in the central cavity of $CadC_{pd}$ mainly affect the cooperativity of the interaction between cadaverine and CadC. For better illustration, the impact of these substitutions on the $CadC_2$ activity is visualized in figure 5.4 (green line) as a function of cadaverine, where it is compared to wild type $CadC_2$ activity (black line). While the activity of the latter sharply decreased with increasing cadaverine levels at K_c, the activity of $(CadC\_T229A\_E447Q)_2$ displayed a significantly reduced cadaverine concentration dependancy. Hill exponents smaller than one may indicate a negative cooperative interaction among the remaining binding sites, but also other mechanisms can be envisioned (see Discussion). Another interesting observation is that in our fit also the maximal number of transcriptionally active CadC_T229A_E447Q dimers, C_0, was increased by 30%. Indeed, the output of transcription activity (β-galactosidase activity in Miller Units) for CadC_T229A_E447Q in our *in vivo* analyses of cadaverine-dependent inhibition was also 10% higher than for the wild type (data not shown).

Next, we analyzed the *in vivo* dynamics of a strain harboring CadC_Y453I after a shift to pH 5.8 and 10 mM lysine (figure 5.3A-C; red circles). In this strain the negative feedback via cadaverine should be nearly abolished (figure 5.1C). Nevertheless, we determined a transient *cadBA* expression with a significantly slower decrease in comparison

to strains harboring wild type CadC or CadC_T229A_E447Q (figure 5.3A; red circles). However, despite this significantly widened expression peak both the specific CadA activity (figure 5.3B; red circles) and the external cadaverine level (figure 5.3C; red circles) did not increase beyond the values of the strain harboring CadC_T229A_E447Q. This indicates that the absolute level of *cadBA* transcription in the CadC_Y453I harboring strain is somewhat lower compared to the other strains[1].

To test whether this behavior is consistent with our model, we again fitted the parameters K_c, n_c and C_0, keeping all other parameters fixed to the values of wild type CadC. Our fit data indicated that cadaverine sensing was indeed significantly impaired in the strain harboring CadC_Y453I (figure 5.4; red line), as reflected in a more than ten-fold reduction in cadaverine affinity (K_c = 2.3 mM) and a very low Hill coefficient (n_c = 0.3). Nonetheless, our model predicted a transient *cadBA* expression even for this largely cadaverine-insensitive CadC variant. Within our model the delayed down-regulation of *cadBA* expression arose from the CadBA-dependent depletion of lysine from the medium, which acts as an additional negative feedback directly on the input stimulus. Indeed, during the course of the experiment the cells converted the entire lysine (10 mM initially added) into cadaverine (approx. 10 mM after 150 min, figure 5.3C).

Beside the direct effects on cadaverine binding, also C_0, the maximal number of active $(CadC\_Y453I)_2$ dimers, was reduced by 25%, indicating that the substitution in CadC_Y453I also affected the spontaneous dimerization capability of CadC. The latter effect would be plausible, as the substituted residue is located at the dimerization interface of CadC (figure 5.2). Indeed, absolute β-galactosidase activities at inducing conditions (pH 5.8, 10 mM lysine; no cadaverine) for CadC_Y453I were reduced by 20 – 30% compared to the wild type protein (data not shown). Taken together, these results reveal that $(CadC\_Y453I)_2$ behaves like "cadaverine-blind" CadC, indicating that Tyr453 is of pivotal importance for cadaverine-sensing *in vivo*.

5.5 Discussion

Adaptation of gene expression in response to external stimuli is crucial for survival under stress conditions. Here, we studied feedback inhibition of the acid stress response Cad module by cadaverine. Intriguingly, the sensor and transcriptional activator CadC simultaneously integrates three external input signals, transduces the signals into the cytoplasm and activates transcription. Thus, all signaling capabilities are concentrated in a single polypeptide. In this study we set out to investigate the cadaverine-dependent

contribution to signal perception in CadC. Based on its crystal structure, the finding of an internal cavity and corresponding docking studies we postulated that cadaverine binding occurs in the cavity of the periplasmic domain [Eichinger *et al.*, 2011]. Site directed mutagenesis of amino acids putatively involved in cadaverine binding and subsequent investigation of the inhibition by cadaverine led to the identification of CadC variants with only moderately altered cadaverine-dependent regulation. Altogether 19 variants with amino acid substitutions in the central periplasmic cavity were tested for alterations in cadaverine-dependent inhibition of *cadBA* expression. Among those, cadaverine sensing by CadC_T229A_E447Q and CadC_T229A_Q421A_E447Q was affected significantly, as these variants still activated *cadBA* expression in the presence of 10.4 mM cadaverine (residual activity approx. 20%). Accordingly, higher cadaverine concentrations were needed to reduce *cadBA* expression to 50% (3.6 mM and 3.8 mM cadaverine, respectively). However, combinations of up to three amino acid substitutions within the central cavity did not significantly potentiate the effect. If the internal cavity would be the major cadaverine binding site, one would expect that amino acid substitutions in this area of CadC have a more drastic effect on cadaverine-mediated repression.

Therefore, data from Dell *et al.* [1994] motivated us to test different CadC variants with substitutions in the region of the patch of acidic amino acids previously identified to be crucial for detection of acidification of the external milieu [Haneburger *et al.*, 2011]. As polyamine binding is generally mediated through interaction with acidic and aromatic amino acids [Kashiwagi *et al.*, 1996, 2000; Soksawatmaekhin *et al.*, 2006; Sugiyama *et al.*, 1996; Vassylyev *et al.*, 1998], Tyr453 in this region appeared as a promising candidate for coordination of one amine group of cadaverine. Substitution of Tyr453 against alanine, isoleucine or phenylalanine revealed that the aromatic character at this position is essential for cadaverine-dependent repression. Hence, Tyr453 might represent one binding partner. Since Leu474, Thr475 and Leu479 were also crucial for cadaverine-dependent inhibition it is speculated that these residues may contact the carbon backbone of cadaverine. Thr475 could possibly assist by binding of the second amine group as it is located within the correct range to bind a bent cadaverine molecule (approx. 5 Å). It is also conceivable, however, that there is no pronounced interaction partner for the second amine group. An inhibitory effect on *cadBA* expression has been shown for higher concentrations of agmatine, putrescine and lysine derivatives [Neely *et al.*, 1994], indicating that variable chain lengths of the inhibitory ligands are tolerated. Altogether, amino acids around and within the patch of acidic amino acids are pivotal for cadaverine sensing, since CadC variants carrying substitutions in this area activated *cadBA* expression despite the presence of cadaverine.

To better understand the mechanisms behind cadaverine-dependent inhibition of CadC *in vivo*, we experimentally recorded the dynamics of the Cad module with two cadaverine-sensing impaired CadC variants and employed a mathematical model to extract key molecular parameters for the interaction between cadaverine and CadC. With the help of the model the affinity parameter for cadaverine was determined to be 170 μM. This is in good agreement with the *in vitro* determined affinity of $CadC_{pd}$ for cadaverine [Tetsch *et al.*, 2008]. It should be noted that in our *in vivo* reporter assays significantly higher cadaverine concentrations were needed to inactivate wild type CadC (table 5.1). This discrepancy might be explained by the exogenous addition of cadaverine in our reporter assays, while cadaverine is endogenously produced in our dynamical experiments. Cadaverine reportedly binds to porins and reduces the outer membrane permeability. This effect is even larger when cadaverine exerts its effect from the periplasmic side [Iyer and Delcour, 1997; Samartzidou and Delcour, 1999; Samartzidou *et al.*, 2003]. Therefore, internally produced cadaverine is captured in the periplasm as diffusion is limited through the cadaverine-inhibited porins. As externally added cadaverine is also hindered in diffusing into the periplasm it is conceivable that a much higher cadaverine concentration is needed to exert the inhibitory effect on CadC, when added from the outside.

Among all fit parameters, we found that the Hill exponents were most significantly affected by amino acid substitutions in CadC. Specifically, the Hill exponent of 3.5 determined for wild type CadC points to at least four cadaverine binding sites that might cooperatively inactivate the $CadC_2$ dimer. It is known, that the Hill exponent gives a lower limit on the number of interacting sites in positively cooperating systems [Abeliovich, 2005]. This finding is in agreement with a model according to which one binding site is located in the internal cavity of each monomer and a second binding site at the dimer interface of each monomer, resulting in four distinct cadaverine binding sites in a $CadC_{pd}$ dimer. Interestingly, both the substitutions in the central cavity (CadC_T229A_E447Q) as well as the substitution at the dimer interface (CadC_-Y453I) drastically reduced the Hill exponent to 0.5 and 0.3, respectively. Since the Hill exponents were reduced by more than 2 in each case (as naively expected for the removal of two binding sites per dimer), these results suggested that the four sites do not modulate $CadC_2$ activity in a simple, additive manner. Such non-additive effects may result, e.g., from allosteric effects, for which binding to a single site induces small conformational changes in $CadC_2$ wild type, such that binding to the second site is facilitated. Indeed, when binding to the central cavity was impaired (CadC_T229A_-E447Q), cadaverine only weakly affected the $CadC_2$ activity (green line in figure 5.4; Hill exponent of 0.5). This effect is even stronger if binding to the dimerization site is

impaired (CadC_Y453I), where cadaverine has basically no effect on the proportion of active $CadC_2$ (red line in figure 5.4; Hill exponent of 0.3). It is important to note that this hierarchy was also detected in our reporter assays, where CadC_Y453I was almost immune against cadaverine while CadC_T229A_E447Q still responded to cadaverine. Generally, the Hill exponents smaller than one observed in these mutants may point to a "negative cooperative" interaction between the residual binding sites, such that binding to one site impedes binding to the second site. However, heterogeneous mixtures of binding sites with different affinities can also lead to an apparent Hill exponent smaller unity [Abeliovich, 2005]. From our data it is not possible to discriminate between both effects. It is nevertheless intriguing that changes in single amino acids can change the cooperativity from positive to negative or to no cooperativity [Ikeda *et al.*, 2000; Kolodziej *et al.*, 1996], indicating that the Hill exponent can be readily tuned by evolutionary forces. Thus, we expect that the precise value of the Hill exponent for the interaction between cadaverine and CadC is optimized for its function *in vivo*, for which the sensitivity and range of response have to be carefully adjusted.

Based on our results we propose the following sequential model for the inactivation of CadC by cadaverine: Each CadC monomer has a binding site in its central periplasmic cavity and a binding site at the dimerization interface (figure 5.5, white hexagons). In its activated conformation (low pH and high lysine; figure 5.5A) CadC forms a homodimer, which grants cadaverine access to the central cavities, but buries the binding sites at the dimerization interface. Upon binding of cadaverine to the central cavities (figure 5.5B), small conformational changes in $CadC_2$ expose the binding sites at the dimerization interface. Cadaverine binding to the site at the dimerization interface possibly acts like a wedge that leads to disruption or inhibition of the active $CadC_2$ conformation, and consequently inactivates CadC (figure 5.5C). If the binding sites within the central cavities are missing (CadC_T229A_E447Q; figure S4 B), the initial conformational changes are not induced. Therefore, access to the interface binding sites is sterically hindered, and inhibition of this variant occurs at higher cadaverine concentrations compared to the wild type. In contrast, the variant in which binding at the interface is impaired, binding to the internal cavities occurs at similar cadaverine levels as in the wild type (figure S4 C). However, the final inactivation step occurs at even higher cadaverine concentrations.

In summary, the one-component system CadC integrates the inducing (H^+) and inhibiting signals (cadaverine) at the periplasmic dimerization interface. It was already suggested that upon stimulus perception active dimer formation is promoted, e.g., through charge neutralization of repelling surfaces [Haneburger *et al.*, 2011]. Here we found, that the inhibitor, cadaverine, binds to the same protein area and possibly impairs ac-

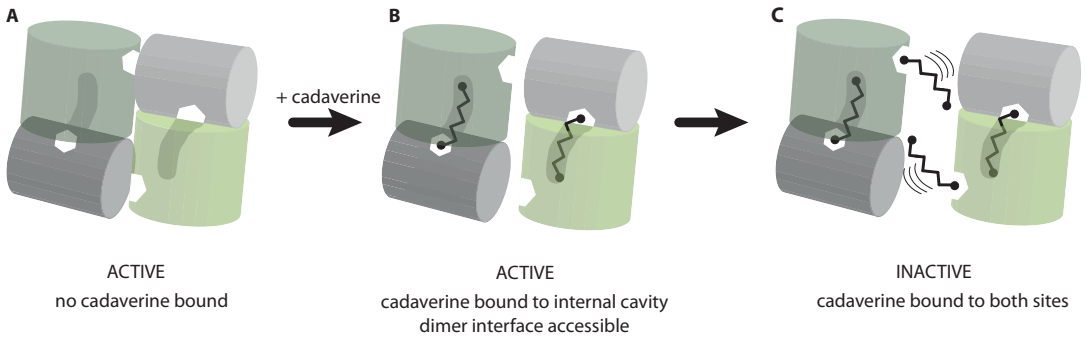

Figure 5.5: Model for cadaverine-dependent regulation of *cadBA* expression. CadC dimers of the periplasmic domain are presented schematically as the two monomers (dark and light colors) in top view. The monomers consist of two subdomains represented by a grey and green cylinder (figure 5.2). Cadaverine binding sites are indicated as hexagons. Cadaverine is depicted as molecule with five carbon atoms and two amine groups (spheres). A. In the presence of both stimuli (low pH and high lysine) CadC is active (as a dimer) and induces *cadBA* transcription. B. As soon as cadaverine reaches the threshold level it binds to the binding site within the internal cavity and evokes (conformational) changes that render the binding sites at the interface accessible. Still, CadC is active. C. After cadaverine binds to the interface binding site, CadC is inactive; probably by disturbance of the dimer conformation. Lysine regulation accomplished by LysP was omitted for simplicity.

tive dimer formation - most likely due to spatial/conformational constraints. As vast amounts of protein are produced upon activation of *cadBA* transcription (CadA increases to up to 2% of cellular proteins [Stim-Herndon *et al.*, 1996]), the sophisticated regulatory mechanism of the Cad module seems to protect the cell against wasting energy under unfavourable conditions. For instance, the cadaverine level might serve as a proxy for the abundance of CadA and CadB, such that the negative feedback via cadaverine asserts homeostatic control of their levels. In that, the transient induction of *cadBA* supports acid stress response by a "produce-to-demand" mechanism that minimizes the cost of *cadBA* induction and saves cellular resources for the investment in other cell-protecting processes.

5.6 Acknowledgements and Footnotes

The authors are grateful to Korinna Burdack for excellent technical assistance. We thank Christoph Küper for providing strain *E. coli* MG155Δ*cadC*.

‡Both authors contributed equally to this work.

This work was supported by the Excellence Clusters "Nanosystems Initiative Munich", "Center for integrated Protein Science Munich (Exc114/1)" and through grants GE 1098/4-1 and JU270/5-1 by Deutsche Forschungsgemeinschaft. I. H. was supported by a research scholarship of the Elite Network of Bavaria and granted financial support by CiPSM women.

[1] Note that figure 5.3A only displays the mRNA levels relative to its maximal value, since those are much better determined than their absolute values.

5.7 References for Manuscript

Abeliovich, H., 2005. An empirical extremum principle for the hill coefficient in ligand-protein interactions showing negative cooperativity. *Biophysical journal*, 89(1):76–79.

Bearson, S., Bearson, B., and Foster, J. W., 1997. Acid stress responses in enterobacteria. *FEMS microbiology letters*, 147(2):173–180.

Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., *et al.*, 1997. The complete genome sequence of *Escherichia coli* K-12. *Science*, 277(5331):1453–1462.

Chung, H. J., Bang, W., and Drake, M. A., 2006. Stress response of *Escherichia coli*. *Comprehensive reviews in food science and food safety*, 5(3):52–64.

Dell, C. L., Neely, M. N., and Olson, E. R., 1994. Altered pH lysine signalling mutants of*cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. *Molecular microbiology*, 14(1):7–16.

Eichinger, A., Haneburger, I., Koller, C., Jung, K., and Skerra, A., 2011. Crystal structure of the sensory domain of *Escherichia coli* CadC, a member of the ToxR-like protein family. *Protein Science*, 20(4):656–669.

Epstein, W. and Kim, B. S., 1971. Potassium transport loci in *Escherichia coli* K-12. *Journal of Bacteriology*, 108(2):639–644.

Fritz, G., Koller, C., Burdack, K., Tetsch, L., Haneburger, I., Jung, K., and Gerland, U., 2009. Induction kinetics of a conditional pH stress response system in *Escherichia coli*. *Journal of molecular biology*, 393(2):272–286.

Gale, E. F. and Epps, H. M. R., 1944. Studies on bacterial amino-acid decarboxylases: 1. L (+)-lysine decarboxylase. *Biochemical Journal*, 38(3):232–242.

Haneburger, I., Eichinger, A., Skerra, A., and Jung, K., 2011. New insights into the signaling mechanism of the pH-responsive, membrane-integrated transcriptional activator CadC of *Escherichia coli*. *Journal of Biological Chemistry*, 286(12):10681–10689.

Ikeda, Y., Taniguchi, N., and Noguchi, T., 2000. Dominant Negative Role of the Glutamic Acid Residue Conserved in the Pyruvate Kinase M1 Isozyme in the Heterotropic Allosteric Effect Involving Fructose-1, 6-bisphosphate. *Journal of Biological Chemistry*, 275(13):9150–9156.

Iyer, R. and Delcour, A. H., 1997. Complex inhibition of OmpF and OmpC bacterial porins by polyamines. *Journal of Biological Chemistry*, 272(30):18595–18601.

Kashiwagi, K., Pistocchi, R., Shibuya, S., Sugiyama, S., Morikawa, K., and Igarashi, K., 1996. Spermidine-preferential uptake system in *Escherichia coli*. *Journal of Biological Chemistry*, 271(21):12205–12208.

Kashiwagi, K., Kuraishi, A., Tomitori, H., Igarashi, A., Nishimura, K., Shirahata, A., and Igarashi, K., 2000. Identification of the putrescine recognition site on polyamine transport protein PotE. *Journal of Biological Chemistry*, 275(46):36007–36012.

Kolodziej, A. F., Tan, T., and Koshland Jr, D. E., 1996. Producing positive, negative, and no cooperativity by mutations at a single residue located at the subunit interface in the aspartate receptor of *Salmonella typhimurium*. *Biochemistry*, 35(47):14782–14792.

Küper, C. and Jung, K., 2005. CadC-mediated activation of the *cadBA* promoter in *Escherichia coli*. *J Mol Microbiol Biotechnol*, 10(1):26–39.

Meng, S. Y. and Bennett, G. N., 1992. Nucleotide sequence of the *Escherichia coli* cad operon: a system for neutralization of low extracellular pH. *Journal of bacteriology*, 174(8):2659–2669.

Miller, J. H. *et al.* *Experiments in molecular genetics*, volume 433. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1992.

Neely, M. N., Dell, C. L., and Olson, E. R., 1994. Roles of LysP and CadC in mediating the lysine requirement for acid induction of the *Escherichia coli* cad operon. *Journal of bacteriology*, 176(11):3278–3285.

Quick, M. and Javitch, J. A., 2007. Monitoring the function of membrane transport proteins in detergent-solubilized form. *Proceedings of the National Academy of Sciences*, 104(9):3603–3608.

Sabo, D. L., Boeker, E. A., Byers, B., Waron, H., and Fischer, E. H., 1974. Purification and physical properties of inducible *Escherichia coli* lysine decarboxylase. *Biochemistry*, 13(4):662–670.

Samartzidou, H. and Delcour, A. H., 1999. Excretion of endogenous cadaverine leads to a decrease in porin-mediated outer membrane permeability. *Journal of bacteriology*, 181(3):791–798.

Samartzidou, H., Mehrazin, M., Xu, Z., Benedik, M. J., and Delcour, A. H., 2003. Cadaverine inhibition of porin plays a role in cell survival at acidic pH. *Journal of bacteriology*, 185(1):13–19.

Soksawatmaekhin, W., Kuraishi, A., Sakata, K., Kashiwagi, K., and Igarashi, K., 2004. Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. *Molecular microbiology*, 51(5):1401–1412.

Soksawatmaekhin, W., Uemura, T., Fukiwake, N., Kashiwagi, K., and Igarashi, K., 2006. Identification of the cadaverine recognition site on the cadaverine-lysine antiporter CadB. *Journal of Biological Chemistry*, 281(39):29213–29220.

Stim-Herndon, K. P., Flores, T. M., and Bennett, G. N., 1996. Molecular characterization of *adiY*, a regulatory gene which affects expression of the biodegradative acid-induced arginine decarboxylase gene (*adiA*) of *Escherichia coli*. *Microbiology*, 142(5):1311–1320.

Sugiyama, S., Vassylyev, D. G., Matsushima, M., Kashiwagi, K., Igarashi, K., and Morikawa, K., 1996. Crystal structure of PotD, the primary receptor of the polyamine transport system in *Escherichia coli*. *Journal of Biological Chemistry*, 271(16):9519–9525.

Taylor, R. G., Walker, D. C., and Mclnnes, R. R., 1993. *E. coli* host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic acids research*, 21(7):1677–1678.

Tetsch, L., Koller, C., Haneburger, I., and Jung, K., 2008. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Molecular microbiology*, 67(3):570–583.

Vassylyev, D. G., Tomitori, H., Kashiwagi, K., Morikawa, K., and Igarashi, K., 1998. Crystal Structure and Mutational Analysis of the *Escherichia coli* Putrescine Receptor. *Journal of Biological Chemistry*, 273(28):17604–17609.

Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L., and Olson, E. R., 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli* cad operon by external pH. *Journal of bacteriology*, 174(2):530–540.

Yanisch-Perron, C., Vieira, J., and Messing, J., 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. *Gene*, 33(1):103–119.

6 Identification of ArgP and Lrp as transcriptional regulators of *lysP*, the gene encoding the specific lysine permease of *Escherichia coli*

Reproduced from The Journal of Bacteriology, (2011), J. Ruiz, I. Haneburger and K. Jung, Identification of ArgP and Lrp as Transcriptional Regulators of *lysP*, the Gene Encoding the Specific Lysine Permease of *Escherichia coli*. Vol. 193, No. 10, p. 2536-2548, doi:10.1128/JB.00815-10, with permission from the American Society for Microbiology, Copyright © 2011, American Society for Microbiology.

JOURNAL OF BACTERIOLOGY, May 2011, p. 2536–2548 Vol. 193, No. 10
0021-9193/11/$12.00 doi:10.1128/JB.00815-10

Identification of ArgP and Lrp as Transcriptional Regulators of *lysP*, the Gene Encoding the Specific Lysine Permease of *Escherichia coli*▽†

Jimena Ruiz,‡ Ina Haneburger, and Kirsten Jung*

Ludwig-Maximilians-Universität München, Munich Center for integrated Protein Science (CiPSM) at the Department of Biology I, Microbiology, Grosshaderner Strasse 2-4, 82152 Martinsried, Germany

Received 12 July 2010/Accepted 14 March 2011

Expression of *lysP*, which encodes the lysine-specific transporter LysP in *Escherichia coli*, is regulated by the concentration of exogenous available lysine. In this study, the LysR-type transcriptional regulator ArgP was identified as the activator of *lysP* expression. At lysine concentrations higher than 25 μM, *lysP* expression was shut off and phenocopied an *argP* deletion mutant. Purified ArgP-His_6 bound to the *lysP* promoter/control region at a sequence containing a conserved T-N_{11}-A motif. Its affinity increased in the presence of lysine but not in the presence of the other known coeffector, arginine. *In vivo* data suggest that lysine-loaded ArgP and arginine-loaded ArgP compete at the *lysP* promoter. We propose that lysine-loaded ArgP prevents *lysP* transcription at the promoter clearance step, as described for the lysine-dependent regulation of *argO* (R. S. Laishram and J. Gowrishankar, Genes Dev. 21:1258-1272, 2007). The global regulator Lrp also bound to the *lysP* promoter/control region. An *lrp* mutant exhibited reduced *lysP* expression in the absence of external lysine. These results indicate that ArgP is a major regulator of *lysP* expression but that Lrp modulates *lysP* transcription under lysine-limiting conditions.

Amino acid transporters play several important roles in bacteria. Besides their function in nutrient supply, these systems are also involved in osmoregulation, pH homeostasis, signal transduction, and detoxification. More than one uptake system normally exists for the transport of a single amino acid, allowing bacteria to adapt to different environmental conditions. *Escherichia coli* has three different transport systems for the uptake of the amino acid L-lysine that differ in transport mechanism, substrate specificity, apparent Michaelis constant (K_m), and regulation of their synthesis. The lysine-arginine-ornithine (LAO) system is encoded by *argT-hisJQMP*. $HisQMP_2$ forms the ABC-transporter; HisJ and ArgT are periplasmic binding proteins that are specific for histidine and lysine, arginine, or ornithine, respectively (40). The cadaverine-lysine antiporter CadB imports lysine and excretes cadaverine (48) but is produced only under conditions of low pH. Last, but not least, LysP is a specific transporter for L-lysine that belongs to the amino acid, polyamine, and organocation (APC) transporter family (8, 49). Considering the important role of lysine and LysP in amino acid metabolism and pH homeostasis in *E. coli*, the aim of this work was to investigate the factors and mechanisms involved in the transcriptional regulation of *lysP*.

The *lysP* gene was originally named *cadR* because its mutants exhibit a pleiotropic phenotype including derepressed levels of lysine decarboxylase CadA, in addition to a reduction in lysine transport (37). The Cad system, which plays a role in pH homeostasis in *Enterobacteria*, comprises the membrane-integrated transcriptional activator CadC and the *cadBA* operon, encoding the lysine decarboxylase CadA and the lysine-cadaverine antiporter CadB (26, 27, 58). This system is induced under conditions of low external pH and the simultaneous presence of exogenous lysine. As a result of the lysine decarboxylation reaction, which consumes one cytoplasmic H^+, cadaverine is produced and subsequently excreted, leading to an increase of the external pH. For a long time it, was unclear how the function of LysP was linked to the regulation of the Cad system. Tetsch et al. (2008) demonstrated that LysP is able to modulate the activity of the membrane-integrated protein CadC (55). According to the proposed model, LysP and CadC interact via their transmembrane domains in the absence of lysine. This interaction blocks CadC-dependent expression of the *cadBA* operon. In the presence of lysine, LysP releases CadC and CadC becomes susceptible to activation by low pH. These findings suggest that LysP has an additional regulatory function, which is typical for so-called trigger transporters (54). Thus, in addition to transport activity, LysP senses lysine availability and transduces the signal to CadC.

Neely and Olson (1996) demonstrated that a high external lysine concentration downregulates *lysP* expression (32). Expression of most of the genes belonging to the lysine biosynthesis pathway in *E. coli* is repressed by lysine, but there are multiple modes of regulation known (2, 34, 50). In *E. coli* and *Bacillus subtilis*, expression of *lysC*, which encodes one of the isoenzymes that catalyze the first step in the lysine biosynthesis pathway, is controlled by direct binding of lysine to a conserved leader sequence in its mRNA (34, 52). Orthologs of *lysP* in Gram-positive bacteria are controlled by lysine-dependent

* Corresponding author. Mailing address: Ludwig-Maximilians-Universität München, Department Biologie I, Bereich Mikrobiologie, 82152 Martinsried, Germany. Phone: 49-89-2180-74500. Fax: 49-89-2180-74520. E-mail: jung@lmu.de.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Intendente Guiraldes 2160 (1428), Ciudad Autónoma de Buenos Aires, Argentina.

▽ Published ahead of print on 25 March 2011.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant genotype or description	Reference or source
E. coli strains		
MG1655	F^- λ^- *ilvG rfb50 rph-1*	1
MG1655-Δ*lacZ*	MG1655 Δ*lacZ*::Tet^r	K. Jahreis (personal gift)
MC4100	F^- *araD139* Δ(*argF-lacZ*)*U169 rpsL150 relA flb-530* Str^r	7
MG16R	MG1655 Δ*lacZ*::Tet^r *rpsL150* Str^r	This work
MG16R4	MG1655 Δ*lacZ*::Tet^r *rpsL150* Δ*lysP*::*rpsL-neo* Km^r Str^s	This work
MG-LR	MG1655 Δ*lacZ*::Tet^r *rpsL150* Δ*lysP* P_{lysP}::*lacZ* Str^r	This work
MG-LR9	MG-LR Δ*lysR*::Km^r Str^r	This work
MG-LR10	MG-LR Δ*yeiE*::Km^r Str^r	This work
JCP95	pop3125 (*dapBp* [−118/+35]-*lacZ*) Δ*argP*::Cam^r	2
MG-LR17	MG-LR Δ*argP*::Cam^r Str^r	This work
MG-LR15	MG-LR Δ*lrp*::Km^r Str^r	This work
MG-LR4	MG-LR $P_{lysPΔTN11A}$::*lacZ* (deletion from position −83 to position −52 in the *lysP* promoter)	This work
MG16R12	MG1655 Δ*lacZ*::Tet^r *rpsL150* Δ*cadBA*::*rpsL-neo* Km^r Str^s	S. Ude (unpublished results)
MG-CR	MG1655 Δ*lacZ*::Tet^r *rpsL150* Δ*cadBA* P_{cadBA}::*lacZ* Str^r	S. Ude (unpublished results)
MG-CR15	MG-CR Δ*lrp*::Km^r Str^r	This work
BL21(DE3) pLysS	F^- *ompT* r^-_B m^-_B	51
Plasmids		
pBAD33	Arabinose-inducible P_{BAD} promoter; pACYC184 ori; Amp^r	18
pBAD*lysP*	*lysP* in pBAD33; Amp^r	55
pBAD24	Arabinose-inducible P_{BAD} promoter, pBR322 ori; Amp^r	18
pBAD*argP*	*argP* cloned in the EcoRI and HindIII sites of pBAD24; Amp^r	This work
pBAD*lrp*	*lrp* cloned in the EcoRI and HindIII sites of pBAD24; Amp^r	This work
pET21a	T7 promoter based expression vector with His tag; Amp^r	Novagen
pET21*argP*	*argP* cloned in the NdeI and XhoI sites of pET21a; Amp^r	This work
pET16b	T7 promoter based expression vector with His tag; Amp^r	Novagen
pET16*lrp*	*lrp* cloned in the NdeI and BamHI sites of pET16b; Amp^r	This work
pRS415	Operon fusion vector	47
pRS*lysP*	pRS415::*lysP* promoter (positions −218 to +28)	This work
pRS*lysP*0	pRS415::*lysP* promoter with a deletion from position −88 to position −53	This work
pRS*lysP*1	pRS415::*lysP* promoter with replacement of A/C at position −53	This work
pRS*lysP*2	pRS415::*lysP* promoter with replacement of T/G at position −65	This work
pRS*lysP*3	pRS415::*lysP* promoter with replacements of A/C and T/G	This work

riboswitches, named LYS or L-box elements (43). However, lysine riboswitches are not found in the 5′ untranslated region of *lysP* mRNA of Gram-negative bacteria (43). Here, we show that *lysP* transcription is subject to two different types of control. The LysR-type transcriptional regulator (LTTR) ArgP was identified as a major regulator of *lysP* transcription. In addition, *lysP* expression is fine-tuned by the global regulator Lrp (leucine-responsive protein). Lrp was also found to be involved in the regulation of *cadBA* expression.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The strains used in this study are listed in Table 1. *E. coli* JM109 (60) was used as a carrier for all plasmids. For the construction of the reporter strains, MG-LR and MG-CR, a method based on *rpsL* counterselection in combination with the Red/ET recombination system was employed (19) according to the protocol recommended by the technical manual of the Quick and Easy *E. coli* deletion kit (Gene Bridges). Briefly, the coding sequence of the target gene (*lysP* or *cadBA*) was replaced by an *rpsL-neo* cassette (Gene Bridges) by Red/ET recombination in strain MG16R (Table 1) to give strains MG16R4 and MG16R12. Afterwards, the *rpsL-neo* cassette was replaced by promoterless *lacZ* using the Red/ET recombination technique, according to the following procedure. Strain MG16R4 or MG16R12 carrying plasmid pRedET (Gene Bridges) was transformed with a linear DNA fragment comprising the promoterless *lacZ* gene flanked with homology sequences for the target genes, and subsequently, clones of interest were selected on LB agar plates (P_{lysP}::*lacZ* fusion) or LB (pH 5.8) agar plates (P_{cadBA}::*lacZ* fusion) containing 50 μg ml^{-1} 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and 50 μg ml^{-1} streptomycin. Blue colonies were tested for kanamycin sensitivity. The Km^s clones were verified by colony PCR followed by DNA sequencing.

E. coli strains MG-LR4, MG-LR9, MG-LR10, and MG-LR15 were constructed using the Quick and Easy *E. coli* deletion kit (Gene Bridges) according to the instructions of the manufacturer. Introduction of the *argP*::Cam^r allele into MG-LR and *lrp*::Km^r into MG-CR was performed by P1*vir*-mediated phage transduction (56), using JCP95 and MG-LR15, respectively, as donor strains. Strain BL21(DE3)pLysS was cultivated in tryptone-phosphate medium (29) and used as a host for pET plasmids (Novagen) for protein overproduction. For determination of *lysP* expression levels, cells were grown in minimal medium (15) supplemented with glucose or fructose at a final concentration of 0.4% (wt/vol). For determination of *cadBA* expression levels, cells were cultivated in glucose minimal medium; the phosphate buffer of the medium was adjusted to either pH 5.8 or pH 7.6. Antibiotics were used at the following concentrations: ampicillin, 100 μg ml^{-1}; kanamycin, 50 μg ml^{-1}; chloramphenicol, 34 μg ml^{-1}; tetracycline, 12.5 μg ml^{-1}; and streptomycin, 50 μg ml^{-1}.

Construction and analysis of *lysP* promoter variants. A sequence of 246 bp encompassing the whole *lysP* promoter/control region (positions −218 to +28) (see Fig. 2) was amplified by PCR with primers lysPupEcoRI (5′G GAA TTC CGC TTT CTG GAC TAT TGC GAT C 3′) and lysPprBamHI (5′ CGG GAT CCA CAA AAA TGC TAT CCA TCT TAA 3′) and cloned upstream of the promoterless *lacZ* gene in vector pRS415 (47). Introduction of deletions or point mutations of the conserved T-N_{11}-A motif was achieved by purchasing the corresponding synthetic DNA fragments (Mr. Gene, Regensburg, Germany), which were subcloned into the EcoRI and BamHI sites of pRS415. *E. coli* strain MG1655-Δ*lacZ* was transformed with the resulting plasmids to test *lysP* expression as described below.

***In vivo lysP* and *cadBA* expression studies.** Expression of *lysP* and *cadBA in vivo* was determined by means of β-galactosidase assays. For the analysis of P_{lysP}::*lacZ* expression, cells of an overnight culture grown in minimal medium were inoculated into fresh medium (supplemented with amino acids where indicated), resulting in an optical density at 600 nm (OD_{600}) of 0.05. Cultures were grown aerobically in Erlenmeyer flasks at 37°C. To determine the expression of

P_{cadBA}::*lacZ*, cells were precultured in minimal medium at pH 7.6 and then inoculated into fresh minimal medium at pH 5.8 or pH 7.6, supplemented with lysine and/or arginine where indicated. Cultures were incubated under microaerobic conditions at 37°C to mid-logarithmic growth phase. β-Galactosidase activity measurements were performed as previously described (55) for at least three independent experiments. Values are given in Miller units (MU), which were calculated according to Miller (28).

Molecular biology techniques. Plasmid DNA and genomic DNA were isolated by using a HiYield plasmid minikit (Sued-Laborbedarf Gauting) and a DNeasy blood and tissue kit (Qiagen), respectively. DNA fragments were purified from agarose gels using a Hi-Yield PCR cleanup and gel extraction kit (Sued-Laborbedarf Gauting). Phusion high-fidelity DNA polymerase or Phire hot-start DNA polymerase (Finnzymes) was used according to the supplier's instructions. Restriction enzymes were purchased from New England Biolabs and used according to the manufacturer's directions.

Purification of Lrp and ArgP. *E. coli* BL21(DE3)pLysS harboring plasmid pET16*lrp* or pET21a*rgP* (Table 1) was grown to exponential phase at 30°C, and expression of genes encoding N-terminally His-tagged Lrp (His_6-Lrp) or C-terminally His-tagged ArgP (ArgP-His_6) was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 3 h of induction, cells were harvested and washed with 100 mM Na-K-phosphate buffer (pH 7.5) at 4°C. The cell pellet was frozen in liquid nitrogen and stored at −80°C until use. Cells were lysed by passage through a high-pressure cell disrupter (Constant Systems). After centrifugation of the disrupted cells, the supernatant containing the His_6 protein was incubated with Ni^{2+}-nitrilotriacetic acid (NTA) resin (Qiagen) preequilibrated with lysis buffer (20 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl, 10% [vol/vol] glycerol, pH 8.0). After 1 h of incubation, the protein-resin complex was washed twice with washing buffer (50 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl, 10% [vol/vol] glycerol, pH 8.0). Finally, the His-tagged protein was eluted in several fractions with buffer containing 250 mM imidazole, 50 mM NaH_2PO_4, 300 mM NaCl, and 10% (vol/vol) glycerol, pH 8.0. His_6-Lrp was dialyzed against Lrp binding buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM dithiothreitol [DTT], 0.1 mM EDTA, 10% [vol/vol] glycerol) and ArgP-His_6 against ArgP-binding buffer (24) at 4°C. The purified proteins were stored in the corresponding binding buffers in the presence of 1 mM phenylmethylsulfonyl fluoride (PMSF) at 4°C and used within 1 week or stored at −20°C. Protein concentration was determined according to Bradford (3).

EMSAs. Probes for the electrophoretic mobility shift assay (EMSA) were amplified by PCR using primers labeled at their 5′ ends with the 6-isomer of carboxyfluorescein (6-FAM) and genomic DNA from *E. coli* MG1655 as a template unless indicated otherwise. To analyze the binding of ArgP-His_6 to the *lysP* promoter/control region, three different fragments were used: (i) a fragment of 276 bp comprising the sequence from position −218 to position +58 (P_{lysP} fragment) (see Fig. 2), which was obtained by PCR using primers lysPup (5′-CGCTTTCTGGACTATTGCGATC-3′) and lysPprlow (5′-CGCTTCTGTGGTTTTAGTTTCG-3′); (ii) a fragment of 142 bp comprising the sequence from position −84 to position +58 (T-N_{11}-A fragment), which was amplified with primers TN11A (5′-TATAATCCCTGGGCGATCATG-3′) and lysPprlow; and (iii) a fragment of 93 bp (−35-10 fragment, comprising positions −35 to +58) obtained by amplification with primers −35 (5′-CGGAAGGATTGCCAATCGT-3′) and lysPprlow. To evaluate the binding of His_6-Lrp to the *lysP* promoter/control region, only the P_{lysP} fragment was used. To determine the binding of His_6-Lrp to the *cadBA* promoter/control region, a fragment comprising positions −150 to +72 upstream of the *cadB* gene was cloned into the EcoRI and BamHI sites of the pUC19 plasmid (60) and amplified with primers 6-FAM uni-24 (5′-ACGACGTTGTAAAACGACGGCCAG-3′) and rev-24 (5′-TTCACACAGGAAACAGCTATGACC-3′). As a control for nonspecific binding, a DNA fragment of 258 bp within the *lysP* coding sequence (control fragment) obtained by amplification with lysPcup (5′-ACATCAGCGTTAGTCCGT-3′) and lysPclow (5′-ATGGAGGTCAGGAAGCACA-3′) was used. After PCR amplification, the obtained DNA fragments were purified by 7% (wt/vol) polyacrylamide gel electrophoresis according to the protocol of the GenElute gel extraction kit (Sigma). ArgP-DNA binding assays were performed by incubating 30 fmol of a DNA fragment with increasing concentrations of ArgP-His_6 in 25 μl ArgP binding buffer supplemented with 12 μg ml^{-1} sonicated salmon sperm DNA as a nonspecific competitor, 50 μg ml^{-1} bovine serum albumin, and, where indicated, 0.1 mM lysine or arginine. Binding of His_6-Lrp to DNA was performed with 30 fmol of a DNA fragment and increasing concentrations of His_6-Lrp in 25 μl Lrp binding buffer supplemented with 20 μg ml^{-1} salmon sperm DNA and 0.1 mM lysine where indicated. After incubation at 25°C for 30 min, complexes were resolved by electrophoresis in a 6.5% (wt/vol) polyacrylamide gel under a constant voltage of 10 V/cm at room temperature for 1.5 h. Gels were scanned with a Typhoon Trio imager (Amersham Biosciences) under an excitation wavelength of 488 nm. Quantification of free DNA and protein-bound DNA was performed by densitometry using the ImageQuant 5.0 analysis software program (Molecular Dynamics). The quantified data were plotted versus the protein concentration to obtain the binding profile. The apparent dissociation constants (K_D) of ArgP-DNA binding assays were determined to be the protein concentration at which the fraction of bound DNA equals 0.5. The binding profiles obtained from the Lrp binding assays were fitted to the Hill equation to determine the K_D value and the Hill coefficient (n).

DNase I footprinting assays. DNase I footprinting analysis was performed according to the method described by Sandaltzopoulos and Becker (44). One hundred fifty nanograms of a P_{lysP} DNA fragment (bases −218 to +58) labeled at its 5′end with 6-FAM as described above was incubated with different concentrations of ArgP-His_6 in ArgP binding buffer supplemented with 12 μg ml^{-1} sonicated salmon sperm DNA as a nonspecific competitor, 50 μg ml^{-1} bovine serum albumin, and, where indicated, 0.1 mM lysine in a final volume of 50 μl. After incubation at 25°C for 30 min, 50 μl of a 5 mM $CaCl_2$ solution was added, and incubation was prolonged for 1 min. Subsequently, 0.25 U of DNase I was added, and after 5 min, the reaction was stopped by adding 0.5 ml DF buffer from the Hi-Yield PCR cleanup and gel extraction kit (Sued-Laborbedarf Gauting). After purification, the DNA fragments were analyzed with an ABI PRISM 377 DNA sequencer, and the data were evaluated with the Peak Scanner software program (Applied Biosystems).

DNA affinity purification assay for the identification of DNA-binding proteins. To isolate putative transcriptional regulators of *lysP*, a DNA affinity purification protocol was applied (17). For this purpose, a biotinylated P_{lysP} fragment was generated by PCR using primers lysPup, labeled with biotin at the 5′end, and lysPprlow. As a control, a biotinylated DNA fragment located within the *lysP* coding sequence (obtained by amplification with biotin-labeled lysPcup and lysPclow) was used. About 600 pmol of the biotin-labeled DNA fragments was immobilized with streptavidin-coated magnetic particles (Chemagen Biopolymer-Technologies) according to the manufacturer's instructions. For the preparation of the cytoplasmic protein extract, *E. coli* MG1655 was cultivated in 800 ml of glucose minimal medium to an OD_{600} of 0.8. Cells were harvested at 4°C, washed with cold protein binding buffer B (PBB) (41), resuspended in 8 ml of the same buffer, and broken with a French press. After centrifugation to remove the cellular debris, the supernatant extract was incubated with DNA-coated magnetic beads (previously equilibrated with PBB) at room temperature for 30 min. Washing to remove unspecific bound proteins and elution of tightly bound proteins was performed as described by Rey et al. (41). Eluted fractions were collected, subjected to sodium dodecyl sulfate-polyacrylamide electrophoresis (SDS-PAGE) (23), and stained with Coomassie blue. Proteins were identified by peptide fingerprint analysis (20) using a matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry (MS) system (Voyager DE STR; Applied Biosystems). Samples were prepared and identified as described previously (59).

RESULTS

***lysP* transcription is negatively regulated by lysine.** Previous work showed that *lysP* mRNA declines within 4 min after addition of 10 mM lysine (32). To analyze the concentration-dependent effect of external lysine on *lysP* transcription in more detail, *E. coli* strain MG-LR, an MG1655 derivative that carries a P_{lysP}::*lacZ* fusion, was constructed. In this strain, the coding sequence of the *lysP* gene was replaced by the reporter gene *lacZ*, so that the *lacZ* gene is located exactly in the same genetic context as the *lysP* gene. Since the MG-LR strain is Δ*lysP*, an expression plasmid carrying the *lysP* gene (pBAD*lysP*) (Table 1) was introduced into this strain to determine whether the expression from the *lysP* promoter (P_{lysP}) was affected by the presence of LysP. β-Galactosidase activities indicated that *lysP* expression was completely repressed by external lysine concentrations of 50 μM and higher (Fig. 1). The availability of LysP did not alter the lysine-mediated repression (Fig. 1).

Identification of ArgP as a regulator of *lysP* expression. Analysis of the *lysP* promoter revealed a conserved T-N_{11}-A motif, a typical binding site of LysR-type transcriptional regu-

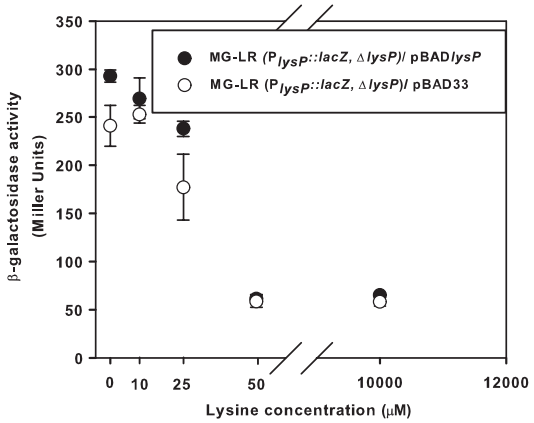

FIG. 1. Effect of the external lysine concentration and LysP on *lysP* expression. Cultures of *E. coli* strain MG-LR (P_{lysP}::*lacZ* Δ*lysP*) with plasmid pBAD33 or pBAD*lysP* were grown in fructose minimal medium supplemented with different concentrations of external lysine and 0.006% (wt/vol) arabinose to induce the expression of the *lysP* gene cloned in the pBAD33 plasmid. When cultures reached an OD_{600} of ~0.5, samples were collected and β-galactosidase activities were determined. The experiment was performed in triplicate, and error bars indicate standard deviations from the means.

lators (LTTRs) that is in close proximity to a sequence with strong similarity to the consensus sequence for σ^{70}-dependent *E. coli* promoters (Fig. 2). Based on this motif and previous data from the literature (2, 8, 16, 50), three LTTRs were selected as putative transcriptional regulators of *lysP*: LysR, YeiE, and ArgP. LysR is the activator protein required for expression of *lysA*, which encodes the enzyme that catalyzes the last step in lysine biosynthesis, the decarboxylation of diaminopimelate (DAP), into lysine (50). LysR is responsive to the intracellular concentration of DAP and lysine. YeiE is encoded by a gene located immediately upstream of *lysP*, and the induction of *yeiE* expression increased the expression of *lysP* (16). Unexpectedly, under the conditions tested (minimal medium at pH 7.6 and 5.8 with or without the addition of external lysine and rich medium), neither a deletion of *lysR* nor a deletion of *yeiE* affected the expression of *lysP* (data not shown). It has been reported that mutations in *argP* affect the uptake of arginine, ornithine, and lysine (8). However, the direct implication of ArgP in the regulation of genes encoding the corresponding transport proteins responsible for the uptake of these amino acids has never been analyzed.

To evaluate whether ArgP was involved in the transcriptional regulation of *lysP*, a nonfunctional *argP* allele (Δ*argP*::Cam^r) was transduced into strain MG-LR carrying the chromosomal P_{lysP}::*lacZ* fusion. The resulting mutant, named MG-LR17, and the *argP*$^+$ parent strain were grown in glucose minimal medium with or without the addition of 0.1 mM lysine and analyzed for β-galactosidase activities (Fig. 3A). The results clearly showed that expression of *lysP* did not occur in the *argP* mutant, either in the absence or in the presence of lysine.

To confirm the role of ArgP in the regulation of *lysP* transcription, the *argP* gene was cloned into plasmid pBAD24 under the control of the arabinose-inducible promoter (18). The *argP* mutant MG-LR17 was transformed with the resulting plasmid, named pBAD*argP* (Table 1). Expression of *lysP* was monitored in cells of the *argP*$^+$ strain (MG-LR) bearing the pBAD24 vector and the *argP* mutant with pBAD24 or pBAD*argP* grown in glucose minimal medium plus arabinose, with and without the addition of lysine. As shown in Fig. 3B, plasmid-carried *argP* fully restored the *lysP* expression pattern in the MG-LR17 mutant. In the absence of lysine, transcription of *lysP* in strain MG-LR increased immediately. In the presence of lysine, induction was prevented within the first hours of growth and slowly increased after prolonged growth, presumably due to lysine limitation. In the *argP* mutant, induction of *lysP* did not occur. However, plasmid-carried *argP* rescued *lysP* expression in mutant MG-LR17. In the absence of lysine, induction of *lysP* occurred after *argP* expression was induced (2.5 h). In the presence of lysine, the expression pattern of this complemented mutant was dependent on the growth phase and probably on lysine availability. These results indicated that ArgP is responsible for the transcriptional activation of *lysP* in the absence of lysine.

Previously, ArgP was identified as a lysine-dependent regulator of *argO*, which encodes an arginine exporter (24, 30), and *dapB*, which encodes an enzyme of the lysine biosynthesis pathway (2). Transcription of these genes was also affected by the presence of arginine in the culture medium, indicating that both lysine and arginine are coeffectors of ArgP (2, 24). Therefore, the effect of arginine on the expression of the P_{lysP}::*lacZ* fusion in strains MG-LR and MG-LR17 was tested. Figure 3C shows the β-galactosidase activities of cells cultivated in glucose minimal medium without amino acids or supplemented with lysine, arginine, or lysine plus arginine. When arginine was added to the growth medium, *lysP* expression in the MG-LR strain was induced to the same extent as in cells that were grown in the absence of amino acids. Importantly, the presence of arginine partially relieved the repressive effect of lysine, as indicated by the >3-fold-higher β-galactosidase activities (251 ± 41.1 Miller units [MU]) of cells that were cultivated in arginine plus lysine medium than of those that were grown in lysine medium (β-galactosidase activity, 67 ± 6.3

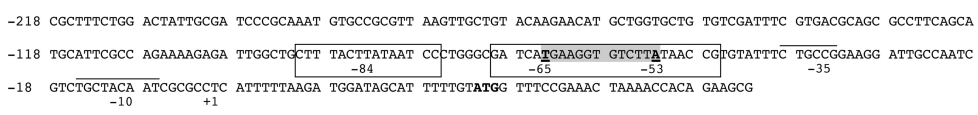

FIG. 2. Nucleotide sequence of the *lysP* regulatory region (positions −218 to +58). Predicted −35 and −10 promoter motifs (BProm; http://linux1.softberry.com/berry.phtml?topic=bprom&group=help&subgroup=gfindb) and the start of transcription (position +1) previously identified by primer extension analysis (32) are indicated. Gray-shaded nucleotides show the potential LTTR binding site identified *in silico* with the conserved T-N_{11}-A motif. Zones protected against DNase I digestion are boxed (see Fig. 5). The start codon is marked in bold letters.

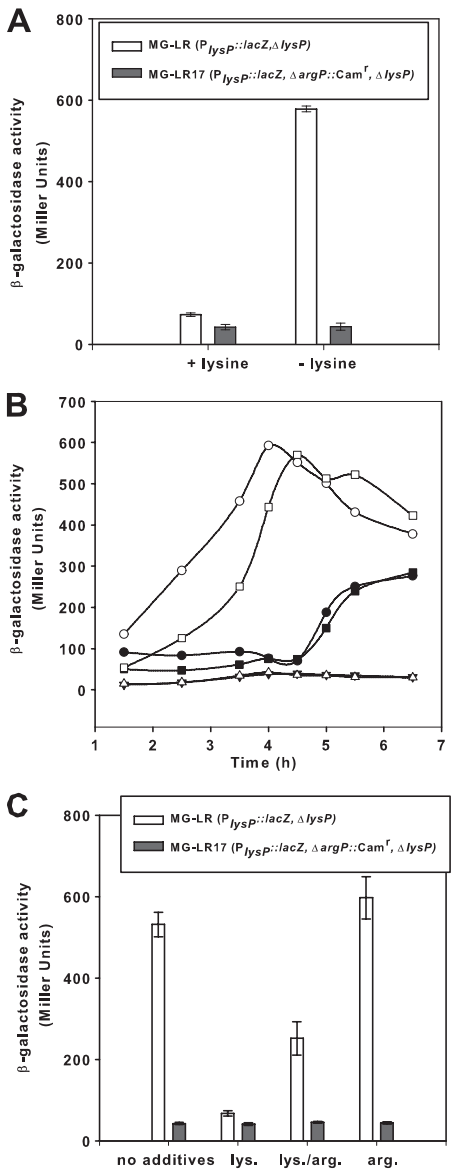

FIG. 3. Regulation of *lysP* expression by lysine, arginine, and ArgP. (A) Effect of the *argP* deletion on *lysP* expression. Parent strain MG-LR (P_{lysP}::*lacZ* Δ*lysP*) and the *argP* mutant MG-LR17 (P_{lysP}::*lacZ* Δ*argP*::Cam^r Δ*lysP*) were grown aerobically in glucose minimal medium with or without the addition of 0.1 mM lysine. When cultures reached an OD_{600} of ~0.8, samples were analyzed for β-galactosidase activity. (B) Complementation of the *argP* mutant with pBAD*argP*. Strains MG-LR/pBAD24 (○, ●), MG-LR17/pBAD24 (△, ▼), and MG-LR17/pBAD*argP* (□, ■) were grown in glucose minimal medium without lysine (open symbols) or with the addition of 0.1 mM lysine (closed symbols). Arabinose at a final concentration of 0.2% (wt/vol) was added to all cultures after 2.5 h of growth. β-Galactosidase activities were determined at different time points during growth. (C) Effect of basic amino acids on *lysP* expression. Strains MG-LR and MG-LR17 were cultivated as described for panel A in glucose minimal medium without supplementation or supplemented with 10 mM lysine (lys.) and/or arginine (arg.) for determination of β-galactosidase activity. All experiments were performed at least three times, and where indicated, error bars represent standard deviations from the means.

MU). As expected, there was no *lysP* expression in the Δ*argP* mutant under all tested conditions.

ArgP binds to the *lysP* promoter/control region at a T-N_{11}-A motif in the presence and absence of lysine. To determine whether ArgP directly regulates *lysP* transcription, we tested binding of ArgP to the *lysP* promoter/control region. For this purpose, ArgP with a C-terminal hexahistidine tag (ArgP-His_6) was purified. In a control experiment, this ArgP derivative complemented strain MG-LR17, indicating that the His_6 tag did not alter the function of ArgP (data not shown). A fluorescently labeled DNA fragment encompassing the *lysP* promoter/control region from position −218 to position +58 (P_{lysP} fragment) was incubated with increasing concentrations of purified ArgP-His_6 in the presence of salmon sperm DNA as a nonspecific competitor (Fig. 4A). A DNA fragment of similar size within the *lysP* coding sequence was used as a control for nonspecific binding (Fig. 4B). ArgP-His_6 specifically bound to the P_{lysP} fragment with an apparent K_D of 125 ± 13 nM (Fig. 4A and F).

To determine whether lysine or arginine affected binding of ArgP to P_{lysP}, these amino acids were added to the binding assay at a final concentration of 0.1 mM (Fig. 4C and D). Whereas arginine did not affect the binding of ArgP to the *lysP* promoter/control region, the presence of lysine increased the binding affinity approximately 2-fold (K_D, 63 ± 9 nM) (Fig. 4C and F). This differential effect of lysine and arginine on DNA affinity of ArgP has already been reported for the *argO* control region (24).

As already mentioned, a potential ArgP-binding site, ATG AAGGTGTCTTAT, is centered at position −59 in the *lysP* promoter/control region (Fig. 2). To evaluate the importance of this sequence for binding of ArgP to the *lysP* control region, a DNA fragment from position −84 to position +58 containing the T-N_{11}-A conserved motif (T-N_{11}-A fragment) and another, from position −35 to position +58, without this T-N_{11}-A sequence (−35-10 fragment) were incubated with ArgP in the presence and absence of lysine (Fig. 4E). A retarded band was apparent when ArgP was incubated with the T-N_{11}-A fragment, whereas in the presence of the DNA fragment without the T-N_{11}-A motif (−35-10 fragment), only a faint retarded band was seen, indicating that the T-N_{11}-A sequence is required for proper binding of ArgP to the *lysP* promoter/control region. Moreover, the affinity of ArgP for the T-N_{11}-A fragment was found to be slightly higher in the presence of lysine (Fig. 4E).

To study the ArgP-binding site in the *lysP* promoter/control region in more detail, DNase I footprinting analysis was performed in the absence and presence of lysine (Fig. 5). ArgP-His_6 protected the stretch from position −91 to position −47, with an intervening unprotected region between positions −76 and −71 (Fig. 2 and 5). Importantly, ArgP bound to the same

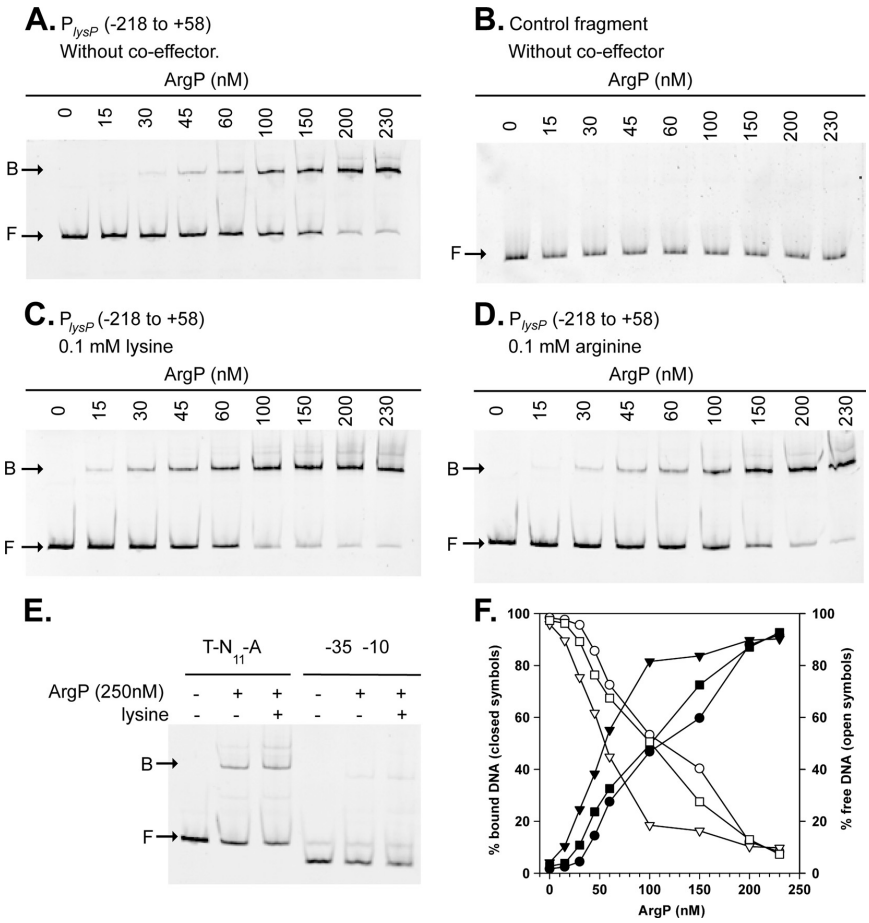

FIG. 4. Binding of ArgP to the *lysP* control region. (A, C, and D) Electrophoretic mobility shift assays (EMSAs) of a fluorescently labeled DNA fragment from bp −218 to bp +58 encompassing the *lysP* promoter/control region (P_{lysP}) with increasing concentrations of purified ArgP-His_6 in the absence of coeffector (A) or in the presence of 0.1 mM lysine (C) or 0.1 mM arginine (D). The positions of free DNA (F) and ArgP-DNA complexes (B) are marked with arrows. (B) A DNA fragment within the *lysP* coding sequence was used as a control for unspecific binding. (E) Binding of ArgP-His_6 to a fragment from position −84 to position +58 bearing the potential ArgP-binding site (T-N_{11}-A) and to a fragment from position −35 to position +58 (−35-10) in the presence or absence of lysine. (F) Binding curves obtained after the quantification of free DNA (open symbols) and ArgP-bound DNA (closed symbols) in EMSA gels without coeffector (○, ●) in the presence of lysine (▽, ▼) or arginine (□, ■).

sites in the presence and absence of its coeffector lysine (Fig. 5B and C).

Elimination or modification of the T-N_{11}-A motif affected *lysP* expression *in vivo*. To evaluate the importance of the T-N_{11}-A motif within the *lysP* promoter/control region for *in vivo* expression of *lysP*, a fragment encompassing this sequence (nucleotides −83 to −52) (Fig. 2) was deleted in the MG-LR strain, resulting in strain MG-LR4. Elimination of this motif completely abolished *lysP* expression under all conditions tested (Fig. 6), indicating its importance for ArgP-mediated transcriptional regulation. To analyze the effect of point mutations within this motif, the whole *lysP* promoter/control region was fused to a promoterless *lacZ* gene in vector pRS415 (Table 2). The deletion of the motif (pRS*lysP*0) or the replacement of both conserved T and A nucleotides (pRS*lysP*3) prevented *lysP* induction under lysine-limiting conditions. The replacements of A (pRS*lysP*1) or T (pRS*lysP*2) did not abolish *lysP* induction but significantly reduced it (Table 2).

Isolation and identification of other proteins that specifically bind to the *lysP* promoter/control region. So far, our results demonstrate that ArgP directly binds to the *lysP* promoter/control region and that the T-N_{11}-A sequence located close to the RNA polymerase binding site is important for binding and crucial for *lysP* transcription. According to the *in vivo* analyses, ArgP is a transcriptional activator of *lysP* in the absence of lysine. The *in vitro* experiments demonstrated that

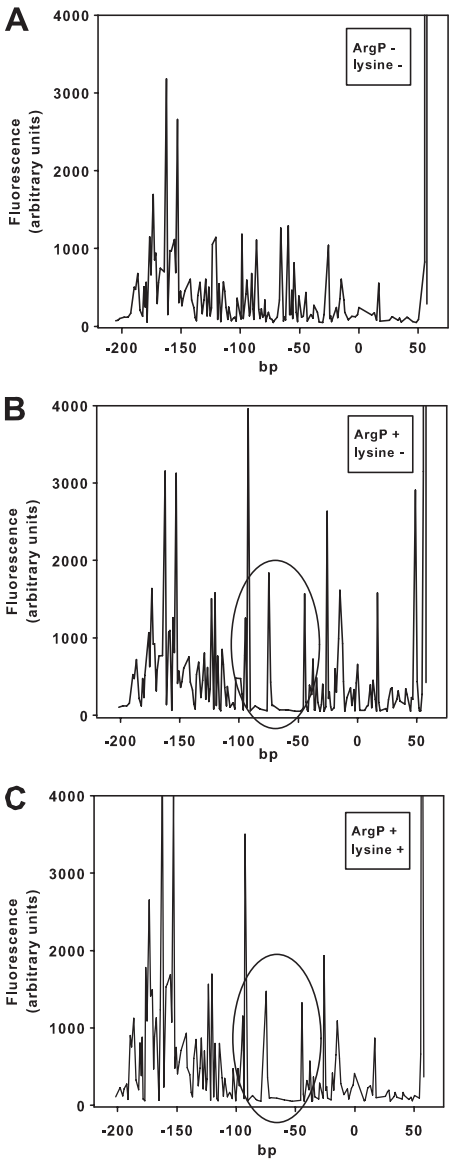

FIG. 5. Determination of the ArgP-binding site within the *lysP* control region. DNase I digestion patterns were determined for a DNA fragment from position −218 to position +58 of the *lysP* control region labeled with fluorescein at the 5′ end of the top strand. Panel A shows the restriction pattern obtained in the absence of purified ArgP-His_6, and panels B and C show the pattern obtained in the presence of 4.3 μM purified protein in the absence (B) or presence (C) of 0.1 mM lysine. Regions protected by ArgP (position −91 to −77 and −70 to −47) are encircled.

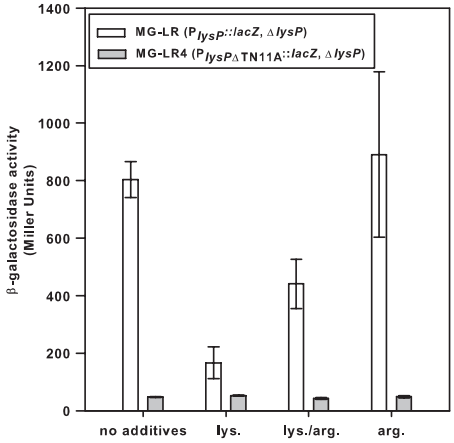

FIG. 6. Effect of elimination of the T-N_{11}-A motif in the *lysP* control region on *lysP* expression. Parent strain MG-LR (P_{lysP}::*lacZ* Δ*lysP*) and strain MG-LR4 ($P_{lysP\Delta TN11A}$::*lacZ* Δ*lysP*) were grown aerobically to an OD_{600} of ~0.8 in glucose minimal medium without supplementation or supplemented with 10 mM lysine (lys.) and/or arginine (arg.) for determination of β-galactosidase activity. All experiments were performed at least three times. Error bars represent standard deviations from the means.

ArgP binds more avidly to the *lysP* control region in the presence of lysine, which might underlie the lysine-mediated prevention of *lysP* expression similarly to the previously described shutoff of *argO* expression by lysine-loaded ArgP (24). Alternatively, another protein might coregulate *lysP* repression. This scenario is plausible, because ArgP and Lrp competitively activate *argO* (35). To search for other proteins that might bind

TABLE 2. Effects of modifications within the *lysP* promoter/control region on *lysP* expression[a]

E. coli strain or strain carrying plasmid	Description of modification within the *lysP* promoter/control region	Induction of *lysP* upon lysine limitation ($lysP_{\text{no lysine}}$/$lysP_{\text{10 mM lysine}}$)[b]
MG-LR	None	7.90
MG1655-Δ*lacZ*/pRS*lysP*	None	3.43
MG-LR4	Deletion of T-N_{11}-A motif (nucleotides −83 to −52)	1.03
MG1655-Δ*lacZ*/pRS*lysP*0	Deletion of T-N_{11}-A motif (nucleotides −88 to −53)	1.28
MG1655-Δ*lacZ*/pRS*lysP*1	Replacement of A/C at position −53	2.77
MG1655-Δ*lacZ*/pRS*lysP*2	Replacement of T/G at position −65	1.85
MG1655-Δ*lacZ*/pRS*lysP*3	Replacement of A/C (position −53) and T/G (position −65)	1.24

[a] Strains were grown aerobically in glucose minimal medium in the absence or presence of 10 mM lysine to an OD_{600} of ~0.8. β-Galactosidase activity was determined and served as a measurement for *lysP* expression.

[b] The ratio between *lysP* expression levels in the absence and presence of lysine indicates the inducibility of the *lysP* promoter. Data were obtained from at least three independent experiments, and average values (the standard deviation was about 15%) were used for calculating the ratios.

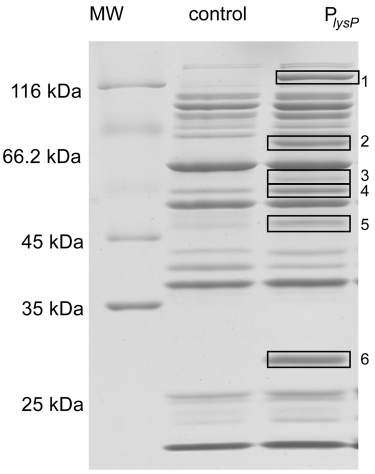

FIG. 7. SDS-PAGE of proteins after DNA affinity purification. A biotin-labeled DNA fragment from position −218 to position +58 encompassing the *lysP* promoter/control region (P_{lysP}) and a DNA fragment within the *lysP* coding region (control) were bound to streptavidin-coated magnetic beads and subsequently incubated with a soluble extract of *E. coli* MG1655 grown in glucose minimal medium. Tightly bound proteins were eluted with a high-ionic-strength buffer and separated by SDS-PAGE. Boxed bands correspond to proteins that specifically bind to the *lysP* promoter/control region (Table 3). MW, molecular mass marker.

to the *lysP* promoter, a DNA affinity purification approach was used. The 276-bp fragment encompassing the *lysP* promoter/control region (positions −218 to +58) was biotinylated, linked to streptavidin-coated magnetic beads, and incubated with a concentrated soluble protein extract from *E. coli* MG1655 grown in glucose minimal medium. The same was done with a fragment of 258 bp within the *lysP* coding sequence, which served as a control. Proteins tightly bound to the DNA fragments were eluted as described in Materials and Methods and analyzed by SDS-PAGE. As shown in Fig. 7, several proteins that specifically bound to the *lysP* promoter/control region and not to the control fragment were detected.

The binding proteins were identified by MALDI-TOF mass spectrometry. With the exception of one, we were able to identify the eluted proteins (Table 3). All of them turned out to be DNA-binding proteins. The most abundant protein was the leucine-responsive-protein, Lrp. As Lrp is a global transcriptional regulator that controls the expression of numerous genes in response to the availability of amino acids and nitrogen bases (6, 9, 33), the role of this protein in *lysP* expression was analyzed in more detail.

Lrp stimulates transcription of *lysP* by direct binding to its control region. To determine whether Lrp influences *lysP* transcription, the *lrp* gene was inactivated in *E. coli* MG-LR, resulting in strain MG-LR15 (Δ*lrp*::Km^r). Expression of the P_{lysP}::*lacZ* fusion in strains MG-LR and MG-LR15 was monitored during growth in glucose minimal medium with and without the addition of lysine. Figure 8A shows that the P_{lysP} activity of the Δ*lrp*::Km^r mutant in the absence of lysine is lower than that of the parent strain. At the late exponential growth phase (optical density at 600 nm [OD_{600}] of about 1.0), the expression level of *lysP* in the MG-LR15 strain was about 50% lower than that in the parent strain MG-LR. The presence of 0.1 mM lysine reduced *lysP* expression in both strains. Therefore, Lrp seems to be an activator of *lysP* transcription in the absence of lysine.

The regulatory effect of Lrp is sometimes modulated by the effector molecule L-leucine (6, 33). Because a high leucine concentration in the culture medium decreases the growth rate due to isoleucine restriction (39), the effect of leucine on *lysP* expression was analyzed by comparing the β-galactosidase activity of bacteria cultivated in glucose minimal medium supplemented with isoleucine and valine to that of cells grown under the same conditions but in the presence of isoleucine, valine, and leucine. We found that L-leucine had no effect on P_{lysP} activity in the Δ*lrp*::Km^r mutant or in the lrp^+ strain (Fig. 8B). To confirm the activator role of Lrp in *lysP* transcription, the pBAD*lrp* plasmid (in which the *lrp* gene was cloned under the control of the arabinose promoter) (Table 1) was introduced into strain MG-LR15, and *lysP* expression levels were monitored before and after the addition of L-arabinose to the culture medium. Figure 8C clearly shows that the induction of *lrp* expression stimulated *lysP* transcription. Together, these results suggest that Lrp potentiates the activator effect of ArgP on *lysP* transcription in the absence of lysine.

Among the regulatory targets of Lrp, there are many transcription factors that participate in the regulation of amino acid metabolism and molecule transport (9). Therefore, it was important to determine whether the activator role of Lrp in *lysP* expression was direct or indirect. Peeters et al. have recently demonstrated that *argP* expression was not regulated by Lrp (35). Thus, EMSAs were performed in the presence of an excess of nonspecific competitor to assess the binding of purified His_6-Lrp to the *lysP* promoter/control region (Fig. 9). One to four different retarded complexes were observed when various concentrations of Lrp were incubated with the 276-bp fragment encompassing the *lysP* promoter/control region (P_{lysP}) (Fig. 9A). The apparent K_D obtained for the binding

TABLE 3. Identification of proteins after DNA affinity purification by peptide fingerprint analysis

Band[a]	Identified protein
1	HsdR; host restriction endonuclease R. Subunit of EcoKI restriction-modification system
2	HsdM; modification methyltransferase component of the EcoKI restriction-modification system
3	HsdS; specificity-determinant component of EcoKI restriction-modification system
4	NadR; transcriptional regulator
5	Nonidentified protein
6	Lrp; leucine-responsive protein

[a] The band numbers correspond to those shown in Fig. 7.

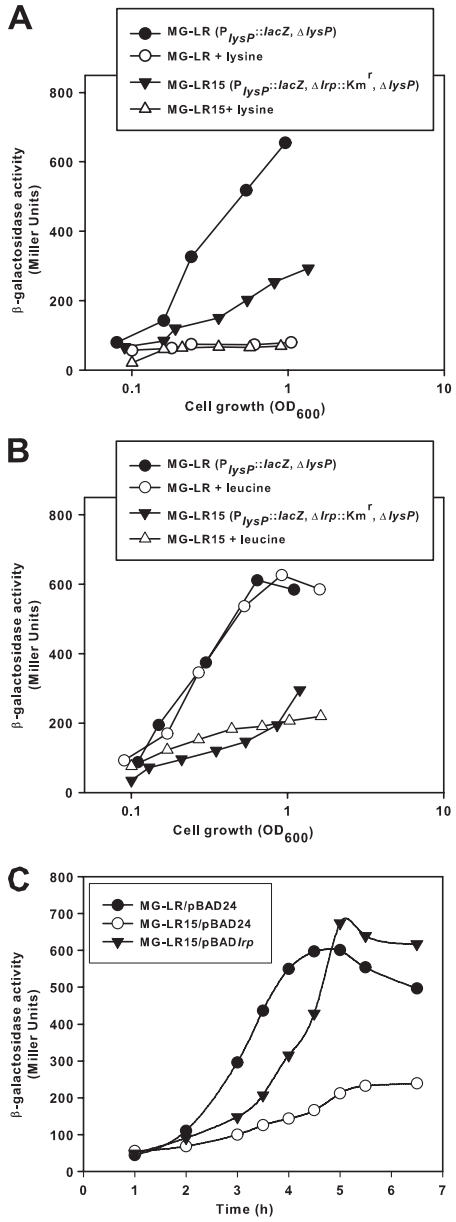

FIG. 8. Effect of *lrp* deletion on *lysP* expression. *E. coli* strains MG-LR (P_{lysP}::*lacZ*, Δ*lysP*) and MG-LR15 (P_{lysP}::*lacZ*, Δ*lysP*, Δ*lrp*::Km^r) were grown aerobically in glucose minimal medium with or without the addition of 0.1 mM lysine (A). To test the effect of leucine on *lysP* expression, the growth medium was supplemented with 0.6 mM valine and 0.4 mM isoleucine, with or without the addition of 10 mM leucine (B). At different times during growth samples were analyzed for β-galactosidase activity. (C) Complementation of the MG-LR15 (P_{lysP}::*lacZ*, Δ*lysP*, Δ*lrp*::Km^r) mutant with the pBAD*lrp* plasmid. Strains MG-LR/pBAD24, MG-LR15/pBAD24 and MG-LR15/pBAD*lrp* were grown in glucose minimal medium, and after 2.5 h 0.2% arabinose (wt/vol) was added to induce the expression of the *lrp* gene cloned in pBAD24. β-Galactosidase activities were determined in samples collected at different times, before and after induction. All experiments were performed in triplicates and mean values are presented. The standard deviations from the mean were less than 10%.

reaction was 610 ± 44 nM (Fig. 9D), and the Hill coefficient was 3.

As previously reported for several Lrp-regulated genes, the concentration-dependent binding of Lrp to *lysP* suggests cooperative binding of Lrp dimers to multiple binding sites resulting in complexes with different stoichiometries. Since the *in vivo* experiments demonstrated that the presence of lysine affected *lysP* expression levels (Fig. 8A), we evaluated whether lysine affects the binding of Lrp to the *lysP* control region *in vitro*. As shown in Fig. 9C and D, the addition of L-lysine to the binding assay neither changed the Lrp affinity nor changed the binding pattern. When His_6-Lrp was incubated with a DNA fragment encompassing the *lysP* coding sequence (control fragment), no retarded bands were observed (Fig. 9B), confirming the specific binding of Lrp to the *lysP* control region.

Lrp modulates expression of the *cadBA* operon. The Lrp regulon comprises several genes involved in amino acid synthesis and degradation. For example, expression of *ldcC*, coding for the constitutive lysine decarboxylase of *E. coli*, is downregulated by Lrp (53). Considering the important role of LysP in regulating the expression of *cadA* (55), which encodes the inducible lysine decarboxylase, we analyzed the effect of Lrp on transcription of the *cadBA* operon, which encodes the inducible lysine decarboxylase CadA and the lysine/cadaverine antiporter CadB. For this purpose, two strains that carry chromosomal promoter-*lacZ* fusions (P_{cadBA}::*lacZ*) and are either *lrp*$^+$ (MG-CR) or *lrp* null (Δ*lrp*::Km^r, MG-CR15) were constructed. Cells were cultivated in glucose minimal medium at physiological (pH 7.6) or low (pH 5.8) pH, with or without the addition of 5 mM lysine, and β-galactosidase activities were determined. As expected, the P_{cadBA} promoter was active only in cells that were exposed to low external pH in the presence of lysine (Fig. 10A), a condition that is known to induce the Cad system (27). However, under inducing conditions, the P_{cadBA} activity in the Δ*lrp*::Km^r mutant strain was >2-fold lower than that in the parent strain, suggesting that Lrp stimulates expression of *cadBA*.

To corroborate these results, the pET16*lrp* plasmid (Table 1) was introduced into the Δ*lrp*::Km^r mutant strain and the P_{cadBA} activity in strains MG-CR/pET16b, MG-CR15/pET16b, and MG-CR15/pET16*lrp* was determined. The pET16b plasmid was chosen as an expression vector in this case, because the T7-driven promoter ensures very low expression levels in strain MG1655. The results presented in Fig. 10B reveal that the reintroduction of *lrp* into the Δ*lrp*::Km^r mutant increases the P_{cadBA} activity to levels similar to those for the parent strain. It has been demonstrated that Lrp does not alter intracellular levels of CadC, the transcriptional activator of the *cadBA* operon (42). Therefore, the possibility of an indirect effect of Lrp via CadC can be discarded.

To evaluate the capacity of Lrp to bind to the P_{cadBA} pro-

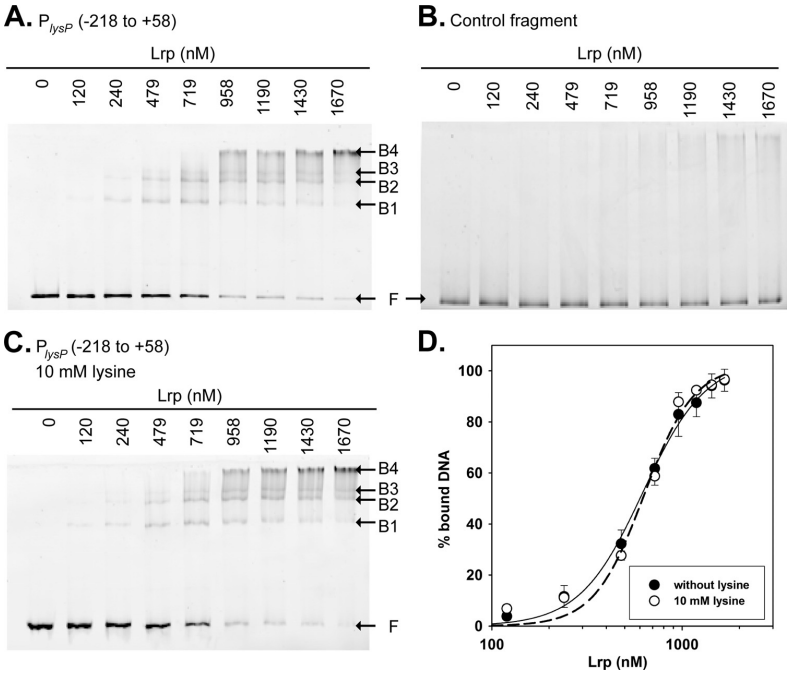

FIG. 9. Lrp binding to the *lysP* promoter region. (A and C) Electrophoretic mobility shift assays (EMSAs) of the fluorescently labeled P_{lysP} fragment (positions −218/+58) with increasing concentrations of purified His_6-Lrp in the absence (A) or presence (C) of lysine. (B) EMSA performed with a DNA fragment within the *lysP* coding sequence as a control for unspecific binding. The positions of free DNA (F) and Lrp-DNA complexes (B1 to B4) are marked. (D) Binding profiles obtained after the quantification of free DNA and Lrp-bound DNA in EMSA gels were fitted using the Hill equation.

moter, EMSAs were performed with various concentrations of the purified His_6-Lrp protein and a DNA fragment encompassing the P_{cadBA} control region. Accordingly, His_6-Lrp is able to bind to the promoter that drives the expression of *cadBA* with an apparent dissociation constant (K_D) of 648 ± 66 nM and a Hill coefficient of 2.15 (Fig. 10C; see also Fig. S1 in the supplemental material). Together, these results indicate that Lrp upregulates expression of *cadBA*. With the identical functions of CadA and LdcC in amino acid catabolism taken into account, the opposite regulation of both genes by Lrp would be an efficient way for the cell to save energy.

DISCUSSION

ArgP is responsible for the lysine-dependent control of *lysP* transcription. While studying the mechanisms involved in the regulation of the Cad system (lysine decarboxylase system) in *E. coli*, Neely and Olson had shown that *lysP* transcription is controlled by the exogenous lysine concentration (32). Here, we identified the LysR-type transcriptional regulator ArgP as the regulator responsible for the control of *lysP* transcription. Our results indicate that under lysine-limiting growth conditions, ArgP functions as a transcriptional activator of *lysP* expression by binding to a sequence located between positions −91 and −47 in the *lysP* promoter/control region. Specifically, a T-N_{11}-A motif (nucleotides −65 to −53), characteristic for LTTR-dependent promoters (45), was identified. *In vitro* EMSAs as well as *in vivo* transcriptional studies indicated that this motif is crucial for ArgP binding. Deletion of the whole motif or substitution of the conserved nucleotides T and A prevented or reduced *lysP* induction in the absence of lysine. ArgP is a member of the LTTR protein family, which binds coeffectors. Arginine and lysine were found to bind to ArgP (24). *In vitro* assays indicated that ArgP bound to the *lysP* promoter/control region irrespective of the presence of lysine and arginine, but lysine increased the affinity of ArgP by a factor of 2.

ArgP of *E. coli* is also responsible for the lysine-dependent regulation of *dapB*, which encodes one of the enzymes of the diaminopimelate and lysine biosynthesis pathway (2), and *argO*, which encodes the arginine exporter ArgO (24, 30). It is important to note that the molecular mechanisms of lysine-dependent regulation by ArgP differ between *argO* and *dapB* (2, 24). In the case of *dapB*, lysine prevents binding of ArgP to its binding site, which is located in the position −118/−81 interval upstream of the transcriptional start site (2). In contrast, ArgP binds to a sequence between positions −85 and −20 of the *argO* operator/promoter and forms a stable binary complex in both its liganded and its unliganded forms. Arg-loaded ArgP binds to the *argO* promoter/control region and

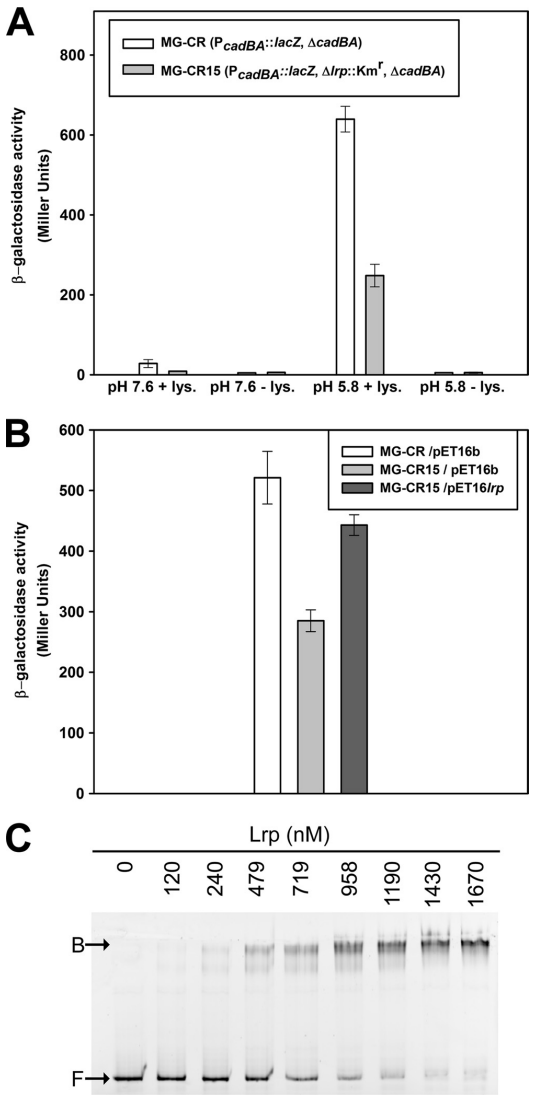

FIG. 10. Lrp stimulates *cadBA* expression. (A) Effect of *lrp* mutation on the expression of the *cadBA* operon under different growth conditions. Strains MG-CR (P_{cadBA}::*lacZ* Δ*cadBA*) and MG-CR15 (P_{cadBA}::*lacZ* Δ*lrp*::Km^r Δ*cadBA*) were grown in glucose minimal medium at pH 7.6 or 5.8 with or without the addition of 5 mM lysine (lys.) under microaerobic conditions. After 7 h of incubation, samples were collected and β-galactosidase activities were determined. (B) Complementation of the Δ*lrp*::Km^r mutant with the pET16*lrp* plasmid. Strains MG-CR/pET16b, MG-CR15/pET16b, and MG-CR15/pET16*lrp* were cultivated in glucose minimal medium (pH 5.8) with 5 mM lysine. After 7 h of incubation, β-galactosidase activities were determined. (C) Binding of His_6-Lrp to the P_{cadBA} promoter/control region. A fragment encompassing the P_{cadBA} promoter/control region was incubated with increasing concentrations of purified His_6-Lrp in the presence of salmon sperm DNA as a nonspecific competitor. The positions of free DNA (F) and the Lrp-DNA complex (B) are marked.

recruits the RNA polymerase, resulting in the induction of *argO*. Lys-loaded ArgP binds to the same sites, albeit with higher affinity, but restrains the polymerase in a molecular complex that is competent for neither productive nor abortive transcription.

Our results suggest that the mechanism for the ArgP-controlled *lysP* expression is similar to the one described for *argO*. Specifically, unloaded or Arg-loaded ArgP induces *lysP* expression, while the Lys-loaded form prevents expression. The lysine K_D value reported for ArgP is 70 μM (24), which is in good agreement with the observed shutoff of the P_{lysP}::*lacZ* activity at an external lysine concentration higher than 25 μM. In contrast to *argO*, external arginine seems not to be essential for transcription of *lysP* by ArgP. This difference might be related to the locations of the ArgP-binding sites within the promoter/control regions. The ArgP-binding site for *argO* extends up to nucleotide −20 and thereby overlaps the −35 promoter motif, whereas the ArgP-binding site for *lysP* extends up to position −47, which is upstream of the −35 promoter site (Fig. 2). Nonetheless, an arginine effect on *lysP* expression was detectable when lysine and arginine were simultaneously added to the cultures. Under this condition, arginine overrode the inhibitory effect of lysine by a factor of 3. Laishram and Gowrishankar (24) demonstrated that arginine and lysine compete for the binding to dimeric ArgP. Therefore, the levels of *lysP* expression measured in cells that were grown in the presence of both amino acids could be attributable to the simultaneous existence of Arg-ArgP and Lys-ArgP complexes, which are effective or ineffective, respectively, in transcriptional activation.

In general, the results obtained in this work underline the importance of ArgP in the transcriptional control of genes involved in basic amino acid transport. ArgP seems to be a versatile regulator, able to control gene expression of various basic amino acid transporters by responding to low-molecular-weight coeffectors in order to maintain a balance in the intracellular concentration of at least lysine and arginine.

Lrp-dependent regulation of *lysP* and *cadBA*. The leucine-responsive protein (Lrp) was identified as another protein that specifically binds to a DNA sequence encompassing the *lysP* control region. Lrp has been designated a physiological barometer (14). Its main function is to control the expression of target genes and operons according to the nutritional status of the cell. Lrp upregulates genes during famine and downregulates genes during feast (6, 9). Most of the genes regulated by Lrp are involved in small-molecule transport and amino acid metabolism (9, 53). The amino acid L-leucine might act as a coeffector of Lrp and potentiates, overcomes, or has no effect on the function of Lrp upon its target genes (6, 33).

According to our results, Lrp potentiates the ArgP-mediated *lysP* expression when cells are cultivated in the absence of lysine, irrespective of L-leucine availability. Thus far, *lysP* has not been identified as a member of the Lrp regulon in the various genome-scale studies performed (9, 21, 53). While Lrp stimulated *lysP* expression in the absence of lysine, this global regulator did not alter the lysine-dependent repression of *lysP*. In addition, Lrp had a positive effect on the expression of the *cadBA* operon when cells were exposed to moderate acidic stress in the presence of lysine. A similar effect of Lrp has been reported for the *cadBA* operon of *Vibrio vulnificus* (42). Lrp binds cooperatively to the DNA at a degenerate consensus

sequence (11), but also with high affinity in a nonspecific manner (36). This last characteristic, together with the high abundance of Lrp in the cell and the fact that binding of Lrp to the DNA causes major conformational changes (13, 57), classifies Lrp as a nucleoid-associated protein (NAP) (14).

The results presented here demonstrate that Lrp has a stimulating effect on the expression of both *lysP* and *cadBA*. As Lrp directly interacts with the *lysP* and *cadBA* control regions in EMSAs performed with a high concentration of nonspecific competitor DNA, the possibility of an indirect effect can be discarded. We have shown that ArgP is the main regulator for *lysP*, and CadC is the major regulator for the *cadBA* operon (12, 58). These regulators activate expression of the corresponding target genes in response to lysine availability (ArgP for *lysP* and CadC via interaction with LysP for *cadBA*) (55, 58) and low pH (CadC) (58). In this scenario, Lrp would impose another level of regulation, adjusting the expression levels of *lysP* and *cadBA* in response to the physiological status of the cell.

The mechanisms employed by Lrp to regulate transcription vary and often involve interaction with other regulator proteins and NAPs. Recently, the implication of Lrp in the regulation of the gene encoding the arginine exporter (*argO*) was reported. It was shown that Lrp and ArgP behave as competitive activators able to activate *argO* expression under different conditions (35). On the other hand, expression of the *artPIQM* operon, one of the systems responsible for arginine uptake, is downregulated by ArgR (5) and Lrp (21). Here, we found that Lrp binds to several sites at the *lysP* control region and is able to potentiate the transcriptional activation mediated by ArgP when lysine becomes limiting. Considering the *in vivo* and *in vitro* data presented in this work, the mechanism by which ArgP regulates *argO* transcription (24), and the known capacity of Lrp to alter the shape of the DNA by inducing bending and wrapping (57), a model for *lysP* regulation in which binding of Lrp to the *lysP* control region may favor and/or stabilize the ArgP-RNA polymerase-DNA complexes or introduce DNA conformational changes is conceivable.

Regarding *cadBA* transcriptional regulation, it is known that H-NS represses *cadBA* expression under noninducing conditions (46), and according to the current model, CadC binding dissolves the repressor complex formed by H-NS (22). The interplay between Lrp and other NAPs, in particular H-NS, in the regulation of several genes is well documented (10, 25, 38). According to the results described here, Lrp participates in the activation of *cadBA* under inducing conditions. It is conceivable that CadC and Lrp dissolve the repressor complex formed by H-NS in a joint action.

In conclusion, the three transporters for lysine (CadB, the LAO system, and LysP) in *E. coli* are produced under different conditions, and the corresponding genes are under the control of various regulators to meet diverse cellular needs. Expression of *cadB*, encoding the lysine/cadaverine antiporter CadB, is induced only at low pH and when external lysine is available to counteract acidic stress (27, 31). *hisJQMP*, encoding the histidine-binding protein and the ABC-type transporter of the LAO system, are repressed by Arg-loaded ArgR (4). In contrast, the arginine-ornithine-lysine binding protein which interacts with the same ABC transporter is induced under nitrogen-limiting conditions and controlled by NtrC (61), suggesting that the LAO system serves as a scavenging system for nitrogen-rich amino acids under conditions of nitrogen starvation. As demonstrated here, *lysP* expression is induced under lysine limitation and requires ArgP and Lrp. Therefore, the main role of LysP seems to be the uptake of lysine for biosynthetic purposes. When lysine is sufficiently available, expression of *lysP* is shut off.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (JU270/5-3 and Exc114/1), the Humboldt Foundation (fellowship to J.R.), and the Elite Network of Bavaria (fellowship to I.H.). J.R. was a researcher on leave from CONICET.

We thank Tobias Kraxenberger and Britta Jungwirth for technical advice, Susanne Ude for constructing strains, and Luitpold Fried for MALDI-TOF measurements. We are grateful to Claude Gutierrez for generously providing *E. coli* strain JCP95.

REFERENCES

1. **Bachmann, B. J.** 1996. Derivations and genotypes of some mutant derivaties of *Escherichia coli* K-12, p. 2460–2488. *In* F. C. Neidhardt et al. (ed.), *Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, DC.
2. **Bouvier, J., P. Stragier, V. Morales, E. Remy, and C. Gutierrez.** 2008. Lysine represses transcription of the *Escherichia coli dapB* gene by preventing its activation by the ArgP activator. J. Bacteriol. **190:**5224–5229.
3. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
4. **Caldara, M., et al.** 2008. Arginine biosynthesis in *Escherichia coli*: experimental perturbation and mathematical modeling. J. Biol. Chem. **283:**6347–6358.
5. **Caldara, M., P. N. Minh, S. Bostoen, J. Massant, and D. Charlier.** 2007. ArgR-dependent repression of arginine and histidine transport genes in *Escherichia coli* K-12. J. Mol. Biol. **373:**251–267.
6. **Calvo, J. M., and R. G. Matthews.** 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. Microbiol. Rev. **58:**466–490.
7. **Casadaban, M. J.** 1976. Transposition and fusion of the *lac* genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. J. Mol. Biol. **104:**541–555.
8. **Celis, T. F., H. J. Rosenfeld, and W. K. Maas.** 1973. Mutant of *Escherichia coli* K-12 defective in the transport of basic amino acids. J. Bacteriol. **116:**619–626.
9. **Cho, B. K., C. L. Barrett, E. M. Knight, Y. S. Park, and B. O. Palsson.** 2008. Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **105:**19462–19467.
10. **Corcoran, C. P., and C. J. Dorman.** 2009. DNA relaxation-dependent phase biasing of the fim genetic switch in *Escherichia coli* depends on the interplay of H-NS, IHF and LRP. Mol. Microbiol. **74:**1071–1082.
11. **Cui, Y., Q. Wang, G. D. Stormo, and J. M. Calvo.** 1995. A consensus sequence for binding of Lrp to DNA. J. Bacteriol. **177:**4872–4880.
12. **Dell, C. L., M. N. Neely, and E. R. Olson.** 1994. Altered pH and lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. Mol. Microbiol. **14:**7–16.
13. **de los Rios, S., and J. J. Perona.** 2007. Structure of the *Escherichia coli* leucine-responsive regulatory protein Lrp reveals a novel octameric assembly. J. Mol. Biol. **366:**1589–1602.
14. **Dillon, S. C., and C. J. Dorman.** 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat. Rev. Microbiol. **8:**185–195.
15. **Epstein, W., and B. S. Kim.** 1971. Potassium transport loci in *Escherichia coli* K-12. J. Bacteriol. **108:**639–644.
16. **Fujii, T., Y. Aritoku, H. Agematu, and H. Tsunekawa.** 2002. Increase in the rate of L-pipecolic acid production using lat-expressing *Escherichia coli* by *lysP* and *yeiE* amplification. Biosci. Biotechnol. Biochem. **66:**1981–1984.
17. **Gabrielsen, O. S., E. Hornes, L. Korsnes, A. Ruet, and T. B. Oyen.** 1989. Magnetic DNA affinity purification of yeast transcription factor τ-a new purification principle for the ultrarapid isolation of near homogeneous factor. Nucleic Acids Res. **17:**6253–6267.
18. **Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith.** 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J. Bacteriol. **177:**4121–4130.
19. **Heermann, R., T. Zeppenfeld, and K. Jung.** 2008. Simple generation of site-directed point mutations in the *Escherichia coli* chromosome using Red(R)/ET(R) recombination. Microb. Cell Fact. **7:**14.
20. **Henzel, W. J., et al.** 1993. Identifying proteins from two-dimensional gels by

molecular mass searching of peptide fragments in protein sequence databases. Proc. Natl. Acad. Sci. U. S. A. **90:**5011–5015.
21. **Hung, S. P., P. Baldi, and G. W. Hatfield.** 2002. Global gene expression profiling in *Escherichia coli* K12. The effects of leucine-responsive regulatory protein. J. Biol. Chem. **277:**40309–40323.
22. **Küper, C., and K. Jung.** 2005. CadC-mediated activation of the cadBA promoter in *Escherichia coli*. J. Mol. Microbiol. Biotechnol. **10:**26–39.
23. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680–685.
24. **Laishram, R. S., and J. Gowrishankar.** 2007. Environmental regulation operating at the promoter clearance step of bacterial transcription. Genes Dev. **21:**1258–1272.
25. **Levinthal, M., P. Lejeune, and A. Danchin.** 1994. The H-NS protein modulates the activation of the *ilvIH* operon of *Escherichia coli* K12 by Lrp, the leucine regulatory protein. Mol. Gen. Genet. **242:**736–743.
26. **Meng, S. Y., and G. N. Bennett.** 1992. Nucleotide sequence of the *Escherichia coli cad* operon: a system for neutralization of low extracellular pH. J. Bacteriol. **174:**2659–2669.
27. **Meng, S. Y., and G. N. Bennett.** 1992. Regulation of the *Escherichia coli cad* operon: location of a site required for acid induction. J. Bacteriol. **174:**2670–2678.
28. **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, Cold Spring Harbor, NY.
29. **Moore, J. T., A. Uppal, F. Maley, and G. F. Maley.** 1993. Overcoming inclusion body formation in a high-level expression system. Protein Expr. Purif. **4:**160–163.
30. **Nandineni, M. R., and J. Gowrishankar.** 2004. Evidence for an arginine exporter encoded by *yggA* (*argO*) that is regulated by the LysR-type transcriptional regulator ArgP in *Escherichia coli*. J. Bacteriol. **186:**3539–3546.
31. **Neely, M. N., C. L. Dell, and E. R. Olson.** 1994. Roles of LysP and CadC in mediating the lysine requirement for acid induction of the *Escherichia coli cad* operon. J. Bacteriol. **176:**3278–3285.
32. **Neely, M. N., and E. R. Olson.** 1996. Kinetics of expression of the *Escherichia coli cad* operon as a function of pH and lysine. J. Bacteriol. **178:**5522–5528.
33. **Newman, E. B., and R. Lin.** 1995. Leucine-responsive regulatory protein: a global regulator of gene expression in *E. coli*. Annu. Rev. Microbiol. **49:**747–775.
34. **Patte, J. C., M. Akrim, and V. Mejean.** 1998. The leader sequence of the *Escherichia coli lysC* gene is involved in the regulation of LysC synthesis. FEMS Microbiol. Lett. **169:**165–170.
35. **Peeters, E., P. Nguyen Le Minh, M. Foulquie-Moreno, and D. Charlier.** 2009. Competitive activation of the *Escherichia coli argO* gene coding for an arginine exporter by the transcriptional regulators Lrp and ArgP. Mol. Microbiol. **74:**1513–1526.
36. **Peterson, S. N., F. W. Dahlquist, and N. O. Reich.** 2007. The role of high affinity non-specific DNA binding by Lrp in transcriptional regulation and DNA organization. J. Mol. Biol. **369:**1307–1317.
37. **Popkin, P. S., and W. K. Maas.** 1980. *Escherichia coli* regulatory mutation affecting lysine transport and lysine decarboxylase. J. Bacteriol. **141:**485–492.
38. **Pul, U., R. Wurm, and R. Wagner.** 2007. The role of LRP and H-NS in transcription regulation: involvement of synergism, allostery and macromolecular crowding. J. Mol. Biol. **366:**900–915.
39. **Quay, S. C., T. E. Dick, and D. L. Oxender.** 1977. Role of transport systems in amino acid metabolism: leucine toxicity and the branched-chain amino acid transport systems. J. Bacteriol. **129:**1257–1265.
40. **Reitzer, L.** 5 July 2005, posting date. Chapter 3.4.7, Catabolism of amino acids and related compounds. *In* R. Curtiss III (ed.), EcoSal—*Escherichia coli* and *Salmonella*: cellular and molecular biology. ASM Press, Washington, DC. doi:10.1128/ecosal.3.6.1.3.
41. **Rey, D. A., A. Puhler, and J. Kalinowski.** 2003. The putative transcriptional repressor McbR, member of the TetR-family, is involved in the regulation of the metabolic network directing the synthesis of sulfur containing amino acids in *Corynebacterium glutamicum*. J. Biotechnol. **103:**51–65.
42. **Rhee, J. E., K. S. Kim, and S. H. Choi.** 2008. Activation of the *Vibrio vulnificus cadBA* operon by leucine-responsive regulatory protein is mediated by CadC. J. Microbiol. Biotechnol. **18:**1755–1761.
43. **Rodionov, D. A., A. G. Vitreschak, A. A. Mironov, and M. S. Gelfand.** 2003. Regulation of lysine biosynthesis and transport genes in bacteria: yet another RNA riboswitch? Nucleic Acids Res. **31:**6748–6757.
44. **Sandaltzopoulos, R., and P. B. Becker.** 1994. Solid phase DNase I footprinting: quick and versatile. Nucleic Acids Res. **22:**1511–1512.
45. **Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47:**597–626.
46. **Shi, X., B. C. Waasdorp, and G. N. Bennett.** 1993. Modulation of acid-induced amino acid decarboxylase gene expression by *hns* in *Escherichia coli*. J. Bacteriol. **175:**1182–1186.
47. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
48. **Soksawatmaekhin, W., A. Kuraishi, K. Sakata, K. Kashiwagi, and K. Igarashi.** 2004. Excretion and uptake of cadaverine by CadB and its physiological functions in *Escherichia coli*. Mol. Microbiol. **51:**1401–1412.
49. **Steffes, C., J. Ellis, J. Wu, and B. P. Rosen.** 1992. The *lysP* gene encodes the lysine-specific permease. J. Bacteriol. **174:**3242–3249.
50. **Stragier, P., F. Richaud, F. Borne, and J. C. Patte.** 1983. Regulation of diaminopimelate decarboxylase synthesis in *Escherichia coli*. I. Identification of a *lysR* gene encoding an activator of the *lysA* gene. J. Mol. Biol. **168:**307–320.
51. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J. Mol. Biol. **189:**113–130.
52. **Sudarsan, N., J. K. Wickiser, S. Nakamura, M. S. Ebert, and R. R. Breaker.** 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. Genes Dev. **17:**2688–2697.
53. **Tani, T. H., A. Khodursky, R. M. Blumenthal, P. O. Brown, and R. G. Matthews.** 2002. Adaptation to famine: a family of stationary-phase genes revealed by microarray analysis. Proc. Natl. Acad. Sci. U. S. A. **99:**13471–13476.
54. **Tetsch, L., and K. Jung.** 2009. The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. Mol. Microbiol. **73:**982–991.
55. **Tetsch, L., C. Koller, I. Haneburger, and K. Jung.** 2008. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. Mol. Microbiol. **67:**570–583.
56. **Thomason, L. C., N. Costantino, and D. L. Court.** 2007. *E. coli* genome manipulation by P1 transduction. Curr. Protoc. Mol. Biol. **79:**1.17.1–1.17.8.
57. **Wang, Q., and J. M. Calvo.** 1993. Lrp, a major regulatory protein in *Escherichia coli*, bends DNA and can organize the assembly of a higher-order nucleoprotein structure. EMBO J. **12:**2495–2501.
58. **Watson, N., D. S. Dunyak, E. L. Rosey, J. L. Slonczewski, and E. R. Olson.** 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. J. Bacteriol. **174:**530–540.
59. **Weber, A., S. A. Kogl, and K. Jung.** 2006. Time-dependent proteome alterations under osmotic stress during aerobic and anaerobic growth in *Escherichia coli*. J. Bacteriol. **188:**7165–7175.
60. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
61. **Zimmer, D. P., et al.** 2000. Nitrogen regulatory protein C-controlled genes of *Escherichia coli*: scavenging as a defense against nitrogen limitation. Proc. Natl. Acad. Sci. U. S. A. **97:**14674–14679.

7 Concluding Discussion

For a long time, histidine kinase/ response regulator systems were thought to exclusively link environmental signals to intracellular adaptation of gene expression. Only recently it became apparent, that one-component systems are the dominating signal transduction systems in prokaryotes [Ulrich *et al.*, 2005]. Remarkably, the membrane-integrated one-component system CadC of *E. coli* integrates three extracellular signals and alters gene expression in response to these signals. In this study the signal perception mechanism of this membrane-spanning one-component system CadC was analyzed in detail. After mathematical analysis of the different in- and output signals (chapter 2) elaborate studies were accomplished that focused on the different signals processed by CadC (chapter 4, chapter 5). These data were backed up by the resolution of the crystal structure of the sensor domain of CadC, the first structure of a domain of a ToxR-like transcriptional regulator (chapter 3). Completed by an analysis of transcriptional regulation of *lysP* and *cadBA* (chapter 6) all results permit the establishment of a comprehensive model for integration of external signals by the acid stress response system Cad.

7.1 Molecular Insights in Signal Perception

At the beginning of this thesis little was known about the signal perception mechanism of ToxR-like regulators. Since they share the same topology and some of them activate genes in response to acidification, CadC can be used as a model for such sophisticated one-component systems. It has been shown that ToxR-dimers are needed for activation of target gene expression although it is not clear if dimerization occurs signal controlled [Dziejman *et al.*, 1999; Kolmar *et al.*, 1995; Ottemann and Mekalanos, 1996]. In contrast to other ToxR-like regulators, CadC does not depend on an accessory ToxS-like protein for transcriptional activation [Li *et al.*, 2000]. For CadC, low pH and cadaverine (first stimulus and inhibitor, respectively) detection was assigned to the periplasmic domain, while lysine sensing (second stimulus) was allocated to the transmembrane domain [Dell *et al.*, 1994; Tetsch *et al.*, 2008].

With the help of the quantitative model developed in chapter 2, the transient expression behavior of *cadBA* transcription and especially the timing of the shut-off of *cadBA* transcription could be related to cadaverine dependent feedback inhibition. Furthermore, with the aid of detailed quantitative experiments performed at high-time resolution and the model, *in vivo* thresholds for activation (approx. 5 mM lysine and pH 6.2) and inhibition (235 μM cadaverine) were estimated. These data confirmed the results of Neely and Olson [1996], who observed that *cadBA* expression is transient. The threshold concentration for cadaverine is in the same range as the *in vitro*-determined cadaverine affinity for the CadC sensor domain (96 μM, Tetsch *et al.* [2008]). The model also helped to evaluate two different hypotheses for (in-)activation of CadC. Earlier, it was suggested that either conformational changes or cleavage of the DNA binding domain of CadC upon stimuli perception lead to activation of transcription [Küper, 2006]. Since cadaverine was thought to be responsible for the early shut-off of *cadBA* transcription, and cadaverine binds to the periplasmic domain, inactivation by cadaverine would be more likely realized via binding to an activator whose confirmation has been altered rather than binding to the remaining proteolysis product in the membrane. As this remaining peptide would lack the effector domain, inhibition of *cadBA* transcription in this case would only be possible with a high CadC turnover. High turnover would have to be counterbalanced by a strong induction of CadC under inducing conditions, conflicting the observation that CadC is constitutively expressed [Dell *et al.*, 1994; Watson *et al.*, 1992].

To learn more about the main player of the Cad system, CadC, the crystal structure of the periplasmic sensor domain was solved. For this purpose the periplasmic domain was produced as a hybrid protein with the highly soluble thioredoxin [Tetsch *et al.*, 2008]. Following cleavage of the thioredoxin tag and multiple purification steps, the protein was concentrated and used for crystallization attempts. After solving the first structure of $CadC_{pd}$ at 2.3 Å with the help of K_2ReCl_6, the apo structure was solved at 1.8 Å. The sensor domain of CadC ($CadC_{pd}$) appeared to be an irregularly shaped protein, basically composed of two subdomains (chapter 3, Fig. 1A). The first subdomain is a mixed parallel and anti-parallel β-sheet in contact with two α-helices (of a three helix bundle). The second (C-terminal) subdomain is a pure α-helical bundle out of 11 helices. The helices are twisted clockwise and oriented vertically towards the first subdomain. Remarkably, no comparable complete three-dimensional structure has so far been added to the protein databank (PDB). Only helical bundles of three different structures showed similarities to the second subdomain of $CadC_{pd}$ (chapter 3, Fig. 4).

In addition to the two domain composition, the structure revealed another feature of the periplasmic domain. Between the two subdomains an internal cavity was discovered (chapter 3, Fig. 3). Since this cavity is negatively charged and bound $ReCl_6^{2-}$ readily, it may be speculated that this might be the part of the protein that binds the positively charged inhibitor, cadaverine. Concordantly, the docking studies performed suggested that the binding of cadaverine occurs within this cavity. Unfortunately, crystallization attempts with cadaverine did not result in complex structures. An accessory finding was that $CadC_{pd}$ crystallized as dimers (chapter 3, Fig. 2A), supporting the current knowledge that CadC acts (at least) as a dimer, as suggested by two CadC binding sites in the P_{Cad} promoter [Küper and Jung, 2005] and by *in vivo* analyses [Küper, 2006].

Transcription of *cadBA* starts at pH 6.6 and reaches the full induction level at pH 5.8 (chapter 2, Fig. 3B and Slonczewski *et al.* [1987]). Therefore, residues within the C-terminal half of the sensor domain, which are protonable at this pH range, were analyzed for their role in detecting the external pH (histidines, glutamates and aspartates; chapter 4). None of the histidines present in the periplasmic domain was involved in pH-dependent regulation (chapter 4, Fig. 1A), whereas an acidic patch on one side of the periplasmic domain was crucial for pH detection (chapter 4, Fig. 6A). As alterations in phospholipid content did not affect pH dependent activation, it was concluded that the activation mechanism does not involve interactions of the patch with the phospholipid head groups (chapter 4). Intriguingly, the acidic patch is located at the dimer interface of the periplasmic domain (chapter 4, Fig. 8). Furthermore, the patch, stretching between Asp_198, Asp_200, Glu_461, Glu_468 and Asp_471, spans the two subdomains of $CadC_{pd}$. The first subdomain of $CadC_{pd}$ is especially essential for signal transduction across the membrane (chapter 4, Fig. 7). Investigating cadaverine-dependent inactivation of CadC, it became obvious that the internal cavity (chapter 3) is likely of minor importance and instead exhibits a cooperative effect (chapter 5, Fig. 1). In accordance with mutagenesis results published earlier [Dell *et al.*, 1994], a second binding site was identified. As can be deduced from the structural data, this binding site also localizes at the acidic patch at the dimer interface (chapter 5, Fig. 2). Remarkably, pH detection and cadaverine binding occur in the same protein area (at the dimerization interface). Nonetheless, both phenomena seem to function independently of each other. pH detection and cadaverine binding have some residues in common (Thr_475, Leu_479), but, coevally, there are residues that recognize cadaverine but are not involved in pH detection (i.e. Tyr_453) and vice versa (Asp_471 and Glu_468; Dell *et al.* [1994]; chapter 4; chapter 5).

Output Generation by a "Simple" One Component System

How are these signals processed to ensure survival at acidic conditions? $CadC_{pd}$ crystallized as a dimer in neutral crystallization conditions. Therefore, CadC might form dimers unaffected by the surroundings (chapter 3). As mentioned before, a similar behavior has been proposed for ToxR of *V. cholerae*. Hence, at neutral pH, these dimers would represent the inactive state. In addition, LysP is thought to sequester CadC in the inactive state as long as no lysine is available. This inhibitory effect is exerted via the transmembrane domain [Tetsch *et al.*, 2008]. Since the transmembrane domain follows the periplasmic domain in signal transduction, LysP can block activation if no substrate (lysine) is present, even at low pH.
In the case where lysine is available and the surroundings become acidic, LysP-CadC interaction would become loose and CadC would be "free" to respond to the increase in external protons. According to the results of this thesis, the acidic patch would then get protonated. As this is equivalent to neutralization of the acidic, negatively charged patch, this would lead to structural alterations that affect transcriptional activation. The nature of these structural rearrangements, such as approximation of subdomains, rotation and piston-like movements, is so-far-unknown. However, it can be supposed that these rearrangements transduce the information from the first periplasmic subdomain to the cytoplasmic effector domain via the transmembrane helix. Since Neely and Olson [1996] demonstrated that CadC synthesis is essential for activation upon acidification, it is suggested that only newly synthesized CadC can adopt an active conformation. Eventually, protonation of the amino acids residing in the acidic patch during protein maturation (thereby eliminating repelling surfaces) leads to alterations in CadC folding, thus enabling signal transduction through the membrane. As stated before, it is thought that CadC is active as a dimer, since two CadC binding sites exist in the *cadBA* promoter and multimerization has been proven *in vivo* [Küper, 2006; Küper and Jung, 2005].
While CadC is involved in the activation of *cadBA* transcription, it also mediates the inhibition, with no induction occuring in the presence of cadaverine. Cadaverine, the product of decarboxylation, also exerts its inhibitory effect on the periplasmic domain (Dell *et al.* [1994], chapter 5, see below for discussion). At low pH, cadaverine binding occurs in the same protein area as protonation. Consequently, it is likely that cadaverine binding interferes with processes related to pH signaling. It is possible that cadaverine binding prevents CadC monomers located within the dimer from further approximating, rotating and undergoing piston-like transitions, thereby preventing the simultaneous binding to the two CadC binding sites and the initiation of transcription

(figure 7.1B). Consequently, *cadBA* transcription in the presence of external (excreted) cadaverine is inhibited.

7.2 Complex Regulation - A Tribute to Cellular Welfare?

Transcriptional (fine)tuning of *lysP* and *cadBA*

In addition to lysine regulating the transcription of *cadBA*, lysine also represses the transcription of *lysP*, which encodes LysP and exerts lysine-dependent inhibition. *lysP* transcription is repressed after addition of lysine (10 mM) within 4 min [Neely and Olson, 1996]. 50 μM lysine are sufficient to repress *lysP* transcription independent of LysP presence (chapter 6). The LysR-type regulator, ArgP, was identified as responsible for lysine-dependent repression of *lysP* transcription. In addition, ArgP was shown to bind to the *lysP* promoter without ligand as well as with lysine or arginine (chapter 6, Fig. 4). Therefore, it is conceivable that lysine-bound ArgP inhibits *lysP* transcription while uncharged or arginine-bound ArgP induces transcription. A similar regulation by ArgP has been suggested by Laishram and Gowrishankar [2007] for expression of *argO*. ArgP was shown to manipulate the expression of several basic amino acid transporter genes, thus one may postulate that ArgP's effects on gene expression is balanced by the presence of low molecular weight co-effectors. Furthermore, another regulator, Lrp, known to regulate multiple genes in response to the nutritional state of the cell, was shown to affect *lysP* expression. The leucine-responsive protein (Lrp) reinforced the effect of ArgP in the absence of lysine (chapter 6, Fig. 8). Thus, Lrp acted in accordance with its role as a physiological barometer, in which Lrp represses genes during feast and induces genes during famine [Calvo and Matthews, 1994; Cho *et al.*, 2008; Dillon and Dorman, 2010]. Intriguingly, Lrp also promotes *cadBA* expression. In contrast to activation of *lysP* in the absence of lysine, Lrp activates transcription of *cadBA* in the presence of lysine.

Acid stress response - without wasting resources

Considering this newly discovered aspect of *cadBA* expression, it seems that regulation of *cadBA* is aimed at optimizing acid defense and survival. Under non-inducing conditions (physiological pH, no lysine) CadC and LysP are both expressed. CadC is present at a relatively low copy number (approx. 20 proteins/ cell, Tetsch *et al.* [2008]). So, for effective inhibition of CadC, LysP should be present with at least the same number of molecules. In order to avoid accidental binding of the promoter, which would result in transcription activation and energetic waste, H-NS occupies the promoter region of

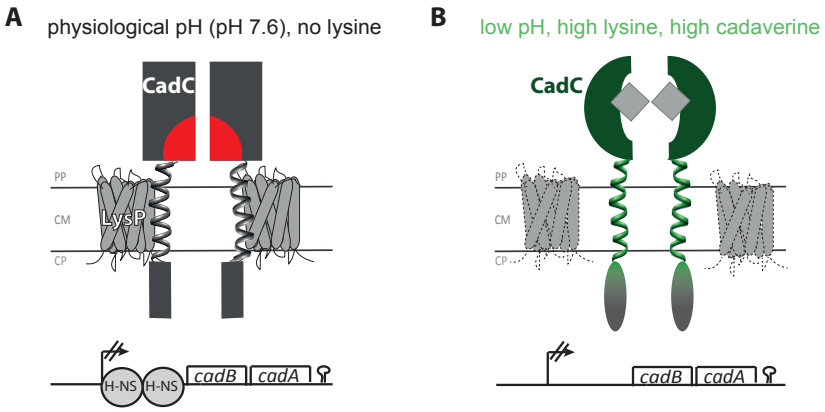

Figure 7.1: Model of inhibition of *cadBA* transcription A. In the absence of stimuli (physiological pH and no lysine). Possibly CadC forms an inactive dimer. The repelling cluster in the periplasmic domain is indicated by red color. The inactive conformation of CadC at physiological pH and lysine deficiency is supported by sequestration by the lysine permease LysP and is schematically represented by a rectangular periplasmic domain of CadC. In addition H-NS binds to the P_{cad} promoter region and thus prevents activator binding. **B.** When the external cadaverine concentration exceeds the threshold, cadaverine (gray squares) binds to newly synthesized CadC molecules and prevents formation of active dimers (conformational rearrangements) that are crucial for transcriptional activation. Thus *cadBA* transcription is inhibited albeit persisting inducing conditions, conceivably due to cadaverine-induced blockage of effector domain positioning. PP: periplasm, CM: cytoplasmic membrane, CP: cytoplasm

cadBA (figure 7.1A).

In the case that the lysine concentration rises to exceed 50 μM (chapter 6), *lysP* transcription is shut off by ArgP•lysine. Thus, presumably the number of LysP molecules (and CadC inhibitors!) decreases. In addition to the presence of less LysP molecules, it is likely that the interaction of transport-active LysP and CadC alters, freeing CadC to be activated by pH. When the cell's surroundings are at physiological pH, CadC remains inactive and *cadBA* is not transcribed (figure 7.2A). In cases where the external pH decreases, a rapid response (within four minutes) is possible, as shown by Neely and Olson [1996].

If the surrounding is acidic but still contains no lysine (figure 7.2B) it is most probable that the fold of the periplasmic domain changes, but that signaling across the membrane is blocked by the persisting interaction with LysP ("semi-active" CadC; figure 7.2B). Unfortunately, trials to "freeze" different CadC conformations by crystallization of different signaling variants were not successful as they all featured the same three-dimensional structure (chapter 3, chapter 4, CadC_D471E: no activation, CadC_D471N: pH independent activation). Inhibition by LysP ensures that *cadBA* is

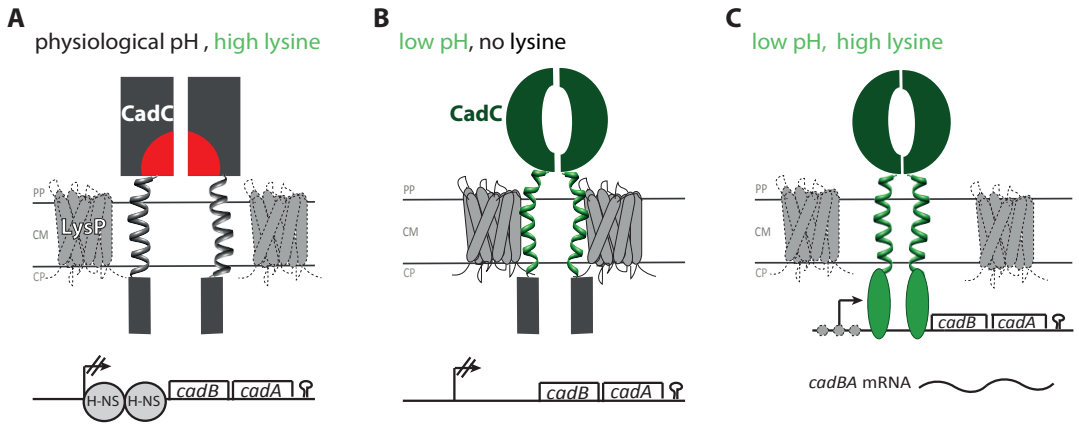

Figure 7.2: Model of CadC action, when either one or both stimuli are present. A. In case of a physiological external pH and lysine present, interaction of LysP and CadC changes and CadC is released for activation. As the milieu is near neutral, CadC remains in its inactive conformation and *cadBA* transcription is off. **B.** In case that the surrounding is acidic but does not contain lysine, the periplasmic domain of newly synthesized CadC adopts the active conformation. Since interaction with LysP persists, activation cannot be transduced over the membrane (*cadBA* is not transcribed). **C.** Both signals (low pH and high lysine concentration) are present: LysP transcription is shut down and interaction of actively transporting LysP with CadC is altered. So, newly synthesized CadC with the active conformation can transduce the signal through the membrane and activates *cadBA* transcription. Concurrently, H-NS is displaced and Lrp (gray hexagons) molecules bind to the *cadBA* promoter and assist transcription activation. PP: periplasm, CM: cytoplasmic membrane, CP: cytoplasm.

not transcribed under lysine limiting conditions. As soon as the lysine concentration passes the threshold, "semi-active" CadC is released and immediately activates *cadBA* transcription (fast response within four minutes; Neely and Olson [1996]).
In the event of both signals being present simultaneously, newly synthesized CadC adopts the active conformation. In this scenario, signal transduction can occur since interaction with transport-active LysP has changed (in the absence of inhibition) and since *lysP* transcription has been shut down. It is assumed that active CadC displaces the global regulator H-NS, binds to the P_{Cad} promoter and activates transcription. In addition, *cadBA* transcription is stimulated by Lrp (figure 7.2C).

As a result, all of these regulatory features seem to guarantee that *cadBA* expression only occurs under acidic conditions when lysine is available. This is especially important if one takes into account that CadA represents over 2% of the total cellular protein under maximal induction [Stim and Bennett, 1993]. Hence, if expressed under the wrong conditions, it takes a high energetic toll on the cell. In addition, cost-intensive transcription (and translation) is shut off when the external cadaverine concentration

reaches a certain threshold level. As CadA and CadB are very stable (in chapter 2, Fig. 4), lysine decarboxylation continuous even though expression of *cadBA* has stopped. This way, the cells are protected without wasting too much energy for transcription and translation. This is also reflected by the transcriptional control by Lrp, which only enhances transcription if lysine is available (chapter 6, Fig. 10A and B). Additionally, Lrp and ArgP regulate transcription of the trigger transporter, LysP, in response to the nutritional state of the cell. As another level of control, the enzymatic activity of CadA is also regulated. At times when cells are experiencing amino acid starvation, when it could be lethal to use amino acids for decarboxylation, CadA is inactivated by the alarmone, ppGpp [Kanjee *et al.*, 2011].

7.3 Outlook

Of the remaining questions regarding signal perception, some of the most pressing are the following: What is the (molecular) difference between the active and the inactive state of CadC? How is the signal transduced (monomer/dimer transition, rotation, piston-like movement)? What happens in the cytoplasmic part upon activation? How does LysP exerts its inhibitory effect?
In order to determine whether CadC turnover is involved in *cadBA* regulation or if the ratio between LysP and CadC plays a crucial role, the signal-dependent levels of CadC and LysP need to be determined. This could be realized by shifting the cells to inducing conditions and comparing the protein levels of CadC and LysP with the help of specific antibodies and Western blotting. In an attempt to better glimpse on transmembrane signaling, a library of single cysteine variants of CadC will be constructed. This library features CadC variants that contain a single cysteine instead of the native amino acid at every position within the transmembrane domain. These variants can be used for cysteine-specific crosslinking [Lee *et al.*, 1995] of CadC variants or to determine distances via EPR (electron paramagnetic resonance) spectroscopy. For this purpose single cysteines are labeled with a cysteine-specific label that features a single electron. Distances between these single electrons can then be determined in a magentic field. Single-cysteine variants could also be used to study the nature of CadC - LysP interaction via EPR. As one could determine signal-dependent distances between a label in CadC and one in LysP [Banham *et al.*, 2008; Hilger *et al.*, 2005]. In order to clarify the terminal number of cadaverine binding sites, a cadaverine-bound crystal structure would be indispensable. Also, to understand the molecular mechanism of CadC to a greater degree, a full-length structure of the membrane-integrated protein would be necessary.

As CadC is highly unstable in solution, one could possibly determine its structure using cryo-electron tomography in the native membrane environment [Trépout *et al.*, 2010]. This technique could also be used to study the nature of the LysP/CadC interaction. Finally, for a complete understanding of signal transduction, the role of the cytoplasmic loop (see chapter 1) needs to be determined. Therefore, a series of different CadC variants, containing single amino acid substitutions, deletions or insertions within this loop region, could prove useful in this endeavor.

7.4 References for Concluding Discussion

Banham, J. E., Baker, C. M., Ceola, S., Day, I. J., Grant, G. H., Groenen, E. J. J., Rodgers, C. T., Jeschke, G., and Timmel, C. R., 2008. Distance measurements in the borderline region of applicability of CW EPR and DEER: a model study on a homologous series of spin-labelled peptides. *Journal of Magnetic Resonance*, 191(2): 202–218.

Calvo, J. M. and Matthews, R. G., 1994. The leucine-responsive regulatory protein, a global regulator of metabolism in *Escherichia coli*. *Microbiology and Molecular Biology Reviews*, 58(3):466–490.

Cho, B. K., Barrett, C. L., Knight, E. M., Park, Y. S., and Palsson, B.Ø., 2008. Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. *Proceedings of the National Academy of Sciences*, 105(49):19462–19467.

Dell, C. L., Neely, M. N., and Olson, E. R., 1994. Altered pH lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. *Molecular microbiology*, 14(1):7–16.

Dillon, S. C. and Dorman, C. J., 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. *Nature Reviews Microbiology*, 8(3):185–195.

Dziejman, M., Kolmar, H., Fritz, H. J., and Mekalanos, J. J., 1999. Toxr co-operative interactions are not modulated by environmental conditions or periplasmic domain conformation. *Molecular microbiology*, 31(1):305–317.

Hilger, D., Jung, H., Padan, E., Wegener, C., Vogel, K. P., Steinhoff, H. J., and Jeschke, G., 2005. Assessing oligomerization of membrane proteins by four-pulse DEER: pH-dependent dimerization of NhaA Na+/H+ antiporter of *E. coli*. *Biophysical journal*, 89(2):1328–1338.

Kanjee, U., Gutsche, I., Alexopoulos, E., Zhao, B., El Bakkouri, M., Thibault, G., Liu, K., Ramachandran, S., Snider, J., Pai, E. F., and Houry, W. A., 2011. Linkage between the bacterial acid stress and stringent responses: the structure of the inducible lysine decarboxylase. *The EMBO Journal*, 30:931–944.

Kolmar, H., Hennecke, F., Götze, K., Janzer, B., Vogt, B., Mayer, F., and Fritz, H. J., 1995. Membrane insertion of the bacterial signal transduction protein ToxR and requirements of transcription activation studied by modular replacement of different protein substructures. *The EMBO Journal*, 14(16):3895–3904.

Küper, C., 2006. Charakterisierung der transkriptionellen Aktivierung des *cadBA*-Operons durch den Transmembranregulator CadC aus *Escherichia coli*. *Doctoral thesis, Ludwig-Maximilians-Universität München.*

Küper, C. and Jung, K., 2005. CadC-mediated activation of the *cadBA* promoter in *Escherichia coli*. *J Mol Microbiol Biotechnol*, 10(1):26–39.

Laishram, R. S. and Gowrishankar, J., 2007. Environmental regulation operating at the promoter clearance step of bacterial transcription. *Genes & development*, 21(10): 1258–1272.

Lee, G. F., Dutton, D. P., and Hazelbauer, G. L., 1995. Identification of functionally important helical faces in transmembrane segments by scanning mutagenesis. *Proceedings of the National Academy of Sciences*, 92(12):5416.

Li, C. C., Crawford, J. A., DiRita, V. J., and Kaper, J. B., 2000. Molecular cloning and transcriptional regulation of ompT, a ToxR-repressed gene in *Vibrio cholerae*. *Molecular microbiology*, 35(1):189–203.

Neely, M. N. and Olson, E. R., 1996. Kinetics of expression of the *Escherichia coli cad* operon as a function of pH and lysine. *J Bacteriol*, 178(18):5522–5528.

Ottemann, K. M. and Mekalanos, J. J., 1996. The ToxR protein of *Vibrio cholerae* forms homodimers and heterodimers. *Journal of bacteriology*, 178(1):156–162.

Slonczewski, J. L., Gonzalez, T. N., Bartholomew, F. M., and Holt, N. J., 1987. Mu d-directed *lacZ* fusions regulated by low pH in *Escherichia coli*. *Journal of bacteriology*, 169(7):3001–3006.

Stim, K. P. and Bennett, G. N., 1993. Nucleotide sequence of the adi gene, which encodes the biodegradative acid-induced arginine decarboxylase of escherichia coli. *Journal of bacteriology*, 175(5):1221–1234.

Tetsch, L., Koller, C., Haneburger, I., and Jung, K., 2008. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol Microbiol*, 67(3):570–583.

Trépout, S., Taveau, J. C., and Lambert, O., 2010. Structure determination of membrane protein by both cryo-electron tomography and single particle analysis. *Methods in molecular biology (Clifton, NJ)*, 654:207–220.

Ulrich, L. E., Koonin, E. V., and Zhulin, I. B., 2005. One-component systems dominate signal transduction in prokaryotes. *Trends in microbiology*, 13(2):52–56.

Watson, N., Dunyak, D. S., Rosey, E. L., Slonczewski, J. L., and Olson, E. R., 1992. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J Bacteriol*, 174(2):530–540.

8 Supplemental Material - chapter 2

Supplementary Material to: Induction kinetics of a conditional pH stress response system in *Escherichia coli*

Georg Fritz [a], Christiane Koller [b], Korinna Burdack [b], Larissa Tetsch [b], Ina Haneburger [b], Kirsten Jung [b], and Ulrich Gerland [a,*]

[a] *Arnold-Sommerfeld Center for Theoretical Physics, LMU München, Theresienstr. 37, 80333 München, Germany*

[b] *Munich Center of integrated Protein Science CiPSM at the Department of Biology I, Microbiology, LMU München, Großhaderner Straße 2-4, 82152 Martinsried, Germany*

[*] *Corresponding author, Email address:* gerland@lmu.de

Correlation analysis between parameters

From Fig. 5 in the main paper we found that for small values of χ^2 some parameters still display a wide variation over their physiological range, as e.g. the transcription and translation rate ν_m and ν_p, respectively. Therefore we reasoned, whether these 'sloppy' parameters do not influence the model behavior at all, or whether not single parameters, but rather combinations of those parameters are constrained by our experiments. To discriminate between these alternative explanations, we selected the best fits with $\chi^2 < 3.22$ from Fig. 5 in the main paper and analyzed the correlations between the fitted parameters, see Fig. S1. This data clearly displays strong pairwise correlations between the sloppy parameters, demonstrating that in fact only their combinations are well constrained. For instance, the transcription and translation rates ν_m and ν_p, which display high variation individually, are strongly anticorrelated, such that their product is kept at a constant level. This is reasonable, since the absolute level of the CadA activity - characterized by $\nu_m \times \nu_p$ - is well determined in our experiments, whereas the mRNA level - characterized by ν_m alone - is only measured in relative units. Likewise, the parameters for lysine turnover, v_{max} and K_m, display a linearly correlated variation, indicating that only their ratio is well constrained by our data. Also among the parameters that control the promoter activity function (K_m/C_0, the promoter fold-change f, K_c and

K_l; cf. Eqs. (2) and (3) in the main paper) pairwise correlations, albeit less pronounced, were detected.

Comparison of estimated parameters with literature values

In the following we discuss the estimated parameters of the Cad module listed in Table 2 of the main paper in the light of previously published literature values.

$\mathsf{K_l}$: activation threshold of CadC by lysine. CadC senses lysine indirectly via interaction with LysP [1–4]. There is recent evidence that CadC is inhibited at low lysine concentrations via a transmembrane domain interaction with the lysine permease LysP, whereas the interaction is released at high lysine levels [5]. In the present work we determined the effective *in vivo* lysine activation threshold to be $K_l = 3.6\,\mathrm{mM}$. This result is somewhat surprising, since the *in vitro* Michaelis constant K_M for lysine transport by LysP is much lower at $\sim 10\,\mu\mathrm{M}$ [6]. However, Neely *et al.* also found that the Cad system is sensitive with respect to lysine in the millimolar range [3], suggesting that our result is plausible under *in vivo* conditions.

$\mathsf{n_l}$: Hill exponent of indirect CadC-lysine interaction. The estimated Hill exponent of $n_l = 1.1$ for the indirect activation of CadC by lysine via LysP suggests that the underlying binding reactions do not involve cooperative mechanisms of CadC activation. We are not aware of a literature value for the Hill exponent for this interaction to which we could compare our estimate.

$\mathsf{K_c}$: inactivation threshold of CadC by cadaverine. For the inactivation of CadC by external cadaverine we found a threshold of $K_c = 235\,\mu\mathrm{M}$. This value is surprisingly close to the *in vitro* binding constant of $96\,\mu\mathrm{M}$ for the interaction of cadaverine with the periplasmic domain of CadC [5].

$\mathsf{n_c}$: Hill exponent of direct CadC-cadaverine interaction. It is noteworthy that the best-fit value for the effective Hill coefficient n_c for the regulation of CadC by cadaverine is 2.8, and is relatively well constrained by the data. This suggests that a molecular mechanism for cooperativity is at work, possibly a multimerization of CadC proteins in the membrane. To the best of our knowledge, this question has so far not been studied *in vivo*.

$\mathsf{pH_0}$ and $\mathsf{\Delta pH}$: activation threshold of CadC by low pH. Our analysis yields for the activation threshold of CadC $\mathrm{pH}_0 = 6.2$ and for the width of the activation range $\Delta\mathrm{pH} = 0.5$. These values agree reasonably well with early experiments of Auger *et al.* in Falkow medium, who found activity of a *cadA-lacZ* fusion at pH 6 and below [7]. Also from a physiological point of view it seems reasonable to

activate the Cad system in this pH regime, since the enzyme activity of CadA is peaked around an optimal pH of 5.7 [8].

$\mathsf{C_0/K_c}$: relative CadC abundance. The level of CadC was found to be just sufficient to activate the promoter P_{cad}, such that multiple copies of a $P_{cad} - lacZ$ fusion were not fully induced [2]. This result suggests, that *in vivo* the number of CadC molecules is on the order of the dissociation constant for CadC-promoter binding, i.e., $\mathsf{C_0} \approx \mathsf{K_C}$. In our analysis we found indeed $C_0/K_c = 1.1$, which is in excellent agreement with the observations of Watson *et al.* [2].

ν_m and ν_p: transcription and effective translation rate of *cadBA*. As inferred from Fig. S1, only the product of transcription and effective translation rate is well constrained by our data. From our analysis we found $\nu_m \times \nu_p = 18\,\mathrm{U/mg/min^2}$. We are not aware of any reference values for these parameters in the literature.

f: fold-change between basal and maximal transcription rate of P_{cad}. For the fold-change we found $f \approx 700$, which indicates that the promoter is highly inducible. Again, we are not aware of any reference value in the literature.

τ_m: half-life of *cadBA* mRNA. From our analysis we found an mRNA half-life of 13.8 min. In contrast, a global analysis of RNA half-lifes in *Escherichia coli* [9] found in an extremely short half-life of less than 2 min for the *cadBA* mRNA in LB medium, suggesting that an active degradation mechanism is involved. Our Northern blot data for the *lysP211* mutant, shown in Fig. 4 (f) of the main paper, does indeed suggest a rapid decay of the mRNA at high levels directly after the peak, followed by a slower decay at lower levels. Our quantitative model only allows for a single degradation rate, which leads to the intermediate half-life of 14 min as a best-fit value. The changing degradation rate could be rationalized under the assumption that the *cadBA* mRNA has a relatively weak binding affinity to the degrading enzyme, such that active degradation only contributes significantly at high mRNA levels, whereas at low mRNA levels a slower passive decay is at work. Fig. S2 of the Supplementary Material illustrates that this mechanism could indeed account for the observed shape of the mRNA curve in Fig. 4(f). However, this explanation would raise the question why the rapid active decay is not observed in the data of the wild-type in Fig. 4(a). Possibly, the kinetics of LysP unbinding from CadC is slow (contrary to our model assumption of a rapid binding equilibrium), such that it interferes with the reception of the cadaverine signal and thereby broadens the onset of transcriptional down-regulation. Further experiments would be required to test these hypotheses.

τ_p: protein half-life. For the protein half-life in our effective model we found $\tau_p = 29\,\mathrm{h}$, indicating high protein stability. This finding is in agreement with a biochemical study which found CadA to be stable [8].

v_{max} and K_m: effective turnover rate and Michealis constant for lysine turnover by CadA and CadB. As inferred from Fig. S1, only the ratio of v_{max} and K_m is well constrained by our data: $v_{max}/K_m = 5 \times 10^{-5}\,(\mathrm{min} \times \mathrm{U/mg})^{-1}$. To the best of our knowledge, there is no report of these parameters in the literature.

References

[1] P. S. Popkin and W. K. Maas. *Escherichia coli* regulatory mutation affecting lysine transport and lysine decarboxylase. *J. Bacteriol.*, 141:485–492, 1980.

[2] N. Watson, D. S. Dunyak, E. L. Rosey, J. L. Slonczewski, and E. R. Olson. Identification of elements involved in transcriptional regulation of the *Escherichia coli cad* operon by external pH. *J. Bacteriol.*, 174:530–540, 1992.

[3] M. N. Neely, C. L. Dell, and E. R. Olson. Roles of LysP and CadC in mediating the lysine requirement for acid induction of the *Escherichia coli* cad operon. *J. Bacteriol.*, 176:3278–3285, 1994.

[4] C. L. Dell, M. N. Neely, and E. R. Olson. Altered pH and lysine signalling mutants of *cadC*, a gene encoding a membrane-bound transcriptional activator of the *Escherichia coli cadBA* operon. *Mol. Microbiol.*, 14:7–16, 1994.

[5] L. Tetsch, C. Koller, I. Haneburger, and K. Jung. The membrane-integrated transcriptional activator CadC of *Escherichia coli* senses lysine indirectly via the interaction with the lysine permease LysP. *Mol. Microbiol.*, 67:570–583, 2008.

[6] B. P. Rosen. Basic amino acid transport in *Escherichia coli*. *J. Biol. Chem.*, 246:3653–3662, 1971.

[7] E. A. Auger, K. E. Redding, T. Plumb, L. C. Childs, S.-Y. Meng, and G. N. Bennett. Construction of *lac* fusions to the inducible arginine-and lysine decarboxylase genes of *Escherichia coli* K-12. *Mol. Microbiol.*, 3:609–620, 1989.

[8] D. L. Sabo, E.A. Boeker, B. Byers, H. Waron, and E. H. Fischer. Purification and physical properties of inducible *Escherichia coli* lysine decarboxylase. *Biochemistry*, 13:662–670, 1974.

[9] D.W. Selinger and C. Rosenow. Global RNA half-life analysis in *Escherichia coli* reveals positional patterns of transcript degradation. *Genome Res.*, 13:216–223, 2003.

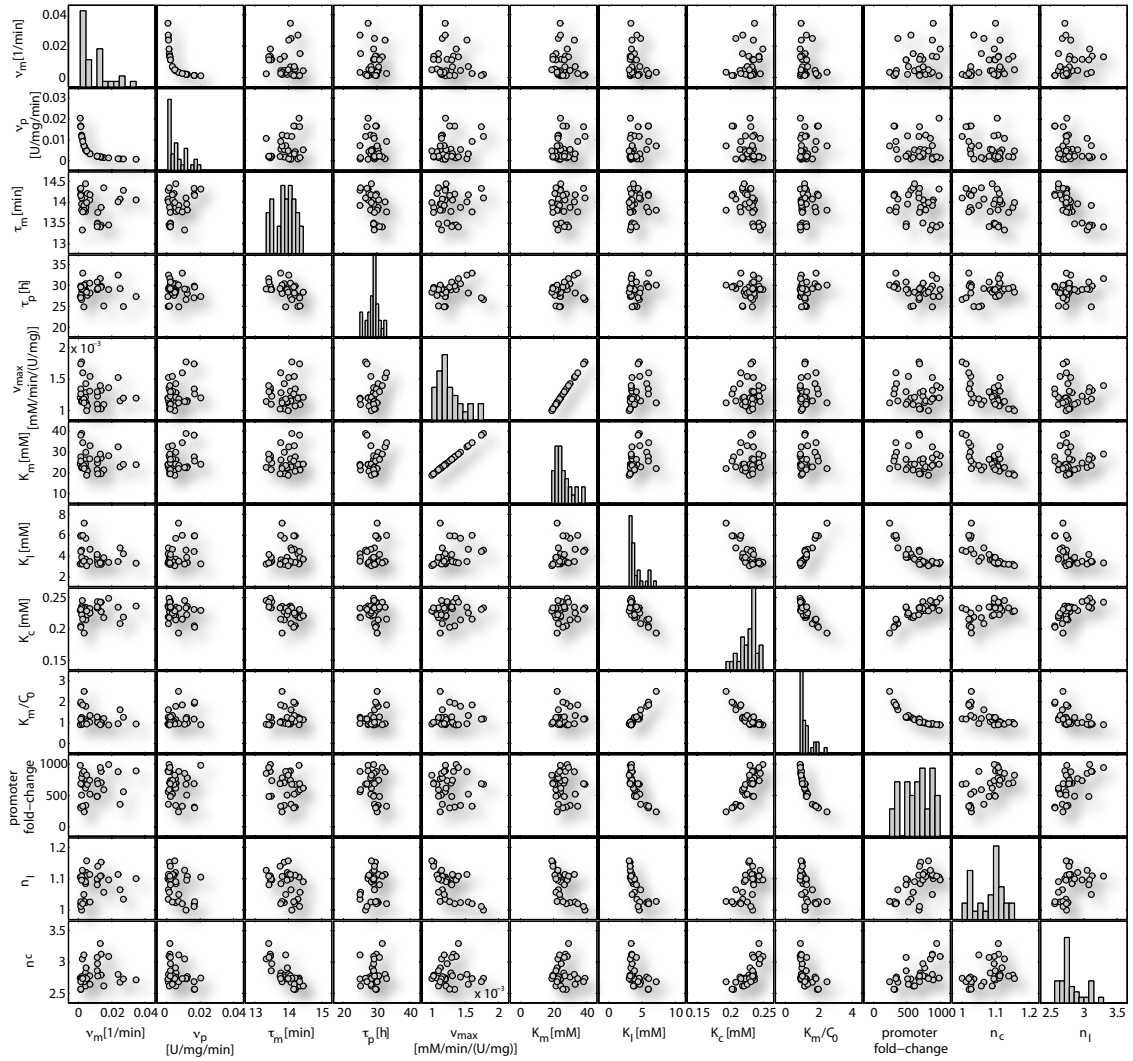

Fig. S1. Pairwise correlations between the fit parameters. For this correlation analysis, the fits of highest quality ($\chi^2 < 3.22$) were selected from Fig. 5 (main paper) and the correlations of the resulting parameters are displayed in pairwise scatterplots (*off-diagonal*). The figure also shows the histograms of the individual parameters (*diagonal*).

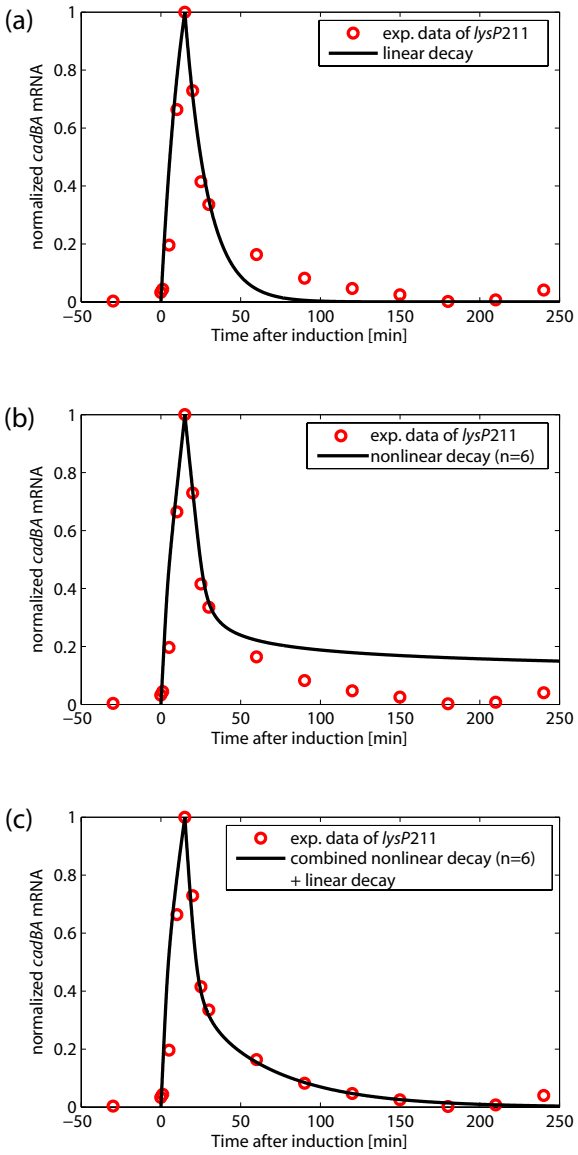

Fig. S2. Schematic influence of different mRNA degradation mechanisms on the dynamics of *cadBA* mRNA. The solid lines correspond to a model which assumes mRNA (m) production at constant rate α until $T = 15\,\text{min}$, followed by a period of pure mRNA decay: $\frac{d}{dt}m = \alpha\Theta(T-t) - g(m)$, where $\Theta(T-t)$ is the Heaviside function and $g(m)$ is the degradation term. In (a) the mRNA dynamics with simple linear decay ($g(m) = \lambda m$) shows major differences to the experimental data of the *lysP211* mutant after $t = 60\,\text{min}$. In (b) the mRNA dynamics is shown for a degradation mechanism which is active when the mRNA level exceeds a threshold value K, i.e. $g(m) = \lambda m^n/(K^n + m^n)$. The Hill exponent was chosen to be large ($n = 6$) in order to obtain a switch-like dependence of the mRNA degradation rate on the mRNA level: below $K = 0.4$ degradation turns off and hence the long-term behavior of the experimental data ($t > 60$ min) is not captured. Only a combination of nonlinear active decay and slow linear decay in (c) yields the required short *and* long-term behavior.

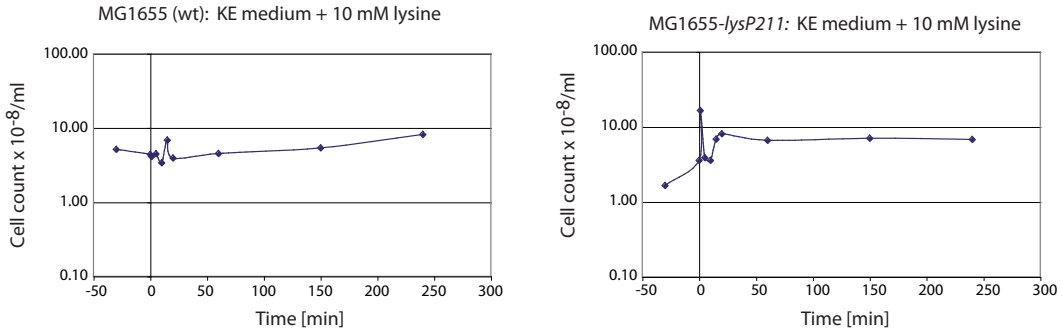

Fig. S3. Dynamics of the life cell counts of *E. coli* MG1655 (*left*) and MG1655-*lysP211* (*right*) upon exposure to pH 5.8 and 10 mM lysine in minimal KE medium.

9 Supplemental Material - chapter 4

Supplementary Data –

NEW INSIGHTS INTO THE SIGNALING MECHANISM OF THE PH-RESPONSIVE, MEMBRANE-INTEGRATED TRANSCRIPTIONAL ACTIVATOR CADC OF *ESCHERICHIA COLI* *

Ina Haneburger[1,2], Andreas Eichinger[1,3], Arne Skerra[1,3] and Kirsten Jung[1,2]

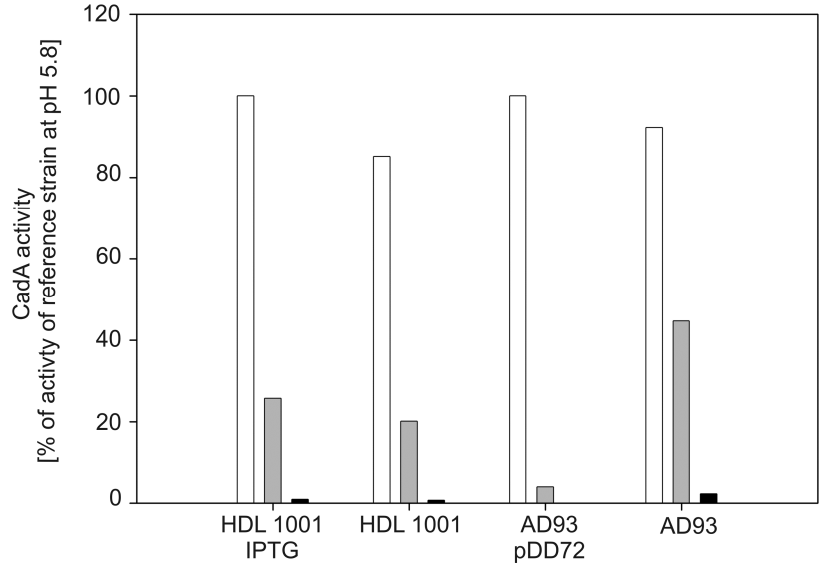

Fig. S1. CadA activity in *E. coli* phospholipid mutants at different external pH values. CadA activity was determined as a measure for *cadBA* expression in the strains indicated. *E. coli* strain HDL1001 has a lower content (< 2%) of phosphatidylglycerol (PG) and cardiolipin (CL), that can be rescued to *E. coli* K12 level through expression of phosphatiylglycerophosphate synthase (*pgsA*) by induction with IPTG (1). *E. coli* strain AD93 is deficient for phosphatidylethanolamine and therefore contains much more phosphatidylglycerol and cardiolipin. Expression of the *pss* gene from a plasmid (plasmid pDD72) restores a normal PG and CL content. Strains were grown in buffered LB medium at pH 5.8 (white bars), pH 7.0 (grey) and pH 7.6 (black). To support growth of strain AD93, these cultures were supplemented with 50 mM $MgCl_2$. CadA activity was determined as described in *Experimental procedures*. 100 % corresponds to the activity measured for HDL1001/IPTG and AD93/pDD72 after growth at pH 5.8.

supplementary Tab. 1: Plasmids and oligonucleotides used in this study.

Plasmid	Oligonucleotide	Oligonucleotide Sequence (5'- 3')	Second Oligonucleotide	RE	Source or Reference
pET16b					Novagen
pET16b-cadC2					Tetsch *et al*., 2008
pET16b-cadC3	CadC-XhoI-sense	GGCATTGATCTCGAGATGTCCTGG	CadC_BamHI_stop_anti	BamHI/SpeI	this work
	CadC-XhoI-anti	GCCAGGACATCTCGAGATCAATGC	5?_CadC_SOE3_SpeI		
pET16b-cadC4	CadC-XmaI_sense	ATTTACGCCCCGGGGCAAACACC	CadC_BamHI_stop_anti	BamHI/XhoI	this work
	CadC-XmaI_anti	TTTGCCCCGGGGCGTAAATTAAAGG	CadC_XhoI_sense		
pET16b-cadC5	CadC_SacI_sense	CATCGTGGAGCTCTATTAACTAATTTTTATC	CadC-XhoI-anti	XhoI/SpeI	this work
	Cad_SacI_anti	TAGTTAATAGAGCTCCACGATGC	5?_CadC_SOE3_SpeI		
pET16b-cadC6		combination of pET16b-cadC4 and pET16b-cadC5		XhoI/BlpI	this work
pET16b-cadC7	CadC_SacII_long	TGTGTAGCACTAGTAGCGTTTTCAAGTCTTGATACACGTCTTCCTATGAGCAAATCGC GTATTTTGCTCAATCCGCGGGGATATTGAC	CadC-XhoI-anti	XhoI/SpeI	this work
pET16b-cadC7_D198A	CadC_D198A	TGCTCAATCCGCGGGCTATTGACA	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC7_D198E	CadC_D198E	TGCTCAATCCGCGGGAAATTGACAT	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC7_D198N	CadC_D198N	TGCTCAATCCGCGGAATATTGAC	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC7_D200A	CadC_D200A	TGCTCAATCCGCGGGATATTGCCATTAATATG	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC7_D200E	CadC_D200E	TGCTCAATCCGCGGGATATTGAAATTAATATG	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC7_D200N	CadC_D200N	TGCTCAATCCGCGGGATATTAACATTAATATG	CadC-XhoI-anti	XhoI/SacII	this work
pET16b-cadC5_H240L	CadC_H240L_sense	TTTATGGTGCTTGACAAAATCAAC	Cad_SacI_anti	SacI/SpeI	this work
	CadC_H240L_anti	TGATTTTGTCAAGCACCATAAAGG	5?_CadC_SOE3_SpeI		
pET16b-cadC5_H240Q	CadC_H240Q_sense	TTTATGGTGCAAGACAAAATCAAC	Cad_SacI_anti	SacI/SpeI	this work
	CadC_H240Q_anti	TGATTTTGTCTTGCACCATAAAGG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_H240R	CadC_H240R_sense	TTTATGGTGAGAGACAAAATCAAC	CadC-XhoI-anti	XhoI/SpeI	this work
	CadC_H240R_anti	TGATTTTGTCTCTCACCATAAAGG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_K242A	CadC_K242A_fw	GTGCATGACGCAATCAACTAC	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_K242A_rev	GTAGTTGATTGCGTCATGCAC	5?_CadC_SOE3_SpeI		
pET16b-cadC6_K242Q	CadC_K242Q_fw	GTGCATGACCAAATCAAC	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_K242Q_rev	GTTGATTTGGTCATGCAC	5?_CadC_SOE3_SpeI		
pET16b-cadC6_K242R	CadC_K242R_fw	GTGCATGACCGAATCAACTAC	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_K242R_rev	GTAGTTGATTCGGTCATGCAC	5?_CadC_SOE3_SpeI		
pET16b-cadC6_Y245I	CadC_Y245I_fw	CAAAATCAACATCAACATTG	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_Y245I_rev	CAATGTTGATGTTGATTTTG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_E249A	CadC_E249A_fw	CATTGATGCACCGAGCAG	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_E249A_rev	CTGCTCGGTGCATCAATG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_E249D	CadC_E249D_fw	CATTGATGATCCGAGCAG	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_E249D_rev	CTGCTCGGATCATCAATG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_E249Q	CadC_E249Q_fw	CAACATTGATCAACCGAGC	CadC-XhoI-anti	XhoI/SacII	this work
	CadC_E249Q_rev	GCTCGGTTGATCAATGTTG	5?_CadC_SOE3_SpeI		
pET16b-cadC5_H332L	CadC_H332L	GTTAATAGAGCTCCACGAAGCGGCAAAATTTTCTGG	5?_CadC_SOE3_SpeI	SacI/SpeI	this work
pET16b-cadC5_H332Q	CadC_H332Q	GTTAATAGAGCTCCACGTTGCGGCAAAATTTTCTGG	5?_CadC_SOE3_SpeI	SacI/SpeI	this work
pET16b-cadC5_H344L	CadC_H344L	ATCGTGGAGCTCTATTAACTAATTTTTATCAGGCACTTGATTATTTAC	CadC_BamHI_stop_anti	SacI/BamHI	this work
pET16b-cadC6_H344Q	CadC_H344Q	ATCGTGGAGCTCTATTAACTAATTTTTATCAGGCACAAGATTATTTAC	CadC_BamHI_stop_anti	SacI/BamHI	this work
pET16b-cadC6_H349L	CadC_H349L	ATCGTGGAGCTCTATTAACTAATTTTTATCAGGCACATGATTATTTACTGCTTGGCGATG	CadC_BamHI_stop_anti	SacI/BamHI	this work
pET16b-cadC6_H349Q	CadC_H349Q	ATCGTGGAGCTCTATTAACTAATTTTTATCAGGCACATGATTATTTACTGCAAGGCGATG	CadC_BamHI_stop_anti	SacI/BamHI	this work
pET16b-cadC5_H387L	CadC_H387L	TAGTTGATATCGTGCGCCTTTCTCAAC	CadC_BamHI_stop_anti	BamHI/EcoRV	this work
pET16b-cadC5_H387Q	CadC_H387Q	TAGTTGATATCGTGCGCCAATCTCAAC	CadC_BamHI_stop_anti	BamHI/EcoRV	this work
pET16b-cadC5_H390L	CadC_H390L	TAGTTGATATCGTGCGCCATTCTCAACTTCCTTTAG	CadC_BamHI_stop_anti	BamHI/EcoRV	this work
pET16b-cadC5_H390Q	CadC_H390Q	TAGTTGATATCGTGCGCCATTCTCAACAACCTTTAG	CadC_BamHI_stop_anti	BamHI/EcoRV	this work
pET16b-cadC3_D434N	CadC-D434N_as	GGACATCTCGAGATCAATGCCAGTATTTATCGCCTGGTAAGACTCATTTGTTTTACC	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC3_D434R	CadC_D434R_as	GGACATCTCGAGATCAATGCCAGTATTTATCGCCTGGTAAGACTCACGTGTTTTACC	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC5_E435Q	CadC_E435Q_as_wobble	GGACATCTCGAGATCAATGCCAGTATTTATCGCTTGGTAAGACTGATCTGTTTTACC	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC3_E435R	CadC_E435R_as	GGACATCTCGAGATCAATGCCAGTATTTATCGCCTGGTAAGACCGATCTGTTTTACC	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC3_D445N	CadC_D445N_as	AGGACATCTCGAGATTAATGCCAG	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC3_D445R	CadC_D445R_as_2	AGGACATCTCGAGTCTAATGCCAG	5?_CadC_SOE3_SpeI	XhoI/SpeI	this work
pET16b-cadC2_E447Q	CadC_E447Q_sense	CATTGATCTTCAAATGTCCTGGC	CadC_BamHI_stop_anti	BamHI/SpeI	this work
	CadC_E447Q_anti	GCCAGGACATTTGAAGATCAATG	5?_CadC_SOE3_SpeI		
pET16b-cadC2_E447R	CadC_E447R_sense	CATTGATCTTCGAATGTCCTGGC	CadC_BamHI_stop_anti	BamHI/SpeI	this work
	CadC_E447R_anti	GCCAGGACATTCGAAGATCAATG	5?_CadC_SOE3_SpeI		
pET16b-cadC6_Y453I	CadC_Y453I	GCATTGATCTCGAGATGTCCTGGCTAAATATTGTGTTGC	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_V454A	CadC_V454A	GCATTGATCTCGAGATGTCCTGGCTAAATTATGCGTTGCTTG	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_K458A	CadC_K458A	GCATTGATCTCGAGATGTCCTGGCTAAATTATGTGTTGCTTGGCGCGGTTTATGAAATG	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_K458Q	CadC_K458Q	GCATTGATCTCGAGATGTCCTGGCTAAATTATGTGTTGCTTGGCCAAGTTTATGAAATG	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_K458R	CadC_K458R	CATTTCATAAACACGGCCAAGCAACACATAATTTAGCCAGGACATCTCGAGATCAATGC	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_Y460I	CadC_Y460I	GGTGAGATATGCATCAGCTGCTTCCCGGTTCATCCCCTTCATTTCAATAACCTTG	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC6_E461A	CadC_E461A	GGTGAGATATGCATCAGCTGCTTCCCGGTTCATCCCCTTCATTGCATAAACC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC3_E461D	CadC_E461D_2	TTGATCTCGAGATGTCCTGGCTAAATTATGTGTTGCTTGGCAAGGTTTATGATATGAAGG	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC3_E461Q	CadC_E461Q_s	TTGATCTCGAGATGTCCTGGCTAAATTATGTGTTGCTTGGCAAGGTTTATCAAATGAAGG	CadC_BamHI_stop_anti	BamHI/XhoI	this work
pET16b-cadC4_E461R	CadC_E461R_s_2	TTGATCTCGAGATGTCCTGGCTAAATTATGTGTTGCTTGGCAAGGTTTATAGAATGAAGG	CadC_BamHI_stop_anti	BamHI/XhoI	this work
pET16b-cadC6_R467A	CadC_R467A	TGAGATATGCATCAGCTGCTTCCGCGTTCATCC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC6_R467E	CadC_R467E	TGAGATATGCATCAGCTGCTTCCTCGTTCATCC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC6_R467K	CadC_R476K	GGTGAGATATGCATCAGCTGCTTCCTTGTTCATCCCC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC6_R467Q	CadC_R467Q	TGAGATATGCATCAGCTGCTTCCTCGTTCATC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC6_E468A	CadC_E468A	TGAGATATGCATCAGCTGCTGCCCGGTTC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC5_E468D	CadC_E468D	TGAGATATGCATCAGCTGCGTCCCGGTTC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC3_E468Q	CadC_E468Q_as	TGAGATATGCATCAGCTGCTTGCCGGTTC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC2_E468R	CadC_E468R_as	TGAGATATGCATCAGCTGCTCGCCGGTTC	5?_CadC_SOE3_SpeI	NsiI/SpeI	this work
pET16b-cadC2_D471A	CadC_D471A	GTGTTTGCCCCGGGGCGTAAATTAAAGGCGGTGAGATATGCTGCAGCTGCTTC	5?_CadC_SOE3_SpeI	XmaI/SpeI	this work
pET16b-cadC2_D471E	CadC_D471E_sense	GGAAGCAGCTGAAGCATATCTCCAC	CadC_BamHI_stop_anti	BamHI/SpeI	this work
	CadC_D471E-anti	GTGAGATATGCTTCAGCTGGCTTCC	5?_CadC_SOE3_SpeI		
pET16b-cadC4_D471N	CadC_D471N_as	GTTTGCCCCGGGGCGTAAATTAAAGGCGGTGAGATATGCATTAGCTGC	5?_CadC_SOE3_SpeI	XmaI/SpeI	this work
pET16b-cadC4_D471R	CadC_D471R_as	GTTTGCCCCGGGGCGTAAATTAAAGGCGGTGAGATATGCACGAGCTGC	5?_CadC_SOE3_SpeI	XmaI/SpeI	this work
pET16b-cadC6_L474A	CadC_L474A	AGCAGCTGATGCATATGCCACCGCC	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_F477A	CadC_F477A	AGCAGCTGATGCATATCTCACCGCCGCTAATTTAC	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_F477I	CadC_F477I	AGCAGCTGATGCATATCTCACCGCCATTAATTTAC	T7 Terminator_as	BlpI/XhoI	this work
pET16b-cadC6_N478A	CadC_N478A	GTGTTTGCCCCGGGGCGTAAAGCAAAGGCGG	5?_CadC_SOE3_SpeI	XmaI/SpeI	this work
pET16b-cadC6_R480Q	CadC_R480Q	GTGTTTGCCCCGGGTTGTAAATTAAAGG	5?_CadC_SOE3_SpeI	XmaI/SpeI	this work
pET16b-cadC4_E490Q	CadC_E490Q_s	TTACGCCCCGGGGCAAACACCCTTTACTGGATTCAAAATGGTATATTCC	CadC_BamHI_stop_anti	BamHI/XmaI	this work
pET16b-cadC4_E490R	CadC_E490R_s	TTACGCCCCGGGGCAAACACCCTTTACTGGATTCGAAATGGTATATTCC	CadC_BamHI_stop_anti	BamHI/XmaI	this work
pET16b-cadC3_D506N	CadC_D506N_as_2	GCCGGATCCTTATTCTGAAGCAAGAAATTTATTGAGATAAGG	CadC_XhoI_sense	BamHI/XhoI	this work
pET16b-cadC3_D506R	CadC_D506R_as_2	GCCGGATCCTTATTCTGAAGCAAGAAATTTCCTGAGATAAGG	CadC_XhoI_sense	BamHI/XhoI	this work
pET16b-cadC3_E512Q	CadC_E512Q_as	GCCGGATCCTTATTGTGAAGCAAG	CadC_XhoI_sense	BamHI/XhoI	this work
pET16b-cadC3_E512R	CadC_E512R_as	GCCGGATCCTTATCGTGAAGCAAG	CadC_XhoI_sense	BamHI/XhoI	this work
pET16b-cadC_D198E_D471E		combination of pET16b-cadC7_D198E and pET16b-cadC2_D471E,		BssHII	this work
pET16b-cadC_D198E_D471N		combination of pET16b-cadC7_D198E and pET16b-cadC4_D471N,		BssHII	this work
pET16b-cadC_K242R_D471E		combination of pET16b-cadC6_K242R and pET16b-cadC2_D471E,		BssHII	this work
pET16b-cadC_K242R_D471N		combination of pET16b-cadC6_K242R and pET16b-cadC4_D471N,		BssHII	this work
	5?_CadC_SOE3_SpeI	CTGTGTAGCACTAGTAGCGTTTTC			Tetsch *et al*., 2008
	CadC_BamHI_stop_antisense	AGTGGATCCTTATTCTGAAGCAAGAAATTTGTCG			Kueper and Jung, 2005
	T7 Promoter Primer	TAATACGACTCACTATAGGG			Invitrogen
	T7 Terminator antisense	CCGTTTAGAGGCCCCAAGG			Invitrogen

10 Supplemental Material - chapter 5

Supplemental Data

The feedback-inhibitor cadaverine suppresses the pH response by binding to the pH susceptible site of CadC in *Escherichia coli*

Ina Haneburger‡, Georg Fritz‡, Nicole Jurkschat, Larissa Tetsch, Andreas Eichinger, Arne Skerra, Ulrich Gerland and Kirsten Jung

‡ Both authors contributed equally to this work

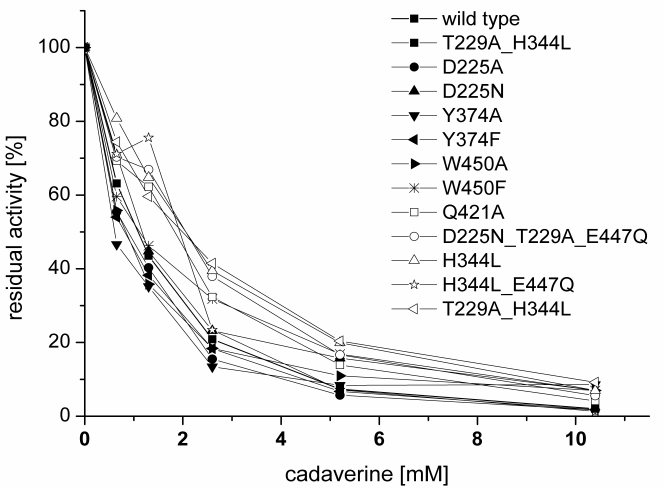

FIGURE S1. Inhibitory effect of cadaverine on *cadBA expression.* Reporter gene assays were performed with *E. coli* EP314 (*cadC1*::*Tn10, cadA'*::*lacZ*) that was complemented with plasmid-encoded CadC (wild type) or the indicated CadC variants. An overnight culture (pH 7.6) was shifted to pH 5.8 with the indicated cadaverine concentrations (0.65; 1.3; 2.6; 5.2; 10.4 mM) and harvested in the mid-logarithmic growth phase. ß-galactosidase activity was measured and the percentage of residual activity was calculated in relation to the same condition without cadaverine. Results are given as mean from at least three independent experiments. For clarity, the corresponding standard deviations are omitted.

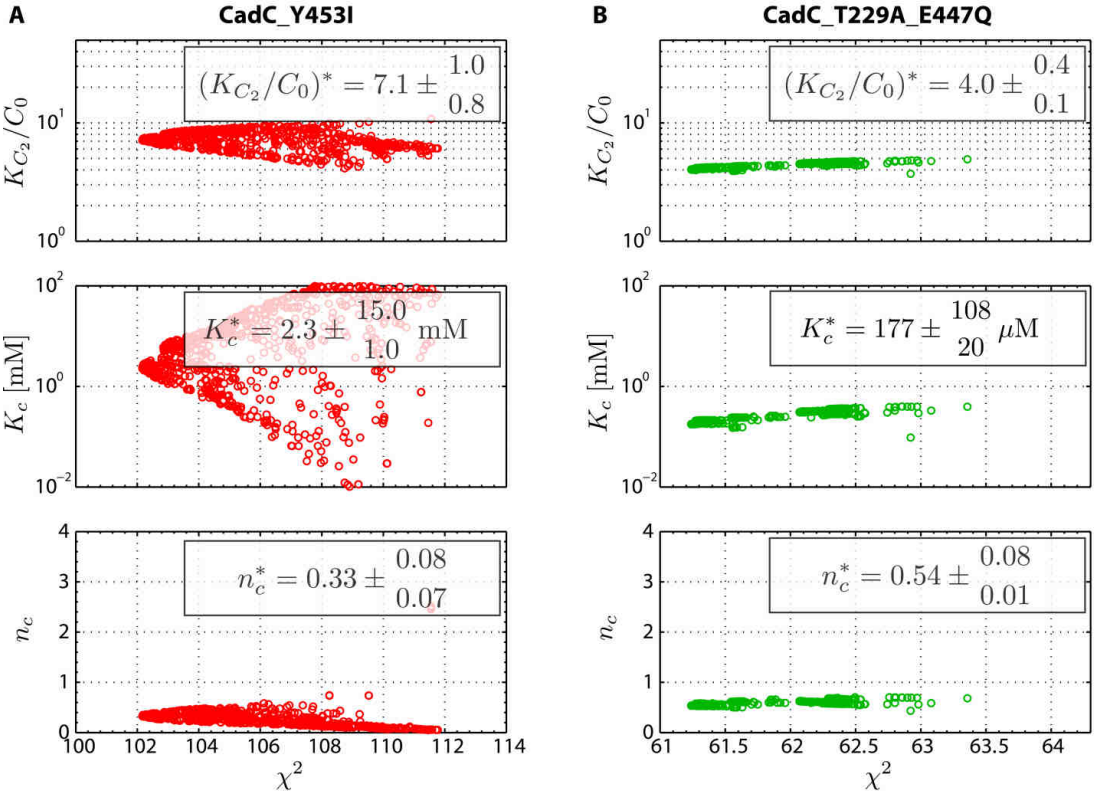

FIGURE S2. Correlations between the goodness of fit and the estimated parameters for (A) the strain harboring CadC_Y453I and (B) the strain harboring CadC_T229A_E447Q. The points correspond to local optima in the parameter space, for which the difference between the quantitative model and the experimental data in Fig. 4 of the main text is minimized (see Materials and Methods). As the fit quality increases (lower χ^2), most parameters are confined to narrow intervals, indicating that their values are well constrained by the experimental data. However, some parameters display significant variation even for the lowest χ^2 values, and we find from parameter-parameter correlation analysis in Fig. S3, that only combinations of those are well confined by our data. In all cases the y-axis corresponds to the full range of allowed parameter values.

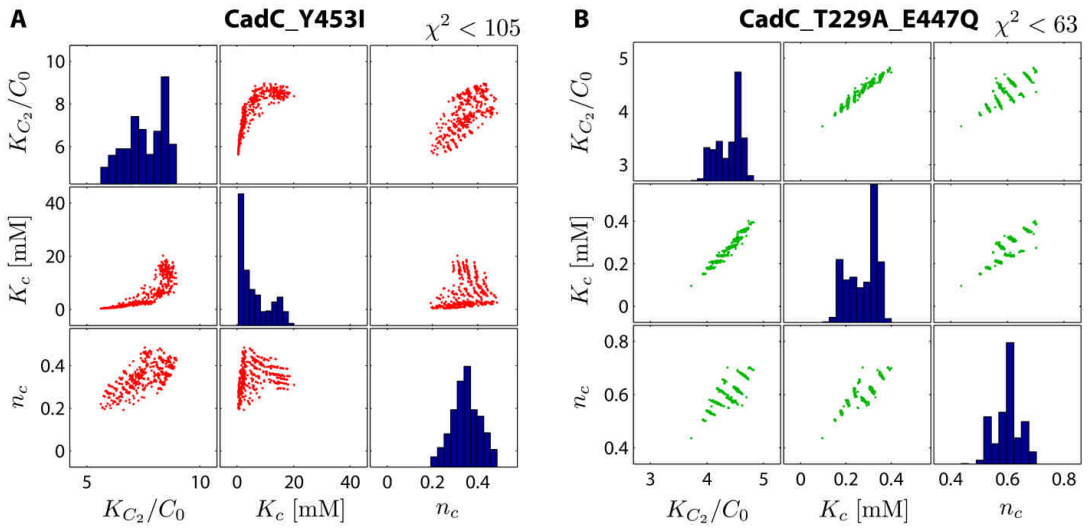

FIGURE S3. Parameter-parameter correlations between the estimated parameters for (A) strain harboring CadC_Y453I and (B) strain harboring CadC_T229A_E447Q.

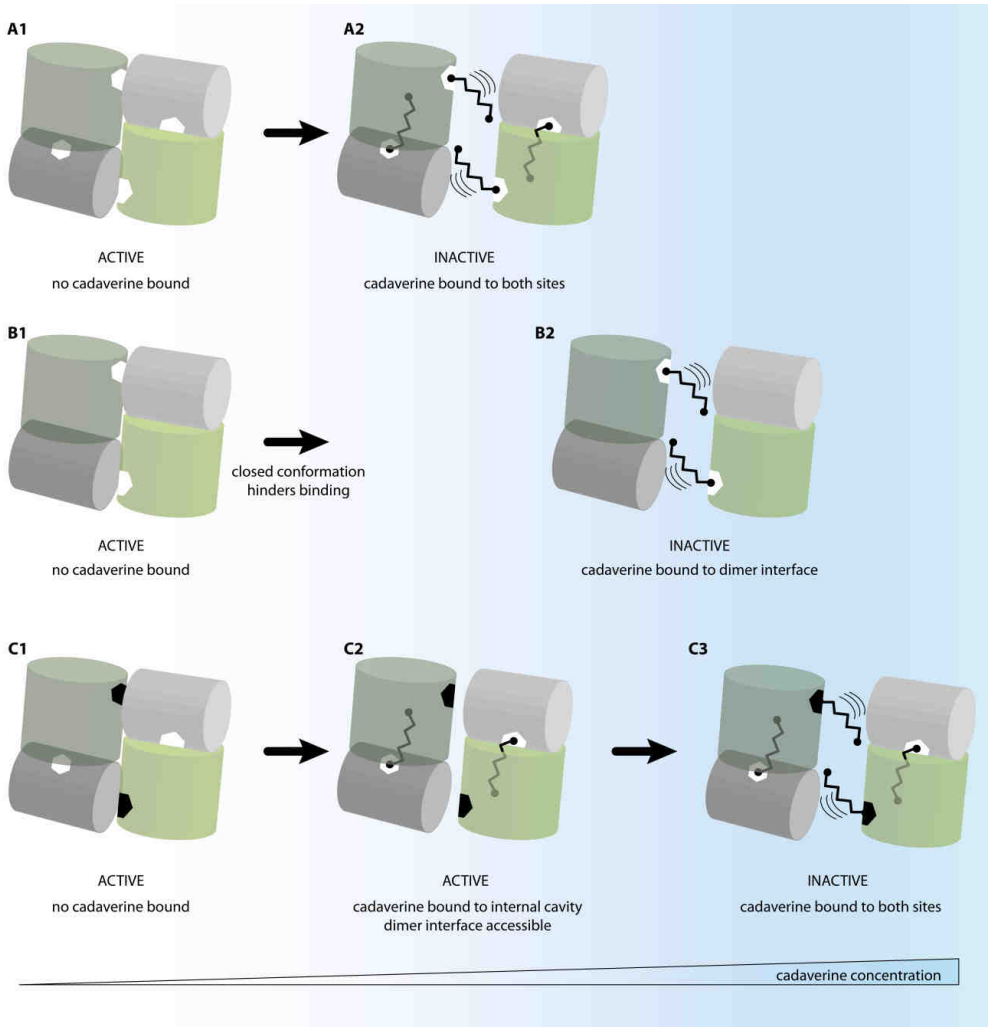

FIGURE S4. Model for cadaverine-dependent regulation of *cadBA* expression. An Increase in cadaverine concentration is indicated by blue shading. CadC dimers of the periplasmic domain are represented schematically as the two monomers (light and dark colors) in top view. The monomers consist of two subdomains represented by a grey and a green cylinder (cf. Fig. 3). Cadaverine binding sites are indicated as hexagons (white: wild type, black: binding impaired due to substitution). Cadaverine is depicted as molecule with five carbon atoms and two amine groups (spheres). A. Inactivation cycle as proposed for the wild type (cf. Fig. 6). B. Effect of substitutions in the central cavity (CadC_T229A_E447Q). Cadaverine binding to the internal cavity is significantly reduced. Therefore, initial conformational changes that would lead to exposure of the binding sites at the dimerization interface are not induced. Nonetheless they may occur at a lower rate simply due to the flexibility of proteins. In consequence, cadaverine-binding and -dependent inhibition occur at higher cadaverine concentration (B2). C. Substitutions that impair binding of cadaverine to the sites at the dimer interface (indicated by the black hexagons at the interface) would lead to a significant decrease in cadaverine-binding, in turn (nearly) prevent CadC inactivation or shift it to even higher cadaverine concentrations (C3). The initial binding step to the binding sites within the central cavities would occur at the same concentration as in the wild type (C2).

TABLE S1. Plasmids and oligonucleotides used in this study. If not indicated pET16b-cadC7 (1) was used as template.

Plasmid	Oligonucleotide	5' - 3' Sequence	Source/Reference
pET16b			Novagen
pET16b-cadC7			(1)
pET16b-cadC_D225A	CadC_D225A_sense CadC_D225A_anti	GGCGTGGGTGCTTTGGTGGCGACATC GATGTCGCCACCAAAGCACCCACGCC	This work
pET16b-cadC_D225N	CadC_D225N_sense CadC_D225N_anti	GGCGTGGGTAATTTGGTGGCGACATC GATGTCGCCACCAAATTACCCACGCC	This work
pET16b-cadC_D225W	CadC_D225W_sense CadC_D225W_as	ATAGGCGTGGGTTGGTTGGTGGCGAC GTCGCCACCAACCAACCCACGCCTAT	This work
pET16b-cadC_T229A	CadC_T229A_sense CadC_T229A_as	TTTGGTGGCGGCATCACTTAAC GTTAAGTGATGCCGCCACCAAA	This work
pET16b-cadC_H344L			(1)
pET16b-cadC_Y374A	CadC_Y374A_sense CadC_Y374A_anti	CCAGAATTTACCGCCGCGAGAGCAG This work CTGCTCTCGCGGCGGTAAATTCTGG	
pET16b-cadC_Y374F	CadC_Y374F_sense CadC_Y374F_anti	CCAGAATTTACCTTCGCGAGAGCAG CTGCTCTCGCGAAGGTAAATTCTGG	This work
pET16b-cadC_Q421A	CadC_Q421A_sense CadC_Q421A_as	CATTATATATGCAATAAAAGCGG CCGCTTTTATTGCATCACTTAAC	This work
pET16b-cadC_E447Q			(1)
pET16b-cadC_W450A	CadC_W450A_anti	CGGTGAGATATGCATCAGCTGCTTC CCGGTTCATCCCCTTCATTTCATAAA CCTTGCCAAGCAACACATAATTTAGC GCGGACATTTCAAGATC	This work
pET16b-cadC_W450F	CadC_W450F_anti	CGGTGAGATATGCATCAGCTGCTTC CCGGTTCATCCCCTTCATTTCATAAA VVTTGCCAAGCAACACATAATTTAGG AAGGACATTTCAAGATC	This work
pET16b-cadC_T229A_E447Q	combination of pET16b-cadC_T229A and pET16b-cadC_E447Q by NcoI and SacI digestion		This work
pET16b-cadC_H344L_E447Q	PCR with with oligonucleotide CadC_H334L (1) and pET16b-cadC_E447Q as template		This work
pET16b-cadC_T229A_H344L	combination of pET16b-cadC_T229A and pET16b-cadC_H344L by BamHI and SacI digestion		This work
pET16b-cadC_D225N_E447Q	PCR with oligonucleotides CadC_D225N_sense and CadC_D225N_anti with pET16b-cadC_E447Q as template		This work
pET16b-cadC_T229A_Q421A_E447Q	PCR with oligonucleotides CadC_D225N_sense and CadC_Q421A_as and pET16b-cadC_T229A_E447Q as template		This work
pET16b-cadC_D225N_T229A_E447Q	PCR with oligonucleotides CadC_D225N_sense and CadC_Q421A_as and pET16b-cadC_T229A_E447Q as template		This work
pET16b-cadC_T229A_H344L_E447Q	combination of pET16b-cadC_T229A and pET16b-cadC_H344L_E447Q by BamHI and SacI digestion		This work
pET16b-cadC_Y453A	CadC_Y453A	GCATTGATCTCGAGATGTCCTGGCTAAA TGCTGTGTTGC	This work
pET16b-cadC_Y453I			(1)
pET16b-cadC_Y453F	CadC_Y453F	GCATTGATCTCGAGATGTCCTGGCTAAA TTTTGTGTTGC	This work
pET16b-cadC_T475A	CadC_T475A_sense CadC_T475A_anti	CATATCTCGCCGCCTTTAATTTACGCCCC GGGGCAAAC GTTTGCCCCGGGGCGTAAATTAAAGGCG GCGAGATATG	This work
pET16b-cadC_T475S	CadC_T475S	GTGTTTGCCCCGGGGCGTAAATTAAAGG CGGAGAGATATGC	This work
pET16b-cadC_H240L			(1)
pET16b-cadC_E468D			(1)
pET16b-cadC_D471A			(1)
pET16b-cadC_D471N			(1)
pET16b-cadC_L474A			(1)
pET16b-cadC_N478A			(1)
pET16b-cadC_F477A			(1)
pET16b-cadC_F477I			(1)
pET32a			Novagen
pET32a-cadC188–512	his 6 -cadC 188–512 in pET32a		(2)
pET32a-cadC188–512_Y453I	his 6 -cadC 188–512_Y453I in pET32a		This work
pET32a-cadC188–512_T475A	his 6 -cadC 188–512_T475A in pET32a		This work
pET32a-cadC188–512_T229A_E447Q	his 6 -cadC 188–512_T229A_E447Q in pET32a		This work

TABLE S2. Parameters of the quantitative model for the wild type Cad module. LB indicates lower bound; UB, upper bound. The estimated parameter values are shown as (best-fit value) $^{+\sigma^+}_{-\sigma^-}$, where σ^+ and σ^- indicate the asymmetric standard errors in the positive direction and in the negative direction, respectively, see (3) for details.

	Parameter	LB	UB	Estimated Value
	Sensory module			
K_I	Threshold for CadC activation by lysine	1	20	$4.8 \pm {}^{3.4}_{0.7}$ mM
K_c	Threshold for CadC inactivation by cadaverine	50	1000	$172 \pm {}^{18}_{20}$ μM
n_I	Hill coefficient for CadC regulation by lysine	1	5	$1.6 \pm {}^{0.1}_{0.2}$
n_c	Hill coefficient for CadC regulation by cadaverine	1	5	$3.5 \pm {}^{0.6}_{0.6}$
pH_0	pH threshold for CadC activation	-	-	6.2
ΔpH	width of the transition from active to inactive CadC	-	-	0.5
	Expression module			
C_0/K_C	Total CadC per cell in relation to the threshold for CadC-promoter binding	0.1	10	$0.19 \pm {}^{1.20}_{0.06}$
ν_m	basal transcription rate	0.001	0.1	$2.4 \pm {}^{2.7}_{1.7} \times 10^{-2}$ min^{-1}
f	fold-change between basal and maximal transcription rate	10	1000	$80 \pm {}^{89}_{42}$
ν_p	effective translation rate	10^{-4}	10^{-1}	$7 \pm {}^{62}_{3} \times 10^{-4}$ U/min
τ_m	mRNA half-life ($= \ln 2/\lambda_m$)	1	50	$14.2 \pm {}^{1.6}_{0.7}$ min
τ_p	protein half-life ($= \ln 2/\lambda_p$)	1	10^4	$35 \pm {}^{118}_{10}$ h
v_{max}	maximal rate for lysine turnover via CadA and CadB	10^{-4}	10	$1.1 \pm {}^{1.5}_{0.3} \times 10^{-3}$ min^{-1}
K_m	effective Michaelis constant for lysine turnover via CadA and CadB	1	100	$20 \pm {}^{38}_{7}$ mM

REFERENCES

1. Haneburger, I., Eichinger, A., Skerra, A., and Jung, K. (2011) *J Biol Chem* **286**, 10681-10689
2. Tetsch, L., Koller, C., Haneburger, I., and Jung, K. (2008) *Molecular microbiology* **67**, 570-583
3. Fritz, G., Koller, C., Burdack, K., Tetsch, L., Haneburger, I., Jung, K., and Gerland, U. (2009) *J Mol Biol* **393**, 272-286

11 Supplemental Material - chapter 6

Supplemental Material

Identification of ArgP and Lrp as transcriptional regulators of *lysP*, the gene encoding the specific lysine permease of *Escherichia coli*

Jimena Ruiz[1,2], Ina Haneburger[1] and Kirsten Jung[1*]

[1]Ludwig-Maximilians-Universität München, Munich Center for integrated Protein Science (CiPSM) at the Department of Biology I, Microbiology, Großhaderner Straße 2-4, 82152 Martinsried, Germany

[2]current address: Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Intendente Guiraldes 2160 (1427), Ciudad Autónoma de Buenos Aires, Argentina.

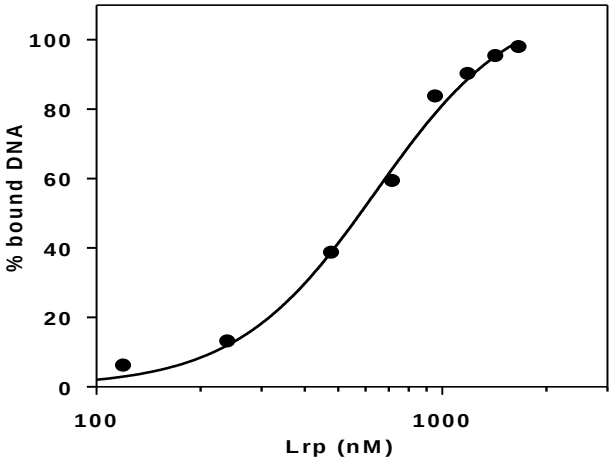

Figure S1. Binding profile obtained after the quantification of free DNA and Lrp-bound DNA in Figure 10C fitted using the Hill equation.

Acknowledgment

I am heartily thankful to my supervisor, Prof. Dr. Kirsten Jung, whose encouragement, supervision and support from the first until the last day enabled me to develop this thesis. Due to the many conferences I was allowed to attend, I got to know the international micro-biological world, opening my eyes still wider. Especially interesting was the chance to work interdisciplinary and to learn more about connected science fields.
With this I would like to thank my collaborators Prof. Dr. Arne Skerra, Andreas Eichinger from Weihenstephan and especially Joseph Danzer for his patience and expertise. As well, I would like to thank Georg Fritz and Ulrich Gerland from the physics department for fruitful discussions and the close collaboration.
Many thanks go to my thesis committee, especially to Prof. Dr. Dirk Schüler for accepting to be the second examiner.
I am extremely grateful to the Elitenetzwerk Bayern and the Center of Integrated Protein Science Munich for the support. Beside the financial help, I could attend their extraordinary soft skill seminars.
Last but not least I offer my blessings to all of those who supported me in many respects during the course of this project: Korinna for excellent help (words cannot express my sincere gratitude and the respect I owe her for her expertise), Jimena for being a mentor, Poldi for lively discussion and all three for being friends and encouraging me. Thanks to the Cad-Mädels and Jürgen for the big support. As well, I would like to thank all including former members of the K. and H. Jung labs for the friendly atmosphere and fun, especially my doctoral fellows Tobi, Günther, Daniel, Michi, Poldi, Claudia, Susi, Martina, Sophie, Stefan, Matthias, Hannah and Nicola. It was great to experience that team spirit can help you bridging troubled water.
My friends from the "Kaffeekränzchen" and Kathrin & Co stayed at my side and encouraged me – thank you so much!
Berni, thank you for staying at my side and supporting me through all these years - you are wonderful!
Finally, I would like to thank my mother who always backed me up. Without her help (in any respect) I would not have gotten so far.

Ina-Maria-Kann-Schon alias Ina Haneburger