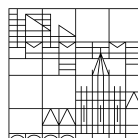


**Novel insights into amino acid catabolism and the
characterisation of a ncRNA regulating BCAA
biosynthesis in bacteria**

**Dissertation zur Erlangung des
akademischen Grades eines Doktors der Naturwissenschaften
(Dr.rer.nat.)**

vorgelegt von
Sebastian Knorr
an der

Universität
Konstanz



Mathematisch-Naturwissenschaftliche Sektion
Fachbereich Chemie

Konstanz, 2019

Tag der mündlichen Prüfung: 10.07.2020

1. Referent/Referentin: Prof. Dr. Jörg Hartig

2. Referent/Referentin: Prof. Dr. David Schleheck

3. Referent/Referentin: Prof. Dr. Andreas Marx

Notes to the reader

This PhD thesis is divided into four main chapters. **Chapter 1** contains a general introduction providing background knowledge relevant for the following parts. It introduces the reader into bacterial responses to nutrient starvation and other stresses emphasising on amino acid catabolism. Furthermore, **Chapter 1** gives an overview on relevant RNA based regulatory mechanisms involved in amino acid metabolism in bacteria.

The following chapters present the main work of this thesis.

Chapter 2 and 3 cover new insights into amino acid metabolism including the characterisation of a novel lysine degradation pathway and a screening system for novel enzyme functions in *E. coli*.

Chapter 4 contains the characterisation of the *ilvH*-motif as a regulatory active RNA element involved in branched chain amino acid biosynthesis. **Materials and methods** that were applied through the main work, can be found at the end of this thesis.

All experiments were conducted in the laboratory of Prof. Dr. Jörg Hartig (Department of Chemistry, University of Konstanz) from June 2016 to August 2019.

Chapter 2 (Biochemical Description of a Novel Lysine-Degradation Pathway in *E. coli* Proceeding via Glutarate and L2HG) is mainly based on the paper “*Wide-spread bacterial lysine degradation proceeding via glutarate and L-2-hydroxyglutarate*” (Knorr & Sinn *et al.* 2018, Nature Communications)¹. It partly contains text, graphs and figures from the publication in consent with authors and journal.

Results presented in Chapters 3 and 4 contain unpublished results.

Table of Contents

<i>Abstract</i>	9
<i>Zusammenfassung</i>	11
Chapter 1	13
<i>General Introduction</i>	13
Amino acid degradation in bacterial stress response	13
Nitrogen starvation and amino acid degradation	14
Cellular responses to carbon limitation	16
Interconnections between nitrogen- and carbon-starvation response	19
Amino acid degradation under various stress conditions	20
Regulation of amino acid metabolism via RNA regulatory elements	21
Ribosome mediated bacterial attenuation	23
T-box regulatory system	24
Riboswitches involved in amino acid metabolism	27
Amino acid metabolism regulation by RNA-binding proteins (RBPs)	29
Chapter 2	31
<i>Biochemical Description of a Novel Lysine Degradation Pathway in E. coli Proceeding via Glutarate and L2HG</i>	31
Introduction	31
Results & Discussion	33
Biochemical characterization of CsiD	33
Structural and functional characterisation of CsiD	35
Taxonomic distribution of CsiD	38
A pair of aminotransferase/dehydrogenase actions mediates catabolism from lysine to glutarate	40
LhgO couples the degradation pathway to the respiratory chain	42
Complete pathway validation by metabolic tracing	47
Physiological role of the novel lysine degradation pathway	49
Conclusion	51
Chapter 3	55
<i>An ASKA- & KEIO-collection Based Screening-Approach for Identifying New Protein Functions in Amino Acid Catabolism</i>	55
Introduction	55
Results and Discussion	58
ASKA screening	59
KEIO Screening	60
Conclusion	63
Chapter 4	65
<i>The ilvH-motif RNA Regulates Gene Expression at the Crossroad of BCAA Biosynthesis</i>	65
Introduction	65
Results and Discussion	69
ilvH-motif RNA mediates gene expression in response to isoleucine and α -ketobutyrate	69
In vitro analysis of ilvH-motif RNA as potential riboswitch	71
BCAA biosynthesis gene knockouts influence reporter activity	73

Functional assignment of the <i>ilvH</i> -motif RNA	74
Conclusion	76
<i>Materials</i>	79
Buffers and media	79
Antibiotics	82
Chemicals and kits	83
Equipment and Consumables	84
Bacterial strains	86
Oligonucleotides	86
Plasmids	89
Software used	90
<i>Methods</i>	91
General methods	91
Methods Chapter 2	97
Methods Chapter 3	106
Methods Chapter 4	110
<i>Reference List</i>	114
<i>Supplementary Information</i>	129
Supplementary Figures	129
Supplementary Tables	134
<i>List of Abbreviations</i>	136
<i>Danksagung</i>	139

Abstract

Amino acid metabolism is not only of interest because of its impact on protein biosynthesis as the motor for catalytic processes in the cell, but also catabolism of amino acids themselves presents a very important biochemical process in bacterial metabolism. Bacteria can exploit amino acids as alternative carbon sources by directly utilizing the carbon backbone for energy metabolism. Furthermore, amino acid degradation presents an important defense strategy against various stress factors. However, white spots on the map still exist for many bacterial degradation pathways of amino acids. For example, complete catabolic pathways of the basic amino acid lysine, the branched chain amino acids and most of the aromatic amino acids are still missing for *E. coli*, the best studied organism.

This thesis presents the complete characterisation of a novel lysine degradation pathway in *E. coli* which can also be found in other bacterial species. We discovered a putative α -ketoglutarate dependent dioxygenase, CsiD, that hydroxylates glutarate producing succinate thus connecting lysine degradation to central carbon metabolism. Furthermore, we showed that the reaction product of CsiD, *L*-2-hydroxyglutarate, is further metabolised via LhgO to channel electrons into the membrane for respiratory energy metabolism. We present evidence that lysine is degraded, in *E. coli*, under nutrient starvation conditions contributing to bacterial adaptation to stationary phase conditions. Furthermore, because of the association of *L*-2-hydroxyglutarate to human diseases such as cancer and organic acidurias, we discuss a potential relation between the human gut microbiome as a potential source of the oncometabolite.

To shed light on other possibly unknown amino acid degradation pathways in *E. coli*, this thesis contains the implementation of a screening assay for the identification of novel enzyme functions involved in amino acid catabolism. The screening results offer new insights into what processes influence the utilization of amino acids and derived metabolites.

Different RNA based regulatory elements, like transcription attenuators and riboswitches, can be found in bacteria regulating important processes in amino acid metabolism.

Abstract

This thesis presents the characterisation of the putative non-coding RNA motif *ilvH* as a *cis*-regulatory acting RNA in branched-chain-amino-acid (BCAA) biosynthesis. We were able to show that the *ilvH*-motif RNA regulates downstream gene expression in response to isoleucine and α -ketobutyrate, an intermediate of isoleucine biosynthesis. However, we could not validate the motif RNA as a classical riboswitch. Mutational analysis of a conserved tetraloop of the *ilvH*-motif indicated that this sub-motif is involved in the regulation mechanism of the small ncRNA. It seems likely that the *ilvH*-motif RNA presents a novel regulation mechanism sensed by BCAA biosynthesis metabolites. However, further investigations are needed to fully understand the regulatory mechanism underlying the *ilvH*-motif RNA.

Zusammenfassung

Der Aminosäure-Stoffwechsel ist nicht nur aufgrund seiner Auswirkungen auf die Proteinbiosynthese als Motor für katalytische Prozesse für Bakterienzellen relevant. Auch der Abbau von Aminosäuren selbst stellt einen wichtigen biochemischen Prozess im bakteriellen Stoffwechsel dar. Bakterien können Aminosäuren als alternative Kohlenstoffquellen nutzen, indem sie das Kohlenstoffgerüst direkt für den Energiestoffwechsel verwenden. Darüber hinaus stellt der Aminosäureabbau eine wichtige Abwehrstrategie gegen verschiedene Stressfaktoren dar. Allerdings sind viele bakterielle Abbauwege von Aminosäuren noch größtenteils unbekannt. So fehlen beispielsweise für *E. coli*, den am besten untersuchten Organismus, noch Evidenzen für die kompletten Abbauwege der basischen Aminosäure Lysin, der verzweigtkettigen Aminosäuren und der meisten aromatischen Aminosäuren.

Diese Arbeit beinhaltet die vollständige Charakterisierung eines vorher unbekanntes Lysin-Abbauweges in *E. coli*, der auch in anderen Bakterienarten zu finden ist. Wir entdeckten, dass die putative Fe(II)/ α -ketoglutarate abhängige Dioxygenase, CsiD, Glutarate hydroxyliert. In dieser Reaktion wird Succinat produziert und so der Lysinabbau mit dem zentralen Kohlenstoffstoffwechsel verknüpft. Darüber hinaus konnten wir zeigen, dass ein weiteres Reaktionsprodukt von CsiD, *L*-2-Hydroxyglutarat, durch LhgO oxidiert wird. Über diese Reaktion werden Elektronen in die Elektronentransportkette geschleust, weswegen LhgO als Bestandteil der aeroben Zellatmung agiert. Außerdem präsentieren wir Hinweise dafür, dass Lysin in *E. coli* unter Nährstoffmangel abgebaut wird und dies zur Anpassung an die Bedingungen während der bakteriellen stationären Phase geschieht. Darüber hinaus diskutieren wir das menschliche Darm-Mikrobiom als mögliche Quelle von *L*-2-Hydroxyglutarat, ein Metabolit, der mit der Entstehung von Krebs und anderen Stoffwechselerkrankungen assoziiert ist.

Weiterhin enthält diese Arbeit die Entwicklung und Implementierung eines Screening-Assays zur Identifizierung neuartiger Enzymfunktionen in *E. coli*, welche mit dem Aminosäureabbau assoziiert sind. Die Screening-Ergebnisse bieten neue Erkenntnisse darüber, welche Prozesse den Katabolismus von Aminosäuren und verwandten Metaboliten beeinflussen.

RNA-basierte regulatorische Elemente, wie Transkriptionsattenuatoren und Riboswitches, regulieren wichtige Prozesse im Aminosäurestoffwechsel von Bakterien. Wir konnten zeigen, dass die putative *non-coding(nc)*RNA *ilvH* als *cis*-regulatorisches RNA-Element für Gene aus dem Biosyntheseweg von verzweigtkettigen Aminosäuren (BCAA) agiert. Das *ilvH*-RNA Motiv reguliert dabei die Genexpression des strangabwärts liegenden Gens (*ilvH*) in Abhängigkeit von Isoleucin und α -Ketobutyrat, ein Zwischenprodukt der Isoleucin-Biosynthese. Unsere Analysen zeigen, dass es sich bei dieser ncRNA vermutlich nicht um einen Riboswitch handelt. Reporter-Konstrukte, die Mutationen im konservierten Tetraloop des *ilvH*-Motivs enthielten, lassen zudem den Schluss zu, dass dieses Submotiv am Regulationsmechanismus der ncRNA beteiligt ist. Es ist möglich, dass das *ilvH*-RNA-Motiv einen neuartigen Regulationsmechanismus darstellt, der von Intermediaten aus der BCAA Biosynthese gesteuert wird. Um den zugrunde liegenden Regulationsmechanismus vollständig aufzuklären, sind jedoch weitere Untersuchungen dieser ncRNA notwendig.

Chapter 1

General Introduction

Amino acid degradation in bacterial stress response

Bacterial cells have developed various adaptations to overcome environmental stress conditions occurring in natural habitats. These stresses can be varying pH conditions, antibiotic- and temperature-stress, osmotic stress, accumulation of toxic compounds or nutrient depletion. Under harsh environmental influences internal programs are activated to protect the cell from stress factors. One of these programs is initiated at the beginning of what in cell culture is observed as stationary phase which mimics a condition that is predominantly found in bacteria living in their natural habitat². In this phase, cells stop growing and genes linked to cell division and biosynthesis of amino acids are highly down regulated whereas a subset of about 20% of all genes in *E. coli* is upregulated. However, it was shown that protein production maintained at high levels even days after entering stationary phase². It was suggested that degradation of proteins and reincorporation of their amino acids are responsible for this high yield of protein production³. Catabolism of amino acids under starved conditions might also be an important contributor for maintaining energy and nitrogen metabolism during stationary phase.

RpoS is an alternative RNA polymerase sigma factor and one of the key regulators of gene expression during stationary phase conditions. During nutrient rich and low stress conditions RpoS is degraded rapidly by proteases^{4,5}. Besides protease degradation, RpoS is regulated post-transcriptionally via regulator proteins like Hfq or small molecules like UDP-glucose⁶. But also the 5'-UTR of *rpoS* RNA is highly structured and under positive control of non-coding RNAs⁷. Furthermore, RpoS expression is regulated by nutrient starvation signals like the stringent response mediator (p)ppGpp or the carbon starvation molecule cAMP^{8,9}. When nutrient depletion occurs RpoS activates gene expression of a variety of genes enabling the cells to utilize amino acids as carbon- (C-) and nitrogen- (N-)sources. On the other hand, house-keeping genes from the TCA cycle (tricarboxylic acid cycle) are down regulated by RpoS via competitive binding of RNA with RpoD, the key sigma factor during exponential growth^{10,11}. Interestingly, mutations in the *rpoS* gene

lead to an induction of the growth advantage in stationary phase phenotype (GASP). Cells that have a GASP phenotype are a subpopulation that outgrow other bacterial populations of the same strain during stationary phase. It was shown that *E. coli* growth on amino acids as sole C-sources is more efficient when GASP mutations are present indicating the importance of amino acid catabolism as energy source during stationary phase conditions¹². Also gene functions involved in the biosynthesis of sugars and amino acids as well as fatty acids can be induced via RpoS¹³. Additionally, RpoS regulates the expression of oxidative and osmotic stress response genes.

Nitrogen starvation and amino acid degradation

Non-nitrogen fixing microorganisms like *E. coli* that live in the intestine of mammals and also in other environments are confronted with nitrogen limited conditions most of the time¹⁴. This is because animals themselves are struggling to gain sufficient nitrogen by capturing it from food. However, microorganisms in the intestine of mammals have even higher nitrogen requirements than their hosts, leading to the situation that bacteria and hosts compete for nitrogen¹⁵. Thus, evolution equipped bacteria with systems allowing them to adapt to varying levels of nitrogen in their natural habitats. Usually, *E. coli* prefers ammonium as N-source leading to fast accumulation of biomass during growth. Under starved conditions, bacteria are able to assimilate nitrogen from organic sources like amino acids and their derivatives (e.g. polyamines like cadaverine or putrescine) predominantly by aerobic metabolism. Catabolism of these N-sources must generate the major intracellular nitrogen donors glutamate (Glu) and glutamine (Gln)^{16,17}. Under nitrogen-rich conditions Glu is produced by direct addition of ammonium to α -ketoglutarate (α KG) catalysed by glutamate dehydrogenase (GDH)¹⁸. When ammonia levels decrease, degradation of other amino acids like serine can lead to the formation of ammonia. But Glu can also be synthesised by transaminases that transfer amino groups of other amino acids (e.g. aspartate) to α KG or when it is produced directly as an end product (e.g. proline or arginine degradation)¹⁹. An overview of amino acids that can be degraded and used as N-sources by *E. coli* and how they contribute to the Glu and ammonium pool is provided in Table 1.

Table 1: Amino acids as N-sources for *E. coli*. Column 2 describes the ability of *E. coli* to grow with the indicated amino acid present as sole N-source. Columns 3, 4, 5 show whether enzyme reactions involving the amino acid produce ammonium (3), Glutamate (Glu) via transamination (4) or Glu as an end product. All data is derived from the EcoCyc database²⁰.

L-amino acid	sole N- source growth	NH ₄ ⁺ generating	Glu via transamination	Glu as end product
alanine	+	+	- ^b	-
arginine	+	+	+ ^a	+
asparagine	+	+	-	-
aspartate	+	+	+	-
cysteine	+	+	+ ^c	-
glutamate	+	+		
glutamine	+	+	+	+
glycine	+	+	-	-
histidine	-	-	- ^b	-
isoleucine	-	- ^b	- ^b	-
leucine	-	-	- ^b	-
lysine	+ ^c	-	+ ^{a,c}	-
methionine	+	-	-	-
phenylalanine	-	-	+ ^c	-
proline	+	-	-	+
serine	+	+	- ^b	-
threonine	+	+	-	-
tryptophan	+ ^d	+	+ ^c	-
tyrosine	-	-	+ ^c	-
valine	-	-	- ^b	-

^a degradation pathway via a polyamine

^b anabolic enzymes of the amino acid are shown to perform reversed reactions *in vitro*, contribution to the Glu pool is not shown

^c described in this thesis

^d dependent on sole carbon source (growth with other than glucose)²¹

^e reactions are shown to be catalysed by AspC (aspartate aminotransferase) and/or TyrB (tyrosine aminotransferase) both with a broad substrate promiscuity, contribution of this reactions to the Glu pool remains elusive

When *E. coli* faces energy-rich but low-nitrogen conditions, nitrogen can be assimilated by glutamine synthase (GS) adding ammonia to Glu in an ATP dependent manner. Additionally, Glu can be synthesised from Gln by transferring the amino group of Gln to α KG by glutamate synthase (GOGAT) producing 2 molecules of Glu. The GS-GOGAT system is more suitable when cells are ammonia depleted than the GDH since GS has a much lower K_m for ammonium¹⁹. Both donors provide nitrogen to a variety of compounds in the cell. On the one hand, Gln provides nitrogen for purines, pyrimidines, asparagine, tryptophan, histidine, arginine, glucosamine and *p*-aminobenzoate¹⁷. On the other hand, the amino group of Glu is transferred by transaminases to synthesize most of the amino acids needed to be incorporated into new biomass while α KG flows back into the TCA²².

The central regulators of nitrogen limitation are nitrogen-regulated response (Ntr) transcription factors. Ntrs induce GS, transport systems for potential N-sources and catabolic enzymes for amino acid and polyamine degradation²³. They repress GDH and GOGAT under low ammonium and Gln conditions that occur when other amino acids are used as sole N-sources²³. This scavenges the remaining ammonia pool into Gln whereas Glu is predominantly synthesized via amino-group transfer from other N-sources. Thus, the Ntr system is responsible for increased ammonia assimilation and the coordination of nitrogen depleted metabolism.

Cellular responses to carbon limitation

In many bacteria, glucose is the preferred C-source and leads to the fastest accumulation of biomass. The diauxic shift, where lactose is only metabolized when glucose is already depleted, reflects this preferential metabolic behaviour. The underlying system is glucose catabolite repression which leads to an inactivation of operons that degrade other C-sources when glucose is present^{24,25}. However, when the cell gets depleted of glucose and other easily degradable sugars, energy metabolism and protein biosynthesis are heavily impaired. This is emphasized by the observation that upon entry into stationary phase slower cell division rates are accompanied by a strong reduction of protein biosynthesis²⁶. Consequently, two important signalling and regulatory programs are activated, the catabolic response under energy limitation and the stringent response under amino acid limitation.

As a result of carbon limitation, important metabolic pathways are disrupted that are responsible for the production of amino acid precursors and energy metabolites. Under glucose starvation glycolysis and the pentose phosphate pathway are interrupted²⁷. This leads to decreasing synthesis rates of all amino acids resulting in rising levels of uncharged tRNAs which form the trigger for the stringent response by activating RelA a GDP/GTP pyrophosphokinase. RelA catalyses the reaction of $ATP + GDP \rightarrow AMP + ppGpp$ and $GTP + ATP \rightarrow pppGpp + AMP$ where (p)ppGpp is the global regulator of the response. The alarmone is supposed to bind to RNA polymerases²⁸. Accumulation of (p)ppGpp inactivates genes involved in translation processes resulting in declining protein biosynthesis and growth rates of cells²⁸. Furthermore, (p)ppGpp activates the alternative

sigma factor RpoS which then competes with RpoD down-regulating TCA gene products^{6,29}. Upscaling the amino acid pool is maintained by degradation of ribosomal proteins promoted by (p)ppGpp³⁰. But also recycling amino acids from exponential phase proteins enable bacteria to keep protein biosynthesis running during stationary phase^{3,31}. The resulting increase in charged tRNAs then leads to the activation of SpoT, the antagonist of RelA that hydrolyses (p)ppGpp, thus interrupting the stringent response³².

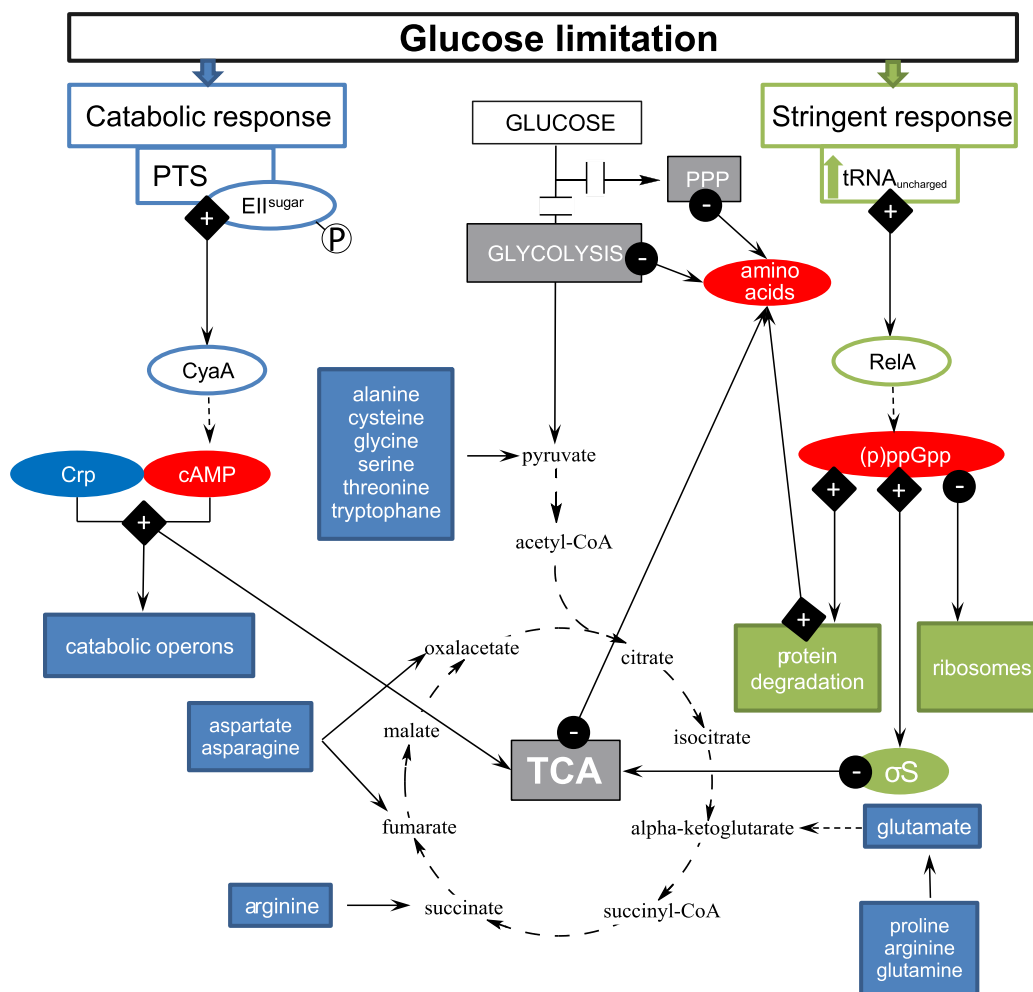


Figure 1: Carbon starvation response of *E. coli*. A simplified overview on cellular processes occurring during glucose limitation. **Catabolic response**, Glucose starvation leads to phosphorylation of parts of the phosphotransferase system (PTS). This induces CyaA which synthesises cAMP the global regulator molecule of the catabolic response. Only when cAMP is bound to the transcription factor Crp catabolic operons are activated that lead to the degradation of amino acids. **Stringent response**, If glucose is depleted important metabolic pathways like the pentose-phosphate-pathway (PPP) as well as glycolysis and within that the TCA are impaired. This leads to a shortage of amino acid precursors resulting in reduced levels of amino acid and increasing levels of uncharged tRNAs. High levels of uncharged tRNAs activate RelA which synthesises (p)ppGpp. (p)ppGpp then induces and inhibits a variety of cellular processes. A more detailed description of these processes is depicted in the main text. Solid lined arrows present catabolic pathways. Dashed lined arrows stand for single enzyme reactions. Arrows with + at the beginning stand for upregulation of the following process, whereas - stands for inhibition. Red filled ellipses present metabolites. Other filled ellipses are transcription factors, whereas outlined ellipses define catalytic enzymes. Blue coloured processes are part of the catabolic response, whereas green colour stands for stringent response.

A more offensive strategy, when carbon is limited and energy supply has to be perpetuated, is the catabolic response. This response is coupled to the phosphotransferase system (PTS) enabling the cell to use alternative C-sources. At low glucose concentrations the soluble components of the system EI and phosphohistidine carrier protein (Hpr) transfer phosphoryl groups of phosphoenol pyruvate to the sugar specific and membrane associated parts EII (EII^{Glc} in case of glucose, but also EIIs for mannitol, fructose, mannose and others are known)³³. It is suggested that phosphorylated EII then activates CyaA, an adenylate cyclase, which catalyses the reaction $ATP \rightarrow cAMP + PPi$. In the next step, the main alarmone cAMP of the catabolic response binds to Crp (cAMP receptor protein) activating this transcription factor. The Crp-cAMP complex then induces three important answers: 1) chemotaxis to increase cell mobility for the exploitation of new C-sources, 2) activation of catabolic operons that enable the cell to degrade alternative C-sources and 3) enhancing TCA and repressing glyoxylate shunt gene expression to utilize precursors produced by catabolism of other carbon metabolites³⁴⁻³⁶. As a part of this metabolic response the degradation of amino acids as C-source leads to the formation of the TCA precursors malate, oxaloacetate, succinate (SA) and α KG in *E. coli* (Figure 1).

Catabolism of free amino acids as C-sources therefore is involved in maintaining TCA precursors and energy supply. In *E. coli* all known amino acid catabolic pathways produce compounds of the TCA cycle and are referred to as glucogenic amino acids since they can contribute to the formation of glucose by gluconeogenesis. Usually, ketogenic amino acid degradation is described as the formation of acetyl-CoA that does not contribute to the production of glucose. However, in bacteria performing the glyoxylate shunt acetyl-CoA can also be channelled into TCA intermediates making the differentiation between glucogenic and ketogenic degradation pathways difficult. Still, in many bacterial species amino acid degradation can be performed in an acetyl-CoA dependent manner (e.g. L-lysine). Stringent- and catabolic-response are strongly interconnected at the level of energy and amino acid metabolism. Whereas many catabolic operons regulated by Crp-cAMP are additionally induced by (p)ppGpp, RelA itself is upregulated by Crp-cAMP³⁷. The additional effect of (p)ppGpp on catabolic operons can be explained by the direct induction of RpoS by the alarmone. RpoS regulates the stationary phase induced genes and therefore also catabolic amino acid pathways⁸.

Interconnections between nitrogen- and carbon-starvation response

Cell programs that respond to nitrogen and carbon starvation are linked at several levels. The main effector of this interconnected regulation is the ratio between Gln and α KG which reflects nitrogen and carbon levels in the cell.

When nitrogen limitation occurs, levels of Gln sink under those of α KG. High levels of α KG occur through Glu depletion caused by transamination producing other amino acid³⁸⁻⁴⁰. As a consequence, the high α KG over Gln ratio blocks adenylation of GS thus activating the enzyme⁴¹. The adenylation state of GS is regulated by Ntr-P_{II}-1 that itself is post-translationally regulated through its uridylylation state by α KG and ATP/ADP^{42,43} (Figure 2). Moreover, the α KG-Ntr-P_{II}1 complex regulates Ntr phosphatases initiating the nitrogen starvation response^{44,45}. Additionally, under nitrogen limitation, α KG inhibits EI of the PTS thus preventing sugar uptake. Furthermore, this results in reduced EII phosphorylation which decreases cAMP concentrations and thus the catabolic response^{44,45}.

In case of carbon limitation, the ratio of Gln to α KG remains stable in the cell. Low α KG concentrations lead to adenylation of the GS by the P_{II}-Ntr system inactivating GS. Since the P_{II}-Ntr system binds both ATP/ADP, it may serve as sensor for cellular adenylate energy charge which is the measure of energy available for metabolism^{46,47}. Thus, under carbon-limitation ATP and α KG levels decrease which results in increased adenylation of GS and the downregulation of the Ntr system. Consequently, α KG is not drained into Glu and can be used in the TCA for energy metabolism.

Interestingly, this feedback regulation pattern can be observed in amino acid catabolic operons under the control of Crp-cAMP. Expression of nitrogen assimilation genes like Glu-synthesising aminotransferases is regulated on a second level by Ntr and differing sigma factors. An example for this regulation at the intercourse of nitrogen and carbon assimilation is the *csiD-lhgO-gabT/D-gabP* operon.⁴⁸ The role of the *csiD* operon in nitrogen and carbon assimilation as well as its contribution to a novel lysine degradation pathway discovered by our group is part of thesis.

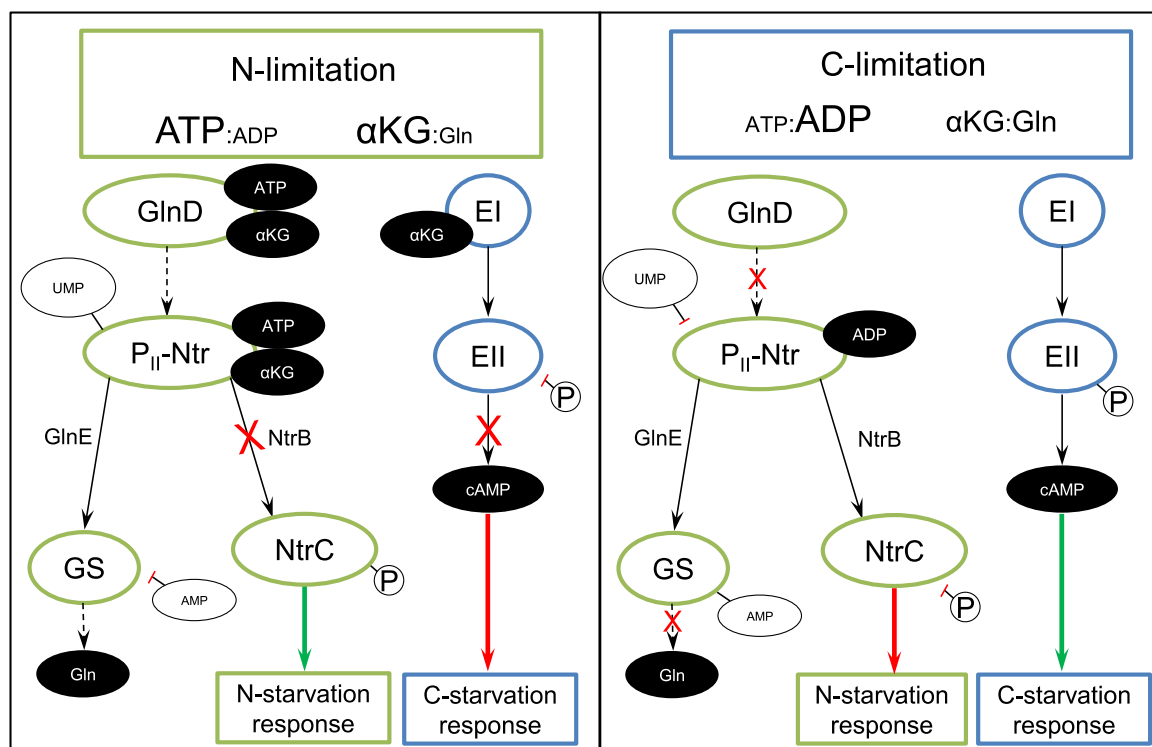


Figure 2: Interconnection between nitrogen- and carbon-starvation response. N-limitation, High ATP and α KG over Gln levels lead to activation of the P_{II} uridylyltransferase (GlnD) transferring UMP to P_{II} -Ntr. P_{II} -Ntr-UMP interacts with Gln synthetase (GS) via adenylyltransferase (GlnE) catalysing the removal of AMP from GS thus activating GS resulting in the formation of Gln. Additionally, P_{II} -Ntr-UMP binds to the sensory histidine kinase NtrB inhibiting phosphatase activity resulting in NtrC-P. Phosphorylated NtrC is one of the main activators of nitrogen-starvation response operons. Both actions of P_{II} -Ntr-UMP are enhanced by allosterically binding of ATP and α KG. Furthermore, α KG binds to EI preventing phosphorylation of EII. Dephosphorylated EII prevents biosynthesis of cAMP and, within that, carbon starvation response processes. **Carbon-limitation,** High ADP levels and an equalized pool of α KG and Gln do not stimulate uridylylation activity of GlnD resulting in unmodified P_{II} -Ntr. Unmodified P_{II} -Ntr and additional binding of ADP instead of ATP and α KG prevent GlnE dependent deadenylation of GS. Modified GS-AMP is catalytically inactive. Additionally, NtrB phosphatase activity is not inhibited by unmodified P_{II} -Ntr leading to dephosphorylation of NtrC and an inactivation of nitrogen-starvation response. Low levels of α KG lead to an activation of EI kinase activity resulting in EII-P. This leads to activation of CyaA and the synthesis of cAMP which triggers the C-starvation response processes. Colours & shapes: Green coloured shapes present components of the N-starvation response, blue coloured shapes C-limitation response. Dashed arrows represent single enzyme reactions. Solid arrows define processes with more than one enzyme involved. Green arrow stands for activation, red arrow for inhibition. Coloured, outlined ellipses define catalytically active enzymes. Black outlined ellipses are protein modifications: AMP = adenylation, UMP = uridylylation, P = phosphorylation.

Amino acid degradation under various stress conditions

Varying oxygen levels or changes in acidity of the surrounding present stress conditions that are often faced by a gastrointestinal tract inhabiting bacterium like *E. coli*. But also under laboratory conditions, for example in liquid culture, these stressors occur and impact cellular responses. Severe acidic stress can be observed when medium's pH is about 2–2.5 and damage is induced to proteins, nucleic acids and the oxidative respiration system⁴⁹. Under these conditions *E. coli* induces different acidic resistance mechanisms in

which amino acid degradation plays an important role. The main parts of the system rely on pH-induced PLP (pyridoxal 5'-phosphate) dependent decarboxylases specific for the amino acids Arg, Glu and Lysine and their cognate inner membrane antiporters. Thus, arginine, Glu and lysine are decarboxylated to agmatine, γ -aminobutyrate (GABA) and cadaverine (catalysed by SpeA, GadA/B, CadA) which produces CO₂ and consumes protons. The decarboxylated reaction products are then exported whereas new amino acid substrates are imported by the specific antiporter⁵⁰. On a transcriptional level the genes are under the control of the RpoS sigma factor, but also decreasing pH induces the decarboxylases allosterically. Additionally, Gln can be amidohydrolysed to Glu by YbaS which can be channelled into the acid resistance system by decarboxylation. Ammonia formed by the amidohydrolase can then sequester protons, too⁵¹.

Degradation products of arginine and lysine catabolism also contribute to other important stress responses by the production of biogenic polyamines. Lysine is decarboxylated to cadaverine, a five-carbon backbone polyamine by both, the acid induced enzyme CadA⁵² and the constitutively expressed LdcC⁵¹. Arginine can be metabolized to putrescine via agmatine in *E. coli*. By addition of 1,3-diaminopropane to the aminobutyl polyamine putrescine, spermidine and spermine are formed⁵³. Through their polycationic nature, polyamines can bind to RNA and DNA helices as well as proteins and within that regulate gene expression and enzyme function⁵⁴. In *E. coli* a complete polyamine modulation could be revealed where putrescine is responsible for the regulation of various genes involved in amino acid metabolism, cofactor biosynthesis, energy metabolism and many other cellular processes⁵⁵. Furthermore, polyamines positively regulate the expression of stress related genes like the cAMP producing enzyme Cya, the stress response sigma factor RpoS⁵⁶ and other transcription factors for peroxide detoxification and superoxide radical response⁵⁷.

Regulation of amino acid metabolism via RNA regulatory elements

Gene expression can be regulated at different levels: at transcription-initiation, transcription-elongation and post-transcription. Transcription-initiation in prokaryotes is dependent on DNA promoter elements like -10 and -35 regions which are bound by different sigma factors associated with the core RNA polymerase and are itself affected by signalling molecules. Additional upstream elements on the DNA interact with

transcription factors enhancing or inhibiting gene expression. The role of different sigma factors and the Crp-cAMP system on gene expression of genes involved in amino acid catabolism was already described in the first part of this introduction. Many more examples of transcription factors regulating amino acid metabolism genes are known. A wide-spread example is Lrp (Leucine-responsive regulatory protein) which is a regulator of about 10% of *E. coli* genes. It mainly acts as an activator or repressor of genes and operons involved in amino acid biosynthesis and catabolism. Moreover, leucine binds to Lrp enabling allosteric regulation of the transcription factor itself^{58,59}.

On the level of transcriptional-elongation and post-transcription regulation, non-coding RNAs (ncRNA) like antisense small-RNAs (sRNAs) and microRNAs are examples of important effectors. Antisense sRNAs regulate gene expression via binding to a complementary mRNA strand thus blocking translation. They can be transcribed *in trans* but also occur *in cis* to their antisense RNA strand. Some sRNAs were also shown to bind proteins thus regulating their enzyme-activity⁶⁰. ncRNAs are often located in intergenic regions 5' or 3' of genes they regulate *in cis*. Unique characteristics of RNA makes this molecule very suited for gene regulation. ncRNAs, but also transcripts containing leader peptide sequences, occur single stranded which enables the RNA to form complex secondary and tertiary structures which can be folded in altering conformations. This enables RNA to form binding pockets for small metabolites. Binding of additional factors like small molecules or RNA binding proteins (RBP) could therefore lead to masking or releasing of important parts of the transcription-elongation or ribosomal-binding machinery. Furthermore, these unique characteristics enable RNA to function as thermosensors which is especially important for gene regulation in bacteria living under extreme environmental conditions⁶¹.

Due to its relevance for all processes in bacterial cells, amino acid metabolism is designated for various gene expression regulation mechanisms enabling adaption to new environmental conditions. Since bacteria are able to import amino acids, synthesize them from simpler compounds or degrade them for energy production, amino acid metabolism is tightly controlled by feedback mechanisms that also include ncRNAs and leader mRNAs. The three most important regulatory elements concerning amino acid metabolism are ribosome mediated attenuators, T-box RNAs and riboswitches. But also,

regulatory mechanisms that include Rho-dependent termination or involve other RBPs occur in amino acid metabolism regulation.

Ribosome mediated bacterial attenuation

One of the best studied examples of bacterial attenuation gene expression control is the regulation of the *trp* operon in *E. coli*⁶¹. In this regulatory system a leader region in the 5'-UTR of tryptophan biosynthesis genes regulates expression by sensing tryptophan levels in the cell. This is mediated by pausing of the ribosome when tryptophan levels are low resulting in conformational changes of the RNA structure. The leader mRNA usually contains four segments that form altering hairpins with each other. Furthermore, the first segment contains a short open reading frame (ORF) including two codons for tryptophan which is referred to as leader peptide. When Trp levels are low the ribosome pauses at the tryptophan codons and the formation of an antiterminator hairpin leads to transcription read through. Under high intracellular tryptophan levels ribosome velocity remains stable until reaching the stop codon of the leader peptide. At this position the ribosome prevents formation of the antiterminator hairpin which leads to the formation of a terminator hairpin (Figure 3).

Since transcription and translation occur simultaneously in bacteria a polyU track downstream of the hairpin leads to pausing of the transcription elongation complex. Additionally, the DNA-RNA hybrid is destabilized due to the weak rU-dA base pairing resulting in dissociation of the RNA polymerase (RNAP) and transcription termination. Ribosome mediated transcription attenuation is wide-spread in almost all clades of bacteria. To date, amino acid attenuators have been found that regulate gene expression in dependency of isoleucine, valine, leucine, tryptophan, cysteine, methionine and histidine resembling the importance of this regulatory mechanism for amino acid metabolism⁶².

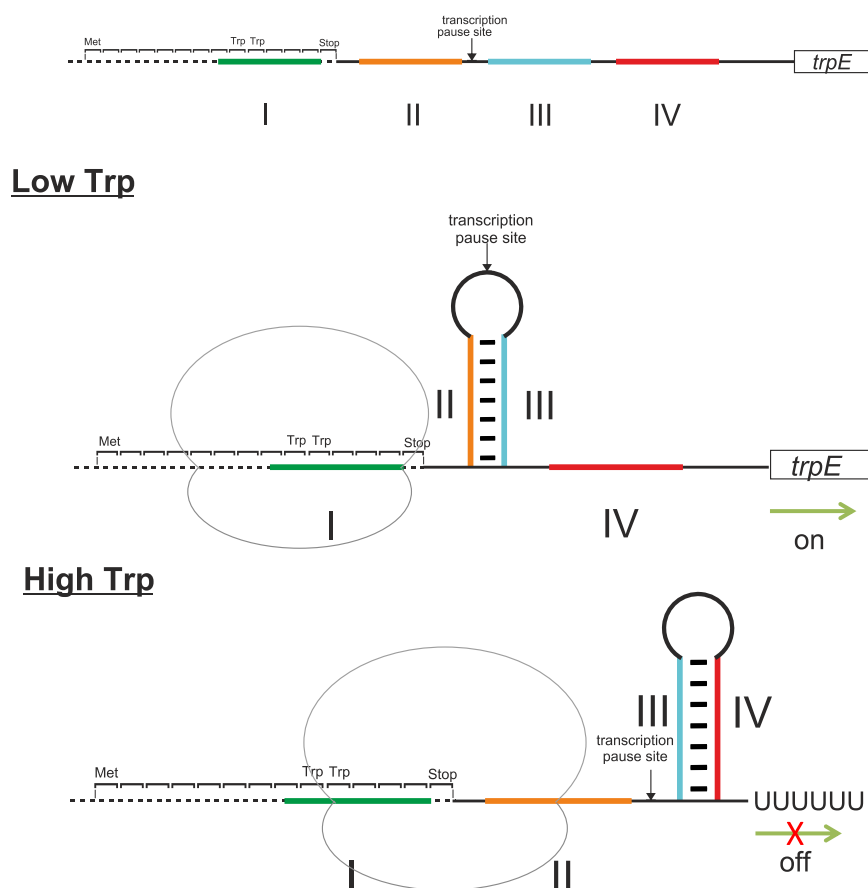


Figure 3: Ribosome mediated transcription attenuation of *trpE* in *E. coli*. Upper image shows the organisation of the regulator elements upstream of *trpE*. Elements I – IV are able to form stable hairpin structures among each other. Element I is part of a leader peptide and contains two codons translating for tryptophan. Adjacent to Element I a stop-codon is located whereas at the 3'end of Element II a transcription pause site can be found. **Low Trp** levels in the cell lead to slower translation velocity of the ribosome enabling the formation of a stable stem between Elements II and III. This releases the stalled RNAP leading to full transcription of *trpE*. Under **High Trp** levels the ribosome is not stalled at the two tryptophan codons and translation occurs rapidly until the ribosome reaches the stop-codon of the leader peptide. Due to higher translation velocity the ribosome prevents formation of the II-III hairpin. The RNAP is now stalled at the III-IV stable hairpin which describes a terminator hairpin leading to destabilization of the mRNA-RNAP complex. This results in dissociation of the RNAP and premature transcription termination. Trp = tryptophan.

T-box regulatory system

T-box regulators define a type of bacterial *cis*-regulatory RNA elements that act via binding of tRNAs. T-boxes have been shown to be wide-spread in Firmicutes and Actinobacteria often occurring with multiple T-boxes per genome. Because of the modular nature of the tRNA-ncRNA interaction, T-box regulators have been found to regulate biosynthesis, transporter and aminoacyl-tRNA synthetase genes of all natural amino acids in these phyla⁶³. T-boxes usually contain a conserved stem loop (specifier hairpin) followed by a variable region, the conserved T-box sequence and a transcription terminator or Shine-Dalgarno sequence (SD) sequester stem. The specifier hairpin holds a

sequence with a base triplet defining the amino acid related gene context. The specifier sequence therefore defines which cognate tRNA is able to bind the RNA structure via its codon sequence.

T-box RNAs regulate gene expression on a transcriptional or translational level. In case of transcriptional regulation, the T-box sequence interacts with the uncharged tRNA leading to the formation of an antiterminator hairpin by base pairing with the 5' part of the transcription terminator stem (Figure 4)⁶⁴. The T-box sequence as part of the antiterminator stem forms a 7 nucleotides long bulge containing a highly conserved UGGN motif. This motif is complementary to the NCCA sequence of the acceptor module of tRNAs and thus forms a second and unspecific binding site for uncharged tRNAs⁶⁵. Under amino acid limited conditions cognate uncharged tRNAs accumulate leading to antiterminator formation. When charged tRNAs are abundant, binding only occurs at the specifier sequence resulting in terminator stem formation and premature transcription termination.

If T-boxes regulate gene expression on a translational level the binding mechanism between 3' acceptor of the tRNA and T-box sequence remains the same. Binding of an uncharged tRNA now leads to the formation of an antisequester stem. The antisequester unmasks the SD initialising translation. When a charged tRNA is present the SD is masked by an SD sequester hairpin preventing ribosome binding and translation⁶⁵.

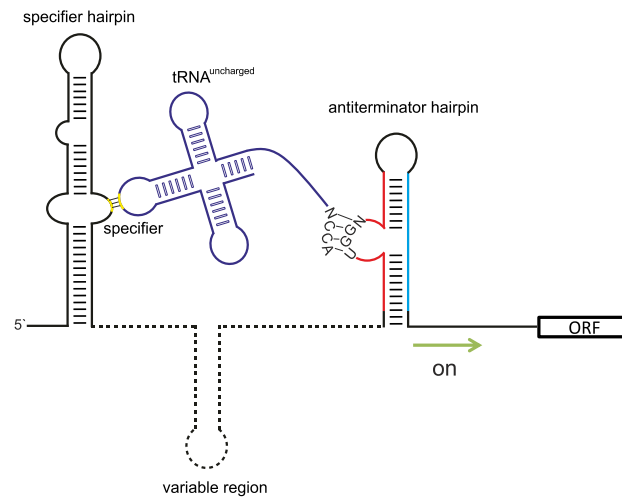
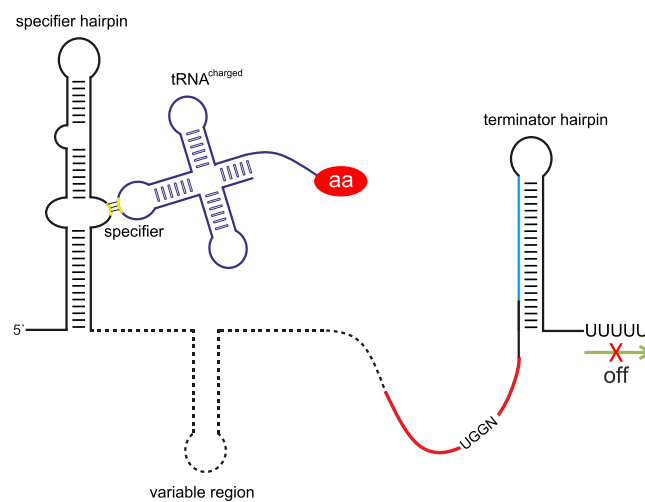
Low levels of charged tRNAs**High levels of charged tRNAs**

Figure 4: Simplified scheme of T-box mediated transcription regulation. The T-box RNA forms a conserved specifier hairpin containing the complementary codon sequence to the anticodon of the respective tRNA. Within that, a tRNA is able to bind to the T-box RNA via base pairing to the codon sequence (yellow) located in a loop of the specifier hairpin. Separated via a variable, less conserved region the antiterminator/terminator hairpin is located downstream of the specifier hairpin. **Upper panel,** If amino acid levels decrease in the cell the amount of uncharged tRNAs increases. Therefore, the amino acid acceptor module of the tRNAs (NCCA) occurs unmasked and pairs to a complementary sequence (UGGN) of the T-box sequence (shown in red). This interaction promotes the formation of an antiterminator hairpin mediating transcription read through. **Lower panel,** High levels of amino acids lead to more charged tRNAs. In charged tRNAs the amino acid is connected to the acceptor module thus preventing interaction between the module sequence and the T-box sequence. This leads to a conformational change promoting the formation of a terminator hairpin. Formation of a terminator hairpin leads to premature transcription termination.

Riboswitches involved in amino acid metabolism

Riboswitches are regulatory active and structured ncRNAs that perform gene regulation upon direct interaction between RNAs and small molecules. They can be found in most bacterial phyla with the highest count per genome found in Firmicutes⁶⁶. Similar to T-boxes and attenuators, riboswitches regulate gene expression via the interplay of alternating conformations. They are located in 5'-UTRs of bacterial mRNAs and are associated with various gene functions. Most commonly, riboswitch regulation mechanisms respond upon binding of nucleotide related coenzymes such as FMN or TPP and nucleotide derivatives like guanine, adenine, ADP and CDP. But also ions (e.g. F^+ , Co^{2+} , Mg^{2+}), signalling molecules (cyclic dinucleotides, ZTP, (p)ppGpp) and other small molecules like guanidine or PRPP are sensed by these ncRNAs⁶⁶. Thus, riboswitches represent genetic tools that participate in the regulation of many different cellular processes like metabolic pathways and the transport of related metabolites. In general, recognition of small molecules by riboswitches is mediated by two domains: An aptamer domain that specifically binds the ligand and an expression platform that mediates conformational changes. The metabolites interact with the nucleobases but also with the phosphate backbone of the RNA. Since altering structures are induced in the presence or absence of the ligand, riboswitches function as on- or off-switches. Thereby, regulation is mediated on a transcriptional or translational level⁶⁷. In some cases, regulation occurs on both levels, e.g. when ribosomal binding sites are part of transcription terminator stems⁶⁸. Riboswitches are also involved in the metabolism of amino acids. To date, three classes are known that directly sense lysine, glycine and glutamate.

The lysine sensing aptamer domain consists of a five-way helical junction followed by an expression platform forming a regulatory hairpin⁶⁹. The expression platform can function as transcriptional terminator in case of the lysine riboswitch in the 5'-UTR of the *asd* (coding for aspartate-semialdehyde dehydrogenase) gene of *T. maritima*. But also SD sequesters are known as seen in the regulation of the *lysC* (coding for aspartate kinase III) gene in *E. coli*⁷⁰. In both cases lysine binding to the centre of the five-way junction leads to structural rearrangements disrupting the antiterminator or antisequestor stem^{71,72}. In the first case this leads to premature transcription termination, in the latter to inhibition of translation of the associated lysine biosynthesis genes. Furthermore, lysine riboswitches exist that activate gene expression in the presence of the amino acid. In these cases, the aptamer domain and expression platform overlap, and binding of lysine leads to the

stabilization of an antiterminator stem and transcription readthrough. In general, lysine riboswitches are found upstream of genes involved in both lysine anabolism and catabolism as well as transport⁷³.

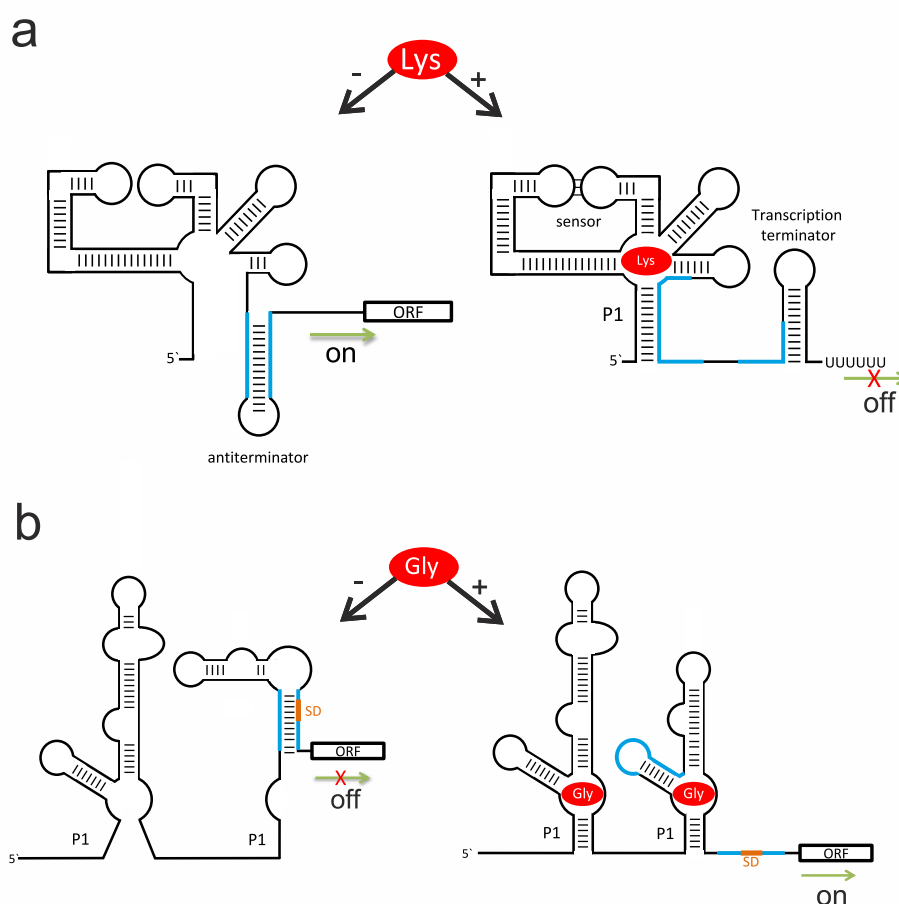


Figure 5: Riboswitch regulation mechanism by amino acid sensing. a, Schematic representation of a lysine sensing riboswitch mediating transcriptional off-switching upon ligand binding. Binding of lysine leads to formation of an alternative structure (blue) that includes a terminator hairpin mediating premature transcription termination. **b,** Simplified presentation of a glycine sensing riboswitch mediating translational on-switching by ligand binding. Upon binding of glycine, the tandem arranged P1 stems get stabilised leading to a conformational change (blue) that is followed by demasking of the Shine-Dalgarno (SD) sequence (orange). The ribosome is now able to bind to the SD and translation is initiated.

Glycine sensing riboswitches modulate gene expression similar to lysine riboswitches. Their mechanism is also based on premature transcription termination and translation inhibition. However, glycine riboswitches activate gene expression upon ligand binding whereas the majority of lysine riboswitches perform downregulation of their downstream gene products⁷⁴. This is accompanied through the localisation of glycine riboswitches upstream of glycine utilization genes like *gcvT*, a glycine aminomethyltransferase that cleaves glycine into CO₂ and ammonium. When glycine binds to the RNA it stabilizes the

P1 stem of the aptamer domain leading to disruption of the terminator or sequester stem thereby activating gene expression⁷⁵. A special feature of the glycine sensing module is that it occurs in a tandem arrangement where two aptamer domains (type I and II) are connected by a linker sequence⁷⁶. Both domains are able to bind to a separate molecule of glycine which, together with the linker, leads to a cooperative functionality of the riboswitch. This is mediated by an increased glycine affinity of domain I when it is already bound to domain II⁷⁷.

The third known class of amino acid sensing riboswitches binds glutamine and is almost exclusively found in Cyanobacteria⁷⁸. However, some representatives also occur in uncultivated marine bacteria. Most glutamine riboswitches are located in 5'-UTRs of genes involved in nitrogen metabolism. Thus, they regulate gene expression of many important proteins for the assimilation of nitrogen like Ntr-P_{II}, GS, GOGAT and ammonium transporters. In Cyanobacteria glutamine aptamers are also found upstream of *gifB* that encodes IF17 an inactivating protein of GS activity. Glutamine riboswitches seem to occupy a major role in the regulation of nitrogen metabolism systems.

But not only direct amino acid sensing contributes to the regulation of amino acid metabolism by riboswitches. The (p)ppGpp riboswitch is mainly found upstream of branched chain amino acid (BCAA) biosynthesis genes⁷⁹. Synthesis of BCAAs in many bacteria is under direct control of the stringent response alarmone (p)ppGpp. In case of the (p)ppGpp riboswitch increasing levels of the alarmone due to amino acid starvation lead to biosynthesis of BCAAs by an on-switching mechanism.

Amino acid metabolism regulation by RNA-binding proteins (RBPs)

Besides intrinsic termination of transcription as seen in riboswitches or ribosome mediated regulation by attenuators, bacteria possess a wide subset of RBPs that regulate gene expression. One of the best studied examples is the Rho-dependent termination. Rho is a homohexameric ring like protein that has two RNA-binding sites per subunit. It binds RNA at Rho-utilization (*rut*) sites. However, consensus sequences for *rut* have not yet been identified for the exception that most binding sites are cytidine rich. Furthermore, the exact mechanism of Rho-dependent regulation remains elusive. Rho-dependent regulation was linked to about 25% of *E. coli* operons, where it acts in diverse roles

including gene silencing or suppression of replication-transcription conflicts^{80,81}. In many cases Rho-dependent regulation is linked to intrinsic terminators as can be seen for riboswitches⁸². Moreover, attenuators are often associated with Rho-dependent regulation. This can be observed for the regulation of the *tna* operon responsible for tryptophan catabolism in *E. coli*. In this attenuation mechanism, a leader transcript containing a tryptophan codon is followed downstream by a *rut* sequence⁸². When the RNAP is stalled at the pause site, the ribosome starts translating the leader peptide. Under tryptophan starved conditions the ribosome dissociates at the tryptophan codon of the leader peptide. Rho now binds to the *rut* site and interacts with the stalled RNAP which leads to termination of transcription. When tryptophan levels are high the ribosome does not dissociate and thus blocks the *rut* site resulting in transcription resumption of the RNAP.

Another RBPs that is involved in tryptophan metabolism is TRAP (*trp* RNA-binding attenuation protein). The β -wheel folded TRAP is composed of 11 subunits each of them able to bind tryptophan and is involved in transcription attenuation of the *trp* operon in *Bacillus subtilis*⁸³. Each of the subunits can bind one of the 11 GAG or UAG triplets of the leader transcript upstream of *trp*⁸⁴. Under tryptophan limited conditions TRAP is not activated by the amino acid and cannot interact with the leader transcript. This enables the RNA to form an antiterminator hairpin which leads to resumed gene expression of the downstream ORFs. When tryptophan levels are high the TRAP-tryptophan complex interacts with the tryptophan codons of the leader transcript. As a consequence, a transcription terminator hairpin is formed resulting in dissociation of the RNAP and premature transcription termination. Furthermore, it was proposed that TRAP interacts directly with the RNAP and other cellular factors involved in transcription termination^{85,86}.

The complexity and variety of gene expression regulation on RNA level involving RBPs but also intrinsic regulators lead to the suggestion that more novel RNA based regulation mechanisms are still waiting to be explored. Given the importance of amino acid metabolism, it is suggested that more ncRNAs will be identified which are triggered directly by amino acids. This PhD thesis includes the characterization of the *ilvH* putative RNA motif that we show to be involved in the regulation of BCAA metabolism.

Chapter 2

Biochemical Description of a Novel Lysine Degradation Pathway in E. coli Proceeding via Glutarate and L2HG

Introduction

Degradation of L-lysine is wide-spread upon many organisms from bacteria to mammals. To date, 11 different catabolic pathways are depicted in the MetaCyc Database⁸⁷ that include a variety of reactions and intermediates. In mammals, lysine is degraded via L-saccharopine and 2-oxoadipate towards glutaryl-CoA which is further utilized to acetyl-CoA. Therefore, lysine is catabolized only in a ketogenic manner in mammals. All known bacterial lysine degradation pathways are proposed to end up in glutaryl-CoA and thus contribute to the acetyl-CoA pool. Interestingly, for *E. coli*, the best studied organism, a complete lysine catabolic pathway is still missing. The only lysine catabolizing reaction known for this organism is the decarboxylation of lysine to cadaverine^{88,89}. This reaction is performed by the acid induced enzyme CadA together with a lysine-cadaverine antiporter and was shown to be involved in the acid stress response of *E. coli*. Cadaverine is a polyamine involved in other stress response mechanisms. Additionally, *E. coli* encodes a constitutively expressed lysine decarboxylase named LdcC. It was already proposed in 1977 that cadaverine might be further metabolized to glutarate (GA), a C5 backbone dicarboxylic acid in pseudomonads⁹⁰. Lysine degradation to GA was revealed in *Pseudomonas putida* to be carried out via 5-aminovalerate (AVA) but without involving cadaverine. The enzymes responsible for the conversion of AVA to GA in *P. putida* are DavT, an AVA aminotransferase and DavD, a glutarate-semialdehyde (GSA) dehydrogenase. Interestingly, both enzymes can also use the C4 backbone equivalents, thus carrying out the reaction from γ -aminobutyrate (GABA) to succinate (SA)⁹¹.

This led our attention to the *E. coli* operon *csiD-lhgO-gabDTP*. The whole operon is specifically induced in stationary phase (carbon starvation) while the *gabDTP* genes are induced also separately in response to nitrogen starvation. Directly downstream of the operon, *csiR* encodes a transcription factor that represses the *csiD* operon⁹². In addition to CsiR, the *csiD* operon is controlled by cAMP-CRP, leu-LRP and H-NS^{93,94}. GabT and GabD encode for GABA aminotransferase and succinate-semialdehyde (SSA) dehydrogenase, able to perform the up mentioned reaction from GABA to SA. Usually, these enzymes are part of the catabolism of arginine via putrescine and GABA possibly

channeling succinate into the TCA⁹⁵. Upstream of *gabDT*, *lhgO* was shown to encode for a L-2-hydroxyglutarate (L2HG) oxidase conducting the oxidation of L2HG to α KG producing H₂O₂ as a byproduct⁹⁶. The function of the *csiD* (carbon starvation induced gene D) gene product remained unknown. CsiD belongs to the non-haem Fe(II)-dependent oxygenase family (protein family PF08943). A crystal structure of CsiD that was solved in a structural genomics effort suggested that CsiD functions as an α KG-dependent dioxygenase⁹⁷. Therefore, we assumed that CsiD performs hydroxylation of GA producing L2HG. To address the conversion from cadaverine to AVA, we hypothesized that two more enzymes, originated from the putrescine degradation pathway, (PatA/PatD) act promiscuously on the C5 diamine. This led to the proposal of a degradation pathway describing the complete catabolism of lysine to SA (Figure 6) of *E. coli* which is experimentally characterised and validated in the following subsections.

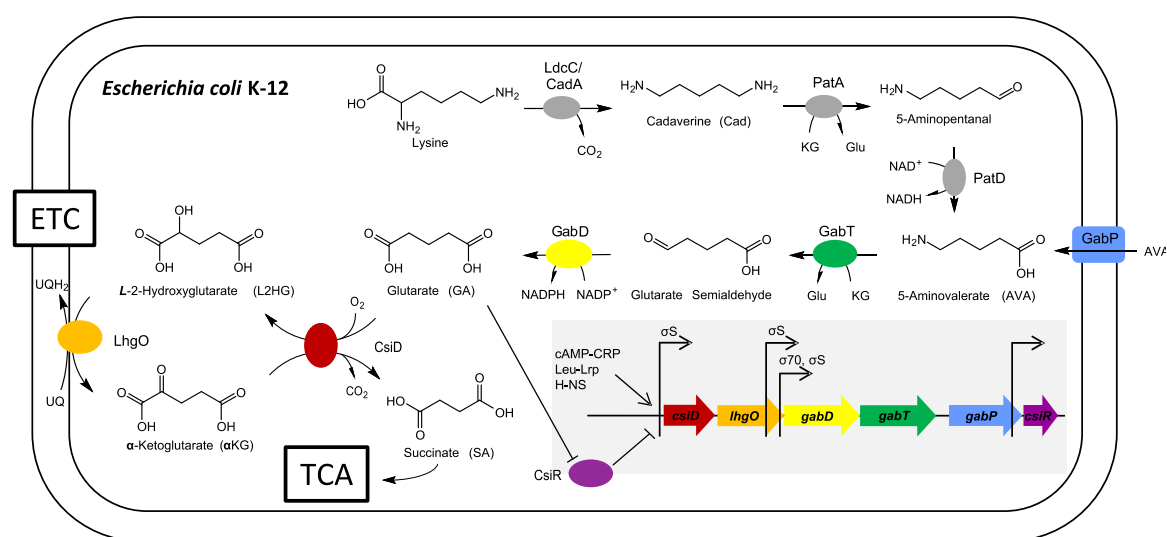


Figure 6: Catabolism of lysine to succinate in *Escherichia coli* K-12. Summary of the central metabolic pathway discovered in this study and of the corresponding gene cluster with its regulation (grey inset) in *E. coli*. We show that *E. coli* CsiD (carbon starvation induced protein D) functions as a α -ketoglutarate- dependent, CO₂- and succinate-forming glutarate hydroxylase, which produces L-2-hydroxyglutarate, and that *E. coli* LhgO (L-2-hydroxyglutarate oxidase) acts as a membrane-bound, ubiquinone-linked, α -ketoglutarate-producing L-2-hydroxyglutarate dehydrogenase. Effectively, glutarate is converted to succinate by these two reactions. We also show that *E. coli* feeds glutarate into these reactions from 5-aminovalerate by promiscuous GabT, when acting as 5-aminovalerate transaminase, and by promiscuous GabD, when acting as glutarate semialdehyde dehydrogenase. Thus, these GabT-GabD-CsiD-LhgO reactions link the central metabolism (TCA cycle) with the yet incompletely known lysine catabolism in *E. coli*, that is, with LdcC and CadA as lysine decarboxylases, PatA (putrescine transaminase) acting as cadaverine transaminase and PatD as 5-aminopentanal dehydrogenase^{88,89,98,99}. The catabolism of lysine by this pathway is specifically induced in the stationary phase of *E. coli* (see text). The inset shows the structure and regulation of the *csiD-lhgO-gabD-gabT* operon, which include also genes *gabP* encoding for a gamma-aminobutyrate (GABA) transporter that also imports the C5-homologue 5-aminovalerate¹⁰⁰ and *csiR* encoding for a ligand-dependent transcription factor. Arrows represent transcription start sites (TSS) of different transcripts. Transcription of the operon is known to be enhanced by regulators cAMP-CRP, leu-LRP and H-NS^{93,94}, and repressed by CsiR⁹².

Results & Discussion

Biochemical characterization of CsiD

Since CsiD might be the key enzyme that connects the degradation of lysine to the TCA, the enzyme was purified and NMR (Figure 7a, Supplementary Figure 1 & 2) and LC-MS experiments were conducted. CsiD was shown to indeed hydroxylate GA to 2-hydroxyglutarate, while the co-substrate α KG is converted to succinate (SA) (and CO₂), as is common for this enzyme class^{101,102}. By derivatisation of the product with diacetyl-L-tartaric anhydride¹⁰³ we demonstrated that L2HG is produced in a highly stereospecific manner; no D-2-hydroxyglutarate (D2HG) was detectable (Figure 7e). This finding is in accordance with the reported specificity for the L-enantiomer of the subsequent LhgO-catalysed reaction⁹⁶. When the reaction was measured with a Clark O₂ electrode, we determined a specific activity of 53 +/- 3 $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ and an apparent $K_m = 650 +/- 20 \mu\text{M}$ for GA and $K_m = 100 +/- 7 \mu\text{M}$ for α KG (Figure 7b, c). Other dicarboxylic acids (oxalate, malonate, SA, adipate, and pimelate tested) were not converted as the main substrate.

Other studies confirm the GA hydroxylating activity of CsiD in different bacterial species. Zhang *et al.*¹⁰⁴ determined a K_m of 145,7 +/- 1,5 μM for GA and a specific activity of 177 +/- 3 $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{mg}^{-1}$ for *P. putida* CsiD. Additionally, CsiD from *P. putida* was found inactive when AVA, SA, 2-aminoadipate, 2-oxoadipate, lysine or glutaryl-CoA were tested as main substrates. Another study confirmed inactivity with substrates tested in our study and added propionate, butyrate, pentanoate and hexanoate to the list of tested compounds¹⁰⁵. This reflects the high specificity for GA as the primary substrate of CsiD.

The physiological role of CsiD as GA-metabolising enzyme in the stationary phase of *E. coli* was confirmed when we tested its *csiD* knockout strain (Δ *csiD*) by LC-MS analysis of small-molecule extracts. With carbon starvation and entry into the stationary phase, the intracellular concentration of GA accumulated to much higher levels in the Δ *csiD* strain compared to the wildtype (Figure 7d). We were able to detect small amounts of GA already at early-exponential phase. GA has also been detected at mid-exponential phase when *E. coli* was grown in minimal medium supplemented with glucose as C-source¹⁰⁶.

Moreover, a *csiD* knockout strain of *P. putida* showed decreased fitness when grown on AVA as sole C-source^{104,107}

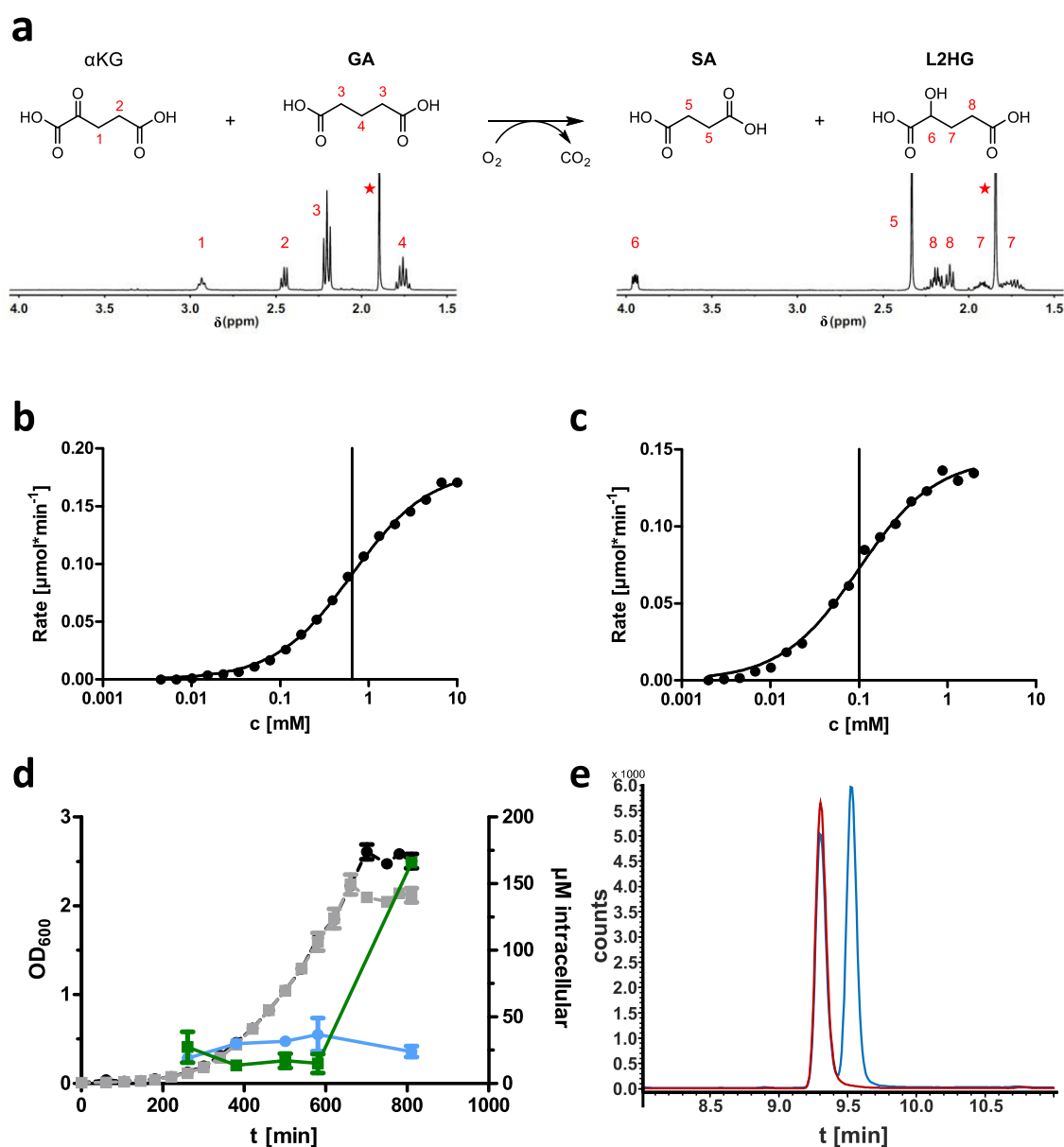


Figure 7. CsiD is an αKG -dependent dioxygenase converting GA to L2HG. **a** Reaction as catalysed by CsiD and its characterisation by NMR. CsiD converts GA and αKG to L2HG and SA, respectively (and carbon dioxide). ^1H -NMR spectra of educts (left) and products (right) after overnight incubation with CsiD in ammonium acetate buffer are shown, asterisks: signals from acetate buffer. **b, c**, Reaction rate, measured as consumption of O_2 in a Clark-type oxygen electrode, is plotted against the concentration of GA (**b**) and αKG (**c**). Data were fitted with Michaelis-Menten equation (solid line). KM is indicated by the vertical line. **d**, Intracellular GA concentrations in different growth phases of *E. coli* M9 cultures. Circles (WT *E. coli* strain) and squares (ΔcsiD *E. coli* strain) represent intracellular GA concentrations (blue: WT, green: ΔcsiD , right y-axis). Bacterial cell growth is indicated by optical density at 600 nm (OD₆₀₀) in black and grey (WT and ΔcsiD , left y-axis). Shown are means of biological duplicates with error bars representing standard deviation including technical duplicates. **e**, *L*-2-hydroxyglutarate produced by CsiD and in a mixture with *D*-2-hydroxyglutarate was derivatized with DATAN¹⁰³ and analyzed by LC-MS. Chromatograms of the derivatized reaction product (red) and the mixture (blue) were recorded in negative SIM at a m/z ratio of 363. The reaction product elutes as a single peak clearly separated from the second peak that represents derivatised *D*-2-hydroxyglutarate.

Structural and functional characterisation of CsiD

Both substrate (GA) and co-substrate (α KG) of CsiD are structurally similar which makes the GA hydroxylation reaction interesting on a mechanistic level. Furthermore, the product of the reaction, L2HG, is known as an oncometabolite inhibiting α KG-dependent dioxygenases such as TET- and JmjC-type demethylases^{108,109}. We observed weak inhibition of CsiD by its product L2HG at concentrations of 5 mM (Figure 8a) in contrast to the aforementioned representatives of the same enzyme class. The reported K_i of L2HG for α -KG dependent dioxygenases is 628 \pm 0.036 μ M, in contrast the estimated K_i of L2HG for CsiD is > 5 mM. The estimated K_i for NOG and D2HG match the reported K_i for α KG-dependent (Figure 8a)¹⁰⁸.

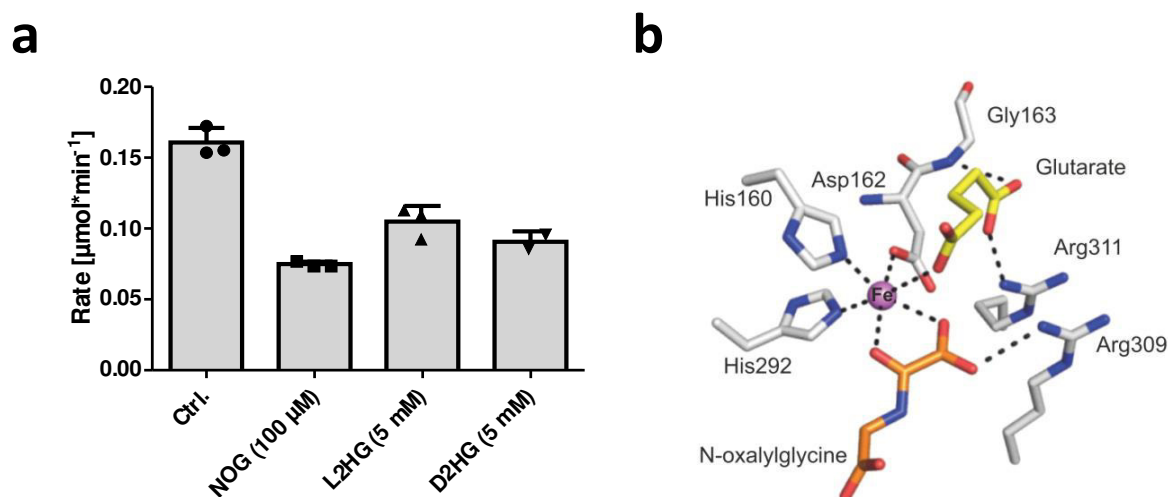


Figure 8: CsiD inhibition and active site. **a**, Reaction rate was determined in a Clark-type oxygen electrode monitoring the oxygen consumption. The rate was determined in the presence of the α KG-dependent dioxygenase inhibitor N-oxalylglycine (NOG), L2HG, D2HG. In the control reaction (Ctrl.) no inhibitor was present. Data represent the mean of triplicate measurements with error bars representing standard deviations. (D2HG; n=2). **b**, Crystal structure of CsiD. The active site of a protomer with substrate GA (yellow) and the α KG co-substrate analog N-oxalylglycine (orange) is shown. The model is a reconstruction obtained by superimposing individual crystal structures of CsiD bound to each of the ligands (the CsiD enzyme is equivalent in both structures; RMSD = 0.16 Å). The Fe ion is shown in magenta. Dashed lines represent interactions between ligands and residues in the binding pocket.

In order to clarify the binding mechanisms of CsiD, atomic structures of the enzyme in complex with its substrate (GA), its product (SA) as well as in complex with NOG were solved by X-ray crystallography (Figure 8b, Figure 9). The obtained crystal form (with symmetry P4₂12) contains two molecular copies of CsiD in its asymmetric unit. These two non-crystallographic copies are identical (RMSD = 0.082 Å). The biological tetrameric form of CsiD is generated by the symmetry of the crystallographic lattice, as it

was previously the case in structures lacking the by then unknown substrate^{97,110}. The enzyme protomer adopts a distorted jelly-roll fold composed of a β -sheet core flanked by α -helices, as previously described. The iron ion is bound to the active site of CsiD by residues His160, Asp162, and His292 and three solvent molecules that complete an octahedral coordination sphere (Figure 9a). In ligand-bound structures, one or more solvent molecules become replaced by interacting oxygen atoms from the ligands.

Ligand-bound CsiD structures revealed two binding sites located at opposite sides of the Fe^{2+} ion (Figure 8b and Figure 9b-d). In the substrate site (here termed site I), GA is directly coordinated by the Fe^{2+} ion via an oxygen atom from one of its terminal carboxyl groups. The other carboxyl group forms a salt bridge with a conserved residue Arg311 and the main chain nitrogen of Gly163 (Figure 9b). The co-substrate site (here termed site II) is seen in crystal structures in this work occupied by (a) the non-processable α KG-analogue NOG (where atom C3 is substituted by nitrogen compared to α KG) or (b) the co-product SA. NOG binds the Fe^{2+} ion via the oxo- and a terminal oxygen atom of its oxalyl moiety as well as forming a salt bridge with residue Arg309 (Figure 9c). SA is bound to Fe^{2+} via a terminal carboxyl group in a bidentate fashion (Figure 9d) and further interacts with Arg309 through a low-occupancy solvent molecule that occupies the site that is generated by decarboxylation of α KG to SA, mimicking the interaction to the NOG oxygen in that complex. Conformational changes in the CsiD protein were not observed in any of the complexed structures, independently of whether ligand site I or site II was occupied. Importantly, in structures with site I or II occupied, always the cognate ligand for the respective site was observed with the other site unoccupied. This finding demonstrates a high degree of specificity of CsiD regarding the binding of the structurally similar substrate, co-substrate, and products.

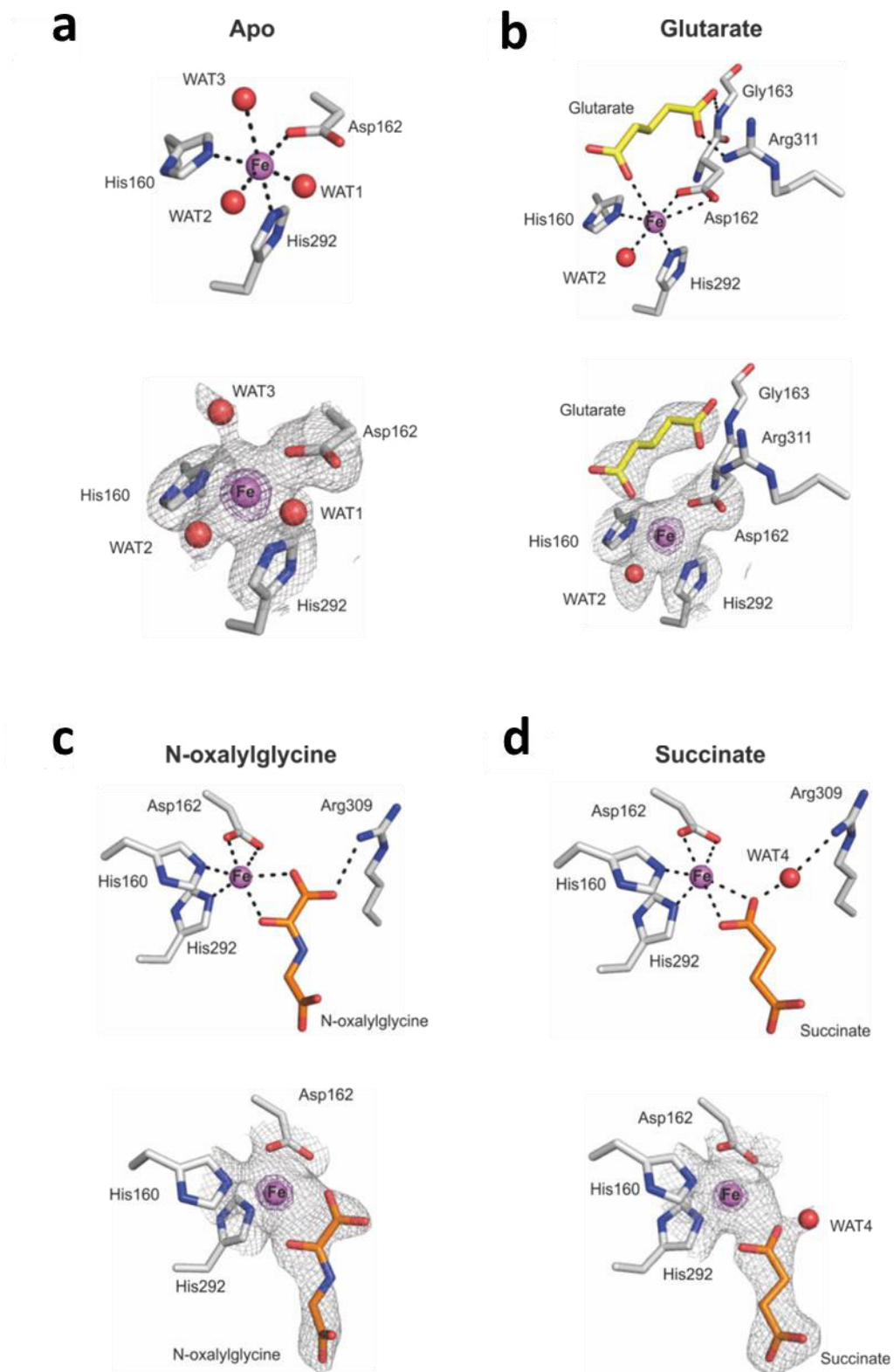


Figure 9: Ligand complexation by CsiD. Detailed view of CsiD in complex with its ligands. Upper panels of **a**, **b**, **c**, and **d**, show interactions (dashed lines), lower panels display $(2F_o - F_c)$ electron density maps contoured at 1σ (grey) and 7σ (purple).

Most representatives of the enzyme class coordinate α KG via two arginine residues at the active center. One arginine forms a hydrogen bond with the proximal carboxylate of α KG and a second one interacts with the distal carboxylate of the co-substrate via salt bridging¹¹¹. For CsiD, we only observed one salt bridge between Arg309 and the proximal carboxylate of NOG. Thus, we hypothesized that binding of α KG by CsiD might be different compared to other members of the enzyme family. This might lead to an expansion of possible co-substrate or inhibitors for CsiD. We measured the effect of different α -ketoacids with and without a distal carboxyl group on enzyme activity of CsiD in a Clark O₂ electrode. We could not determine any activity with α -ketopimelate, oxalate, oxaloacetate, α -ketovalerate and α -ketobutyrate as co-substrates. Only α -ketoadipate was found to have weak activity as an alternative co-substrate (unquantified observations) which was shown earlier for members of the non-haem Fe(II)-dependent oxygenase family^{105,112} and recently for CsiD of *P. putida*¹¹³. Furthermore, we did not detect any inhibitory effect when above mentioned metabolites were inserted as competitors together with α KG in the CsiD reaction. These results hint for a more classical coordination of α KG in the active centre confirming the high co-substrate specificity that is also seen in other members of this class. By re-evaluating our crystal structures, Herr *et al.* (2019)¹⁰⁵ argued that the distal carboxylate of α KG is possibly coordinated by very long hydrogen bonds of 4.0 Å with Arg294 and Arg305. This might explain the strong discrimination between α KG and other co-substrates observed in our results. Still, it remains elusive why L2HG as inhibitor of other enzymes of this class is only showing weak inhibition of CsiD activity.

Taxonomic distribution of CsiD

New insights into the binding mechanism of CsiD obtained by crystal structures might also allow the re-evaluation of the taxonomic distribution of the enzyme. At first, we performed a standard pBLAST search based on the *E. coli* K12 CsiD amino acid sequence retaining sequences under a certain coverage and identity threshold (Figure 10a). We obtained a distribution of possible CsiDs mainly in β - and γ -proteobacterial species, but also many Firmicutes encode CsiD on their genomes. However, under the artificial threshold more possible CsiD candidates emerged, mainly *Aspergilli*, which are also listed in the pfam database¹¹⁴ as CsiD.

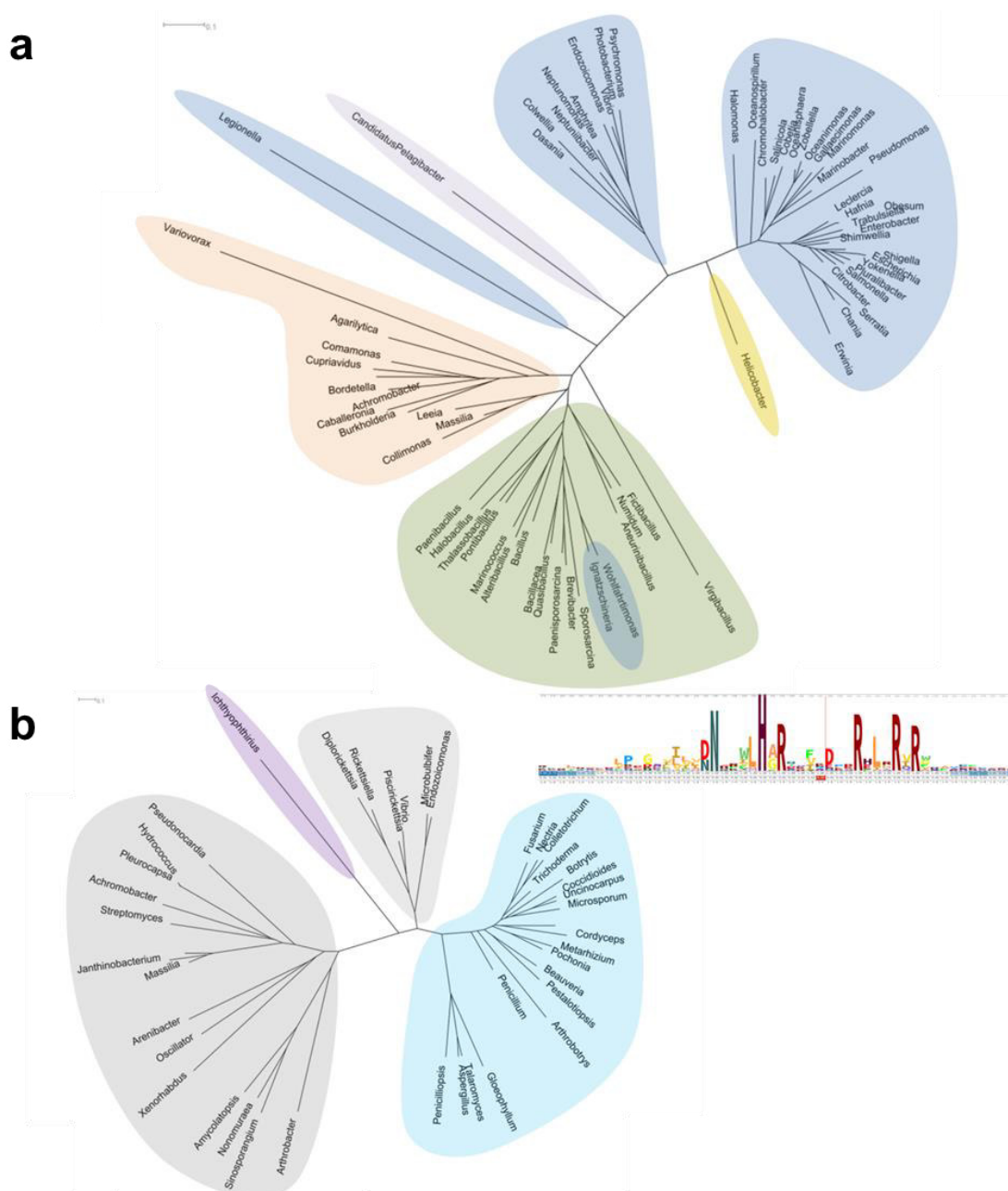


Figure 10: Phylogenetic distribution of CsiD. **a**, A neighbourhood joining phylogeny of CsiD based on *E. coli* K-12 CsiD was built retaining sequences with more than 70% coverage and 25% identity. Homologs of *E. coli* CsiD were searched in the refseq database (last access May 2018)¹ using pBLAST. Bacterial classes are as indicated: α - (purple), β - (orange), γ - (blue), and ϵ - (yellow) proteobacteria and bacilli (green). **b**, Neighbourhood joining phylogeny of sequences obtained by phiBLAST search with *E. coli* K12 CsiD as query using the defined search pattern (RXLXRXR). The tree shows eubacterial genes (grey) and fungal genes (blue) that carry potential CsiD candidates identified by the phiBLAST search. The inset displays a logo hidden Markov model based on a sequence alignment obtained by all BLAST searched sequences.

We next had a closer look at the conservation rates of amino acid residues involved in the active centre binding of GA. Relevant arginines that might execute short- (Arg309, 311) or long- (Arg294, 305) distance interactions with GA or α KG are 100% conserved in the alignment obtained with our pBLAST search. We therefore initiated a Pattern Hit

Initiated (phi) BLAST search using the conserved residues as entry. We obtained sequences from more bacterial species that were not included in the original alignment (Figure 10b). Furthermore, many fungal sequences under these *Aspergilli* and *Penicillia* showed up. All received sequences might be CsiD candidates. They contain all conserved residues for the binding of Fe(II) and α KG and thus belong to the class of non-haem Fe(II) dependent dioxygenases.

Furthermore, the conserved pattern with residues eventually involved in GA binding (Figure 10b inlay) was present in these potential CsiD candidate sequences. The most distinct sequence related to *E. coli* CsiD from the original alignment was from *Legionella pneumophila* with 31% identity. However, query coverage of this sequence was still at 93% which resembles the phylogenetic relationship between these two sequences. All sequences obtained from the phiBLAST only had around 50% coverage compared to the *E. coli* protein. In most sequences, only the C-terminal half, containing the dioxygenase determining residues, were related to CsiD. This overall dissimilarity might contradict the assumption that sequences from the phiBLAST could function as GA hydroxylase. However, it might be reasonable to experimentally validate some of the potential CsiD candidates to see whether similarities in substrate specificities can be observed.

A pair of aminotransferase/dehydrogenase actions mediates catabolism from lysine to glutarate

In the next step we wanted to find out which enzymes are responsible for the production of GA. Therefore, we conducted *in vitro* experiments with the purified recombinant *E. coli* enzymes GabT and GabD. We hypothesized whether GabT may convert also the C5-homolog AVA to GSA and whether GabD converts GSA to GA (Figure 6). In a coupled enzyme assay we were able to demonstrate that GabT and GabD use AVA with comparable efficiency to GABA ($K_M^{GABA} = 197 \pm 27 \mu\text{M}$, $K_M^{AVA} = 439 \pm 29 \mu\text{M}$) (Figure 11b). The reaction was additionally analysed via high-resolution mass spectrometry. In here we detected both the reaction product of GabT (GSA) and GabD (GA) when AVA was inserted as substrate (Figure 11a). This confirms the promiscuous action of both enzymes originally involved in the GABA metabolism.

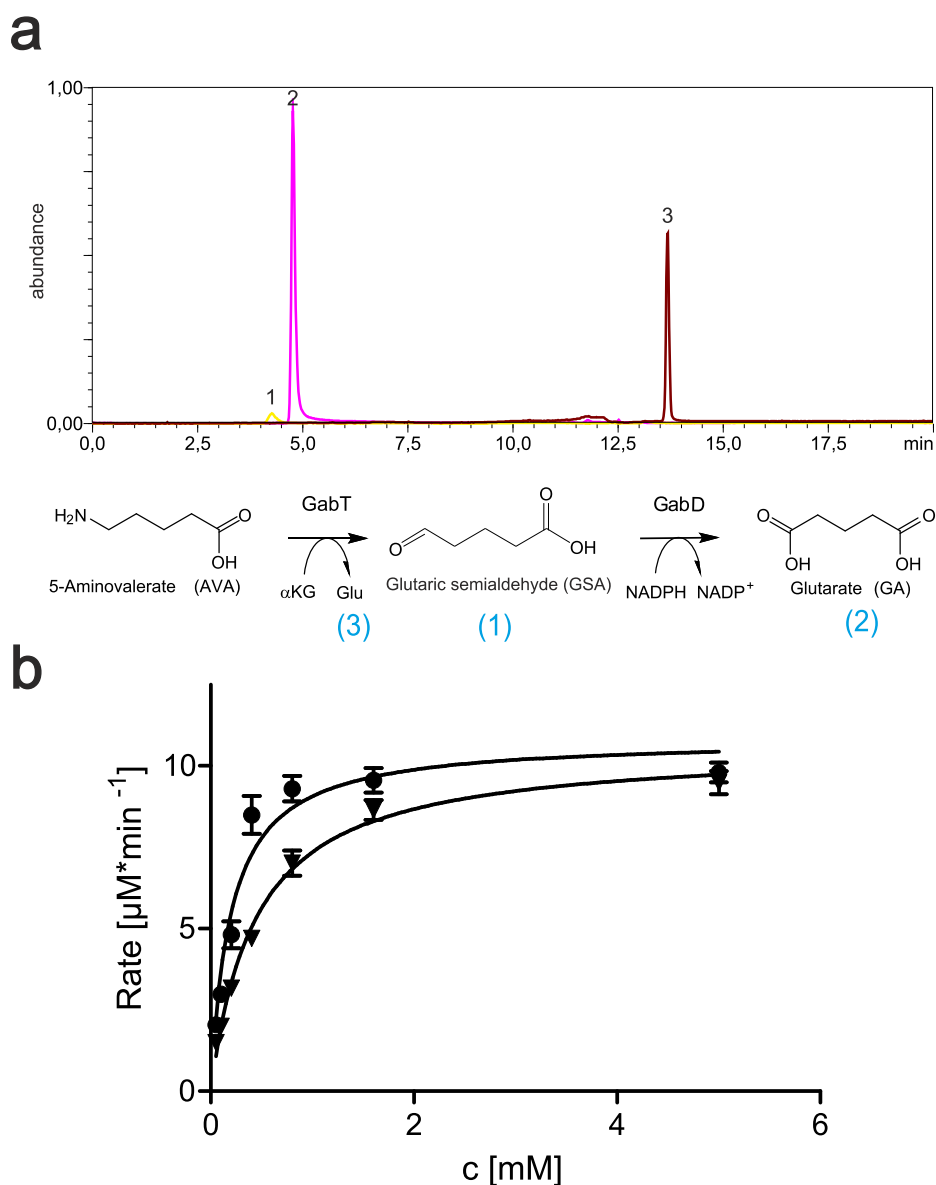


Figure 11: Promiscuous activity of enzyme pair GabT/D. **a**, LC/MS chromatogram of GabT/D enzyme reaction products. (1) Glutarate semialdehyde ($m/z = 115.0$; RT = 4.3 min; yellow). (2) Glutaric acid ($m/z = 131.0$; RT = 4.9 min; magenta). (3) Glutamic acid ($m/z = 148.1$; RT = 13.7 min; brown). **b**, Coupled GabT (GABA transaminase) and GabD (succinic semialdehyde dehydrogenase) reaction: Comparison of the conversion of the C4 substrate GABA (circles) to SA with the C5 substrate AVA (triangles) to GA. Reaction rate ($\mu\text{M}/\text{min}$) is shown as a function of substrate concentration. Rate is given as $\mu\text{M NADP}^+/\text{NAD}^+$ reduced per min, monitoring the dehydrogenase activity. Data represent the mean of triplicate experiments with error bars representing standard deviations.

We further suspected that decarboxylation of lysine to cadaverine, followed by processes like the degradation of putrescine lead to the formation of AVA in *E. coli*. Similar reactions have already been described in Pseudomonads⁹⁰. Two *E. coli* lysine decarboxylases (CadA, LdcC) are responsible for the formation of cadaverine. The gene product of *pata* was originally described as putrescine aminotransaminase catalysing the

reaction from the C4 diamine to 4-aminobutanal. PatD is responsible for the formation of GABA from 4-aminobutanal^{98,115,116}. It has already been shown that the *E. coli* transaminase PatA can process cadaverine with the same activity as putrescine⁹⁸. Since there is no data available concerning PatD producing AVA in *E. coli*, we compared coupled PatA/D activities for putrescine and cadaverine as described for GabT/D before. The coupled PatA/D reaction with cadaverine showed a K_M^{Cad} of 0.37 ± 0.05 mM and a V_{max} of $11.8 \mu\text{M}\cdot\text{min}^{-1}$, whereas PatA/D reaction with putrescine revealed a higher K_M^{Put} of 1.43 ± 0.07 mM and a higher V_{max} of $29.1 \mu\text{M}\cdot\text{min}^{-1}$ (Figure 12d). Furthermore, we confirmed the production of 5-aminopentanal (APA), which spontaneously cyclises to piperidine catalysed by PatA (Figure 12a), via mass spectrometry analysis. In a coupled PatA/D reaction we detected AVA as the final product as well as the reaction intermediates APA and piperidine (Figure 12b).

Interestingly, tetrahydroanabasine was also found under the reaction products which is formed by the dimerization of piperidine (Figure 12b, c). Our findings are further substantiated by the observation that overexpression of *ldcC*, *patA*, and *patD* from *E. coli* can be used for the production of AVA from lysine in *C. glutamicum*⁹⁹. An alternative sequence of reactions that could produce AVA from lysine degradation is the putrescine utilization pathway (*puuABCD*) of *E. coli*, which is proposed to degrade extracellular putrescine¹¹⁷.

LhgO couples the degradation pathway to the respiratory chain

At next, our attention was guided to the second gene of the operon, *lhgO*. LhgO was originally described as a cytosolic FAD-dependent L2HG oxidase using molecular oxygen to produce αKG and H_2O_2 ⁹⁶. We were suspicious concerning the possible function of LhgO since production of hydrogen peroxide in a carbon metabolism coupled pathway seemed unusual. Moreover, homologs of LhgO are very widespread not only in bacterial species. Most archaea and eukaryotes carry the enzyme where it is then often described as L2HG dehydrogenase (L2HGDH). Protein conservation is surprisingly high between *E. coli* LhgO and human L2HGDH with 41% identity over the full amino acid sequence^{118,119}. L2HGDH of mammals and plants were shown to localize to the inner membrane of mitochondria where for the latter a coupling to the electron transport chain

and within that to respiration was proposed. We further analysed the amino acid sequence of *E. coli* LhgO with a prediction algorithm for transmembrane helices in proteins. We found that the first 3 residues of LhgO had a higher probability to be located inside the periplasm together with ~25 N-terminal residues that might be part of transmembrane helices. However, the overall algorithm did not annotate LhgO as a transmembrane protein. Indeed, respiratory chain coupled single dehydrogenases can be found in *E. coli* that are associated with the membrane by only one or few hydrophobic stretches like glycerol-3-P (GlpD), lactate (DldD, LctD) or pyruvate (PoxB) dehydrogenases¹²⁰. It was shown that coupling of these dehydrogenases to the membrane is needed to donate electrons to quinones.

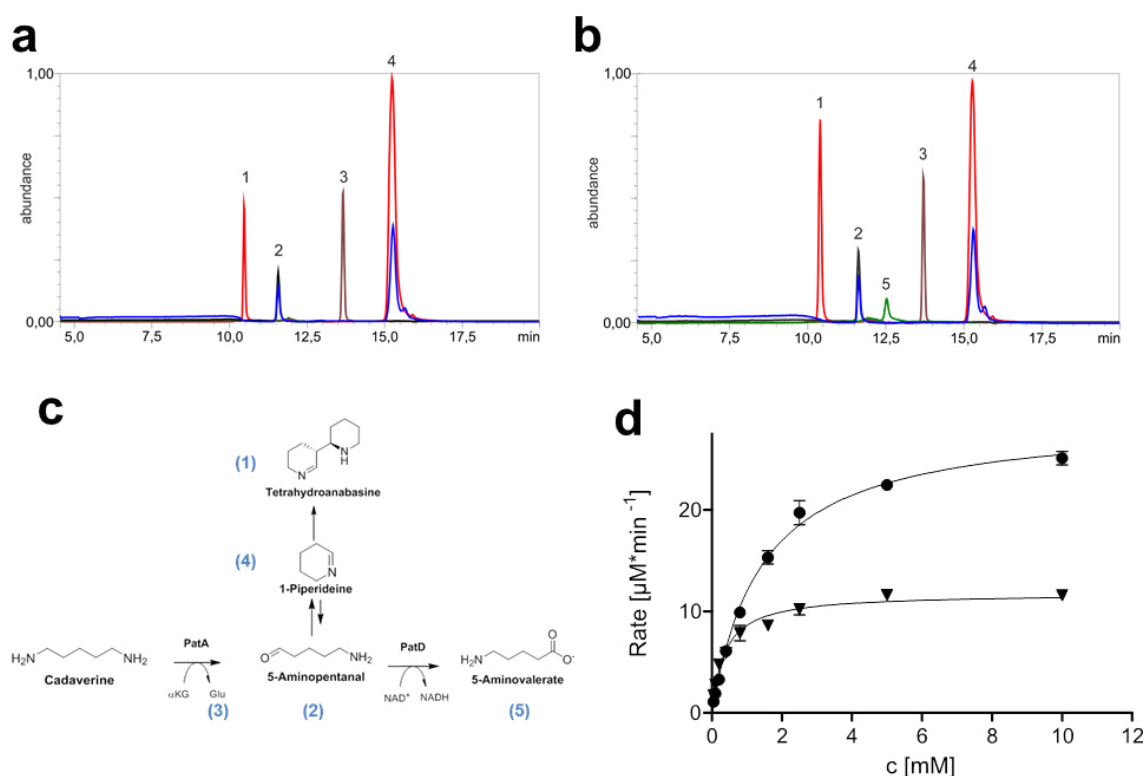


Figure 12: Coupled enzyme pair PatA/D catabolises cadaverine to AVA. **a**, PatA single reaction. PatA transfers the amino group of cadaverine to α KG, producing 5-aminopentanal (2) and glutamic acid (3). **b**, in a coupled PatA/D reaction 5-aminopentanal (2) is oxidised to 5-aminovaleate (5) by PatD. Aminopentanal cyclized spontaneously to 1-piperideine (4) during MS analysis (likely during ionisation), which spontaneously dimerises to yield tetrahydroanabasine. For this reason, 5-aminopentanal is detected at RT=11.6 min together with 1-piperideine. At RT=15.3 min the 1-piperideine monomer is detected together with its dimer. **c**, reactions catalyzed by PatA/D and spontaneous reactions of 5-aminopentanal. (1) Piperideine dimer (Tetrahydroanabasine) ($m/z=167.2$; RT=10.5 min; red). (2) 5-aminopentanal ($m/z=102.1$; RT=11.6 min; black). (3) glutamic acid ($m/z=148.1$; RT=13.7 min; brown); (4) 1-Piperideine ($m/z=84.1$; RT=15.3 min; blue); (5) 5-Aminovaleate ($m/z=118.1$; RT=12.5 min; green). **d**, Coupled PatA (putrescine transaminase) and PatD (γ -aminobutyraldehyd dehydrogenase) reaction: Comparison of the conversion of the C4 substrate putrescine (circles) to GABA with the C5 substrate Cad (triangles) to AVA. Reaction rate ($\mu\text{M}/\text{min}$) is shown as a function of substrate concentration. Rate is given as $\mu\text{M NADP}^+/\text{NAD}^+$ reduced per min, monitoring the dehydrogenase activity. Data represent the mean of triplicate experiments with error bars representing standard deviations.

We were not able to detect oxygen consumption of LhgO in the presence of L2HG even though we could accomplish activity of the enzyme by reduction of 2,6-chloroindophenol (DCPIP) in the same assay (DCPIP turns from blue to colourless when reduced). Therefore, we decided to test whether the *E. coli* LhgO may be a dehydrogenase that feeds the high-energy electrons (reported $E^0 = 19 \pm 8$ mV)⁹⁶ from the oxidation of L2HG via ubiquinone to the respiratory electron transport chain (Figure 6 & Figure 13a). Purified recombinant LhgO directly reduced ubiquinone (Figure 13b) in a L2HG-dependent manner with a specific activity of 0.33 ± 0.002 $\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{mg}^{-1}$ ($K_M^{\text{L2HG}} = 162$ μM ; $V_{\text{max}} = 5.1$ $\text{nmol} \cdot \text{min}^{-1}$). Enzyme activity was not altered by supplementing FAD to the reaction. Together with the yellow colour of purified LhgO this indicates that bound FAD was co-purified saturating the reaction. We separated membrane bound proteins from soluble cytosolic fractions by ultracentrifugation. Activity of native LhgO in *E. coli* was found exclusively in the membrane fraction but not in the soluble protein fraction (Figure 13c). A ΔlhgO strain was treated the same and completely lacked the activity. Additionally, redox activity of membranes upon L2HG addition was restored in a *lhgO+* complementation strain, but the activity was lower compared to WT *E. coli* membranes. This may be explained by inefficient incorporation of recombinant LhgO into the membrane, possibly caused by the His-Tag. Furthermore, we were able to show that the natural quinones of *E. coli* can serve as electron acceptors for LhgO. Membrane fractions of *E. coli* WT were incubated with L2HG or SA (testing succinate dehydrogenase as positive control) as substrates for the respective dehydrogenases. The ratio between ubiquinol to ubiquinone increased in the presence of both SA and L2HG compared to the control reaction without added substrate, indicating that electron transfer from these substrates to ubiquinone occurred in the membrane fractions (Figure 13d). The ubiquinol/ubiquinone pool was unaffected by L2HG in a membrane fraction prepared from an *E. coli* strain lacking LhgO (ΔlhgO). We exclusively detected menaquinone and demethylmenaquinone in their oxidised but not reduced forms. However, many ETC-coupled dehydrogenases in *E. coli* show a wider substrate range for various quinone species and also under aerobic conditions, naphthoquinones can be used as electron acceptors^{120,121}. Notably, since menaquinole and demethylmenaquinole reoxidise quickly¹²² and might not be detectable in our assay, we cannot exclude that LhgO may accept also these quinones as electron acceptors. However, when 2,6-chloroindophenol (DCPIP) was used as artificial electron acceptor both native LhgO in the membrane

fraction and recombinant LhgO revealed higher activities in the presence of ubiquinone₁ compared to menaquinone₄ (Figure 13e). Addition of the respiratory chain inhibitor 2-heptyl-4-hydroxyquinoline n-oxide (HQNO) in concentrations higher than 80 μ M to membrane fractions abolished O₂ consumption and DCPIP activity, again confirming coupling of LhgO to the respiratory chain (Figure 13f). Taking into account that for *E. coli* ubiquinone is the main electron acceptor under aerobic conditions¹²⁰, these results support the conclusion that LhgO acts as a quinone-dependent oxidoreductase located in the cytoplasmic membrane.

It still remains surprising that our determination of LhgO as dehydrogenase is discrepant to the former characterisation as oxidase. Herr *et al.* (2019)¹⁰⁵ argued that differences in the used overexpression plasmid might have led to these diverging results. Another explanation could be differences in the purification method of recombinant LhgO. Possibly, membrane integrity of LhgO was altered due to different methods used. In our case purified LhgO would have remained in a more intact membrane structure leading to the observed dehydrogenase function coupled to ubiquinone. In the purification by Kalliri *et al.* disruption of the membrane would result to LhgO functioning as an oxidase. However, purification processes in our study compared to Kalliri *et al.*⁹⁶ are very similar except for adding FAD to the cell lysis buffer and reaction buffer by Kalliri. Eventually, excess amounts of FAD alter enzyme functionality in a way that supports oxidase activity *in vitro*. Nonetheless, we were able to allocate LhgO as part of the *E. coli* membrane whereas Kalliri *et al.* did not determine a specific location of the enzyme experimentally. Furthermore, the reduction of cell extract derived ubiquinone together with the membrane association hints for a dehydrogenase function *in vivo*. Observed oxidase activity might therefore present a side reaction of LhgO.

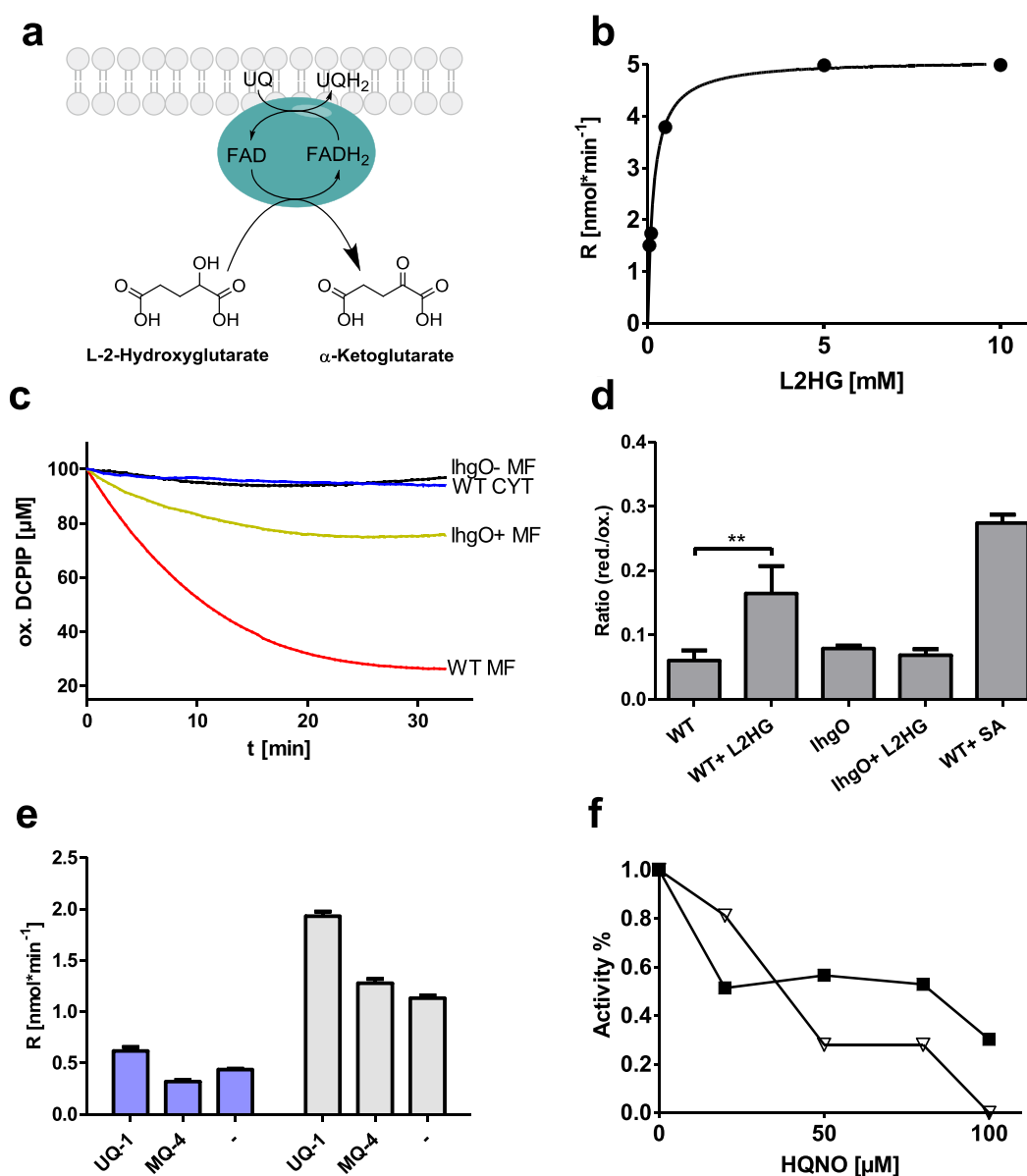


Figure 13: LhgO as a membrane-associated, L2HG:quinone oxidoreductase in *E. coli*. **a**, Reaction scheme of LhgO dehydrogenase activity including cofactors FAD and UQ. **b**, Reaction rate (R) of purified LhgO was measured at different L2HG concentrations and is given as nmol UQ1 reduced per time. **c**, Activity of *E. coli* WT membrane fraction (red; WT MF) transferring electrons from L2HG to DCPIP compared to a WT cytosolic fraction (blue; WT CYT), a membrane fraction isolated from a $\Delta lhgO$ knockout strain (black; $lhgO^-$ MF) and a complementation strain of $\Delta lhgO$ with recombinant LhgO (green, $lhgO^+$ MF). Depletion of the oxidized form of DCPIP over time was measured photometrically. **d**, Ubiquinol to ubiquinone ratio (red./ox.) was measured in *E. coli* membrane fractions in the presence or absence of L2HG or SA. Quinones were extracted from membrane fractions from either *E. coli* WT or $\Delta lhgO$ after the incubation with either L2HG or SA and measured by HPLC. Error bars represent standard deviations of triplicate experiments. Significance was assessed by an unpaired one-tailed Student's t -test (** P value = 0.0083). **e**, Reaction rate (R) of purified LhgO (blue columns) and *E. coli* WT membrane fraction (grey columns) measured by DCPIP reduction per time in dependence of ubiquinone 1 (UQ1) or menaquinone-4 (MQ4) after addition of L2HG, (-) no quinone added. Error bars represent standard deviations of triplicate experiments. **f**, Inhibition of *E. coli* WT membrane activity by HQNO titration. HQNO was added before L2HG and rates were normalised to the activity without added HQNO. DCPIP reduction (spectrophotometrically) is indicated with black squares and oxygen consumption of membranes (Clark electrode) by white triangles.

Complete pathway validation by metabolic tracing

Since we successfully revealed the enzyme functions of the proposed pathway in Figure 6 we further validated it by tracing fully ^{13}C - and ^{15}N -labelled lysine in growth experiments. Fully isotope labelled amino acids present a valuable tool for the determination of metabolic pathways because clear mass shifts are easier to detect by MS based methods. However, tracing the labelled carbon backbone of lysine is challenging in *E. coli*, since it cannot grow on this amino acid as sole C-source. We therefore supplemented standard minimal glucose containing medium with isotope-labelled lysine during growth. Since the complete *csiD* operon is induced under carbon starvation we argued that possible lysine catabolism intermediates might be predominantly produced at late growth phases. Thus, metabolite extraction of *E. coli* cultures was only conducted after cells reached mid- to late-stationary phase (24 hours) when glucose as C-source should be completely exhausted.

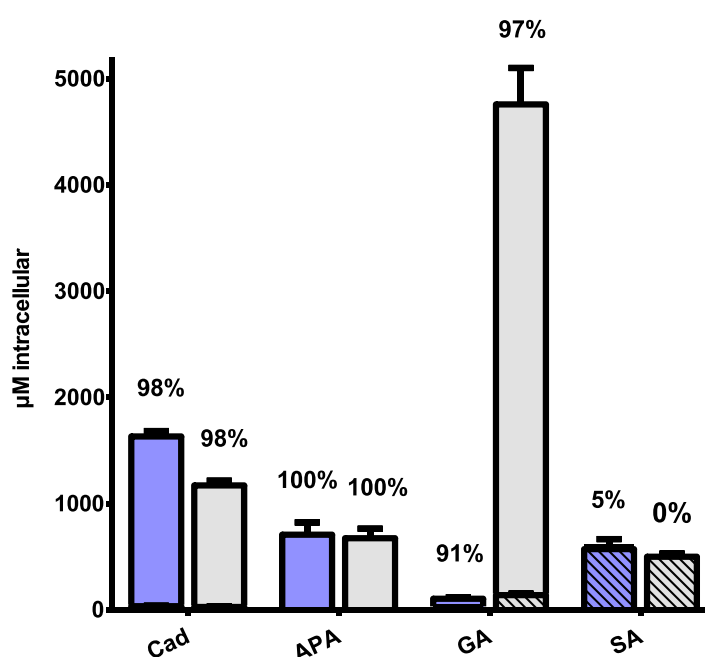


Figure 14: 5 Isotope tracing of the lysine degradative pathway in *E. coli*. Intracellular concentrations of cadaverine (Cad), 5-aminopentanal (APA), glutarate (GA) and succinate (SA) of *E. coli* K-12 wt (blue) and Δ csiD (grey) in stationary phase cultures. Non-isotope-labelled compounds are indicated as shaded bars. The sum of ^{13}C -labelled metabolites are shown as blank bars. Fractions of labelled compounds relative to the sum of labelled and non-labelled compounds are indicated above the respective bar (in percent). For isotope labelling patterns see Supplementary Table 3. Data represent the means of biological triplicates with error bars representing standard deviations. Each replicate was measured at least once.

Indeed, we found high intracellular concentrations of cadaverine, APA, and GA predominantly in the completely labelled form with more than 90% labelling of cadaverine, APA and 86% of GA (Figure 14). The high degree of labelling indicates that lysine is the main source of these metabolites under conditions investigated.

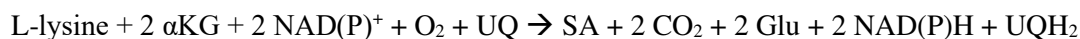
Fully ^{13}C -labeled AVA was only detectable at 640 μM in the ΔgabT strain. AVA could neither be detected in WT or ΔcsiD cells, indicating that AVA is rapidly utilized under conditions investigated. Of note, we could also detect ^{13}C -labeled GA in a ΔgabT strain indicating that other transaminases besides GabT are able to process AVA. In *E. coli*, there are isoenzymes for both GabT and GabD known, namely the GABA transaminase PuuE and the succinate semialdehyde dehydrogenase Sad. We demonstrate for a coupled PuuE/Sad reaction with purified enzymes that, similar to the findings with GabT/D (Figure 11), AVA can be used as a substrate (Supplementary Figure 3).

Although GA was detected only at 180 μM in WT *E. coli*, GA levels accumulated to almost 5 mM in a ΔcsiD strain. Merely 3 % of the accumulated GA was unlabelled indicating that the majority of GA was produced from imported lysine. The proposed catabolic pathway in our study provides evidence that lysine is metabolised to αKG and SA. Both metabolites can be channelled into the TCA and lysine degradation might therefore contribute to energy production under starved conditions. αKG and SA were detected predominantly in non-labelled form which is not surprising since both are central metabolites that can be derived from multiple sources. Still, SA as the proposed end product of the lysine degradation pathway was observed [M+4] and [M+2]-labelled. Furthermore, in the ΔcsiD strain we found highly elevated levels of ^{13}C -labeled GA but no ^{13}C -labeled metabolites downstream of the CsiD reaction (SA, αKG). These results confirm that lysine is metabolized to succinate via the CsiD pathway. However, it remains difficult to predict an impact of the lysine degradation pathway to central energy metabolism on basis of the here presented results. Labelled succinate might already be further metabolised in the TCA resulting in a drainage of ^{13}C by decarboxylation reactions.

Taken together, the labelling pattern of the detected intermediates support an unbranched pathway as depicted in Figure 6. Isotope labelling experiments furthermore proved that

lysine degradation is the source of GA in *E. coli*. A detailed list of the intracellular concentrations and labelling patterns of all intermediates of the pathway can be found in Supplementary Table 2.

Taking together all biochemical and metabolic experiments we can sum up the lysine catabolic pathway as:



Physiological role of the novel lysine degradation pathway

Finally, we wanted to get an overview of the physiological role of this novel lysine degradation pathway. We compared growth of knockout strains of genes directly involved in the degradation pathway. Moreover, we investigated whether growth on different sole C- and N-sources is influenced by the different gene knockouts (Figure 15a). We confirmed that wildtype *E. coli* BW25113 is not able to grow on lysine as sole C-source in M9 minimal medium. In addition, other tested sole C-sources from the pathway (cadaverine, AVA, GA) did not lead to measurable growth of *E. coli*, too. Single knockouts of the pathway genes did not alter growth of the strains on the tested sole C-sources.

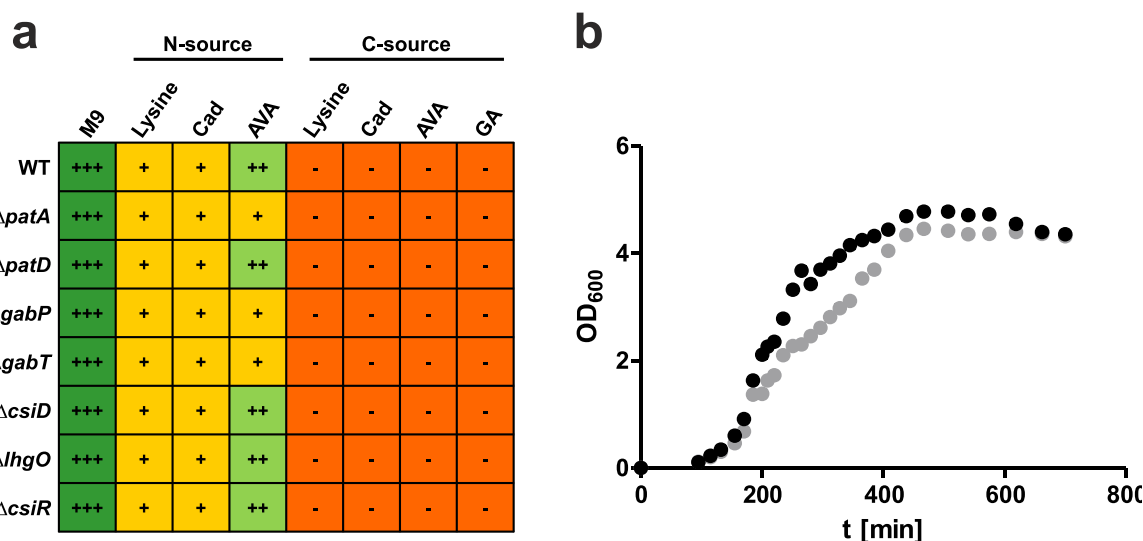


Figure 15: Growth analysis of genotypes involved in the pathway. **a**, Growth analysis of *E. coli* knockout strains. *E. coli* strains were grown in minimal medium containing 10 mM of the respective nitrogen (N-) or carbon (C-) source. Growth was assessed by measuring OD₆₀₀ and distinguished between good growth (dark green, +++), intermediate growth (light green, ++), low growth (orange, +) and no growth (red, -), see Methods for OD thresholds. Evaluated were the mean of independent triplicates. **b**, Growth curve of *E. coli* BW25113 compared to the $\Delta csiD$ strain in LB medium. Growth was assessed by recording the optical density at 600 nm. $\Delta csiD$ (grey) exhibits a growth defect around OD₆₀₀ \approx 2 compared to the wildtype strain. The shown curve is a representative of 5 individual experiments. Growth (means of triplicates) was rated as high growth (0.47 > OD₆₀₀ > 0.43 a.u.); intermediate growth (0.35 > OD₆₀₀ > 0.25 a.u.), low growth (0.15 > OD₆₀₀ > 0.07 a.u.) or no growth (0.02 a.u. > OD₆₀₀).

We observed a growth defect of $\Delta csiD$ compared to wildtype *E. coli* BW25113 in LB medium (Figure 15b). Growth of the knockout strain diminished at $OD_{600} \approx 2$ which correlates with entrance of the culture into late-exponential or early-stationary phase. In glucose minimal medium *E. coli* WT grew up to $OD_{600} \approx 2.5$ and then completely stopped growing probably due to the exhaustion of glucose as C-source (Figure 7d). A $\Delta csiD$ strain stopped growing even earlier at $OD_{600} \approx 2$ which is comparable to the OD we observed a growth defect in LB medium. Furthermore, Herr *et al.*¹⁰⁵ found increased fitness of *E. coli* BL21 (DE3) when it was supplemented with GA during growth. This effect was especially observed during long stationary phase conditions. When the cells were grown in the presence of 5 mM GA an increasing OD_{600} was even observed after 40 hours of incubation. The fitness relevant effect of GA was not observed in a $\Delta csiD$ strain. This states that the final part of the catabolic pathway responsible for utilizing the carbon backbone of lysine depicts a growth promoting effect for *E. coli*. It seems likely that metabolizing GA might be an important metabolic strategy to adapt to conditions that occur at entering stationary phase. This is eventually performed by starting to degrade lysine. However, it could also be possible that GA, available from other sources, promotes growth via the CsiD pathway since many microorganisms produce GA via benzoate degradation¹²³. In addition, free AVA produced by other organisms could be imported into the cell by GabP for further utilization. Nonetheless, it remains elusive why *E. coli* K12 possesses a carbon starvation induced lysine catabolic pathway with elements coupled to energy production and is still not able to grow on lysine. However, arginine is also not utilized as sole C-source by *E. coli*¹²³ even if two catabolic pathways exist for this amino acid. A big part of the arginine degradation pathway via putrescine is shared with the here presented lysine catabolic pathway. The coupling of both pathways might at least hint for a common trigger that prevents both amino acids to be used as sole C-sources.

There are diverging observations in the literature concerning the usage of lysine as sole N-source in *E. coli* K12. However, most studies decline lysine as a sole N-source in K12 MG1655¹²³. The genotype of *E. coli* K12 BW25113 that was used in our study is almost identical to MG1655¹²³. We show evidence that *E. coli* BW25113 is able to grow on lysine as the sole source of nitrogen. Furthermore, *E. coli* growth was promoted on AVA as N-source. Growth of knockout strains $\Delta patD$, $\Delta lhgO$ and $\Delta csiR$ did not affect growth

on the N-sources. We provide evidence that GabP influences growth when AVA is given as N-source probably due to its function as AVA transporter. In addition, *ΔgabT* impacts growth on the non-proteinogenic amino acid AVA as N-source. Knockdown of PatA which should prevent endogenous AVA production also decreases growth on AVA as sole N-source. Taken together, our results support the theory that the two transaminases (PatA and GabT) involved in lysine catabolism contribute to overall nitrogen assimilation of *E. coli*.

Conclusion

The here presented results provide evidence for the first complete catabolic pathway of L-lysine in *E. coli*. Moreover, they describe the first glucogenic lysine degradation pathway by the conversion of GA to SA and CO₂ via αKG-dependent hydroxylation and subsequent oxidation to αKG. Lysine was formerly seen as one of the amino acids that could only be utilized by the production of acetyl-CoA as an endproduct. The key enzymes CsiD and LhgO have been shown to participate in lysine degradation also in *P. putida*^{104,113}. We provide evidence that GA and therefore lysine degradation participates in a cellular stress response under stationary phase conditions. It seems likely that lysine is recycled by *E. coli* in stationary phase for carbon and energy regeneration. Moreover, we could reveal that lysine can be exploited by *E. coli* as alternative (sole) N-source contributing to the assimilation of nitrogen. Since many bacteria are confronted in their natural habitats with conditions that resemble nitrogen or carbon starvation, the found degradation pathway could lead to new insights into how bacteria adapt to certain conditions. Furthermore, CsiD is widest distributed in Enterobacteriaceae and can also be found in pathogenic members of this family like *E. coli* O157:H7 (an enterohemorrhagic strain), pathogenic *Salmonella* or potentially pathogenic *Shigella* species like *S. sonnei* or *S. dysenteriae*. It would be interesting to see whether the complete pathway or single reactions involved might impact pathogenicity of these strains under starvation conditions.

Metabolites AVA and GA are used as precursors for the production of polymers like polyesters. Thus, the presented pathway implements potential implications for biotechnology. Both compounds have been produced from engineered *E. coli*

before^{124,125}. Considering the presented pathways and activities, it is likely that yields could be improved significantly by utilizing knockout strains or adopting conditions such as oxygen limitation preventing degradation of the target compounds by GabT/D and CsiD/LhgO. Indeed, Zhang and co-workers demonstrate increased production of GA by knocking out both the GA hydroxylation as well as the glutaryl-CoA dehydrogenase pathway in *P. putida*¹²⁶.

Both L2HG and GA are known to be involved in human diseases like cancer¹²⁷ and genetic organic acidurias¹²⁸. The taxonomic distribution of CsiD (Figure 10) reveals that it is not conserved on the human genome. However, L2HG is produced by malate and lactate dehydrogenases in a promiscuous way acting on α KG. Acidic conditions caused by e.g. hypoxia and certain mutations of these enzymes lead to increased production of the metabolite¹²⁹. L2HG is a potent inhibitor of TET- and Jmjc-type α KG-dependent oxygenases responsible for nucleobase and histone demethylation, resulting in epigenetic deregulation of gene expression and thereby progression in certain cancers^{108,109}. Moreover, gene inactivation of L2HGDH leads to L2HG aciduria, a rare neurometabolic disease affecting the central nervous system which also correlates to certain kinds of tumours^{130,131}. Malfunctioning glutaryl-CoA dehydrogenase and glutaryl-CoA transferase lead to GA aciduria by increasing GA levels^{132,133}. Noteworthy, human L2HG and GA metabolism is still under investigation. GA producing enzymes have not yet been found. The human genome encodes enzymes that are closely related to LhgO and GabT/D (homologs of these genes are encoded on the mitochondrial genome). It would be of interest whether human ABAT (GABA aminotransferase, mitochondrial) and ALDH5A1 (SSA dehydrogenase, mitochondrial) are also able to utilize the C5 homologs and produce GA. It could also be possible that GA hydroxylating enzymes are encoded on the human genome that are different in the reaction mechanism apart from CsiD. Also in bacteria GA was thought to be a dead end metabolite¹³³ until the discovery of the *csiD* gene function. However, L2HG processing enzymes are much wider distributed in organism than GA utilizing. This hints for an evolutionary process in where LhgO (and also GabT/D) might be conserved due to new functions acquired (L2HG detoxification by LhgO, GABA as neurotransmitter by GabT/D) whereas CsiD might have been lost. In one case study, patient carrying glutaric aciduria responded to antibiotic treatment after standard treatments did not impact the disease¹³⁴. This might hint for a possible

involvement of the human gut microbiome into elevating levels of GA. Many members of bacterial geni carrying *csiD* are inhabitants of the non-pathogenic human gut microbiome. Under these, mostly members of *Enterobacteriaceae* like *Shigella*, *Escherichia*, *Enterobacter* and *Klebsiella*. Enterobacteria as part of the phylum Proteobacteria mainly colonize the small intestine together with *Firmicutes* like some *Bacillus* species^{135,136} that also carry *CsiD* frequently. However, these bacterial geni only contribute to a minor proportion to the overall count of bacteria in the gut. They colonize regions of the gut that contain small amounts of oxygen. Thus, they are believed to still have certain influence on metabolic processes in the gut because of the big diversity of compounds produced by their aerobic metabolism¹³⁶. It is still under debate whether additional sources of GA and L2HG exist in humans and if the microbiome is involved¹²⁷. In case of cadaverine, which is also an intermediate of the lysine degradation pathway it was shown that it might perform as an anticancer metabolite in breast cancer carcinogenesis. Moreover, a correlation between cadaverine produced by the gut microbiome and breast cancer proliferation was observed¹³⁷. Thus, investigating the contribution of metabolites produced by the here presented lysine degradation pathway in bacteria to overall processes in the human gut should be undertaken.

Chapter 3

An ASKA- & KEIO-collection Based Screening-Approach for Identifying New Protein Functions in Amino Acid Catabolism

Introduction

Most knowledge concerning amino acid degradation pathways in *E. coli* was achieved more than 50 years ago. Nonetheless, descriptions of complete catabolic pathways for valine, isoleucine and leucine, the aromatic amino acids tyrosine, histidine and phenylalanine as well as methionine are still missing for *E. coli* (Figure 1). Usually, metabolic studies exploit the fact that during growth on specific C- or N-sources, enzymes of the catabolic pathways are upregulated and can be monitored on a transcriptomic or proteomic level. Since the growth on sole nutrient sources strongly affects the expression levels of catabolic enzymes, it is possible to identify the encoding genes. The fact that *E. coli* is not able to grow on a big proportion of the proteinogenic amino acids as sole N- or C-source hampers this analysis strategy.

Of all 4,497 *E. coli* genes 32 % are not assigned to a specific gene function whereas it is suggested that ¼ of known biochemical reactions are not linked to a specific gene product¹³⁸. Thus, new unbiased screening methods and comparative genomic approaches might lead to the identification of new metabolic pathways including the missing amino acid catabolism in *E. coli*. An interesting approach developed by Sèvin *et al.*¹³⁸ uses high-throughput MS for the identification of uncharacterized *E. coli* proteins. They used overexpression lysates or purified proteins of all 1,275 unidentified proteins and incubated them with metabolome extracts. The pre-incubation extract was compared with the extract post incubation analysing accumulating or depleting metabolites via non-targeted metabolomics. Within that, they were able to identify and validate 12 new gene functions, most of them involved in nucleic acid metabolism and some acetyltransferases. The gene function of *ybiC* was assigned to a hydroxycarboxylic acid dehydrogenase, catalysing the conversions of α KG to 2-hydroxyglutarate (R/L2H), phenylpyruvate to phenyllactate and 4-hydroxyphenylpyruvate to 4-hydroxyphenyllactate, all in a NAD(P)⁺ dependent manner. Metabolites of the latter reactions are involved in degradation pathways of phenylalanine in other organisms. YbiC therefore represents a potential candidate for the missing phenylalanine degradation route in *E. coli*. However, the substrate promiscuity of YbiC determined by the approach could also hint for other

natural substrates. Besides that, there was no obvious evidence for more amino acid degradation enzymes even in the list of 241 candidate reactions that were not further validated in the study.

The fact that more than one reaction is catabolised by a single enzyme is wide-spread among enzyme classes and is called enzyme promiscuity. Studying these promiscuous actions of enzymes might therefore lead to the discovery of new metabolic routes. The group of Paolo Picotti¹³⁹ focused directly on metabolite binding sites of proteins with known function to unravel potential promiscuous interactions. They used proteomics to identify differing cleavage patterns of proteins from cell extracts that were supplemented with an unmodified metabolite. Within that, they were able to identify and assign functional effects of new metabolite-protein interactions. Concerning amino acid metabolism, they showed that methionine and phenylalanine bind to TCA enzymes and that IlvC (isoleucine, valine, leucine biosynthesis) also binds phenylalanine.

A screening approach that discovered a new D-lysine degradation pathway in *Pseudomonas putida* was applied by Thompson *et al.*¹¹³ They used a random barcode TnSeq (RB-TnSeq) library for *P. putida* and performed fitness profiling during growth on D-lysine as C-source. Several genes emerged that were fitness relevant under these conditions including gene products of *ydcJ* and *ydiJ*. *In vitro* validations showed that these enzymes catalyse novel reactions involved in the degradation of D-lysine via L-2-aminoadipate and 2-oxoadipate to α KG. However, this approach is again dependent on the ability of an organism to grow on specific sole nutrient sources. But also approaches using *in vitro* screenings for enzyme identification underlie disadvantages since some reactions or interactions mapped might be artificial and mask the physiological role of the protein in the cell. The development and implementation of a new screening approach for the identification of novel enzymes involved in amino acid catabolism is presented in the following subsections.

The presented screening approach in this study is based on the theory that overexpression or knockouts of *E. coli* BW25113 single endogenous genes might enable cells to grow on non-sole C-sources. Thus, the ASKA-collection¹⁴⁰, a complete set of *E. coli* ORFs on overexpression plasmids and the KEIO-collection¹⁴¹, a single gene knockout library of *E. coli*, were used to identify gene products involved in the catabolism of potential C-sources. A purified and pooled fraction of all 4,217 ASKA-plasmids was used as basis for transformation into *E. coli* BW25113 via electroporation. Pooled ASKA-plasmid

containing cells were plated on agar dishes containing single C-source metabolites. Single colonies emerging on the plates were regrown and purified plasmids were sequenced to identify genes of interests (GOIs) (Figure 16).

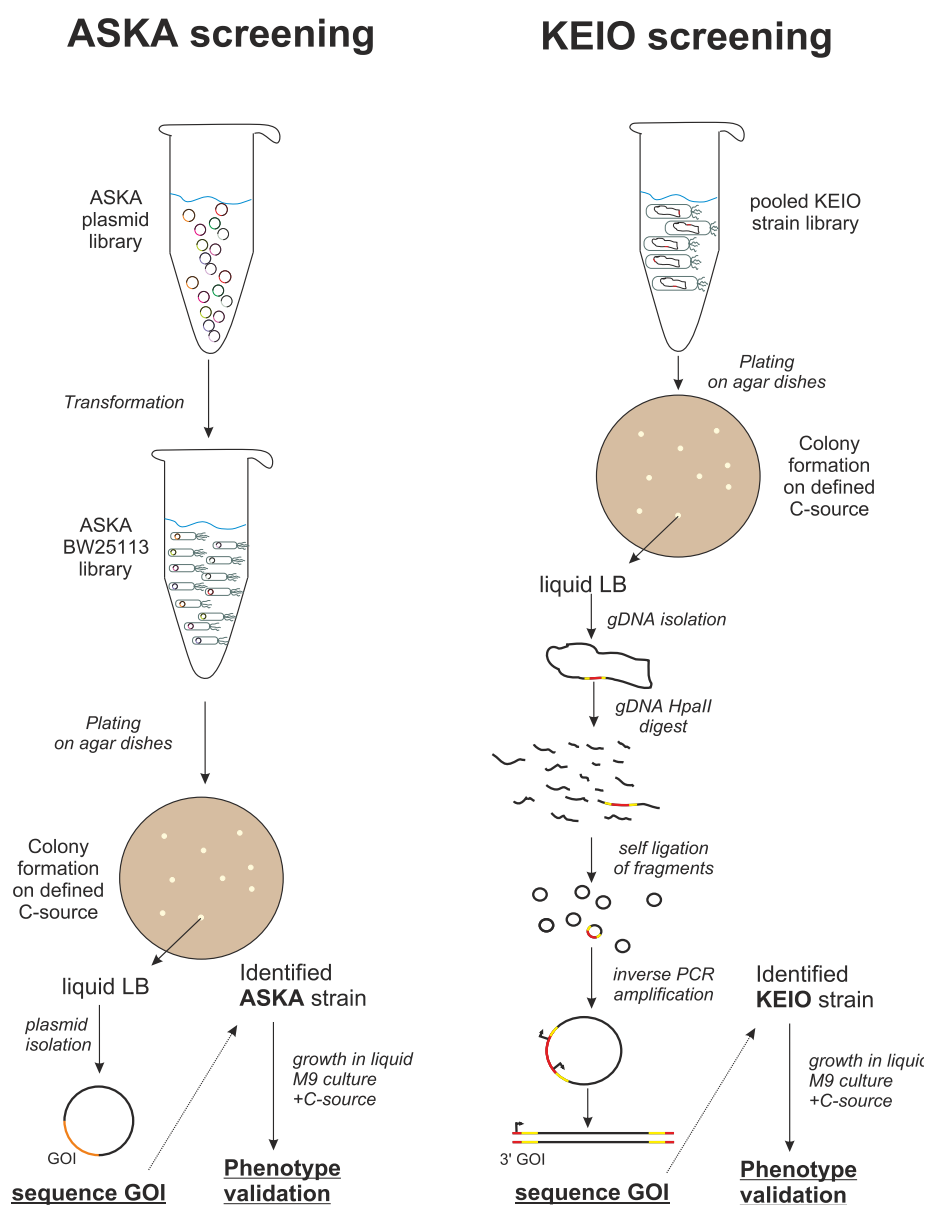


Figure 16: Screening strategy for the identification of proteins involved in carbon source catabolism. ASKA Screening, Pooled ASKA plasmids were transformed into *E. coli* BW25113 via electroporation and cells were incubated in SOC medium for 30 min. 10^7 cells were plated on minimal medium agar dishes containing 10 mM of each C-source tested. Dishes were incubated at 37 °C until first colonies were visible and were further incubated for 1 week. Single colonies were inoculated in LB medium and after incubation plasmids were isolated and sequenced to identify potential GOIs. **KEIO screening,** 10^8 cells of a pooled KEIO strain fraction were plated on agar dishes containing 10 mM of respective carbon sources. If first colonies were visible during incubation the cells were regrown in LB medium to isolate gDNA. gDNA of single colonies was HpaII digested and resulting fragments were self-ligated. GOI was identified by an inverse PCR technique followed by sequencing. Growth phenotypes of GOIs were validated by re-growing single ASKA and KEIO strains in M9 medium containing the respective carbon source. GOI = gene of interest, gDNA = genomic DNA.

The KEIO knockout collection was constructed by replacing ORF regions with a kanamycin-resistance-gene cassette. Overnight cultures of all 3,985 single-gene knockout KEIO strains were pooled and the resulting culture was plated on agar dishes containing defined C-sources. Single colonies were isolated, regrown and extracted genomic DNAs (gDNAs) were digested via HpaII, an restriction enzyme cutting at 24,237 sites in the BW25113 genome. Fragments of gDNA were circularized via self-ligation and an inverse PCR technique was used to capture the kanamycin-resistance cassette disrupted DNA fragment. The amplified DNA product containing the transition between kanamycin resistance gene and ORF was sequenced to identify the GOI. Within that, identified KEIO strains were regrown from original KEIO stock cultures in the respective C-source containing minimal medium. If an increase in OD₆₀₀ was detectable during incubation the KEIO genotype was recognized as trigger for utilizing the C-source.

Results and Discussion

Potential C-sources tested in the approach were non-growth promoting amino acids as well as derivatives of these amino acids that are known intermediates of degradation pathways in other bacteria (Table 2). Thus, the three basic amino acids arginine, histidine and lysine were included as potential C-sources as well as the ketogenic lysine degradation products cadaverine, 5-aminovalerate (AVA) and glutarate (GA). *E. coli* possesses two arginine degradation pathways. The first proceeds via agmatine to putrescine (which is further degraded to SA) and the second via addition of succinyl to arginine producing glutamate and succinate (SA) as end products (AST pathway). Histidine catabolism is resolved for many Pseudomonads where it is degraded via uroconate to glutamate by the *hut* pathway (*hutF/G/H/I/U* in *P. putida* KT2440) whereas no homologs of these genes can be found on the *E. coli* genome. In general, *Ps. putida* KT2440 is able to utilize all L-amino acids except methionine, threonine, leucine and tryptophan as sole C- and N-source. Whereas *P. putida* is able to degrade BCAAs beginning with the reverse catalytic action of IlvE producing acetyl-CoA as end product, *E. coli* has no known BCAAs catabolism. The aromatic amino acid tryptophan, however, is a known sole C-source for *E. coli* by conversion to indole and 2-aminoprop-2-enoate which spontaneously reacts to pyruvate. We speculated that the carbon backbone of indole might be further degraded by unknown catalytic reactions.

We tested two more aromatic amino acids, phenylalanine and tyrosine as well as two related compounds 4-hydroxyphenylpyruvate (4-HPP) and 4-hydroxyphenylacetate (4-HPA) as potential C-sources. In *E. coli*, the benzoate derivate 2,3-dihydroxybenzoate (2,3-DHBA) is synthesized via the chorismate pathway using phenylalanine, tryptophan and tyrosine as precursors. 2,3-DHBA itself is then used as siderophore for complexation and uptake of iron in *E. coli*. We suspected that the carbon backbone of 2,3-DHBA, when imported back into the cell, might be further catabolised to recycle the compound. We thus tested 2,3-DHBA and benzoate as possible C-sources.

ASKA screening

Growth was observed on agar plates containing 2,3-DHBA, 4HPA and GA (Table 2). However, we failed to identify potential GOIs since no sequence could be obtained after single colonies were analysed by plasmid extraction and sequencing. Eventually these cells have lost the ASKA plasmid and evolved other mechanisms to utilize the C-source. Interestingly, these colonies were still able to grow in full medium containing chloramphenicol. This hints for a resistance mechanism derived independently from the chloramphenicol resistance gene containing ASKA plasmid. However, due to the very long incubation time of 2 months on agar plates a contamination with another organism than *E. coli* seems plausible, too. We did not try to identify the organism itself. However, colony shape and colour did not differ substantially from the expected *E. coli* colony form.

In general, we observed almost no growth on the sole C-sources tested in the ASKA screening. A possible reason might be that the IPTG induced overexpression system itself presents a burden for cell proliferation. It has been shown that the chemical properties of IPTG induce cell stress additionally to the already existing possible negative effects of multicopy plasmids coding additional proteins¹⁴². The ASKA screening approach should therefore be improved by reducing the concentration of IPTG in the medium. Additionally, C-sources were inserted in a standard concentration of 10 mM. However, also these metabolites can be toxic in high concentrations as has been shown for the BCAA valine and leu. Thus, the approach must be adapted to individual physiological properties of used metabolites.

Similar approaches performed by other groups conducted the assay in liquid culture instead of solid medium^{143,144}. We suspected that growth of the overexpression strains might be better in liquid culture. Thus, we reinitiated the experiment and inoculated the ASKA strain library into M9 liquid medium containing indicated C-sources and only 100 μ M IPTG. Nonetheless, after 1 month of incubation, no growth in any of the C-source containing media was detectable indicating that the screening approach is not applicable for initiating growth of *E. coli* on the tested C-sources.

Table 2: Colony formation visible on agar plates containing indicated C-sources.

Carbon-source	Growth after 2 months of incubation	
	ASKA screening	KEIO screening
L-Lysine	-	-
AVA	-	++
Glutarate	++	++
Cadaverine	-	++
Arginine	-	++
L-Tryptophan (pos. control)	+++	+++
Indole	-	-
L-Histidine	-	+
L- Leucine	-	-
L-Isoleucine	-	+
L-Valine	-	-
Benzoate	-	-
2,3-Dihydroxybenzoic acid	+++	-
4-Hydroxybenzoic acid	-	-
L-Tyrosine	-	-
L-Phenylalanine	-	+
4-Hydroxyphenylpyruvate	-	-
4-Hydroxyphenylacetic acid	+++	-

+++ : > 100 colonies

++ : > 10 colonies

+ : < 10 colonies

KEIO Screening

We observed growth on agar plates that contained GA, cadaverine, AVA, histidine, isoleucine and phenylalanine. The first colonies occurred on the GA containing plate after 2 weeks and were regrown in LB medium for further identification. For cadaverine, AVA, histidine, isoleucine and phenylalanine containing plates growth was observed after

~1 month of incubation. For an overall of 41 colonies analysed from all plates we were only able to assign 7 GOIs from the KEIO collection (Table 3).

Table 3: Identified KEIO strains promoting growth on the indicated sole C-source.

Carbon-source	Identified KEIO strain	Annotated or putative gene function	Relevant literature
GA	<i>ΔhokB</i>	Small toxic membrane polypeptide	145,146,147
	<i>ΔyhiI</i>	Putative membrane fusion protein	148,149
	<i>ΔyggT</i>	Transporter: 'YggT or fanciful K ⁺ uptake-B family'	150
	<i>ΔyggX</i>	Putative Fe ²⁺ -trafficking protein	151,152,153
	<i>ΔlldR</i>	DNA-binding transcriptional dual regulator	154
	L-Phenylalanine	<i>ΔygdG</i>	flap endonuclease
<i>ΔogrK</i>		Prophage P2 late control protein	157

When the KEIO strain with the *ΔhokB* genotype was regrown in minimal medium containing GA, *E. coli* grew to an OD₆₀₀ ≈ 1 within 3 days of incubation. This marks the HokB protein as a potential impact factor for cell growth promoted by GA as C-source. HokB is assigned as a small toxic membrane polypeptide as part of the type 1 toxin-antitoxin system HokB-SokB. Additionally, *hokB* gene expression is upregulated by Obg an GTPase which is induced by the alarmone (p)ppGpp during starvation processes¹⁴⁵. Increased HokB expression leads to a collapse of the membrane potential due to ATP leakage. Rising levels of HokB have been assigned to be an inducer of dormancy in bacterial cells leading to metabolic inactivity¹⁴⁶. A potential destabilization of the membrane potential by HokB might prevent the functionality of membrane coupled dehydrogenases and complexes of the ETC. This would also impair catabolic pathways responsible for energy production. Knocking out *hokB* might then result in a better maintenance of metabolic processes in starved cells that get exposed to potential C-sources. This would explain why GA is metabolised in a *ΔhokB* environment in a way that produces enough energy to initiate cell proliferation. ATP and amino acid precursors produced by GA degradation might be sufficient to initiate protein biosynthesis and cell division because HokB is not able to interrupt these membrane-coupled processes. It is also thinkable that other metabolites serve as additional sole C-sources for *E. coli* growth when HokB expression is prevented. However, this theory must be validated experimentally.

Moreover, two additional KEIO strains with genotypes $\Delta yhiI$ and $\Delta yggT$ were able to grow in GA medium. Both knocked out genes originally encode membrane associated proteins. Not much is known about the putative membrane fusion protein YhiI besides its predicted localization to the inner membrane and that it is involved in biofilm formation^{148,149}. YggT is annotated as an inner membrane transporter that might be involved in K⁺-uptake¹⁵⁰. Besides the connection to HokB through their membrane localisation, no obvious explanation can be made why the knock down of YhiI or YggT resulted in growth of *E. coli* with GA as C-source.

YggX is proposed to be involved in the protection of iron-sulfur clusters under oxidative stress and belongs to a subclass of the treble clef protein family that has zinc finger like structural features but lost ability to bind zinc^{151,152}. Furthermore, YggX has been shown to induce aconitase activity when overexpressed. Aconitase is the second enzyme of the TCA and catalyses the reversible isomerization of citrate and iso-citrate via *cis*-aconitate as intermediate. However, when aconitase lacks an iron-sulfur cluster it becomes a regulatory protein able to bind mRNA. It thus influences gene expression on a posttranscriptional level impacting oxidative stress systems as well as TCA enzyme expression^{158,159}. One could speculate that missing YggX might influence the catalytic or regulatory activity of aconitase thus altering carbon utilization processes of the cell. In this context, YggX was shown to be upregulated under glucose limitation¹⁵³ whereas aconitase AcnA of *E. coli* plays a role in surviving nutritional stress¹⁶⁰.

Interestingly, also a $\Delta lldR$ strain was able to grow on GA as sole C-source. LldR is a transcriptional regulator of L-lactate catabolism and transport¹⁵⁴. It might be possible that LldR as a member of the GntR transcription factor family is also able to bind to other DNA regions on the genome. Thus, it could negatively influence the ability to utilize other C-sources whereas a knockdown could re-enhance growth on C-sources.

The only additional potential C-source leading to GOI identification was phenylalanine. Knockdowns of *ogrK* (encoding a prophage P2 late control protein) and *ygdG* (a putative flap endonuclease) led to growth on the amino acid as C-source. However, too little is known about the function of these two genes to discuss possible roles in C-source utilization. Still, it could be interesting to further analyse the knockout strains. The ability to grow on phenylalanine as sole C-source could be exploited to identify genes upregulated during this process thus extracting potential catalytic enzymes.

Overall, of 11 sole C-source plates that showed colony formation, we only could identify genotypes for 2 C-sources (GA & phenylalanine) and only in case of the KEIO screening approach. Whereas for GA as sole C-source in the KEIO experiment we obtained colonies within two weeks, on all other plates colonies only arose after more than 1 month of incubation. Furthermore, after 1 month of incubation we could observe growth on a control plate that did not contain any C-source indicating that these cells were even able to utilise carbon contaminations in the agar. For one colony we were able to obtain a sequence of a kanamycin cassette that was not inserted by the KEIO collection procedure. This cassette was similar to the KEIO kanamycin resistance gene enabling capturing by our screening procedure. However, it lacks the FLP recognition sites for resistance cassette excision and furthermore sequence of the obtained cassette fragment was not 100% identical to the KEIO cassette. This hints for other evolutionary processes occurring during the very long incubation time of the screening. Other mutations developed and not the KEIO knockouts might have led to the ability of these cells to utilise the present C-source. Similar screening approaches had much shorter incubation times between 1 - 7 days. This clearly describes a critical disadvantage of the screening approach restricting its applicability.

Conclusion

The presented screening approach was designed to identify enzymes possibly involved in previously unknown catabolic pathways of non-sole C-sources of *E. coli*. The ASKA screening approach should find catalytically active enzymes via the overexpression of recombinant proteins. However, we could not obtain any target genes and growth on the IPTG and C-source containing agar plates might be impaired by toxic side effects. Thus, the ASKA screening method without improvement seems not to be applicable for the identification of catalytic enzymes involved in C-source utilisation. The KEIO screening approach was designed to uncover proteins with possible novel regulatory roles that have influence on C-source utilization. We were only able to identify candidate genes involved in GA and phenylalanine catabolism. Based on the annotated gene functions it was difficult to develop specific theories how these genes could influence catabolism of the indicated C-sources. However, we saw that genes thought to be involved mainly in distinctly related processes seem to impact C-source utilisation in a way that even growth of *E. coli* was enabled. These results indicate that more complex regulatory relationships

are responsible for the catabolism of C-sources and that more research is necessary to unravel these connections.

Potential effects on C-source utilization triggered by KEIO knockouts were highly covered possibly by other mutational processes leading to not identifiable genotypes. Therefore, the high false positive rate of the screening approach strongly reduces its applicability as an identification platform of novel gene functions.

Chapter 4

The ilvH-motif RNA Regulates Gene Expression at the Crossroad of BCAA Biosynthesis

Introduction

The discovery of riboswitches more than 15 years ago^{161–163} gave hope that these gene regulatory elements represent genetic tools with various applications. The modular nature of aptamer domains by coupling it to self-cleaving ribozymes made it a promising tool in gene therapy as has already been shown for treating age-related-macular degeneration in humans¹⁶⁴. But also riboswitches as potential antibiotic targets have been discussed due to the possibility to mimic ligands of the 38 known riboswitch classes^{165–167}. Implementing these projects has been shown to be more difficult than expected. However, discovering novel functional ncRNAs and within that riboswitches still resulted in unexpected insights. An important example of these findings was the discovery of a riboswitch class sensing guanidine, a metabolite that has not been shown to play a role in natural systems^{168–170}. It is still under investigation what role guanidine could play in nature that entails such a wide-spread distribution of its cognate riboswitch classes in bacteria. Within that, finding novel riboswitch classes could expand knowledge concerning bacterial virulence, stress adaptations and antibiotic resistance mechanisms. However, unravelling new riboswitch classes has become more and more difficult. Whereas the most wide-spread riboswitch classes could be annotated relatively easy through high sequence and structure conservation of 5'-UTRs of bacterial genes, rarer classes are difficult to detect via bioinformatic or genetic approaches. The most promising approach for identifying new riboswitches was shown to be computational, comparative sequence analysis of 5'-UTRs in bacterial genomes⁶⁶. This bioinformatical method also includes secondary structure and covariational properties of sequences, thus being most applicable for the identification of not only riboswitches, but also other structure ncRNAs. Applying this method led to the discovery of more than 300 novel structured, conserved ncRNAs and within that to the characterisation of various new riboswitch and ribozyme classes^{171,172}. However, one of the big challenges of the characterization of riboswitches remains unravelling the cognate ligand and within that proofing functionality of potential ncRNA motifs. In many cases, ligands of riboswitches could be revealed because association of the putative RNA motif with its downstream gene's function gave

references to the type of small molecule binding to the RNA. However, the majority of putative RNA motifs do not have such a clear association. This was stated for the *ykkC*-motif which is mainly located upstream of different transporter families and genes coding for enzymes involved in nitrogen metabolism. Only an unbiased screening approach using a *lacZ*-reporter system led to the discovery of guanidine as the natural ligand¹⁶⁸. However, screening of big metabolite libraries is often less effective because many relevant small molecules are not commercially available. Additionally, aforesaid approaches are often laborious and expensive. Thus, rational approaches via educational guesses still present a suitable method for characterising putative functional RNA-motifs.

In this context, the putative *ilvH*-motif RNA was identified in 2017 by Weinberg *et al.*¹⁷² via aforementioned comparative sequence analysis of 5'-UTRs. The *ilvH*-motif is one of 224 structured candidate RNAs that very likely consist of new RNA classes with riboswitch- and ribozyme-candidates but also RNAs with novel functions. The motif can only be found in β -Proteobacteria and even there it is only annotated in 80 species. Thus, it represents one of the rare RNA-motifs that might have evolved recently or got distinct during evolution. Weinberg *et al.* determined the motif as likely to be a non-coding, structured and *cis*-regulatory RNA. However, until now no experimental evidence was given for the functionality of the ncRNA.

Our attention was additionally guided on this putative ncRNA-motif because it is associated with genes that have a specific annotated function (Figure 17). The *ilvH*-motif RNA is located upstream of the name giving *ilvH* ORF which codes for an acetohydroxyacid synthase (AHAS) small regulatory subunit. In most organisms containing an *ilvH*-motif RNA, it is located downstream of *ilvI* coding for the catalytic subunit of AHAS. Downstream of the *ilvH* gene *ilvC* encodes a ketol-acid reductoisomerase (Figure 17).

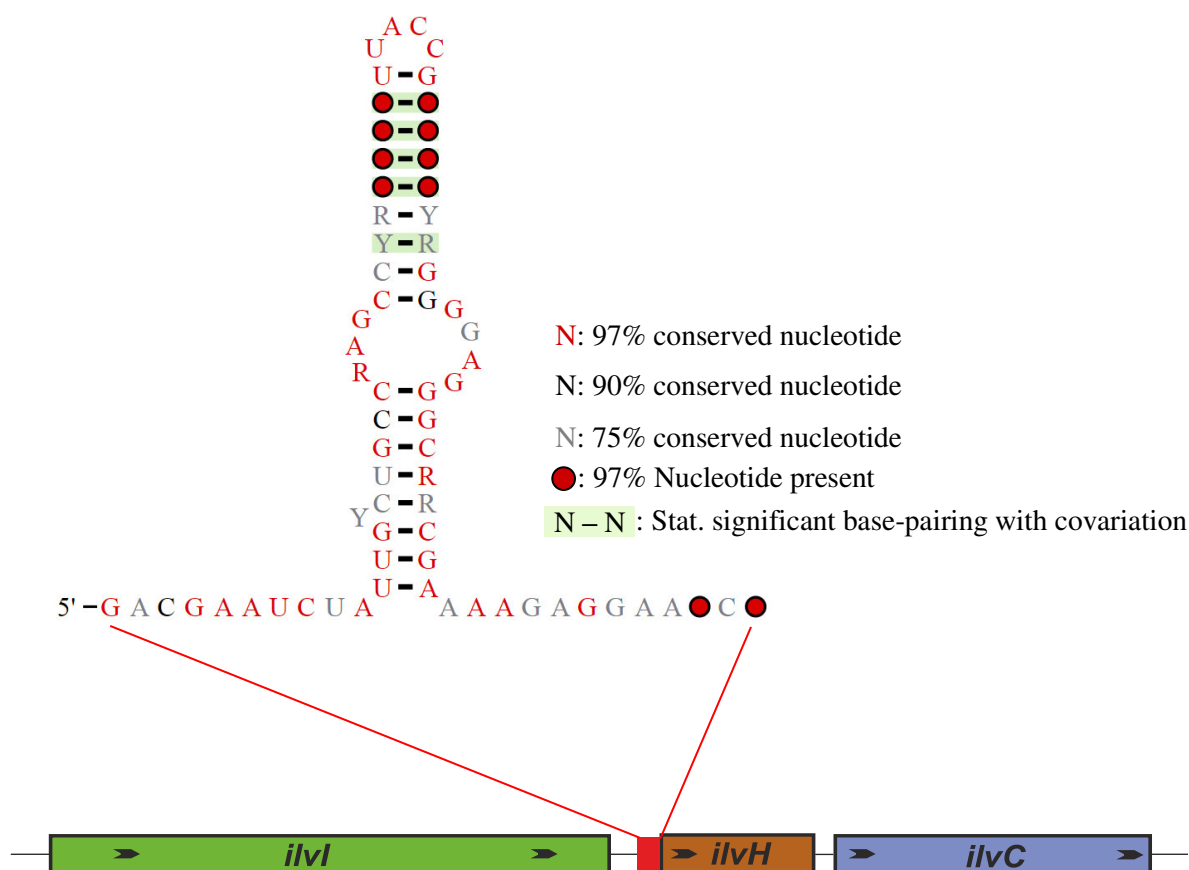


Figure 17: Consensus sequence with secondary structure and genomic neighbourhood of the *ilvH*-motif RNA. Consensus sequence and structure image derived from the rfam^{173,174} database including 90 sequences.

The genetic neighbourhood strongly suggests an association with regulation of BCAA biosynthesis. In BCAA metabolism, one subset of enzymes: AHAS encoded e.g. by *ilvI/H* (bacterial genomes often contain several homologues of these genes), *IlvC*, *IlvD* and *IlvE* is responsible for the biosynthesis of leucine, valine and isoleucine. *IlvA* is only participating in isoleucine biosynthesis as first enzyme of the pathway catalysing the reaction from threonine to α -ketobutyrate (α KB). *IlvD* produces the leucine and valine precursor α -ketoisovalerate, which is metabolised to leucine by an additional subset of enzymes (Figure 18). Most of the genes involved in BCAA biosynthesis and their products are feedback regulated by BCAAs or their intermediates. For example, binding of valine to the regulatory subunit of AHAS leads to inhibition of the enzyme. In *E. coli*, three different AHAS are known where AHAS I is encoded by *ilvB/N*, AHAS II by *ilvG/M* (non-functional) and AHAS III by *ilvI/H*. AHAS III shows a preferential activity as it converts around 40-fold more α KB to acetoxybutyrate (isoleucine

intermediates) than pyruvate to acetolactate (valine intermediates) - whereas AHAS I shows no preference in *E. coli*^{175,176}. The varying substrate preferences of these two isozymes are underlined by more efficient binding of the respective regulatory subunit to its holoenzyme¹⁷⁷. A regulatory active ncRNA upstream of *ilvH* would thus present an additional level of regulation for BCAA metabolism. We therefore tested whether the *ilvH*-motif RNA might be a riboswitch involved in the regulation of BCAA biosynthesis.

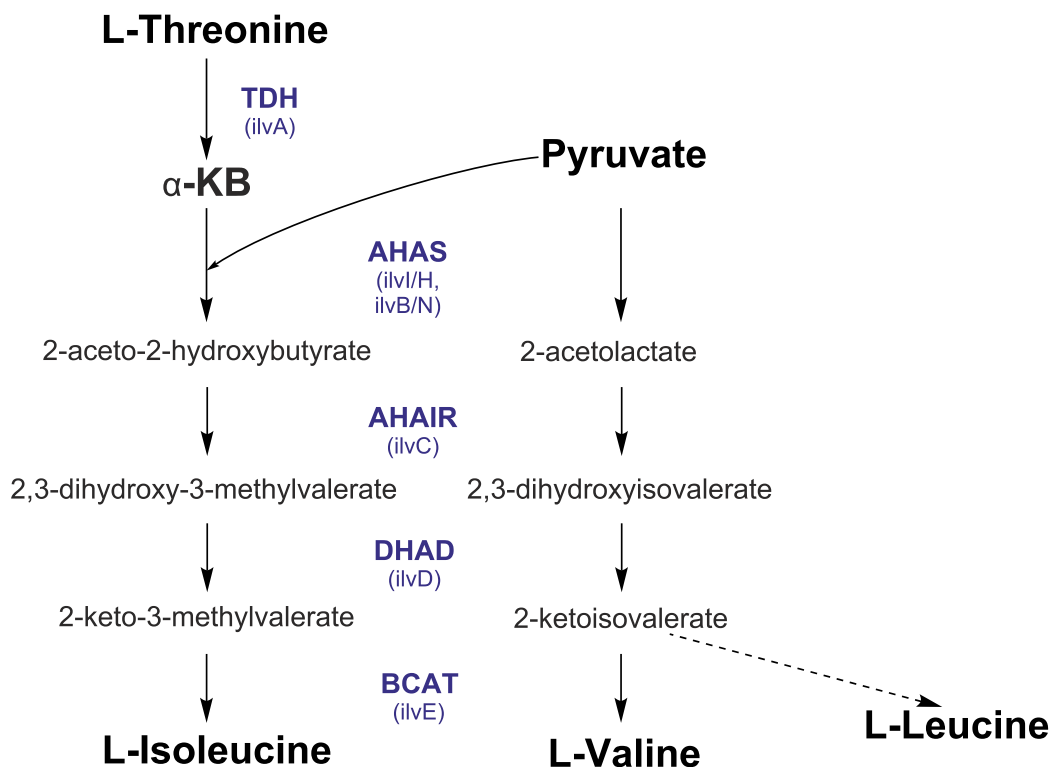


Figure 18: Schematic representation of bacterial BCAA biosynthesis. Enzymes are shown in bold capital letters and their respective genes in brackets. TDH = threonine dehydratase, AHAS = acetohydroxyacid synthase, AHAIR = acetohydroxy-acid reductoisomerase, DHAD = dihydroxy acid dehydratase, BCAT = branched-chain amino acid aminotransferase.

Results and Discussion

***ilvH*-motif RNA mediates gene expression in response to isoleucine and α -ketobutyrate**

We first constructed a reporter plasmid by inserting the conserved *ilvH*-motif RNA of *Comamonas testosteroni* KF-1 in frame with *lacZ* (Figure 19). A purine rich sequence is located between the structured stem-loop and the annotated start codon of *ilvH* and was supposed to be a potential ribosomal binding site (SD). This close proximity of structured RNA and SD might hint for a regulation mechanism on translational level. We thus also included the start codon and eight additional codons of the *Comamonas* sequence upstream of *lacZ*.

Reporter activity of the *ilvH*-motif:*lacZ* fusion diminished with increasing concentrations of isoleucine. A 30-fold reduction of reporter activity was observed when 20 mM isoleucine was present in the assay. A control reporter construct (*mykkC*:*lacZ* fusion) did not respond to increasing concentrations of isoleucine. This indicates that the *ilvH*-motif RNA off-switches downstream gene expression with isoleucine as a potential trigger. We additionally tested valine and leucine in reporter assays and did not observe comparable changes in reporter activity. However, valine could only be tested up to a concentration of 1.2 mM. Higher levels of valine inhibited cell growth due to toxic effects as has already been depicted for *E. coli*¹⁷⁸. Also pyruvate was tested as a potential trigger but did not lead to any change in reporter activity (not shown). When α KB was titrated into the reporter assay we saw an inhibition of reporter activity at a concentration of 10 mM. Moreover, an up to 70-fold change in reporter activity was observed at much lower α KB concentrations when 1 mM isoleucine was already present in the assay. The presence of isoleucine in the assay also led to an increased inhibition of reporter activity at lower levels of α KB. Thus, α KB might be an additional metabolite involved in the regulation mechanism of the *ilvH*-motif RNA. However, α KB is a precursor of isoleucine and might have been metabolised further to isoleucine. The additive effect in presence of isoleucine on reporter activity might also hint for isoleucine being the actual trigger of regulation.

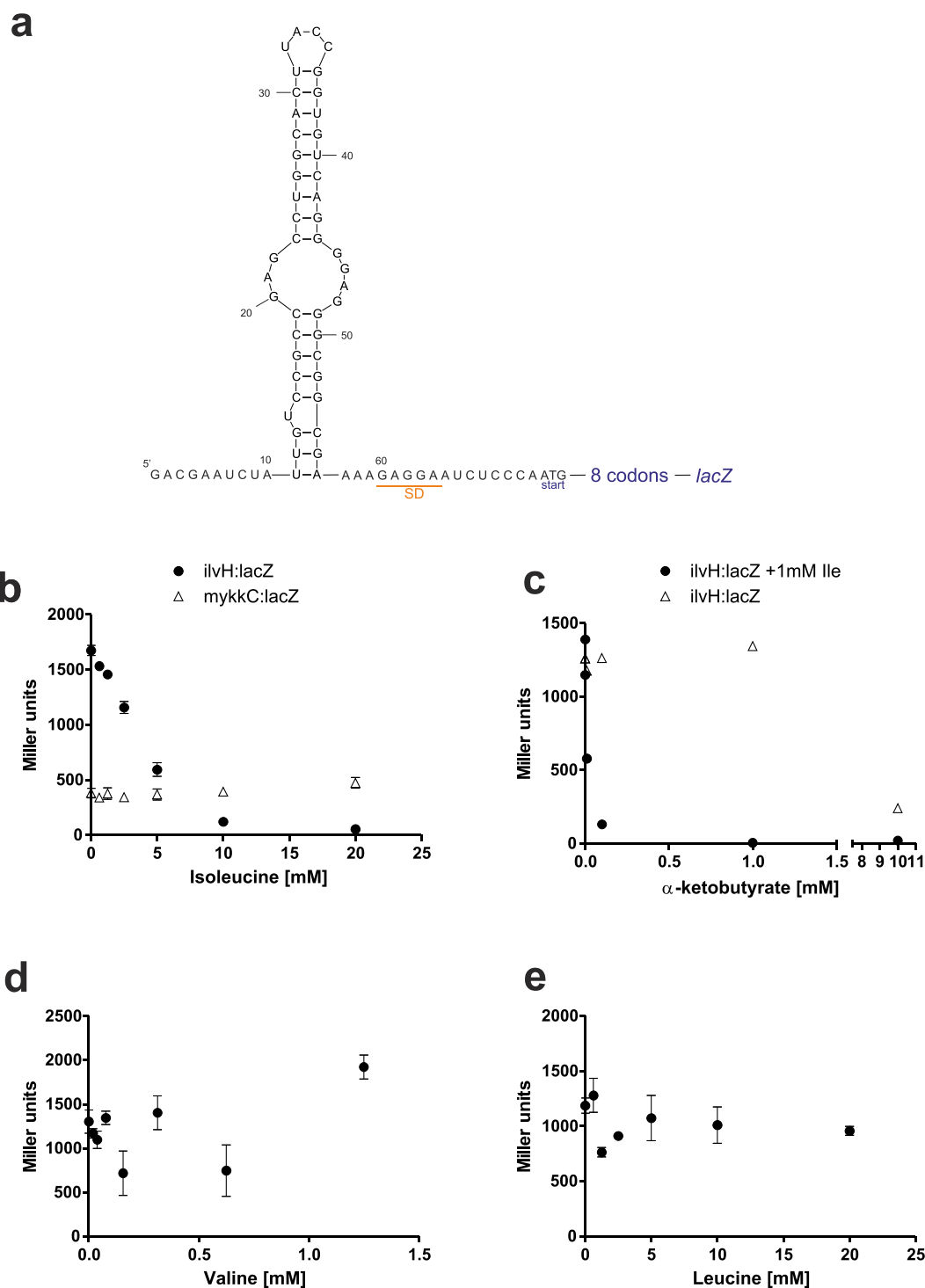


Figure 19: The *ilvH*-motif RNA regulates gene expression in response to Ile and α KB. a, Translational fusion between *ilvH*-motif RNA including proposed SD (orange), start codon (start), 8 codons of the native sequence and a *lacZ* reporter gene. Structure of the *ilvH*-motif RNA was predicted by the mfold¹⁷⁹ algorithm and matched manually to the consensus structure provided by rfam. **b**, Reporter activity of *E. coli* cultures carrying *ilvH*-motif:*lacZ* and *mykKC:lacZ* (control) reporter plasmids incubated with different concentrations of Ile. Activity is given in Miller units monitored by the ONPG assay. **c**, Reporter activity of *ilvH*-motif:*lacZ* monitored in response of increasing α KB concentrations in the presence (black circles) or absence (white triangles) of 1 mM Ile. **d**, Reporter activity of *ilvH*-motif:*lacZ* response to differing Val concentrations. **e**, Reporter assay testing different Leu concentrations. All reporter assays are shown as mean \pm standard deviation of triplicate measurements. **c** presents unique measurements.

***In vitro* analysis of *ilvH*-motif RNA as potential riboswitch**

To investigate whether the *ilvH*-motif RNA is a riboswitch sensing involved metabolites, we used the in-line probing technique to investigate whether a potential regulation mechanism is based on direct RNA-small molecule interaction. In-line probing is a widespread method for the characterisation of riboswitches. It exploits the ability of RNA to undergo self-cleavage. This is caused through a nucleophilic attack of the 2' oxygen on the adjacent P-center when the RNA is in an in-line conformation. Single stranded RNA occurs in in-line conformation and is thus cleaved more rapidly than double stranded RNA. When small molecules directly interact with the RNA, the structure of the RNA is altered which would lead to changes in the cleavage pattern of radioactively labeled RNA.

We tested valine, leucine, threonine, α KB and various concentrations of isoleucine in in-line reactions (performed by Anna Heiler) containing the *ilvH*-motif RNA (Figure 20). No obvious changes in the cleavage pattern upon the addition of any of the ligands could be observed. This indicates that a direct ligand-RNA interaction and within that conformational changes or stabilisation of the RNA seems unlikely. In addition, a RNA construct including the start codon and 8 downstream codons of the native RNA sequence (Figure 19) did not show any ligand dependent changes in in-line probing reactions (Supplementary Figure 5).

Our data suggest that gene expression regulation observed in reporter assays is probably not caused by a riboswitch-like mechanism involving isoleucine or α KB as ligands. Furthermore, the predicted consensus secondary structure folds do not show any obvious structural variations like SD masking caused by e.g. pseudoknot formation that are often found in riboswitches. Also of interest: A RNase T1 digest (T1 cleaves at single-stranded RNA after G residues) did not cleave most of the G residues of the RNA indicating very stable stem structures which is also atypical for riboswitches. However, we cannot exclude that the *ilvH*-motif RNA is a riboswitch requiring additional co-factors or sensing a ligand not tested in the experiments. To integrate possible additional effectors like co-factors, an *in vitro* transcription-translation assay (performed by Malte Sinn) was conducted that resembles the cellular surrounding better than an in-line probing. We tested isoleucine and α KB as possible effectors but could not see any changes in downstream gene product regulation (Supplementary Figure 6).

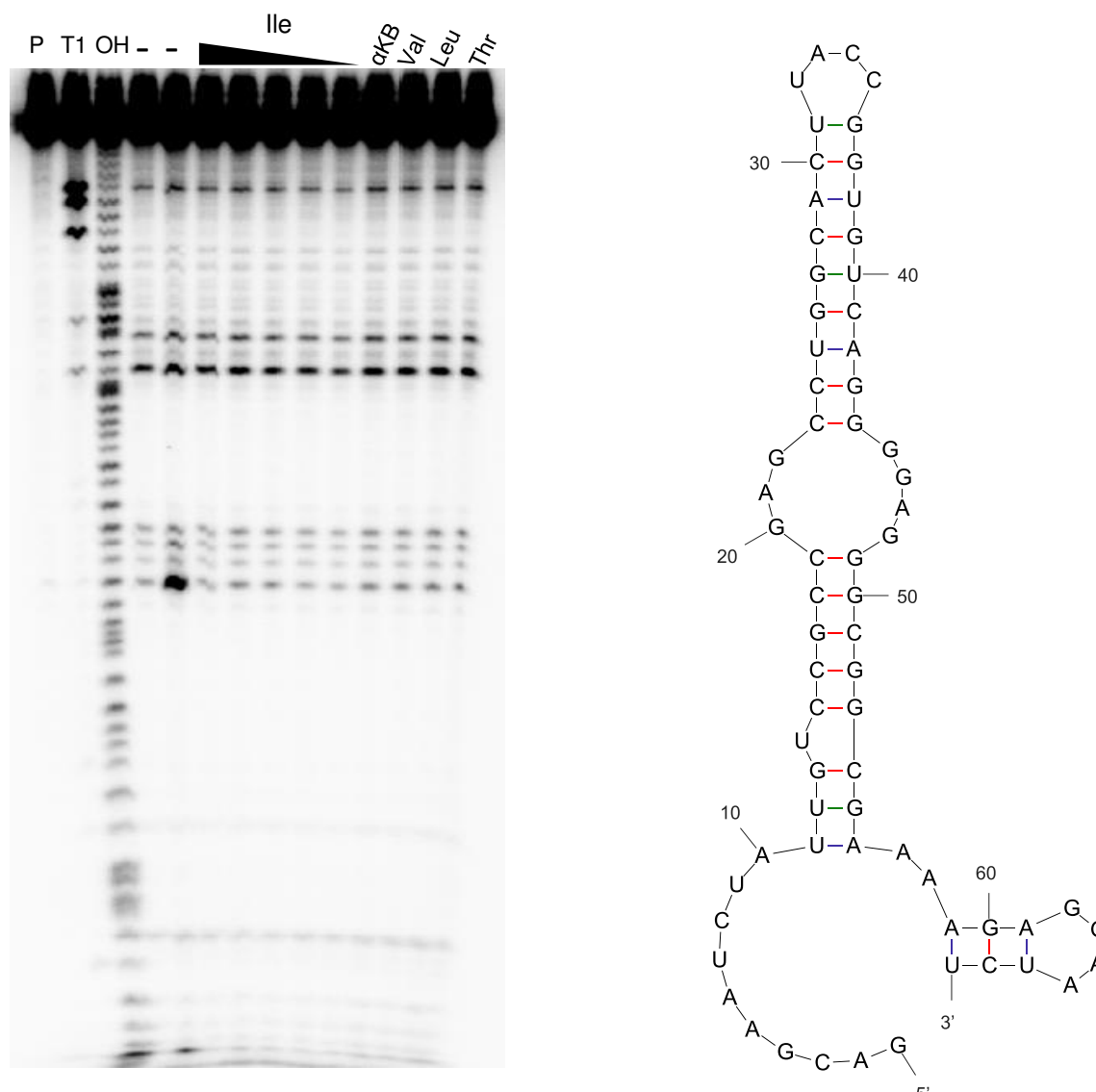


Figure 20: In-line probing of the *ilvH*-motif RNA. Left panel, PAGE analysis of in-line probing reactions of a 68-nucleotide construct of *Comamonas testosteroni* KF-1 *ilvH*-motif RNA. Reactions were conducted in the absence (-) or presence of potential ligands (Ile=isoleucine, αKB=α-ketobutyrate, Val=valine, Leu=leucine, Thr= threonine). P describes the precursor RNA without reaction, T1 stands for the partially digested RNA by RNase T1 and OH stands for a complete digest due to alkaline conditions. Val, Leu, Thr and αKB were inserted in the reactions in a concentration of 100 μM. Isoleucine concentrations tested were 100 μM, 50 μM, 10 μM, 1 μM and 0.1 μM. **Right panel,** Secondary structure of *Comamonas testosteroni* KF-1 68-nucleotide *ilvH*-motif RNA predicted by free energy calculations performed by the mfold algorithm.

This supports a mechanism where isoleucine and αKB are not directly sensed by the *ilvH*-motif RNA and other cellular processes might have led to gene expression changes found in *in vivo* reporter assays. Supplementation of isoleucine or αKB could result in changes of cellular metabolism promoting the production or decay of metabolites that are the actual trigger of gene regulation. It has been shown that BCAAs levels have a direct effect on the amount of (p)ppGpp present in the cell. Binding of BCAAs to RelA activates hydrolase activity of the enzyme leading to an increased decay of (p)ppGpp¹⁸⁰.

Recently, a (p)ppGpp sensing riboswitch was discovered in bacteria that predominantly regulates genes involved in BCAA transport and metabolism. Thus, we tested (p)ppGpp as ligand for a potential *ilvH*-motif RNA riboswitch (conducted by Malte Sinn). However, these preliminary results (not shown in this thesis) indicate that (p)ppGpp is also not bound by the ncRNA.

BCAA biosynthesis gene knockouts influence reporter activity

BCAAs regulate enzyme activity and gene expression of their own biosynthesis pathway products by feedback mechanisms. Supplementation of isoleucine and α KB might thus have led to the shutdown of BCAA biosynthesis enzymes in the *in vivo* reporter assays. A possible result would be the accumulation of BCAA precursors that could also be direct effectors of *ilvH*-motif RNA dependent gene regulation. Since most of the intermediates of BCAA biosynthesis are not commercially available we conducted *in vivo* reporter experiments with knockout strains of BCAA biosynthesis genes carrying the *ilvH*-RNA:*lacZ* reporter plasmid. Since most of single knockouts of BCAA synthesis genes lead to auxotrophy of *E. coli* for BCAAs valine, leucine and isoleucine were supplemented in the assays. For all three amino acids a concentration of 1 mM was used which did not: i) inhibit reporter activity (in case of isoleucine) and ii) negatively affect cell growth (in case of valine). We saw diminished reporter activity for all tested genetic knockouts (Figure 21a) when grown in minimal medium. When the assay was conducted in complex medium only minor genotype dependent effects were observed. The strongest reduction of reporter activity was observed when *IlvA*, the first enzyme in isoleucine biosynthesis, was knocked out. *IlvA* catalyses the reaction from threonine to α KB. We can exclude that Δ *ilvA* mediated accumulation of threonine is the trigger for observed reporter activity since supplementation of external threonine did not alter reporter gene activity (Supplementary Figure 4). An *IlvA* knockout would additionally result in decreased levels of α KB. However, we saw that increasing levels of supplemented α KB reduce reporter activity which is contradictory to the observed results. Thus, it remains elusive which effects evoked by missing *IlvA* reduce reporter activity. In general, we observed genotype dependent differences only in minimal medium supplemented with valine, leucine and isoleucine. Whereas the strongest reduction in reporter activity was around 3-fold in minimal medium (for Δ *ilvA*) only a minor change was observed in LB

medium ($\Delta ilvA$: 468 Miller units to WT: 604 Miller units). These data suggest that growth in minimal medium might enhance the effector that downregulates gene expression *in vivo*.

Functional assignment of the *ilvH*-motif RNA

To further characterize the *ilvH*-motif regulation mechanism, we started to investigate which sub-structures of the *ilvH*-motif RNA might influence its regulatory action. Within that, our attention was guided to the 100% conserved tetraloop (uUACCg) observed at the terminal stem-loop of the *ilvH*-motif RNA. The rfam database assigns the *ilvH*-motif RNA to be related to three different sub-structures: i) a GNRA tetraloop, ii) an UNCG tetraloop and iii) a T-loop. All three substructures represent important features for mediating tertiary structure interactions in many structured RNAs like riboswitches, tRNAs and ribosomal RNAs. Both assigned tetraloop motifs do not match the UACC sequence of the *ilvH*-RNA tetraloop. T-loops were originally described in the T-arm of tRNAs where they mediate interactions with the ribosome¹⁸¹. Furthermore, T-loops mediate RNA-small molecule interactions as shown for the FMN and cobalamin riboswitch. But this structure motif can also be involved in RNA-RNA and other RNA-protein interactions. Common structural features of T-loops are U-turns flanked by non-canonical base-pairs which would match to the uUACCg loop of the *ilvH*-RNA. Thus, we decided to investigate the influence of the tetraloop on regulatory functionality by mutational studies. We constructed *ilvH*-motif RNA reporter plasmids containing mutations in the UACC tetraloop interchanging purines or pyrimidines. Thus, we created different tetraloops variants CACC (Mut1), UGCC (Mut2), AUUC (Mut3), AUCU (Mut4) and a complete exchange of nucleotides CGUU (Mut5) that do not alter the original *ilvH*-RNA secondary structure based on free energy calculations. Reporter assays with different mutant variants revealed that single mutations in the UACC loop lead to neutralization of the isoleucine induced inhibition of reporter activity (Mut1, 3 and 5) (Figure 21b).

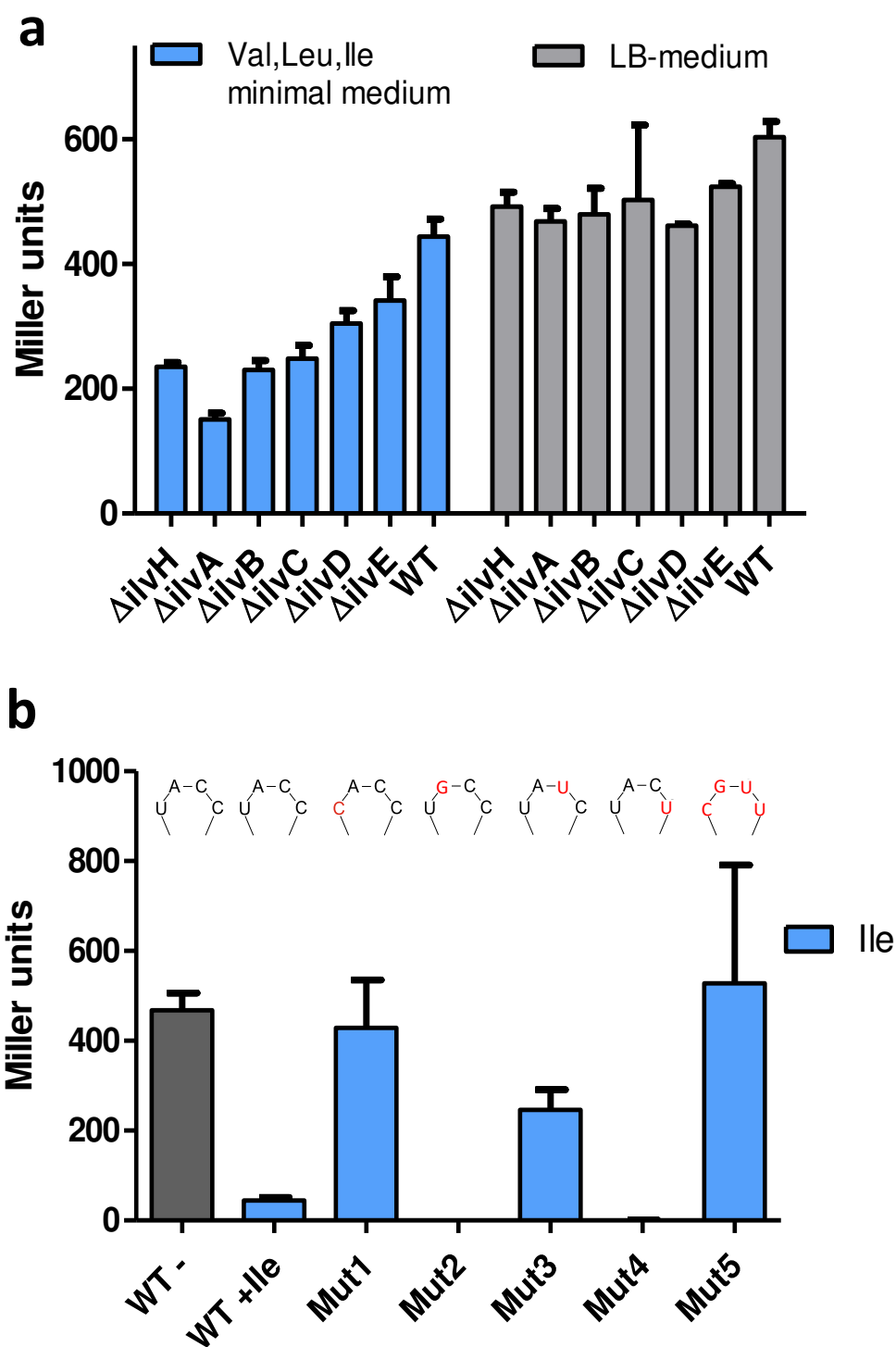


Figure 21: Reporter assays of the *ilvH*:motif RNA. **a**, LacZ activity for strains carrying the reporter plasmid and a single gene knockout of genes involved in BCAA biosynthesis. Strains were grown in liquid minimal medium containing 1 mM Ile, Leu and Val (blue bars) and LB medium (grey bars). **b** LacZ activity assays for reporter strains carrying different mutations in the tetraloop of the *ilvH*:motif RNA in absence (grey bar) or presence (blue bar) of 1 mM Ile. Mutations are indicated above activity bars where UACC is the wildtypic sequence and red letters stand for base exchanges.

However, Mut2 and Mut4 showed no reporter activity which could be caused by mutation-induced misfolding of the RNA or inactivation of the reporter by recombinational events in course of the assay. The results indicate that the UACC tetraloop might be involved in the isoleucine mediated regulation mechanism of the *ilvH*-motif RNA. Interestingly, another conserved structured ncRNA found by comparative sequence analysis also contains a UACC tetraloop, which is 100% conserved. The *ilvB*-OMG-RNA motif was found in the poorly characterized OMG group of marine bacteria. It is predominantly located upstream of *ilvB* (sometimes *ilvH*) and might thus be involved in BCAA metabolism, too. The predicted consensus secondary structure is more complex compared to *ilvH*-RNA. It consists of five conserved stem-loops. The first stem-loop contains the UACC tetraloop motif and has been proposed to be part of a variable secondary structure fold (Supplementary Figure 7). It would be interesting to investigate whether the *ilvB*-OMG-RNA is mediating gene expression regulation in a similar manner compared to the *ilvH*-RNA.

Conclusion

We present evidence that the putative *ilvH*-motif RNA is an active *cis*-regulatory RNA element. It mediates gene expression regulation in dependency of intracellular isoleucine levels whereas the other BCAAs leucine and valine seem not to be involved in the regulatory mechanism. An intermediate of isoleucine biosynthesis, α KB, was also shown to downregulate gene expression via the *ilvH*-motif RNA. However, we could not figure out the exact mechanism on how isoleucine and α KB levels are sensed by the ncRNA. We can exclude that both metabolites bind directly to the RNA which would hint for a classical riboswitch mechanism. Still, it seems possible that the *ilvH*-motif RNA is a riboswitch sensing a metabolite which intracellular levels are indirectly influenced by isoleucine or α KB supplementation. In general, the *ilvH*-motif RNA presents a very stable stem structure that shows only little structural diversity when analysed computationally. Since riboswitches often undergo relevant conformational changes, this might also hint for the *ilvH*-RNA not being a riboswitch. It could thus be possible that the stable structure of the motif is a platform for protein binding. Upon binding of a potential RBP to the motif, the SD, which is in close proximity to the conserved stem could be masked leading to inhibition of translation. It will thus be important to further analyse the RNA-motif concerning its ability to bind proteins. We could already show that the conserved UACC

tetraloop is probably involved in the regulation mechanism of *ilvH*-RNA. Since another ncRNA motif, *ilvB*-OMG, also contains this feature further investigations should be expanded in regard of a potential related functionality to the *ilvH*-RNA. A recent bioinformatic study analysing intergenomic regions regarding ncRNAs found an additional motif which is located in the 5'-UTR of BCAA biosynthesis genes. The motif was called protein binding candidate 10 (PBC-10). However, it contains no obvious featural relation to *ilvH* or *ilvB*-OMG¹⁸². The authors discuss whether a palindromic sequence as part of the motif might be a dimeric protein binding site that could lead to an alternative formation containing a transcription terminator stem. But in this case, experimental evidence is still missing. Investigating these ncRNAs might present the possibility to elucidate a new level of regulation of BCAA biosynthesis.

Materials

Buffers and media

Table 4: Buffers prepared and used in this thesis. Individual chemicals were purchased from either Sigma-Aldrich or Roth Pharmaceutical.

Name	Ingredients	Final concentration
2x RNA denaturing PAGE loading buffer	Formamide	80% (v/v)
	Bromophenol blue	0.5 % (w/v)
	Xylene cyanol	0.5 % (w/v)
His-tag lysis buffer (pH 8)	Imidazole	10 mM
	NaH ₂ PO ₄ x 2 H ₂ O	50 mM
	NaCl	300 mM
His-tag wash buffer (pH 8)	Imidazole	20
	NaH ₂ PO ₄ x 2 H ₂ O	50 mM
	NaCl	300 mM
His-tag elution buffer (pH 8)	Imidazole	500
	NaH ₂ PO ₄ x 2 H ₂ O	50 mM
	NaCl	300 mM
Vitamine solution	4-aminobenzoic acid	80 mg/l
	Cyanocobalamin,	100 mg/l
	D-(+)-biotin	20 mg/l
	Niacin	20 mg/l
	Ca-D-(+)-pantothenic acid	10 mg/l
	Pyridoxamine-chloride	30 mg/l
	Thiamindichlorid	20 mg/l
AA mix for Transcr./Transl. Assay	All 20 proteinogenic amino acids except methionine and cysteine	50 mM
10x Salts for Transcr./Transl. Assay	K ⁺ Glu	1.3 M
	NH ₄ ⁺ Glu	100 mM
	Mg ₂ ⁺ Glu	80 mM
10x Mastermix for Transcr./Transl. Assay	ATP	12 mM
	GTP	8.5 mM
	CTP	8.5 mM
	Folinic acid	340 µg/ml
	<i>E. coli</i> tRNA	1.7 mg/ml
2x In-line probing reaction buffer (pH 8.3)	Tris	100 mM
	MgCl ₂ x 6 H ₂ O	40 mM

Materials

	KCl	200 mM
	PEG 20,000	2 g/l
10x Na₂CO₃ buffer (pH 9)	Na ₂ CO ₃	500 mM
	EDTA pH 8	10 mM
10x Sodium citrate buffer (pH 5)	Tri-Sodium Citrate	250 mM
2x Urea loading buffer	Urea	9 M
	Sucrose	20% (w/v)
	SDS	0.1% (w/v)
	Bromophenol blue	0.05% (w/v)
	Xylene cyanol	0.05% (w/v)
Permeabilisation Buffer	Na ₂ HPO ₄	100 mM
	KCl	20 mM
	MgSO ₄	2 mM
	CTAB	0.8 mg/ml
	Sodium deoxycholate	0.4 mg/ml
	β-mercaptoethanol	5.4 % (v/v)
Substrate solution	Na ₂ HPO ₄	60 mM
	NaH ₂ PO ₄ x 2 H ₂ O	40 mM
	ONPG	1 mg/ml
	β-mercaptoethanol	2.7 % (v/v)
Stop solution	Na ₂ CO ₃	1 M
6x Agarose gel loading buffer	Glycerol	30% (v/v)
	Bromophenol blue	0.25% (w/v)
	Xylencyanol	0.25% (w/v)
1x TBE buffer	Tris Base	10.8 g/l
	Boric acid	5.5 g/l
	EDTA (pH 8.0)	2 mM
Crush-Soak buffer (pH 7.5)	NaCl	200 mM
	EDTA	1 mM
	HEPES	10 mM
10X TBE, 9 M urea	Tris base	108 g/l
	Boric acid	55 g/l
	EDTA	20 mM
	Urea	540 g/l
10x sodium citrate buffer	Tri-sodium citrate x 2H ₂ O	250 mM

Table 5: List of bacterial growth media used in this thesis. Individual chemicals were purchased from either Sigma-Aldrich or Roth Pharmaceutical. MiliQ water was used as solvent and media were sterile filtrated with 0.22 μm filters or autoclaved.

Name	Ingredients	Final concentration
Lysogenic broth (LB) medium, liquid	Yeast extract	5 g/l
	Tryptone	10 g/l
	NaCl	10 g/l
solid	Agar-Agar	20 g/l
M9 minimal medium	NaHPO ₄ x 2H ₂ O	8.5 g/l
	KH ₂ PO ₄	3 g/l
	NaCl	0.5 g/l
	NH ₄ Cl	1 g/l
	MgSO ₄ x 7H ₂ O	1 mM
	CaCl ₂	100 μM
	Glucose	0.2% (v/v)
	Vitamin solution	1 ml/l
Super optimal broth with catabolite repression (SOC medium), liquid	Yeast Extract	5% (w/v)
	Tryptone	0.2% (w/v)
	NaCl	10 mM
	KCl	2.5 mM
	MgCl ₂	10 mM
	MgSO ₄	10 mM
	Glucose	20 mM
	M9 no C-source liquid	NaHPO ₄ x 2H ₂ O
KH ₂ PO ₄		3 g/l
NaCl		0.5 g/l
NH ₄ Cl		1 g/l
MgSO ₄ x 7H ₂ O		1 mM
CaCl ₂		100 μM
Vitamin solution		1 ml/l
M8 no N-source liquid		NaHPO ₄ x 2H ₂ O
	KH ₂ PO ₄	3 g/l
	NaCl	0.5 g/l
	MgSO ₄ x H ₂ O	1 mM
	CaCl ₂	100 μM
	Glucose	0.2% (v/v)
	Vitamin solution	1 ml/l
	1x MOPS minimal medium no C-source liquid	K ₂ HPO ₄
MOPS		40 mM
Tricine		4 mM
FeSO ₄		10 μM

Materials

	NH ₄ Cl	18 mM
	K ₂ SO ₄	276 μM
	CaCl ₂	0.5 μM
	MgCl ₂	523 μM
	NaCl	5 mM
	10x Trace elements	1x
solid	Agar-Agar	20 g/l
1x Salts solution	NH ₄ Cl	18 mM
	K ₂ SO ₄	276 μM
	CaCl ₂	0.5 μM
	MgCl ₂	523 μM
	NaCl	5 mM
	10x Trace elements	1x

Antibiotics

Antibiotics used for selective growth of *E. coli* strains were inserted to the growth media in following concentrations:

Table 6: Antibiotics used for selective growth of *E. coli*.

Antibiotic	Stock conc.	Final conc.
Carbenicillin	100 mg/mL in 50% Ethanol	100 μg/ml
Chloramphenicol	34 mg/mL in 100% Ethanol	34 μg/ml
Kanamycin	30 mg/mL in H ₂ O	30 μg/ml

Chemicals and kits

Table 7: Enzymes and Enzyme buffer used in this thesis.

Enzyme	Manufacturer
Taq DNA Polymerase	Thermo Scientific
T7 RNA Polymerase	Fermentas
Phusion® Hot Start II DNA Polymerase	New England Biolabs (NEB)
rSAP	NEB
T4 Quick Ligase	NEB
RNase inhibitor (Ribolock)	Fermentas
DpnI	NEB
PPase	Thermo Scientific
T4 polynukleotide kinase	Thermo Scientific
NdeI	NEB
XhoI	NEB
HpaI	NEB
Enzyme Buffers	
5x HF-buffer	Thermo Scientific
5x GC-buffer	Thermo Scientific
2x Quick ligase buffer	NEB
10x CutSmart	NEB
10x Thermo Buffer	Thermo Scientific
Buffer A PNK	Thermo Scientific

Table 8: Kits, rulers, dyes and other chemicals.

Kits	Manufacturer
Zymo DNA Clean & Concentrator™ Kit	Zymo Research
Zymoclean™ Gel DNA Recovery Kit	Zymo Research
Zyppy™ Plasmid MiniPrep Kit	Zymo Research
DNeasy® Blood & Tissue Kit	Quiagen
Rulers, dyes & other chemicals	
Midori Green Advanced	Nippon Genetics Europe
Ni-NTA agarose beads	Qiagen
Gene Ruler 1kb	Thermo Scientific

Gene ruler 100 bp	Thermo Scientific
Gene ruler ULR	Thermo Scientific
PageRuler™ Plus Prest.	Thermo Scientific
PageRuler™ Prest.	Thermo Scientific
dNTPs/ NTPs	Thermo Scientific
γ - [³² P]-ATP	Hartmann Analytik
[³⁵ S]Met-label	Hartmann Analytik
Oligonucleotides	Sigma-Aldrich
Rotiphorese Gel 40 (37.5:1)	Roth
10x Rotiphorese SDS-PAGE buffer	Roth
IPTG	Sigma-Aldrich
ONPG	Sigma-Aldrich

Equipment and Consumables

Table 9: Equipment used in the studies.

Name	Manufacturer
Agarose and SDS-PAGE gel equipment	BioRad
Amersham Imager 600	GE Healthcare
Autoclave	Systec
Biacore T200	GE Healthcare
Biometra GelDoc	Biometra
Electroporator 2510	Eppendorf
Erlenmeyer flasks (baffled)	Duran, Schott
French press equipment	Glen mills
Gel drier	BioRad
Incubators	Infors, memmert
Laboratory Pipettes	Eppendorf
Laser scanner Typhoon FLA 7000	GE Healthcare
Multichannel Pipettes	Eppendorf, Pipus
Multipipette	Eppendorf
NanoQuant plate	Tecan
Photometer	Eppendorf

Photostimulable phosphor screens	Fuji
Plate reader Tecan infinite M200	Tecan
Prominence HPLC	Shimadzu
Single Quadrupole Mass spectrometer	Shimadzu
Sonicator	Branson Sonifier 450, Heinemann
Table top centrifuge mini spin	Eppendorf
Thermocycler	Biometra
Thermomixer comfort	Eppendorf
Ultracentrifuge	Beckman Coulter
UV-VIS Photodiode Array Detector	Shimadzu
Vortexer	VWR international

Table 10: Consumables used in the studies.

Consumables	Manufacturer
96 deep well plates	Sarstedt
96-well plates	BD Biosciences
Culture plates	Corning
Electroporation cuvettes 1 mm	Roth
Falcon tubes (15/50 mL)	BD Biosciences
Nitrocellulose blotting membrane 0.2 μ M	Roth
Nylon plus blotting membrane 0.45 μ M	Roth
PCR tubes	Thermo Fisher Scientific
Petri dishes	Sarstedt
Plastic pipette tips	Sarstedt
Polypropylen columbs	Quiagen
Reaction tubes (1.5/ 2 mL)	Sarstedt
Serological pipettes	Sarstedt
UV-cuvette semi-micro	Brand
UVette® routine pack	Eppendorf

Bacterial strains

Table 11: Bacterial strains with genotypes used.

Bacterial strain	Genotype
<i>Escherichia coli</i> XL 10 Gold	Tet ^r Δ(<i>mcrA</i>)183;Δ(<i>mcrCB</i> - <i>hsdSMR</i> - <i>mrr</i>)173; <i>endA1</i> ; <i>supE44</i> ; <i>thi-1</i> <i>recA1</i> ; <i>gyrA96</i> ; <i>relA1</i> ; <i>lac</i> ; <i>Hte</i> [F'; <i>proAB</i> ; <i>lacI</i> ^q ΔM15; Tn10;(Tet ^r); Amy; Cam ^r].
<i>Escherichia coli</i> BL21 gold (DE3)	F ⁻ ; <i>ompT</i> ; <i>hsdS</i> (_{RB- MB-}); <i>dcm</i> + Tet ^r ; <i>gal</i> ; λ(DE3); <i>endA</i> ; <i>Hte</i>
<i>Escherichia coli</i> BW25113 The KEIO collection is based on this strain ¹⁴¹	F ⁻ , Δ(<i>araD-araB</i>)567, Δ <i>lacZ</i> 4787(:: <i>rrnB-3</i>), λ ⁻ , <i>rph-1</i> , Δ(<i>rhaD-rhaB</i>)568, <i>hsdR</i> 514
<i>Escherichia coli</i> K12 AG1 The ASKA collection is based on this strain ¹⁴⁰	F ⁻ , <i>recA1</i> ; <i>endA1</i> ; <i>gyrA96</i> ; <i>thi-1</i> ; <i>hsdR</i> 17(_{r_k m_k⁺}); <i>supE44</i> ; <i>relA1</i>
<i>Comamonas testosteroni</i> KF-1	Wildtype

Oligonucleotides

Table 12: List of oligonucleotides used in the experiments. [P] = phosphorylated nucleotide; [Btn] = biotinylated nucleotide.

Name	Sequence	Description
MS161	CCCGATAAAACGGGGCAGATAA	
MS162	ATAGTCCTGGCCTGAATCGACA	
MS170	TGTTGCTTTTGATCACAATAAG	
MS171	TATGAGATGTAGGGTGACATGG	
MS177	[Btn]CTTTTGTCGCGCATTTT	
MS178	CTCATTCGTAGCCATAA	
SP04	TCCCCATCGGTGATGTC	fw sequencing primer for pET28a
SP10	CTAGTTATTGCTCAGCGG	rv sequencing primer for pET28a
SK80	aaaaaCATATGAATGCACTGACCGCCGTACA	fw primer for <i>csiD</i> (<i>E. coli</i>) gene amplification + NdeI site

SK81	ttttCTCGAGTTACTGATGCGTCTGGTAGT	rv primer for <i>csiD</i> (<i>E. coli</i>) gene amplification + XhoI site
SK99	aaaaaCATATGAACAGCAATAAAGAGTTAATGCAG	fw primer for <i>gabT</i> (<i>E. coli</i>) gene amplification + NdeI site
SK100	ttttCTCGAGCTACTGCTTCGCCTCATCAAA	rv primer for <i>gabT</i> (<i>E. coli</i>) gene amplification + XhoI site
SK101	aaaaaCATATGAAACTTAACGACAGTAACTTATCCG	fw primer for <i>gabD</i> (<i>E. coli</i>) gene amplification + NdeI site
SK102	ttttCTCGAGTTAAAGACCGATGCACATATATTT	rv primer for <i>gabD</i> (<i>E. coli</i>) gene amplification + XhoI site
SK121	aaaaaCATATGACCATTACTCCGGCAACTCA	fw primer for <i>sad</i> (<i>yneI</i>) (<i>E. coli</i>) gene amplification + NdeI site
SK122	ttttCTCGAGTCAGATCCGGTCTTTCCACA	rv primer for <i>sad</i> (<i>yneI</i>) (<i>E. coli</i>) gene amplification + XhoI site
SK123	aaaaaCATATGAGCAACAATGAATTCATCAGC	fw primer for <i>puuE</i> (<i>E. coli</i>) gene amplification + NdeI site
SK124	ttttCTCGAGTTAATCGCTCAGCGCATCCT	rv primer for <i>puuE</i> (<i>E. coli</i>) gene amplification + XhoI site
SK137	aaaaaCCATGGGGATGTATGATT	fw primer for <i>lhgO</i> (<i>E. coli</i>) gene amplification + NcoI site
SK138	ttttCTCGAGTTGATTAAATGCGGCGTGT	rv primer for <i>lhgO</i> (<i>E. coli</i>) gene amplification + XhoI site
MS205	cgacGGATCCTCAGTGGTGGTGGTGGTG	rv primer for <i>lhgO</i> (<i>E. coli</i>) gene amplification + BamHI site, <i>Alhgo</i> complementation
Sk145	CATTAAGAGGAGAAATTA ACTATGAGAGG	fw Primer for ASKA ORF (plasmid pCA24N) sequencing
Sk146	CATCTAATTCAACAAGAATTGGGACA ACTC	rv Primer for ASKA ORF (plasmid pCA24N) sequencing
SK152	CCTTCTATCGCCTTCTTGAC	Sequencing primer for KEIO KanR-ORF transition
SK175	AATGGGCTGACCGCTTCCTC	Primer1 for inverse PCR amplifying HpaII digested KEIO KanR-ORF transition fragment
SK176	CACGGGTAGCCAACGCTATG	Primer2 for inverse PCR amplifying HpaII digested KEIO KanR-ORF transition fragment
SK177	GACACCGGTAAGTGCCAGGCTCGGCGACAATAGATT CGTCAAAAATGCTAGCATTATACCTAGGA	fw primer whole plasmid PCR cloning for <i>ilvH</i> :RNA into pQE_J06(1)
SK178	[P]AGGGGAGGGCGGCGAAAAGAGGAATCTCCCAATGA AACACATTATTGCCGTGCTGACCATGATTACGGATTCA CT	rv primer whole plasmid PCR cloning for <i>ilvH</i> :RNA into pQE_J06(1)
SK197	gatcccgcgaaattaatacactcactataggagGACGAATCTATTGTCC GCCGA	fw primer for start-codon (95 nt) containing <i>ilvH</i> :RNA in-line probing DNA template

Materials

		with T7-promoter sequence
SK198	CAGCACGGCAATAATGTGTT	rv primer for start-codon (95 nt) containing <i>ilvH</i> :RNA in-line probing DNA template
SK199	gatcccgcgaaattaatacgaactcactatagGACGAATCTATTGTCCGC CGA	fw primer no-start-codon (69 nt) containing <i>ilvH</i> :RNA in-line probing DNA template with T7-promoter sequence
SK200	AGATTCCTCTTTTCGCCGCC	rv primer for no-start-codon (69 nt) containing <i>ilvH</i> :RNA in-line probing DNA template
SK213	[P]AGGCTCGGCGGACAATAGATTCGTCaaaaatgctagcattat	rv primer <i>ilvH</i> :RNA loop mutation(1-5) whole plasmid cloning into pQE_J06(1)
SK214	GGCACTCACCGGTGTCAGGGGAGGGcggcgaaaagaggaatct	fw primer <i>ilvH</i> :RNA loop mutation(Mut1) whole plasmid cloning into pQE_J06(1)
SK215	GGCACTTGCCGGTGTTCAGGGGAGGGcggcgaaaagaggaatct	fw primer <i>ilvH</i> :RNA loop mutation(Mut2) whole plasmid cloning into pQE_J06(1)
SK216	GGCACTTATCGGTGTCAGGGGAGGGcggcgaaaagaggaatct	fw primer <i>ilvH</i> :RNA loop mutation(Mut3) whole plasmid cloning into pQE_J06(1)
SK217	GGCACTTACTGGTGTTCAGGGGAGGGcggcgaaaagaggaatct	fw primer <i>ilvH</i> :RNA loop mutation(Mut4) whole plasmid cloning into pQE_J06(1)
SK218	GGCACTCGTTGGTGTTCAGGGGAGGGcggcgaaaagaggaatct	fw primer <i>ilvH</i> :RNA loop mutation(Mut5) whole plasmid cloning into pQE_J06(1)

Plasmids

Created with SnapGene®

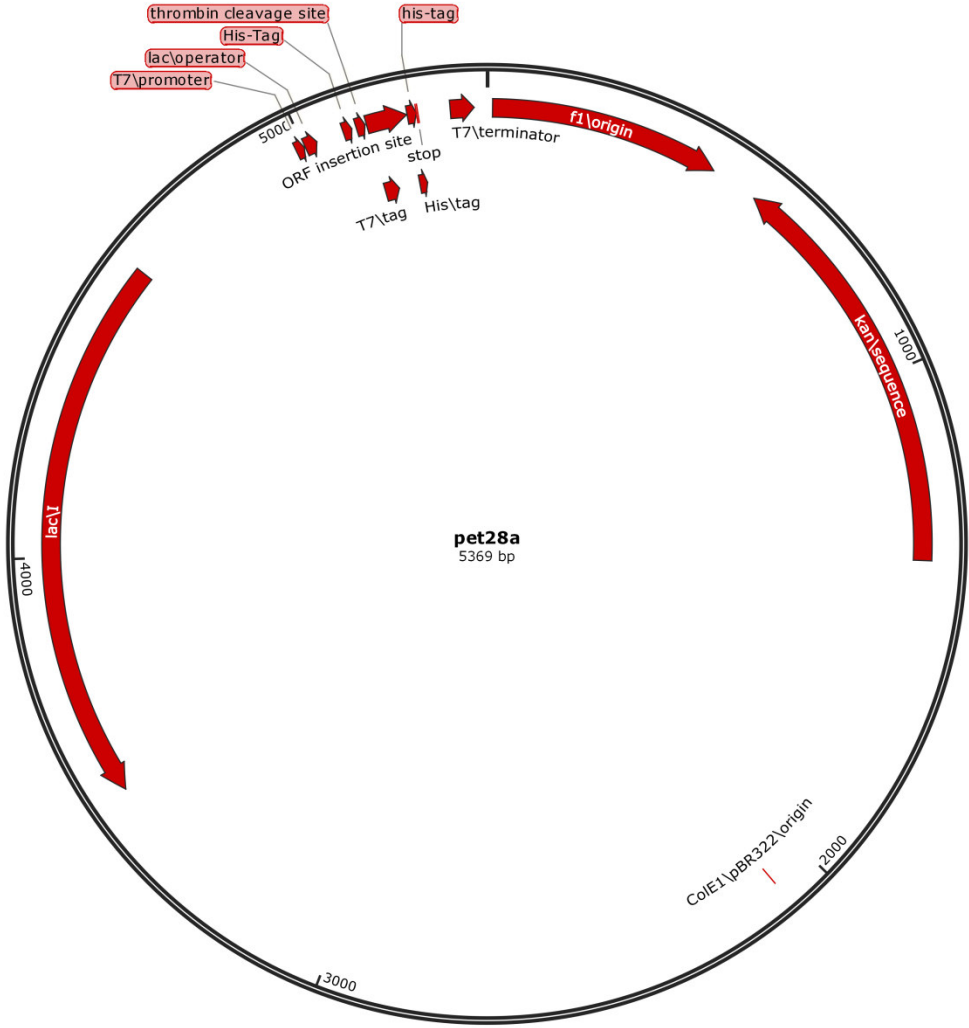


Figure 22: Map of pet28a overexpression vector for recombinant proteins used in these studies.

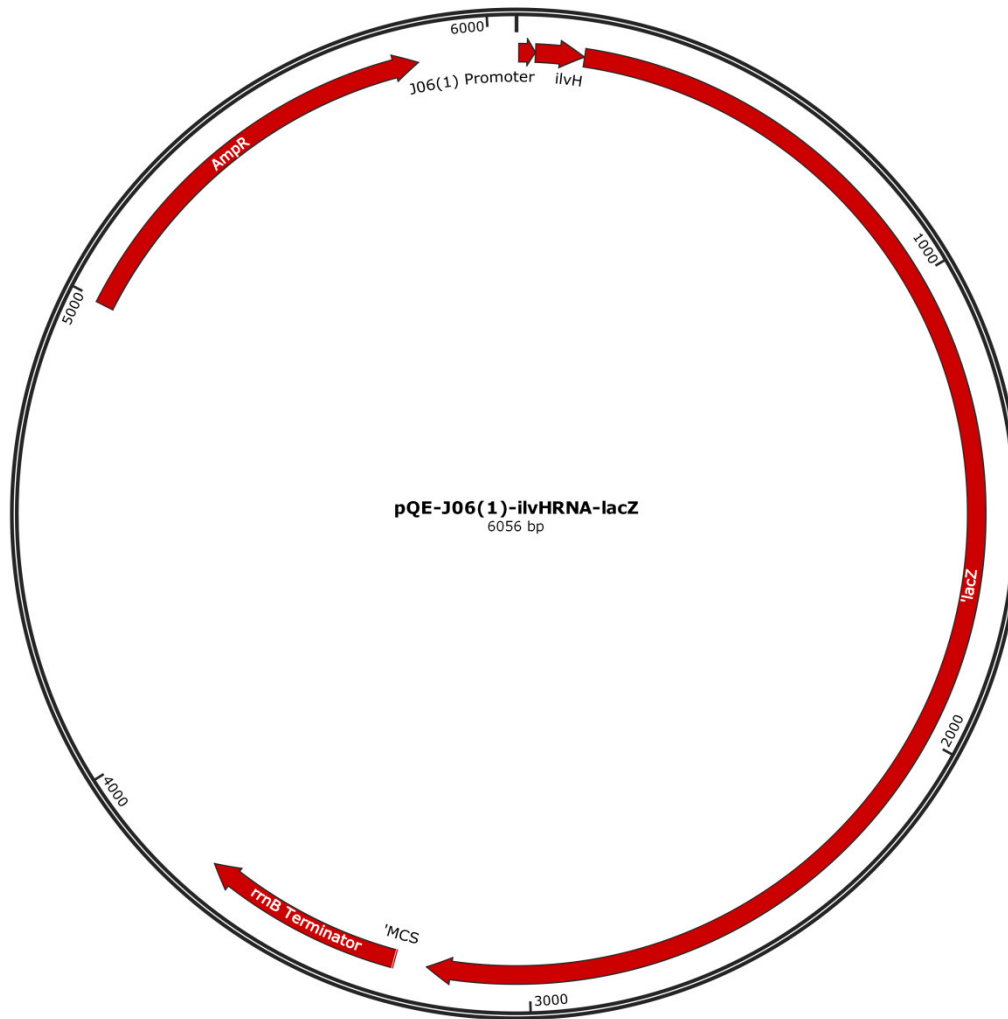


Figure 23: Map of pQE31 derived vector containing J06 promoter, *ilvH*:RNA motif and *lacZ* used for reporter experiments in *E. coli* BW25113.

Software used

Program	Purpose
Graphpad Prism 5.0	Statistics and graph imaging
CorelDraW X5	Imaging of figures
UGENE	Sequence alignments and phylogenetic tree alignment
Bioedit	Checking sequencing chromatograms
TreeGraph	Phylogenetic tree visualisation
Quantity One, ImageQuant	Visualisation of SDS-PAGE and PAGE radiographs

Methods

General methods

Taq DNA Polymerase PCR

PCRs using the *Taq* DNA polymerase were conducted in 50 μ l overall volume using the following reaction setup:

Table 13: Taq PCR standard reaction setup.

Reagent	Volume [μ L]	Stock conc.	Final conc.
Thermostuffer	5	10x	1x
dNTPs	6.5	2 mM	200 μ M
Primer A	2	10 μ M	600 nM
Primer B	2	10 μ M	600 nM
Template	10	1-20 ng/ μ l	0.2 ng/ μ l
DMSO	2	100%	4%
Taq DNA polymerase	0.5	5 U/ μ L	2.5 U
H ₂ O	22		
total	50		

Reactions were conducted in a thermocycler using the following PCR protocol:

Table 14: Taq PCR standard thermocycler protocol.

	Temperature [$^{\circ}$ C]	time	
Initial denaturation	95	5 min	
Denaturation	95	15 s	} 35 cycles
Annealing	55	15 s	
Extension	72	60 s/kbp	
Final extension	72	10 min	

Phusion DNA Polymerase PCR

Table 15: Standard reaction setup for Phusion DNA polymerase PCR.

Reagent	Volume [μ L]	Stock conc.	Final conc.
HF/GC Buffer	10	5x	1x
dNTPs	5	2 mM	200 μ M
Primer A	6	5 μ M	600 nM
Primer B	6	5 μ M	600 nM
Template	1	10 ng/ μ l	0.2 ng/ μ l
DMSO	4	100%	8%
Phusion polymerase	0.5	2 U/ μ l	1 U
H ₂ O	17,5		
total	50		

Table 16: Standard thermocycler protocol for Phusion DNA polymerase PCR.

	Temperature [$^{\circ}$ C]	time	
Initial denaturation	98	30 s	
Denaturation	98	10 s	} 25 cycles
Annealing	58	30 s	
Extension	72	20 s/kbp	
Final extension	72	7 min	

Agarose gel electrophoresis for DNA separation

DNA samples were separated by electrophoresis using gels containing standard agarose. Agarose was solved in 0.5% TBE heated in a microwave and 1 μ l Midori Green Advanced was added for DNA staining. DNA samples in 1x agarose loading dye were loaded together with DNA ladders onto the gels to identify DNA length in base-pairs. Gels were run at 135 V for 30-45 min for qualitative agarose gels and 100 V for 90 min for preparative agarose gels. Detection of DNA bands was conducted under UV-light.

Kit-based DNA isolation and purification techniques

PCR products were directly purified with Zymo DNA Clean & ConcentratorTM Kit using the manufacturer's standard protocol, eluting purified DNA with MiliQ ultra-pure water.

PCR products separated by agarose gel electrophoresis were treated with Zymoclean™ Gel DNA Recovery Kit for purification using the manufacturer's standard protocol. Elution was conducted in MiliQ ultra-pure water.

Plasmid DNA was isolated from *E. coli* cultures applying the standard protocol of the Zyppy™ Plasmid MiniPrep Kit and eluted with MiliQ ultra-pure water.

gDNA was isolated from bacteria using the DNeasy® Blood & Tissue Kit first applying the protocol for "Pretreatment for Gram-Negative bacteria" to lyse bacterial cells. Afterwards the standard protocol "Purification of Total DNA from Animal Tissues (Spin-column Protocol)" was further carried out. gDNA was eluted with MiliQ ultra-pure water.

DNA/RNA concentration determination

DNA/RNA concentration was determined by analysing 2 µl of samples on a NanoQuant plate using plate reader Tecan infinite M200.

DNA sequencing

Sequencing of DNA samples was conducted by GATC Eurofins custom DNA Sanger-Sequencing. DNA concentrations of sequencing samples was adjusted to GATC Eurofins Sanger-Sequencing requirements for PCR products and plasmids.

Plasmid construction using whole plasmid PCR

Short DNA fragments like the *ilvH*:RNA motif (104 nt) were inserted into a template plasmid (Figure 23) by using primers that contained half the sequence of the fragment and a complementary sequence to the insertion site of the plasmid. PCR with these primers leads to the amplification of a linear DNA product containing the split motif at the terminal ends. The standard Phusion polymerase protocol was used adapting the extension time to the expected DNA product length.

After PCR the reaction mixture was incubated with Dpn1 at 37 °C for 1 h to digest the template plasmid.

Methods

The resulting DNA product was separated by agarose gel electrophoresis using high quality agarose at 0.8% and purified by a kit-based gel DNA recovery. 25 ng of DNA was inserted into a ligation reaction incubated for 15 min at 25 °C using the following protocol:

Table 17: Quick Ligase protocol for whole plasmid PCR ligation.

Reagent	Volume [μ L]	Stock	Final conc.
Quick Ligase Buffer	5	2x	1x
DNA	4.5	25 ng	2.5 ng/ μ l
Quick Ligase	0.5		

Ligation product was recovered with a commercial kit and eluted in 7 μ l MiliQ ultrapure water and 1-2 μ l were transformed into electrocompetent cells.

Plasmid construction using restriction enzymes

DNA products were amplified using cell material (1 colony) of bacteria suspended into 100 μ l water as template. 1 μ l of the template was inserted into standard Phusion DNA polymerase PCR protocol. Resulting PCR product was isolated and purified via gel electrophoresis and kit-based clean-up eluting in 43 μ l water. 1 μ l of each restriction enzyme and 5 μ l CutSmart buffer were added and incubated for 1 h at 37 °C. Digested insert DNA was recovered via kit-based purification. 1-3 μ g in 43 μ l water of backbone plasmid DNA was incubated with 1 μ l of each restriction enzyme adding 5 μ l CutSmart buffer for 1 h at 37 °C. Next, 1 μ l of rSAP was added and the mixture was incubated for 1 h at 37 °C for dephosphorylation. Digested and dephosphorylated backbone was recovered via agarose gel electrophoresis followed by kit-based purification of the gel fragment.

Digested insert and backbone were inserted into a ligation reaction for 15 min at 25 °C using following protocol:

Table 18: Quick ligase protocol for restriction enzyme cloning

Reagent	Amount
Quick Ligase Buffer	10 μ l
Insert DNA	0.02 pmol
Backbone DNA	0.1 pmol
Quick Ligase	1 μ l
Water	Up to 20 μ l

Ligated plasmid was recovered via kit-based purification eluting into 7 μ l water and transformed into electropotent cells.

Transformation into electrocompetent cells

1 μ l of purified plasmids was given to 80 μ l of electrocompetent *E. coli* cells thawed on ice. The mixture was rapidly given to a pre-cooled electroporation cuvette and electroporation was conducted in an Eppendorf 2510 Electroporator. The mixture was transferred from the cuvette into pre-warmed (37 °C) SOC medium. After incubation for 30 min the SOC culture was plated on LB-Agar plates containing selective antibiotics.

Ethanol precipitation of DNA and RNA

DNA or RNA samples were supplemented with 1/10 volumes of 3 M NaAcetate and 3 volumes of 100 % pre-cooled (-80 °C) Ethanol and incubated for 1h at -80 °C. Samples were centrifuged for 20 min at 15,000 rpm at 4 °C. The supernatant was discarded and the pellet was resuspended in MiliQ ultrapure water.

In vitro transcription**Table 19: In vitro transcription protocol**

Reagent	Volume [μ L]	Stock	Final conc.
NTPs	15	25 mM	1.875 mM
T7-transcription buffer	40	5x	1x
RiboLock	2	-	-
PPase	1.5	-	-
T7 RNA polymerase	3	-	-
Template DNA	50	From 200 μ l PCR product	-
Water	88.5		
Total	200		

In vitro transcription mixture containing RNA was ethanol precipitated after incubation at 37 °C for 1h.

Polyacrylamide gel electrophoresis (PAGE) for RNA separation and purification

Preparation of a 1.5 mm 5% denaturing PAGE gel was conducting using the following protocol:

Table 20: Protocol for denaturing PAGE gel

Name	Ingredients	Final concentration
Denaturing PAGE gel (5%)	Urea	6.34 M
	10x TBE, Urea buffer	1x
	10% APS	0,08%
	TEMED	0,04%

The gel was allowed to polymerise for at least 1.5 h. RNA samples containing denaturing PAGE loading dye were loaded on the gel. Gel was run at 65 W for app. 2 h. RNA was visualized by UV shadowing or radiograph. Identified RNA band was cut out and crushed manually before suspending with Crush-Soak buffer. The mixture was incubated for 24 h at 4 °C. A syringe filled with glass wool was used as filter and the flow through containing RNA was ethanol precipitated.

SDS-PAGE for protein analysis

Gel electrophoresis was conducted using BioRad SDS gel electrophoresis equipment. A 1 mm gel was prepared consisting of a stacking and resolving gel. The stacking gel contained of 10-16 % (w/v) Rotiphorese Gel 40 (37.5:1), 250 mM Tris-HCl (pH 8.8), 0.1% SDS (w/v), 0.1% APS (w/v) and 0.1% TEMED (w/v), was overlaid with 100% isopropanol and allowed to polymerize for at least 30 min. Isopropanol was removed and a resolving gel consisting of 5% Rotiphorese Gel 40 (37.5:1), 125 mM Tris-HCl (pH 6.8), 0.1% SDS (w/v), 0.1% APS (w/v) and 0.1% TEMED (w/v) was prepared and given on top of the stacking gel. After polymerisation of the resolving gel protein samples were loaded and the gel was run for a maximum of 2 h at 120 V in 1x Rotiphorese SDS-PAGE buffer.

Methods Chapter 2

Biochemical Description of a Novel Lysine-Degradation Pathway in E. coli Proceeding via Glutarate and L2HG

Enzyme nomenclature

Suggestions regarding enzyme nomenclature in the original peer reviewed paper were accepted and disposed in public databases. Formerly *csiD* was renamed to *glaH* encoding glutarate hydroxylase. The formerly annotated L2HG oxidase encoded by *lhgO* was renamed to L2HG dehydrogenase encoded by *lhgD*. The gene encoding the HTH-type transcriptional repressor was renamed from *csiR* to *glaR*.

Bacterial strains and general culture conditions

Escherichia coli BW25113 and its single gene knockout derivatives Δ *csiD*, Δ *csiR*, Δ *gabD*, Δ *gabT*, Δ *glcD*, Δ *lhgO*, Δ *patD* and Δ *patD* from the Keio collection¹⁴¹, that were used for analysis, were all obtained from the National BioResearch Project (NIG, Japan) : *E. coli*. Genes encoding CsiD, GabD, GabT, LhgO, PuuE, Sad were amplified and restriction enzyme sites were introduced with primers listed in Table 12. *Escherichia coli* BW25113 was used as template. PCR fragments were inserted into vector pET28a (Km^r) carrying a T7-promoter under control of *lacI* using standard protocols. Thus, overexpression vectors

Methods

were created for CsiD, GabD, GabT, PuvE and Sad ORFs carrying a 6x N-terminal His-tag and for LhgO carrying a 6x C-terminal His-tag. To create a complemented *lhgO*⁺ strain, *lhgO* was inserted into pET16b using restriction sites and primers listed in Table 12. Recombinant cloning vectors were transformed into *Escherichia coli* BL21 (DE3) (Stratagene, USA, La Jolla) via electroporation. Plasmid constructs were validated by sequencing using primers SP04 and SP10. Bacterial strains were either grown in LB medium or M9 minimal medium containing 8.5 g/l Na₂HPO₄·2H₂O, 3 g/l KH₂PO₄, 0.5 g/l NaCl, 1 g/l NH₄Cl, 2 mM MgCl₂, 100 μM CaCl₂ supplemented with trace elements (0.1 mM EDTA, 0.03 mM FeCl₃, 6.2 μM ZnCl₂, 0.76 μM CuCl₂, 0.42 μM CoCl₂, 1.62 μM H₃BO₃; 0.08 μM MnCl₂) and vitamins (0.1 mg/l cyanocobalamin, 0.08 mg/l 4-aminobenzoic acid, 0.02 mg/l D-(+)-biotin, 0.2 mg/l niacin, 0.1 mg/l Ca-D-(+)-pantothenic acid, 0.3 mg/l pyridoxamine-chloride, 0.2 mg/l thiamindichloride) at 37 °C. As carbon source in minimal medium 0.2 % (w/v) glucose was used. When necessary, medium was supplemented with 30 μg/ml kanamycin.

Growth on different C- and N-sources

E. coli strains were grown in 96 deep well plates in minimal medium as triplicates containing trace elements supplemented with 10 mM C- (lysine, cadaverine, 5-aminovalerate, glutarate) or N-Sources (lysine, cadaverine, 5-amniovalerate). Growth was monitored by measuring OD₆₀₀ in 96 well plates by Tecan plate reader. Growth (means of triplicates) was rated as high growth (0.47>OD₆₀₀>0.43 a.u.); intermediate growth (0.35>OD₆₀₀>0.25 a.u.), low growth (0.15>OD₆₀₀>0.07 a.u.) or no growth (0.02 a.u >OD₆₀₀).

Protein purification

Strains were grown in LB with suitable antibiotic to an OD₆₀₀ of 0.5. Protein expression was induced with 1 mM IPTG and cells were grown for 4-5 h. Cells were harvested by centrifugation at 4000 rpm at 4 °C. Cells were lysed by sonication in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 10 mM imidazole) with Branson sonifier for 3 min duty cycle at 20 % power with 0.5 on and 0.5 s off. Cell debris was removed by centrifugation at 10000 rpm at 4 °C and the cell lysate incubated with Ni-NTA beads (Qiagen) for 1 h at 4 °C at a rotary shaker. Beads were washed twice with wash buffer (50 mM NaH₂PO₄, 300

mM NaCl, 20 mM Imidazole) and the protein finally eluted with elution buffer (50 mM NaH₂PO₄, 300 mM NaCl, 500 mM Imidazole). Purity was checked by SDS-gel. Protein was transferred to the respective buffer system required for the experiment by size exclusion chromatography (PDE-10 columns, GE healthcare).

NMR of the CsiD reaction

The reaction of α -ketoglutarate (10 mM) and glutarate (10 mM) catalysed by CsiD in 20 mM ammonium acetate buffer pH 7.25 was monitored by ¹H NMR spectra without purification. ¹H NMR spectra were recorded at 300 K on Bruker Avance III 400 MHz spectrometers (ayita 400 and isa 400) with a BBFO plus probe for N to F/H or F. Data for NMR spectra were recorded as follows: chemical shift (δ , ppm), multiplicity (s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet), integration, coupling constant (Hz). Acquired data was processed and analysed using MestReNova software.

CsiD reactions analysed by Clark-type O₂ electrode

Reactions were conducted in 100 mM MOPS, 70 mM NaCl and 20 mM KCl pH 7.2 in the presence of 400 μ M Ascorbat and 4 μ M Fe(NH₄)₂SO₄. K_M for glutarate and α -ketoglutarate was determined in the presence of 1 mM α -ketoglutarate and 5 mM glutarate, respectively. 5 mM substrate was used for the C2-C7 analogues of glutarate. 180 nM CsiD were used for the reactions. Reactions were conducted at 30 °C in the sealed reaction chamber with magnetic stirrer. v_0 was determined and plotted against the concentration. The plot was fitted with Michaelis-Menten equation with in GraphPad Prism5 (HYPNOS).

Compounds α -ketopimelate, oxalate, oxaloacetate, α -ketovalerate and α -ketobutyrate and α -keto adipate were tested in reactions as alternative co-substrates in concentrations of 1 mM.

Potential inhibitors of the CsiD reaction were analysed via Clark-type O₂ electrode determining reaction rates in the presence of α -ketopimelate, oxalate, oxaloacetate, α -ketovalerate and α -ketobutyrate, α -keto adipate (all 1 mM), the α KG-dependent dioxygenase inhibitor N-oxalylglycin (NOG) (100 μ M), L2HG (5 mM) and D2HG (5 mM).

Stereospecificity of the CsiD reaction

Evaluation of the 2-hydroxyglutarate enantiomer was performed as described before¹⁰³. The CsiD reaction was performed as before (oxygen electrode) for 1 h at 30 °C. Reaction was incubated for 10 min at 70 °C to heat-inactivate CsiD. Denatured protein was removed by centrifugation (14000 rpm). Supernatant of the CsiD reaction and *L*-2-hydroxyglutarate and *D*-hydroxyglutarate as controls were derivatised with 50 g/l diacetyl-*L*-tartaric anhydride in 4:1 (v/v) dichloromethane/acetic acid for 30 min at 75 °C. Supernatant was evaporated to dryness and residue was dissolved in water. Derivatised products were analysed by LC-MS and identified by comparison to standards.

Crystallisation of CsiD-ligand complexes

Purified CsiD protein was concentrated to 13 mg/mL and its crystallization in liganded form pursued by co-crystallization, crystal soaking or a combination of both approaches, as follows:

Apo CsiD: crystals grew from reservoir solutions containing 80 mM sodium chloride, 12 mM potassium chloride, 20 mM magnesium chloride hexahydrate, 40 mM sodium cacodylate trihydrate pH 6.0, 30 % [v/v] 2-methyl-2,4-pentanediol, 12 mM spermine tetrahydrochloride;

CsiD-GA: crystals from apo CsiD (obtained as described in i) were soaked in 10 mM glutarate prior to flash cooling;

- i) CsiD-SA: CsiD was co-crystallised with succinate by using a mother liquor containing 1.0 M succinic acid pH 7.0, 0.1 M Bis-Tris propane pH 7.0;
- ii) CsiD-NOG: CsiD protein was incubated with NOG at a CsiD:NOG molar ratio of 1:5 for 1 hour on ice and the mixture used in crystallisation trials. Crystals grew from 1.0 M magnesium sulfate, 0.1 M Tris pH 8.5. Prior to flash freezing, crystals were soaked in 20 mM NOG to improve ligand occupancy;

In all cases, crystallization was performed in 96-well Intelliplates (Art Robbins) and used the sitting drop, vapour diffusion format implemented with a Gryphon (Art Robbins) nanolitre dispensing robot. Drops consisted of 200 nL protein or protein/ligand mixture and 200 nL reservoir solution, with reservoirs containing 70 µL mother liquor. All

crystallization trials were incubated at 18°C. For X-ray data collection, crystals were harvested into LithoLoops (Hampton Research) and cryo-protected in mother liquor supplemented with 20% [v/v] glycerol.

X-ray data collection and structure elucidation

X-ray diffraction data were collected at beamline PXI (X06SA) of the Swiss Light Source synchrotron (Villigen, Switzerland) equipped with an Eiger 16 M detector (Dectris, Switzerland). Data were collected at a wavelength of 1.00 Å with 0.1°-0.2° oscillation per frame. Data processing used the XDS/XSCALE package¹⁸³. Phasing was performed by molecular replacement in PHASER¹⁸⁴. First, a CsiD protomer obtained from PDB entry 2R6S¹¹⁰ [10.2210/pdb2R6S/pdb] was used as search model for the elucidation of the apo CsiD structure in this work. The resulting apo model was then used for the phasing of liganded crystal forms of CsiD. Model refinement used phenix.refine¹⁸⁵ and manual model building was performed in COOT¹⁸⁶. Ligand restraints were generated using ELBOW¹⁸⁷. Non-crystallographic symmetry restraints and TLS group identification were calculated automatically in phenix.refine¹⁸⁵.

Phylogenetic distribution of CsiD

A neighbourhood joining phylogeny of CsiD based on *E. coli* K-12 CsiD protein sequence was built retaining sequences with more than 70% coverage and 25% identity. Homologs of *E. coli* CsiD were searched in the refseq database (last access May 2018)¹⁸⁸ using pBLAST. Retained sequences were further analysed by creating a multiple sequence alignment with *E. coli* K-12 CsiD as seed using the ClustalW algorithm implemented in the UGENE software¹⁸⁹. Sequence alignment was analysed concerning residues with 100% conservation. Based conserved residues that are probably involved in glutarate coordination an amino acid pattern (RXLXRXR) was defined. A phiBLAST was initiated with this search pattern using *E. coli* K-12 CsiD sequence as query. Sequences with more than 15% query coverage were retained and a neighbourhood joining phylogeny based on these sequences was created. The web based Skylign HMM logo creator (<http://www.skylign.org/>) was used to visualize the consensus sequence of phiBLAST and pBLAST derived sequences.

Kinetics of GabT/D, Sad/PuuE, and PatA/D

Assays were carried out in a buffer (pH=8.0) containing 5 mM α -ketoglutarate, 500 μ M NADP⁺ (GabT/D & Sad/PuuE) or NAD⁺ (PatA/D), 100 μ M DTT, 100 mM sodiumpyrophosphat and 0.01 mg/ml of both purified GabT/D, Sad/PuuE or PatA/D. Reactions were conducted at room temperature and started by the addition of different concentrations of AVA or GABA (GabT/D & Sad/PuuE) respectively cadaverine or putrescine (PatA/D). Increase of NADPH or NADH was monitored over time measuring the absorbance at 340 nm, thus determining the kinetic of the coupled enzyme reaction. The reduction of NADP⁺ to NADPH corresponds stoichiometrically to the conversion of GABA (or AVA) to succinate semialdehyde (or glutarate semialdehyde) and then to succinate (or glutarate) (GabT/D). In case of PatA/D, NAD⁺ reduction corresponds to the conversion of putrescine (or cadaverine) to aminobutanal (or aminopentanal) followed by oxidation to GABA (or AVA) catalysed by PatD. Decoupling the assay by first starting the transaminase reaction followed by addition of the dehydrogenase did not alter overall velocities concluding that the dehydrogenase reaction must be the rate limiting step in the reaction. Starting velocities (v_0) of the reaction were plotted against substrate concentrations and non-linear regressions were calculated with GraphPad Prism 5 (HYPNOS).

Preparation of *E. coli* membrane fractions

Wildtype and knockout *E. coli* strains were inoculated in 200 ml LB medium and incubated over night at 37 °C at 200 rpm on a rotary shaker. Overnight culture was centrifuged and the pellet was washed twice with ice-cold membrane fraction buffer containing 50 mM HEPES, 10 mM potassiumacetate, 10 M CaCl₂, 5mM MgCl₂ titrated with NaOH to pH=7.5 and resuspended in 10 ml buffer. All following steps were carried out on ice or at 4 °C. Cell suspension was passed four times through a french press cell at 10,000 psi. To remove unlysed cells and insoluble cell debris the lysate was centrifuged at 20,000 x g for 20 minutes. Supernatant was centrifuged for 30 minutes at 100,000 x g. After ultracentrifugation supernatant was stored as cytosolic fraction and the brownish membrane pellet was washed with membrane fraction buffer (5 ml) and centrifuged again. The membrane pellet was resuspended in 500 μ l membrane fraction buffer. Protein

concentration of cytosolic and membrane fraction was determined via BSA-Bradford assay.

Activity of purified LhgO and membrane fractions

E. coli K12 LhgO amino acid sequence was analysed concerning transmembrane helices using TMHMM Server v. 2.0 (<http://www.cbs.dtu.dk/services/TMHMM/>). LhgO activity was assayed spectrophotometrically in a reaction buffer containing 25 mM HEPES, 100 mM NaCl, 5 mM EDTA (pH=7.5). Direct reduction of 100 μ M ubiquinone-1 (UQ₁) (coenzyme Q1, Sigma Aldrich Co.) by purified LhgO was measured by the decrease of absorbance at 278 nm. A final enzyme concentration of 650 nM was used and the reaction was monitored dependent on L-2-hydroxyglutarate concentration. Differential absorption coefficient of UQ₁ at this wavelength in reaction buffer was determined as $\Delta\epsilon = 8.36$ mM. Enzyme activities were determined as UQ₁ reduced per minute. Specific activities were referred to 0.03 mg/ml LhgO used in the assay. To exclude possible effects of H₂O₂ produced by LhgO on redox dyes 100 U/ml catalase (Catalase from bovine liver, Sigma Aldrich Co.) was added. Initial velocities (v_0) of the reaction were plotted against L2HG concentrations and non-linear regressions were calculated with GraphPad Prism 5 (HYPNOS).

Redox activity of 0.1 mg/ml total protein of membrane- and cytosolic-fraction was determined in membrane fraction buffer (50 mM HEPES, 10 mM potassiumacetate, 10 mM CaCl₂, 5 mM MgCl₂, pH=7.5). Reduction of 100 μ M DCPIP by purified LhgO (217 nM) or membrane-bound enzyme was monitored in presence or absence of 50 μ M ubiquinone and menaquinone-4 (MQ₄) (menaquinone K2, Sigma Aldrich Co.) by absorbance measurement at 600 nm. Oxidized DCPIP has an absorbance maximum at 600 nm. Activities of membrane-bound and purified LhgO were determined as DCPIP reduced per minute (nmol/min). DCPIP reduction of *E. coli* WT membranes compared to wildtype cytosolic fraction and membranes isolated from a Δ lhgO strain were conducted in the presence of 50 μ M UQ₁. DCPIP reduction assays were carried out at room temperature and started by the addition of 5 mM L-2-hydroxyglutarate.

Oxygen consumption by membranes and purified LhgO were measured in a Clark-type electrode at room temperature. Activity of membrane-bound LhgO upon titration of the

respiratory chain inhibitor HQNO was determined by oxygen consumption and DCPIP reduction.

Bacterial ubiquinone reduction

Membrane fractions of *E. coli* WT and *ΔlhcO* were prepared as described in the section “Preparation of *E. coli* membrane fractions”. In brief, cells were lysed in a french press and membranes were isolated by ultracentrifugation. Total protein concentration was determined. 150 μg of the membrane fraction in 200 μl membrane fraction buffer (50 mM HEPES, 10 mM potassiumacetate, 10 mM CaCl₂, 5 mM MgCl₂, pH=7.5) were incubated for 1 h at 25 °C under nitrogen atmosphere in the presence or absence of 5 mM L2HG or SA. Quinones were extracted with 1 ml of a 1:3:1 mixture of methanol:hexane:acetone. Samples were dried under a constant flow of nitrogen and analysed by HPLC coupled to a PDA detector (Shimadzu) on a Eurospher RP C18 column (125 x 3 mm; Knauer). Mobile phases contained 100% methanol (A) and 10% heptane in methanol (B). Isocratic elution with 100% A was performed for 10 min followed by a linear gradient from 0% to 50% B in 2 min and 50% to 100% B in 12 min. Spectra were recorded from 200 nm to 400 nm. Area under the curve were determined for ubiquinol (RT: 9 min; 290 ± 4nm) and ubiquinone (RT 16 min; 275 nm ± 4nm). Statistical significance of the data was assessed by an unpaired one-tailed Student’s t-test in GraphPad Prism5 (HYPNOS). The nature of the quinone species and reduction state was confirmed by the corresponding retention times and absorption spectra^{190–192}. Furthermore, ubiquinone from the membrane fractions could be assigned to the mass of ubiquinone-8 by mass spectrometry.

Growth experiments with isotopically labelled lysine

For growth and metabolite analysis of *Escherichia coli* BW25113 WT and *ΔcsiD* cells were inoculated in 100 ml M9 minimal medium supplemented with glucose in the presence of 10 mM lysine in 500 ml baffled flasks. Cultures were grown in duplicates at 37 °C on a rotary shaker at 200 rpm. Cell growth over time was determined spectrophotometrically by measuring optical density at 600 nm in a 1 ml cuvette at indicated intervals. For monitoring intracellular metabolite concentrations during *E. coli*

culture growth samples with a volume equivalent to $OD_{600} = 0.1/\text{ml}$ were taken at indicated time points.

For ^{13}C labeling experiments WT and ΔcsiD cells were inoculated in triplicates supplemented with 10 mM L-lysine- $^{13}\text{C}_6$, $^{15}\text{N}_2$ hydrochloride (Sigma Aldrich Co.) in 5 ml M9 minimal medium containing 0.2 % (w/v) glucose in 50 ml falcons and incubated at 37 °C for 24 hours at 200 rpm. After 24 hours a volume equivalent to $OD_{600} = 0.2/\text{ml}$ was harvested and centrifuged at -9 °C. To rapidly quench metabolic activities the harvested cells were methanol-extracted by resuspending the cell pellet in 2 ml 80 % (v/v) methanol cooled down close to freezing with liquid nitrogen. The suspension was frozen in liquid nitrogen for 1 minute, thawed on ice, centrifuged for 10 minutes at 10,000 x g and the supernatant was stored as metabolite extract. This step was repeated three times. The supernatant fractions were pooled and lyophilized to dryness. Lyophilized metabolite extract was dissolved in 500 μl deionized water and analyzed by LC-MS.

Analysis of lysine degradation metabolites by LC-MS

Identification of metabolites were performed with UltiMate 3000 HPLC system and LTQ Orbitrap Velos (Thermo Scientific). Nucleodur C18 ISIS column (250 mm length x 2 mm i.d., 2.7 μm particle size, Macherey-Nagel, Germany, Düren) was used. The injection volume was 10 μl and the flow rate was 0.25 ml min^{-1} . The mobile phases contained 10 mM ammonium formate (pH 3.2) (A) and 0.1% formic acid in acetonitrile (B). The linear gradient comprised 0 to 30% B for 10 min and 90% B for 2 min. The MS scan ranged from an m/z 100-400 with a resolution of 100,000 at m/z 400 was achieved in positive and negative ionization modes. Accurate mass (± 3 ppm) and retention time (± 0.2 min) values were used for molecular assignment. For the analysis of hydroxyglutarate derivatised with DATAN Prominence HPLC system with LCMS-2020 single quadrupole MS (Shimadzu) was applied. HPLC conditions was as previously described. MS Detection was performed in single ion monitoring (SIM) negative ionization mode at m/z 363. The same instrumentation was used for the quantification of metabolites in cell extracts. Prior to analysis 16 μl of aqueous metabolite extract was mixed with 8 μl phase B (90% acetonitrile, 0.2% formic acid, 10 mM ammonium formate), 5 μl of this solution was injected. Metabolites were separated using Nucleodur HILIC column (250 mm length x 2 mm i.d., 3 μm particle size, Macherey-Nagel), which

was equilibrated with buffer B and eluted with a linear gradient of 45% buffer A (10 mM ammonium formate, pH 3.0) over a 10-min period followed by isocratic step of 45% buffer B for 8 min. Column was operated at $(35.0\pm 0.1)^\circ\text{C}$ with a flow rate of 0.15 ml/min. SIM detection of corresponding protonated ions in positive ionization mode was used for lysine, cadaverine and 5-aminovalerate. Glutarate, succinate and α -ketoglutarate ions were detected in negative SIM mode.

To identify and quantify compounds in cell extracts standard solutions of pure substances were measured under the same conditions. Calibration curves for cadaverine, glutarate and succinate were obtained for the measuring range and used for quantification of these substances. Intracellular metabolite concentrations were calculated according to following equation:

$$C_{avg} = C_{ex} \times V_{ex} \times \frac{DW_{cell}}{DW_{tot} \times V_{cell}} \quad (1)$$

with C_{ex} being the metabolite concentration of the extraction solution determined via external calibration. V_{ex} is the volume of the extraction solution (0.5 ml). DW_{tot} is the experimentally determined total dry weight of the metabolite sample. DW_{cell} is the dry weight per cell (3×10^{-13} g) and V_{cell} is the volume per cell (6.7×10^{-16} l)¹⁹³.

Methods Chapter 3

An ASKA- & KEIO-collection Based Screening-Approach for Identifying New Protein Functions in Amino Acid Metabolism

Production of agar-plates containing C-sources

3.2% (w/v) agar-agar (Bioscience-grade, Roth) was suspended in MiliQ ultrapure water (water-agar) and autoclaved. 2x MOPS medium (containing no C-source and NH_4^+ as N-source) was supplemented with 20 mM of the following C-sources (purchased from Sigma-Aldrich) and sterile filtration was conducted with a 0.22 μm filter:

- L-Lysine
- 5-aminovaleric acid
- Glutaric acid
- Cadaverine

- L-Arginine
- L-Tryptophan
- Indole
- L-Histidine
- L- Leucine
- L-Isoleucine
- L-Valine
- Benzoic acid
- 2,3-Dihydroxybenzoic acid
- 4-Hydroxybenzoic acid
- L-Tyrosine
- L-Phenylalanine
- 4-Hydroxyphenylpyruvic acid
- 4-Hydroxyphenylacetic acid

2x MOPS (containing C-sources) medium used for ASKA-screening was supplemented with 64 µg/ml chloramphenicol (Cm) and 1 mM IPTG, for KEIO-screening 60 µg/ml kanamycin (Km) was added. Per plate 10 ml of 3.2% water-agar (melted in a microwave before use) and 10 ml of 2x MOPS (containing C-sources) medium were mixed and the solution was poured rapidly into petri dishes. Final concentrations of the plates were 10 mM of each C-source and indicated standard concentrations of Cm and Km in 1x MOPS medium and 1.6% (w/v) agar-agar.

ASKA screening experimental setup

The ASKA collection¹⁴⁰ in our lab is stored at -80 °C in 96x deep-well plates containing over-night cultures of each ASKA strain in LB medium. Plates were thawed and 50 µl of each ASKA strain were pooled. Resulting culture was centrifuged and the pellet was treated with a Zyppy™ Plasmid Maxiprep Kit to isolate a pooled fraction of all ASKA plasmids. Next, 115 ng of the pooled ASKA plasmids were transformed into electrocompetent *E. coli* BW25113 cells via electroporation. Cells were resuspended in SOC-medium and incubated for 30 min at 30 °C. Culture was centrifuged for 4 min at 13,000 rpm and the resulting pellet was washed three times and resuspended in 1 ml of 1x Salts Solution. Via measuring OD₆₀₀ and assuming that an OD₆₀₀ = 1 corresponds to 8 x 10⁸ cells/ml, we were able to determine a cell count of 2.8 x 10⁹ cells/ml of the washed culture. We plated 30 µl (~10⁷ cells) on a control agar plate containing 10 mM glucose as C-source and repeated the experiment three times. An average of 30,000 to 40,000

colonies could be determined on the plates which ensures that >95% of the ASKA library is represented under these conditions¹⁹⁴. At next, we plated the pooled ASKA library on agar-plates containing indicated C-sources. Plates were incubated at 37 °C until colony growth was visible. Emerging colonies were regrown in LB-medium and plasmid isolation was performed via Zyppy™ Plasmid MiniPrep Kit. Eluted fractions were sent to sequencing using primers SK145 & SK146 to identify possible ORFs on ASKA plasmids.

KEIO screening experimental setup

50 µl of each KEIO¹⁴¹ strain stored in 96x deep-well plates at -80 °C were pooled. Aliquots of the culture were washed three times in 1x Salts Solution. We could determine an OD₆₀₀ of the suspension that corresponds to 8 x 10⁹ cells/ml. 30 µl (~10⁸) cells were plated on a 10 mM glucose containing control plate to estimate cell growth. More than 100,000 colonies could be observed representing >95% of the KEIO library. Plating was repeated for all agar-dishes containing indicated C-sources. Plates were incubated at 37 °C until colony growth was visible. Colonies were picked and regrown in LB-medium for further identification.

To identify potential KEIO strains, gDNA was extracted from 2 ml LB over-night cultures containing the target strains with a DNeasy® Blood & Tissue Kit. The gDNA was digested with HpaII cutting at 5'-CCGG-3' motifs using 20 ng/µl gDNA, 1 µl HpaII, 1 µl RNaseF1 and 5 µl 10x CutSmart Buffer in an overall volume of 50 µl. The mixture was incubated at 37 °C for 30 – 60 min and after that heat inactivation of the restriction enzyme was conducted at 80 °C for 20 min. Digested fragment were re-ligated using the following protocol:

Table 21: Ligation protocol for KEIO strain digested gDNA.

Reagent	Volume [µL]	Stock	Final conc.
Digested gDNA	1	20 ng/µl	1 ng/µl
ATP	2	10 mM	1 mM
CutSmart Buffer	2	10	1x
T4 DNA Ligase	1	-	-
Water	14		
Total	20		

Ligation was performed for 20 min at 25 °C. Ligation products were purified using Zymo DNA Clean & Concentrator™ Kit and eluted in 20 µl water. 10 µl of the ligation product was inserted into a Taq PCR with primers SK175 & SK176 resulting in an inverse amplification of the Km^R cassette containing ligated fragment using the following protocol:

Table 22: Taq PCR for capturing Kn cassette – ORF transitions.

	Temperature [°C]	time	
Initial denaturation	95	5 min	
Denaturation	95	15 s	} 35 cycles
Annealing	55	15 s	
Extension	72	90 s	
Final extension	72	10 min	

For an extension time of 90 s the Taq DNA polymerase should be able to amplify ~1.5 kb. This extension time was chosen because the average fragment size of the HpaII digest is 191 bp and 99% of the resulting fragments are <1000 bp (Figure 24).

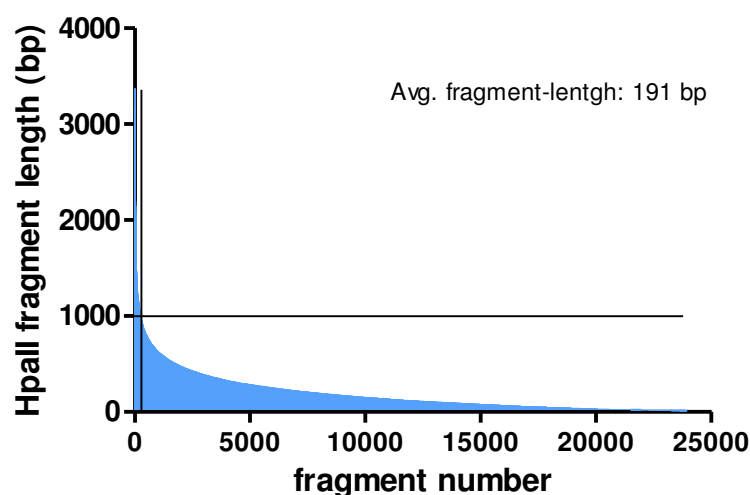


Figure 24: Computational analysis of an HpaII digest of the *E. coli* BW25113 genome. Length of each HpaII digested fragment is shown on the y-axes. The x-axis shows the total count of fragments by HpaII. Area bottom-right of the crosshair indicates the number of fragment smaller than 1 kb. Area left of the vertical line is presenting the number of fragments bigger than 1 kb.

Resulting PCR product was purified via DNA Clean & Concentrator™ Kit. Purified PCR product was sent to sequencing using primer SK152. ORFs disrupted by Km^R cassette were identified using nBLAST. Strains that were ambiguously identified via this procedure were validated by re-growing the original KEIO strain (from -80 °C stock cultures) in 1x MOPS and 1x M9 minimal medium containing the respective C-source. Only when bacterial cell growth could be obtained by measuring an increasing OD₆₀₀, a knockout genotype was assigned to influence growth for the respective C-source.

Methods Chapter 4

The ilvH-motif RNA Regulates Gene Expression at the Crossroad of BCAA Biosynthesis

Reporter construction and growth conditions

A reporter plasmid was constructed by insertion of the *ilvH*:RNA motif from *Comamonas testosteroni* KF-1 into an pQE31 derived vector containing an modified version of the artificial and constitutive active J06 promoter (modified from the Anderson promoter library: <http://parts.igem.org/Promoters/Catalog/Anderson>) and *lacZ* reporter gene (Figure 23 using primers SK177 & SK178 via whole plasmid PCR. The reporter plasmid (WT) was transformed into *E. coli* BW25113 via electroporation. Reporter assays were conducted by inoculating the reporter strain in 1x MOPS medium containing 0.4% glucose in the presence of different concentrations of isoleucine, valine, leucine, threonine and αKB. Additionally, reporter assays were conducted in an knockout environment of $\Delta ilvH$, $\Delta ilvA$, $\Delta ilvB$ $\Delta ilvC$, $\Delta ilvD$ and $\Delta ilvE$. Therefore, KEIO knockout strains were made electrocompetent and the reporter plasmid was transformed into the strains. Reporter plasmid (WT) containing knockout strains and a wildtype *E. coli* were inoculated in 1x MOPS medium containing 0.4% glucose, 1 mM isoleucine, valine and leucine. Mutated *lacZ* reporters were constructed via whole plasmid PCR using primer pairs SK213/214 (Mut1), SK213/215 (Mut2), SK213/216 (Mut3), SK213/217 (Mut4) and SK213/218 (Mut5) and were transformed into *E. coli* BW25113. WT and Mut1-5 reporter strain were inoculated in the presence and absence of 20 mM isoleucine. All reporter assays were conducted in a volume of 500 µl in 48x culture well plates incubated overnight at 37 °C in a rotational shaker at 700 rpm.

Reporter activity measurements (ONPG assay)

Reporter cultures were diluted 1:20 in 1x MOPS minimal medium and OD₆₀₀ was measured. 10 µl of diluted cultures were added to 40 µl of Permeabilisation Solution and pre-incubated at 30 °C for 30 min. Reporter enzyme activity was started by the addition of 300 µl Substrate Solution (containing ONPG as LacZ substrate). Enzyme reaction was stopped by addition of 350 µl 1 M Na₂CO₃ (Stop Solution) and reaction time was monitored. Stopped enzyme reactions were centrifuged for 10 min at 4000 rpm. Supernatant was analysed recording absorption at 420 nm in a plate reader (Tecan infinite M200). Reporter activities were calculating in Miller Units using the following equation:

$$(2) \quad \text{Miller Units} = 1000 * \frac{\text{Abs}_{.420}}{((\text{Abs}_{.600}) * (\text{sampe volume [0.01ml]} * (\text{reaction time [min]}))}$$

In-line probing assays

DNA templates for in-line probing reactions were generated using T7 promoter containing primers SK197/198 (95 nt product) and SK198/199 (69 nt product) via PCR amplification from WT plasmid. DNA templates were purified via Zymo DNA Clean & ConcentratorTM Kit and an *in vitro* transcription reaction was conducted using T7 RNA polymerase. RNA products were purified by PAGE and after gel removal RNA was precipitated with ethanol. 20 pmol of the RNA product were dephosphorylated by rSAP and 172 µl water and 8 µl of 5 M NaCl were added. A phenol/chloroform extraction was performed using Roti-P/C/I for nucleic acid purification (Roth). The formed aqueous phase was again treated with Roti-C/I and RNA was precipitated by ethanol containing 1 µL glycogen instead of NaAcetate. 20 pmol of dephosphorylated and purified RNA products were inserted into an *in vitro* [³²P]-labeling reaction using the following setup:

Table 23: Labeling reaction for RNA.

Reagent	Volume [μL]	Stock	Final amount
γ -[³² P]-ATP	6	-	-
RNA product	-	-	20 pmol
RiboLock	2	-	-
T4 PNK Buffer	4	-	-
T4 PNK	3	200 U/ μ l	3 U/ μ l
Total	Up to 20 μ l		

Radioactive end-labeled RNA was again purified. The amount of RNA equivalent to 1 kBq was supplemented together with the ligand into 1x In-line probing reaction buffer. The mixture was incubated for 40 h at 25 °C. Partial digestion of the RNA by RNase T1 was conducted by incubating 1 kBq of RNA for 5 min at 55 °C in 1x sodium citrate buffer containing 1 U RNase T1 and 1x Urea loading buffer. Complete digest of RNA was conducted by incubating 1 kBq RNA for 5 min at 95 °C in 1x Na₂CO₃ buffer. For precursor samples 1 kBq RNA was supplemented into water. All reactions were conducted in a volume of 10 μ l and quenched by 10 μ l 2x Urea loading buffer. Control- and in-line probing-reactions were analysed by PAGE and radioactivity was imaged with a Typhoon FLA 7000 phosphorimager. Bands were analysed via ImageQuant and QuantityOne imaging softwares.

***In vitro* transcription-translation assay**

An S30 extract was established by culturing *E. coli* K-12 MG1655 in 2x 500 mL LB supplemented with 0.25% (w/v) glucose and 500 μ M MgCl₂ in 5 l baffled Erlenmeyer flasks to OD 2 at 37°C under constant agitation at 210 rpm. Cell culture was rapidly chilled by adding ice into the culture (app. 500 ml beaker). Cells were harvested at 4000x g and 4°C for 20 min. Cells were resuspended in 5 ml chilled S30 buffer per gram wet weight. Cells were harvested again and pellet was re-suspended in 1 ml S30 buffer per gram wet weight. Cells were lysed in a French press at 10,000 psi. Cell lysate was centrifuged at 30,000x g at 4°C for 30 min. Supernatant was carefully transferred into a reaction tube without disturbing the pellet. Ultracentrifugation was repeated. Supernatant was incubated for 60 min at 37°C to remove endogenous mRNA. Supernatant was

centrifuged at 14,000x g at 4°C for 10 min. 100 µl aliquots were frozen in liquid nitrogen and stored at -80°C for up to one year.

Coupled *in vitro* transcription and translation reactions were conducted as followed:

Table 24: Transcription-Translation assay reaction setup.

Reagent	Volume [µL]	Stock	Final amount
10x Salt solution	1.5	10x	1x
10x MasterMix	1.5	10x	1x
AA mix w/o Met and Cys	0.6	50 mM each	2 mM
Pyruvate	0.5	1 M	33 mM
Oxalate	0.6	100 mM	4 mM
Putrescine	0.075	200 mM	1 mM
Spermidine	0.1125	200 mM	1.5 mM
CoA	0.2	20 mM	0.27 mM
NAD	0.05	100 mM	0.33 mM
CTP	0.1125	100 mM	0.75 mM
³⁵ S-label mix	0.3	10 mCi/ml	200 µCi/ml
WT plasmid	-	-	10 µg/ml
H2O	Up to 15 µl	-	

9.9 µL master mix were combined with 1.5 µL ligand, metabolite extract or protein and 3.6 µL S30 extract. Reactions were incubated for 10 h at 37°C. Reactions were quenched by adding SDS-loading dye. Samples were heated to 95°C for 3 min and analyzed on a SDS gel. SDS gels were dried under vacuum in a BioRad gel drier. A photostimulable phosphor screen was exposed overnight. The screen was scanned with a Typhoon FLA 7000 phosphoimager.

Reference List

1. Knorr, S. *et al.* Widespread bacterial lysine degradation proceeding via glutarate and L-2-hydroxyglutarate. doi:10.1038/s41467-018-07563-6
2. Gefen, O., Fridman, O., Ronin, I. & Balaban, N. Q. Direct observation of single stationary-phase bacteria reveals a surprisingly long period of constant protein production activity. *Proc. Natl. Acad. Sci.* **111**, 556–561 (2014).
3. Shaikh, A. S. *et al.* Study of stationary phase metabolism via isotopomer analysis of amino acids from an isolated protein. *Biotechnol. Prog.* **26**, NA-NA (2010).
4. Becker, G., Klauck, E. & Hengge-Aronis, R. Regulation of RpoS proteolysis in *Escherichia coli*: The response regulator RssB is a recognition factor that interacts with the turnover element in RpoS. *Proc. Natl. Acad. Sci.* **96**, 6439–6444 (1999).
5. Schweder, T., Lee, K. H., Lomovskaya, O. & Martin, A. Regulation of *Escherichia coli* starvation sigma factor (σ^s) by ClpXP protease. *J. Bacteriol.* **178**, 470–476 (1996).
6. Hengge-Aronis, R. Signal Transduction and Regulatory Mechanisms Involved in Control of the σ^S (RpoS) Subunit of RNA Polymerase. *Microbiol. Mol. Biol. Rev.* **66**, 373–395 (2002).
7. Shimizu, K. & Kazuyuki. Regulation Systems of Bacteria such as *Escherichia coli* in Response to Nutrient Limitation and Environmental Stresses. *Metabolites* **4**, 1–35 (2013).
8. Jishage, M., Kvint, K., Shingler, V. & Nyström, T. Regulation of sigma factor competition by the alarmone ppGpp. *Genes Dev.* **16**, 1260–70 (2002).
9. Loewen, P. C., Hu, B., Strutinsky, J. & Sparling, R. Regulation in the rpoS regulon of *Escherichia coli*. *Can. J. Microbiol.* **44**, 707–717 (1998).
10. Rahman, M., Hasan, M. R., Oba, T. & Shimizu, K. Effect of rpoS gene knockout on the metabolism of *Escherichia coli* during exponential growth phase and early stationary phase based on gene expressions, enzyme activities and intracellular metabolite concentrations. *Biotechnol. Bioeng.* **94**, 585–595 (2006).
11. Wei, B., Shin, S., LaPorte, D., Wolfe, A. J. & Romeo, T. Global regulatory mutations in csrA and rpoS cause severe central carbon stress in *Escherichia coli* in the presence of acetate. *J. Bacteriol.* **182**, 1632–40 (2000).
12. Zinser, E. R. & Kolter, R. *Escherichia coli* evolution during stationary phase. *Res. Microbiol.* **155**, 328–336 (2004).

13. Vijayakumar, S. R. V, Kirchhof, M. G., Patten, C. L. & Schellhorn, H. E. RpoS-regulated genes of *Escherichia coli* identified by random lacZ fusion mutagenesis. *J. Bacteriol.* **186**, 8499–507 (2004).
14. Reese, A. T. *et al.* Microbial nitrogen limitation in the mammalian large intestine. *Nat. Microbiol.* **3**, 1441–1450 (2018).
15. Sterner, R. W. & Elser, J. J. *Ecological stoichiometry: the biology of elements from molecules to the biosphere.* (Princeton University Press, 2002).
16. Ikeda, T. P., Shauger, A. E. & Kustu, S. Salmonella typhimurium Apparently Perceives External Nitrogen Limitation as Internal Glutamine Limitation. *J. Mol. Biol.* **259**, 589–607 (1996).
17. Reitzer, L. Nitrogen Assimilation and Global Regulation in *Escherichia coli*. *Annu. Rev. Microbiol.* **57**, 155–176 (2003).
18. Helling, R. B. The glutamate dehydrogenase structural gene of *Escherichia coli*. *Mol. Gen. Genet.* **223**, 508–12 (1990).
19. Helling, R. B. Pathway choice in glutamate synthesis in *Escherichia coli*. *J. Bacteriol.* **180**, 4571–5 (1998).
20. Keseler, I. M. *et al.* The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. *Nucleic Acids Res.* **45**, D543–D550 (2017).
21. NG, H. & GARTNER, T. K. Selection of mutants of *Escherichia coli* constitutive for tryptophanase. *J. Bacteriol.* **85**, 245–6 (1963).
22. Kim, M. *et al.* Need-based activation of ammonium uptake in *Escherichia coli*. *Mol. Syst. Biol.* **8**, 616 (2012).
23. Atkinson, M. R. & Ninfa, A. J. Role of the GlnK signal transduction protein in the regulation of nitrogen assimilation in *Escherichia coli*. *Mol. Microbiol.* **29**, 431–447 (1998).
24. Stülke, J. & Hillen, W. Carbon catabolite repression in bacteria. *Curr. Opin. Microbiol.* **2**, 195–201 (1999).
25. Magasanik, B. Catabolite Repression. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 249–256 (1961).
26. Reeve, C. A., Amy, P. S. & Matin, A. Role of protein synthesis in the survival of carbon-starved *Escherichia coli* K-12. *J. Bacteriol.* **160**, 1041 (1984).
27. Pirt, S. J. Maintenance energy: a general model for energy-limited and energy-sufficient growth. *Arch. Microbiol.* **133**, 300–302 (1982).
28. Artsimovitch, I. *et al.* Structural basis for transcription regulation by alarmone

- ppGpp. *Cell* **117**, 299–310 (2004).
29. Chauvier, A. *et al.* Transcriptional pausing at the translation start site operates as a critical checkpoint for riboswitch regulation. *Nat. Commun.* **8**, 13892 (2017).
 30. KURODA, A. A Polyphosphate-Lon Protease Complex in the Adaptation of *Escherichia coli* to Amino Acid Starvation. *Biosci. Biotechnol. Biochem.* **70**, 325–331 (2006).
 31. Rozen, Y., LaRossa, R. A., Templeton, L. J., Smulski, D. R. & Belkin, S. Gene expression analysis of the response by *Escherichia coli* to seawater. *Antonie Van Leeuwenhoek* **81**, 15–25 (2002).
 32. Richter, D. Uncharged tRNA inhibits guanosine 3',5'-bis (diphosphate) 3'-pyrophosphohydrolase [ppGppase], the spoT gene product, from *Escherichia coli*. *MGG Mol. Gen. Genet.* **178**, 325–327 (1980).
 33. Tchieu, J. H., Norris, V., Edwards, J. S. & Saier, M. H. The complete phosphotransferase system in *Escherichia coli*. *J. Mol. Microbiol. Biotechnol.* **3**, 329–46 (2001).
 34. Postma, P. W., Lengeler, J. W. & Jacobson, G. R. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* **57**, 543–94 (1993).
 35. Hogema, B. M. *et al.* Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. *Mol. Microbiol.* **30**, 487–498 (1998).
 36. Park, Y.-H., Lee, B. R., Seok, Y.-J. & Peterkofsky, A. In vitro reconstitution of catabolite repression in *Escherichia coli*. *J. Biol. Chem.* **281**, 6448–54 (2006).
 37. Zhu, M. & Dai, X. Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in *Escherichia coli*. *Nucleic Acids Res.* **47**, 4684–4693 (2019).
 38. Jiang, P. & Ninfa, A. J. *Escherichia coli* PII signal transduction protein controlling nitrogen assimilation acts as a sensor of adenylate energy charge in vitro. *Biochemistry* **46**, 12979–12996 (2007).
 39. Jiang, P. & Ninfa, A. J. Alpha-ketoglutarate controls the ability of the *Escherichia coli* PII signal transduction protein to regulate the activities of NRII (NrB but does not control the binding of PII to NRII. *Biochemistry* **48**, 11514–21 (2009).
 40. Ninfa, A. J. & Jiang, P. PII signal transduction proteins: sensors of alpha-ketoglutarate that regulate nitrogen metabolism. *Curr. Opin. Microbiol.* **8**, 168–73

- (2005).
41. Rhee, S. G., Chock, P. B. & Stadtman, E. R. Regulation of *Escherichia coli* glutamine synthetase. *Adv. Enzymol. Relat. Areas Mol. Biol.* **62**, 37–92 (1989).
 42. Kamberov, E. S., Atkinson, M. R. & Ninfa, A. J. The *Escherichia coli* PII signal transduction protein is activated upon binding 2-ketoglutarate and ATP. *J. Biol. Chem.* **270**, 17797–17807 (1995).
 43. Jiang, P., Peliska, J. A. & Ninfa, A. J. The Regulation of *Escherichia coli* Glutamine Synthetase Revisited: Role of 2-Ketoglutarate in the Regulation of Glutamine Synthetase Adenylylation State †. *Biochemistry* **37**, 12802–12810 (1998).
 44. Jiang, P. & Ninfa, A. J. α -Ketoglutarate Controls the Ability of the *Escherichia coli* PII Signal Transduction Protein To Regulate the Activities of NRII (NtrB) but Does Not Control the Binding of PII to NRII. *Biochemistry* **48**, 11514–11521 (2009).
 45. van Heeswijk, W. C., Westerhoff, H. V. & Boogerd, F. C. Nitrogen Assimilation in *Escherichia coli*: Putting Molecular Data into a Systems Perspective. *Microbiol. Mol. Biol. Rev.* **77**, 628–695 (2013).
 46. Jiang, P. & Ninfa, A. J. *Escherichia coli* PII Signal Transduction Protein Controlling Nitrogen Assimilation Acts As a Sensor of Adenylate Energy Charge *in Vitro* †. *Biochemistry* **46**, 12979–12996 (2007).
 47. Jiang, P. & Ninfa, A. J. Sensation and Signaling of α -Ketoglutarate and Adenylate Energy Charge by the *Escherichia coli* PII Signal Transduction Protein Require Cooperation of the Three Ligand-Binding Sites within the PII Trimer. *Biochemistry* **48**, 11522–11531 (2009).
 48. Metzner, M., Germer, J. & Hengge, R. Multiple stress signal integration in the regulation of the complex σ S-dependent *csiD-ygaF-gabDTP* operon in *Escherichia coli*. *Mol. Microbiol.* **51**, 799–811 (2003).
 49. Gorden, J. & Small, P. L. Acid resistance in enteric bacteria. *Infect. Immun.* **61**, 364–7 (1993).
 50. Kanjee, U. & Houry, W. A. Mechanisms of Acid Resistance in *Escherichia coli*. *Annu. Rev. Microbiol.* **67**, 65–81 (2013).
 51. De Biase, D. & Lund, P. A. The *Escherichia coli* Acid Stress Response and Its Significance for Pathogenesis. *Adv. Appl. Microbiol.* **92**, 49–88 (2015).
 52. Kikuchi, Y., Kojima, H., Tanaka, T., Takatsuka, Y. & Kamio, Y. Characterization

- of a second lysine decarboxylase isolated from *Escherichia coli*. *J. Bacteriol.* **179**, 4486–4492 (1997).
53. Wu, W. H. & Morris, D. R. Biosynthetic arginine decarboxylase from *Escherichia coli*. Purification and properties. *J. Biol. Chem.* **248**, 1687–95 (1973).
54. Gevrekci, A. Ö. The roles of polyamines in microorganisms. *World J. Microbiol. Biotechnol.* **33**, 204 (2017).
55. Yoshida, M. *et al.* A unifying model for the role of polyamines in bacterial cell growth, the polyamine modulon. *J. Biol. Chem.* **279**, 46008–13 (2004).
56. Igarashi, K. & Kashiwagi, K. Polyamine Modulon in *Escherichia coli*: Genes Involved in the Stimulation of Cell Growth by Polyamines. *J. Biochem.* **139**, 11–16 (2006).
57. Chiang, S. M. & Schellhorn, H. E. Regulators of oxidative stress response genes in *Escherichia coli* and their functional conservation in bacteria. *Arch. Biochem. Biophys.* **525**, 161–169 (2012).
58. Brinkman, A. B., Ettema, T. J. G., De Vos, W. M. & Van Der Oost, J. The Lrp family of transcriptional regulators. *Mol. Microbiol.* **48**, 287–294 (2003).
59. Tani, T. H., Khodursky, A., Blumenthal, R. M., Brown, P. O. & Matthews, R. G. Adaptation to famine: A family of stationary-phase genes revealed by microarray analysis. *Proc. Natl. Acad. Sci.* **99**, 13471–13476 (2002).
60. Gottesman, S. & Storz, G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. *Cold Spring Harb. Perspect. Biol.* **3**, (2011).
61. Chowdhury, S., Ragaz, C., Kreuger, E. & Narberhaus, F. Temperature-controlled Structural Alterations of an RNA Thermometer. *J. Biol. Chem.* **278**, 47915–47921 (2003).
62. Lopatovskaya, K. V., Seliverstov, A. V. & Lyubetsky, V. A. Attenuation regulation of the amino acid and aminoacyl-tRNA biosynthesis operons in bacteria: A comparative genomic analysis. *Mol. Biol.* (2010). doi:10.1134/s0026893310010164
63. Vitreschak, A. G., Mironov, A. A., Lyubetsky, V. A. & Gelfand, M. S. Comparative genomic analysis of T-box regulatory systems in bacteria. doi:10.1261/rna.819308
64. Henkin, T. M. tRNA-directed transcription antitermination. *Mol. Microbiol.* **13**, 381–387 (1994).
65. Henkin, T. M., Glass, B. L. & Grundy, F. J. Analysis of the *Bacillus subtilis* tyrS

- gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. *J. Bacteriol.* **174**, 1299–306 (1992).
66. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. & Breaker, R. R. Riboswitch diversity and distribution. *RNA* **23**, 995–1011 (2017).
 67. Serganov, A. & Nudler, E. A decade of riboswitches. *Cell* **152**, (2013).
 68. Breaker, R. R. Riboswitches and Translation Control. *Cold Spring Harb. Perspect. Biol.* **10**, a032797 (2018).
 69. Sudarsan, N., Wickiser, J. K., Nakamura, S., Ebert, M. S. & Breaker, R. R. An mRNA structure in bacteria that controls gene expression by binding lysine. *Genes Dev.* **17**, 2688–97 (2003).
 70. Grundy, F. J., Lehman, S. C. & Henkin, T. M. The L box regulon: Lysine sensing by leader RNAs of bacterial lysine biosynthesis genes. *Proc. Natl. Acad. Sci.* **100**, 12057–12062 (2003).
 71. Garst, A. D., Héroux, A., Rambo, R. P. & Batey, R. T. Crystal Structure of the Lysine Riboswitch Regulatory mRNA Element. *J. Biol. Chem.* **283**, 22347–22351 (2008).
 72. Serganov, A., Huang, L. & Patel, D. J. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* **455**, 1263–1267 (2008).
 73. van Hoek, A. H. A. M. *et al.* Prevalence and characterization of ESBL- and AmpC-producing Enterobacteriaceae on retail vegetables. *Int. J. Food Microbiol.* **204**, 1–8 (2015).
 74. Mandal, M. *et al.* A Glycine-Dependent Riboswitch That Uses Cooperative Binding to Control Gene Expression. *Science (80-.).* **306**, 275–279 (2004).
 75. Lipfert, J. *et al.* Structural Transitions and Thermodynamics of a Glycine-Dependent Riboswitch from *Vibrio cholerae*. *J. Mol. Biol.* **365**, 1393–1406 (2007).
 76. Welz, R. & Breaker, R. R. Ligand binding and gene control characteristics of tandem riboswitches in *Bacillus anthracis*. *RNA* **13**, 573–82 (2007).
 77. Kwon, M. & Strobel, S. A. Chemical basis of glycine riboswitch cooperativity. *RNA* **14**, 25–34 (2008).
 78. Klähn, S. *et al.* A glutamine riboswitch is a key element for the regulation of glutamine synthetase in cyanobacteria. *Nucleic Acids Res.* **46**, 10082–10094 (2018).
 79. Sherlock, M. E., Sudarsan, N. & Breaker, R. R. Riboswitches for the alarmone ppGpp expand the collection of RNA-based signaling systems. *Proc. Natl. Acad.*

- Sci. U. S. A.* **115**, 6052–6057 (2018).
80. Kriner, M. A., Sevostyanova, A. & Groisman, E. A. Learning from the Leaders: Gene Regulation by the Transcription Termination Factor Rho. *Trends Biochem. Sci.* **41**, 690–699 (2016).
81. Peters, J. M. *et al.* Rho directs widespread termination of intragenic and stable RNA transcription. *PNAS September 8*, (2009).
82. Stewart, V., Landick, R. & Yanofsky, C. Rho-dependent transcription termination in the tryptophanase operon leader region of Escherichia coli K-12. *J. Bacteriol.* **166**, 217–23 (1986).
83. Yang, M. *et al.* Alanine-scanning mutagenesis of Bacillus subtilis trp RNA-binding attenuation protein (TRAP) reveals residues involved in tryptophan binding and RNA binding. *J. Mol. Biol.* **270**, 696–710 (1997).
84. Babitzke, P., Bear, D. G. & Yanofsky, C. TRAP, the trp RNA-binding attenuation protein of Bacillus subtilis, is a toroid-shaped molecule that binds transcripts containing GAG or UAG repeats separated by two nucleotides. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7916–20 (1995).
85. Mondal, S., Yakhnin, A. V & Babitzke, P. Modular Organization of the NusA- and NusG-Stimulated RNA Polymerase Pause Signal That Participates in the Bacillus subtilis trp Operon Attenuation Mechanism. *J. Bacteriol.* **199**, e00223-17 (2017).
86. Yakhnin, A. V & Babitzke, P. NusA-stimulated RNA polymerase pausing and termination participates in the Bacillus subtilis trp operon attenuation mechanism invitro. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 11067–72 (2002).
87. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res.* **42**, D459–D471 (2014).
88. Sabo, D. L., Boeker, E. A., Byers, B., Waron, H. & Fischer, E. H. Purification and physical properties of inducible Escherichia coli lysine decarboxylase. *Biochemistry* **13**, 662–670 (1974).
89. Lemonnier, M. & Lane, D. Expression of the second lysine decarboxylase gene of Escherichia coli. *Microbiology* **144** (Pt 3, 751–760 (1998).
90. FOTHERGILL, J. C. & GUEST, J. R. Catabolism of L-Lysine by Pseudomonas aeruginosa. *J. Gen. Microbiol.* **99**, 139–155 (1977).
91. Revelles, O., Espinosa-Urgel, M., Molin, S. & Ramos, J. L. The davDT Operon of Pseudomonas putida, Involved in Lysine Catabolism, Is Induced in Response to the

- Pathway Intermediate -Aminovaleric Acid. *J. Bacteriol.* **186**, 3439–3446 (2004).
92. Metzner, M., Germer, J. & Hengge, R. Multiple stress signal integration in the regulation of the complex sigma S-dependent *csiD-ygaF-gabDTP* operon in *Escherichia coli*. *Mol Microbiol* **51**, 799–811 (2004).
93. Marschall, C. *et al.* Molecular analysis of the regulation of *csiD*, a carbon starvation-inducible gene in *Escherichia coli* that is exclusively dependent on sigma s and requires activation by cAMP-CRP. *J Mol Biol* **276**, 339–353 (1998).
94. Germer, J., Becker, G., Metzner, M. & Hengge-Aronis, R. Role of activator site position and a distal UP-element half-site for sigma factor selectivity at a CRP/H-NS-activated sigma(s)-dependent promoter in *Escherichia coli*. *Mol Microbiol* **41**, 705–716 (2001).
95. Feehily, C. & Karatzas, K. A. G. Role of glutamate metabolism in bacterial responses towards acid and other stresses. *J. Appl. Microbiol.* **114**, 11–24 (2013).
96. Kalliri, E., Mulrooney, S. B. & Hausinger, R. P. Identification of *Escherichia coli* YgaF as an L-2-hydroxyglutarate oxidase. *J. Bacteriol.* **190**, 3793–8 (2008).
97. Chance, M. R. *et al.* Structural genomics: a pipeline for providing structures for the biologist. *Protein Sci* **11**, 723–738 (2002).
98. Samsonova, N. N., Smirnov, S. V, Altman, I. B. & Ptitsyn, L. R. Molecular cloning and characterization of *Escherichia coli* K12 *ygiG* gene. *BMC Microbiol.* **3**, 2 (2003).
99. Jorge, J. M. P., Perez-Garcia, F. & Wendisch, V. F. A new metabolic route for the fermentative production of 5-aminovalerate from glucose and alternative carbon sources. *Bioresour Technol* **245**, 1701–1709 (2017).
100. Brechtel, C. E., Hu, L. & King, S. C. Substrate specificity of the *Escherichia coli* 4-aminobutyrate carrier encoded by *gabP*. Uptake and counterflow of structurally diverse molecules. *J. Biol. Chem.* **271**, 783–8 (1996).
101. Martinez, S. & Hausinger, R. P. Catalytic Mechanisms of Fe(II)- and 2-Oxoglutarate-dependent Oxygenases. *J Biol Chem* **290**, 20702–20711 (2015).
102. Islam, M. S., Leissing, T. M., Chowdhury, R., Hopkinson, R. J. & Schofield, C. J. 2-Oxoglutarate-Dependent Oxygenases. *Annu Rev Biochem* (2018). doi:10.1146/annurev-biochem-061516-044724
103. Struys, E. A., Jansen, E. E., Verhoeven, N. M. & Jakobs, C. Measurement of urinary D- and L-2-hydroxyglutarate enantiomers by stable-isotope-dilution liquid chromatography-tandem mass spectrometry after derivatization with diacetyl-L-

- tartaric anhydride. *Clin Chem* **50**, 1391–1395 (2004).
104. Zhang, M. *et al.* Increased glutarate production by blocking the glutaryl-CoA dehydrogenation pathway and a catabolic pathway involving 1-2-hydroxyglutarate. *Nat. Commun.* **9**, 2114 (2018).
105. Herr, C. Q., Macomber, L., Kalliri, E. & Hausinger, R. P. Glutarate L-2-hydroxylase (CsiD/GlaH) is an archetype Fe(II)/2-oxoglutarate-dependent dioxygenase. in (2019). doi:10.1016/bs.apcsb.2019.05.001
106. Werf, M. J. van der, Overkamp, K. M., Muilwijk, B., Coulier, L. & Hankemeier, T. Microbial metabolomics: Toward a platform with full metabolome coverage. *Anal. Biochem.* **370**, 17–25 (2007).
107. Price, M. N. *et al.* Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* **557**, 503–509 (2018).
108. Xu, W. *et al.* Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17–30 (2011).
109. Chowdhury, R. *et al.* The oncometabolite 2-hydroxyglutarate inhibits histone lysine demethylases. *EMBO Rep.* **12**, 463–469 (2011).
110. Lohkamp, B. & Dobritzsch, D. A mixture of fortunes: the curious determination of the structure of Escherichia coli BL21 Gab protein. *Acta Crystallogr D Biol Crystallogr* **64**, 407–415 (2008).
111. Aik, W. S. *et al.* CHAPTER 2. Introduction to Structural Studies on 2-Oxoglutarate-Dependent Oxygenases and Related Enzymes. in 59–94 (2015). doi:10.1039/9781782621959-00059
112. Martinez, S. & Hausinger, R. P. Biochemical and Spectroscopic Characterization of the Non-Heme Fe(II)- and 2-Oxoglutarate-Dependent Ethylene-Forming Enzyme from *Pseudomonas syringae* pv. *phaseolicola* PK2. *Biochemistry* **55**, 5989–5999 (2016).
113. Thompson, M. G. *et al.* Massively parallel fitness profiling reveals multiple novel enzymes in *Pseudomonas putida* lysine metabolism. doi:10.1101/450254
114. El-Gebali, S. *et al.* The Pfam protein families database in 2019. *Nucleic Acids Res.* **47**, D427–D432 (2019).
115. Samsonova, N. N., Smirnov, S. V, Novikova, A. E. & Ptitsyn, L. R. Identification of Escherichia coli K12 YdcW protein as a gamma-aminobutyraldehyde dehydrogenase. *FEBS Lett* **579**, 4107–4112 (2005).
116. Schneider, B. L. & Reitzer, L. Pathway and enzyme redundancy in putrescine

- catabolism in *Escherichia coli*. *J Bacteriol* **194**, 4080–4088 (2012).
117. Kurihara, S. *et al.* A novel putrescine utilization pathway involves gamma-glutamylated intermediates of *Escherichia coli* K-12. *J Biol Chem* **280**, 4602–4608 (2005).
118. Rzem, R., Van Schaftingen, E. & Veiga-da-Cunha, M. The gene mutated in 1-2-hydroxyglutaric aciduria encodes 1-2-hydroxyglutarate dehydrogenase. *Biochimie* **88**, 113–116 (2006).
119. Hudig, M. *et al.* Plants Possess a Cyclic Mitochondrial Metabolic Pathway similar to the Mammalian Metabolic Repair Mechanism Involving Malate Dehydrogenase and 1-2-Hydroxyglutarate Dehydrogenase. *Plant Cell Physiol* **56**, 1820–1830 (2015).
120. Uden, G. & Bongaerts, J. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochim Biophys Acta* **1320**, 217–234 (1997).
121. Sharma, P., Teixeira de Mattos, M. J., Hellingwerf, K. J. & Bekker, M. On the function of the various quinone species in *Escherichia coli*. *FEBS J.* **279**, 3364–3373 (2012).
122. Bekker, M. *et al.* The ArcBA two-component system of *Escherichia coli* is regulated by the redox state of both the ubiquinone and the menaquinone pool. *J Bacteriol* **192**, 746–754 (2010).
123. Kanehisa, M. *et al.* Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* **42**, D199–D205 (2014).
124. Adkins, J., Jordan, J. & Nielsen, D. R. Engineering *Escherichia coli* for renewable production of the 5-carbon polyamide building-blocks 5-aminovalerate and glutarate. *Biotechnol. Bioeng.* **110**, 1726–1734 (2013).
125. Park, S. J. *et al.* Metabolic engineering of *Escherichia coli* for the production of 5-aminovalerate and glutarate as C5 platform chemicals. *Metab Eng* **16**, 42–47 (2013).
126. Zhang, M. *et al.* Increased glutarate production by blocking the glutaryl-CoA dehydrogenation pathway and a catabolic pathway involving 1-2-hydroxyglutarate. *Nat. Commun.* **9**, 2114 (2018).
127. Ye, D., Guan, K. L. & Xiong, Y. Metabolism, Activity, and Targeting of D- and L-2-Hydroxyglutarates. *Trends Cancer* **4**, 151–165 (2018).
128. Kranendijk, M., Struys, E. A., Salomons, G. S., Van der Knaap, M. S. & Jakobs, C.

- Progress in understanding 2-hydroxyglutaric acidurias. *J Inherit Metab Dis* **35**, 571–587 (2012).
129. Intlekofer, A. M. *et al.* L-2-Hydroxyglutarate production arises from noncanonical enzyme function at acidic pH. *Nat Chem Biol* **13**, 494–500 (2017).
130. Aghili, M., Zahedi, F. & Rafiee, E. Hydroxyglutaric aciduria and malignant brain tumor: a case report and literature review. *J Neurooncol* **91**, 233–236 (2009).
131. Rzem, R. *et al.* A gene encoding a putative FAD-dependent L-2-hydroxyglutarate dehydrogenase is mutated in L-2-hydroxyglutaric aciduria. *Proc Natl Acad Sci U S A* **101**, 16849–16854 (2004).
132. Hedlund, G. L., Longo, N. & Pasquali, M. Glutaric acidemia type 1. *Am J Med Genet C Semin Med Genet* **142C**, 86–94 (2006).
133. Sherman, E. A. *et al.* Genetic mapping of glutaric aciduria, type 3, to chromosome 7 and identification of mutations in *c7orf10*. *Am J Hum Genet* **83**, 604–609 (2008).
134. Wendel, U., Bakkeren, J., de Jong, J. & Bongaerts, G. Glutaric aciduria mediated by gut bacteria. *J Inherit Metab Dis* **18**, 358–359 (1995).
135. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32 (2016).
136. Tropini, C., Earle, K. A., Huang, K. C. & Sonnenburg, J. L. The Gut Microbiome: Connecting Spatial Organization to Function. *Cell Host Microbe* **21**, 433–442 (2017).
137. Kovács, T. *et al.* Cadaverine, a metabolite of the microbiome, reduces breast cancer aggressiveness through trace amino acid receptors. *Sci. Rep.* **9**, 1300 (2019).
138. Sévin, D. C., Fuhrer, T., Zamboni, N. & Sauer, U. Nontargeted in vitro metabolomics for high-throughput identification of novel enzymes in *Escherichia coli*. *Nat. Methods* **14**, 187–194 (2016).
139. Piazza, I. *et al.* A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell* **172**, 358-372.e23 (2018).
140. Kitagawa, M. *et al.* Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12**, 291–9 (2005).
141. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
142. Dvorak, P. *et al.* Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microb. Cell Fact.* **14**, 201

- (2015).
143. Hoegler, K. J. & Hecht, M. H. Artificial Gene Amplification in *Escherichia coli* Reveals Numerous Determinants for Resistance to Metal Toxicity. *J. Mol. Evol.* **86**, 103–110 (2018).
 144. Soo, V. W. C., Hanson-Manful, P. & Patrick, W. M. Artificial gene amplification reveals an abundance of promiscuous resistance determinants in *Escherichia coli*. doi:10.1073/pnas.1012108108
 145. Verstraeten, N. *et al.* Obg and Membrane Depolarization Are Part of a Microbial Bet-Hedging Strategy that Leads to Antibiotic Tolerance. *Mol. Cell* **59**, 9–21 (2015).
 146. Wilmaerts, D. *et al.* The Persistence-Inducing Toxin HokB Forms Dynamic Pores That Cause ATP Leakage. *MBio* **9**, (2018).
 147. Gerdes, K. Hypothesis: type I toxin–antitoxin genes enter the persistence field—a feedback mechanism explaining membrane homeostasis. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20160189 (2016).
 148. Tenorio, E. *et al.* Systematic characterization of *Escherichia coli* genes/ORFs affecting biofilm formation. *FEMS Microbiol. Lett.* **225**, 107–114 (2003).
 149. Lomize, A. L., Lomize, M. A., Krolicki, S. R. & Pogozheva, I. D. Membranome: a database for proteome-wide analysis of single-pass membrane proteins. *Nucleic Acids Res.* **45**, D250–D255 (2017).
 150. ITO, T. *et al.* The Implication of YggT of *Escherichia coli* in Osmotic Regulation. *Biosci. Biotechnol. Biochem.* **73**, 2698–2704 (2009).
 151. Pomposiello, P. J., Koutsolioutsou, A., Carrasco, D. & Demple, B. SoxRS-regulated expression and genetic analysis of the yggX gene of *Escherichia coli*. *J. Bacteriol.* **185**, 6624–32 (2003).
 152. Kaur, G. & Subramanian, S. Classification of the treble clef zinc finger: noteworthy lessons for structure and function evolution. *Sci. Rep.* **6**, 32070 (2016).
 153. Raman, B., Nandakumar, M. P., Muthuvijayan, V. & Marten, M. R. Proteome analysis to assess physiological changes in *Escherichia coli* grown under glucose-limited fed-batch conditions. *Biotechnol. Bioeng.* **92**, 384–392 (2005).
 154. Aguilera, L. *et al.* Dual Role of LldR in Regulation of the lldPRD Operon, Involved in L-Lactate Metabolism in *Escherichia coli*. *J. Bacteriol.* **190**, 2997–3005 (2008).
 155. Hodkinson, M. R. G., Allen, L. M., Thomson, D. P. & Sayers, J. R. Molecular

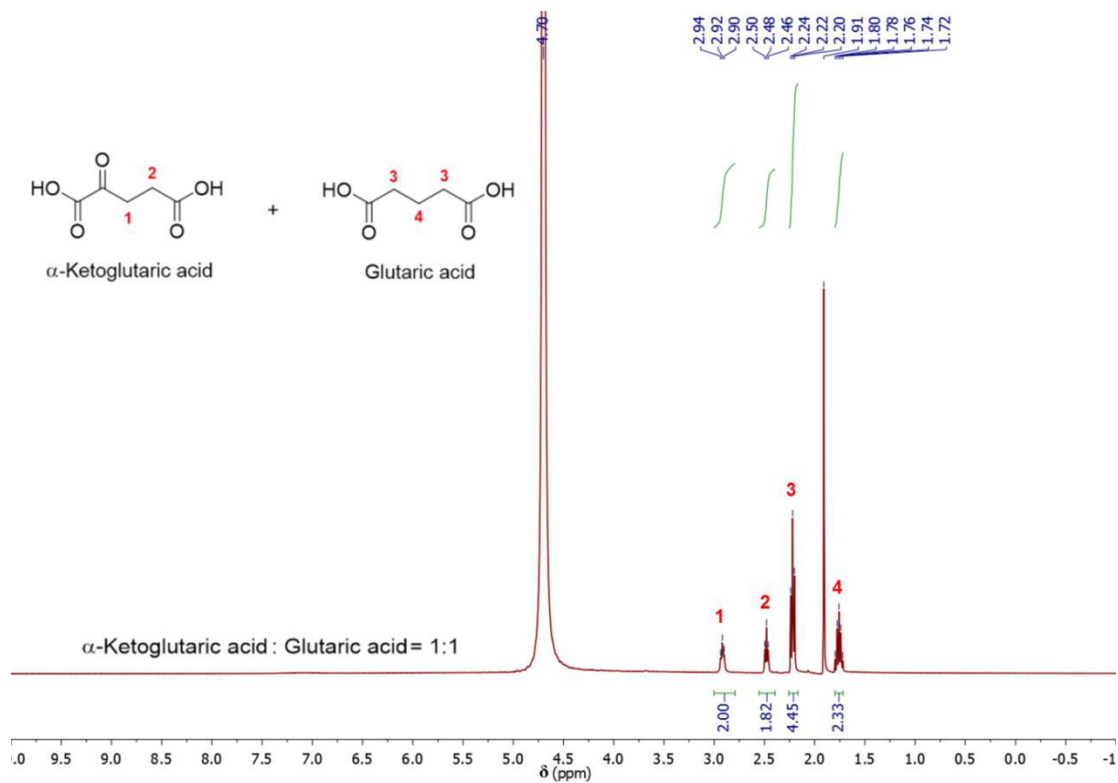
- interactions of *Escherichia coli* ExoIX and identification of its associated 3'-5' exonuclease activity. *Nucleic Acids Res.* **35**, 4094–4102 (2007).
156. Anstey-Gilbert, C. S. *et al.* The structure of *Escherichia coli* ExoIX—implications for DNA binding and catalysis in flap endonucleases. *Nucleic Acids Res.* **41**, 8357–8367 (2013).
157. Maharjan, R. P. *et al.* A case of adaptation through a mutation in a tandem duplication during experimental evolution in *Escherichia coli*. *BMC Genomics* **14**, (2013).
158. Artymiuk, P. J., Green, J., Quail, M. A., Tang, Y. & Guest, J. R. *Escherichia coli* aconitases and oxidative stress: post-transcriptional regulation of *sodA* expression. *Microbiology* **148**, 1027–1037 (2002).
159. Tang, Y. & Guest, J. R. Direct evidence for mRNA binding and post-transcriptional regulation by *Escherichia coli* aconitases. *Microbiology* **145**, 3069–3079 (1999).
160. Cunningham, L., Gruer, M. J. & Guest, J. R. Transcriptional regulation of the aconitase genes (*acnA* and *acnB*) of *Escherichia coli*. *Microbiology* **143**, 3795–3805 (1997).
161. Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956 (2002).
162. Nahvi, A. *et al.* Genetic control by a metabolite binding mRNA. *Chem. Biol.* **9**, 1043 (2002).
163. Mironov, A. S. *et al.* Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**, 747–56 (2002).
164. Parashar, A. Aptamers in Therapeutics. *J. Clin. Diagn. Res.* **10**, BE01-6 (2016).
165. Deigan, K. E. & Ferré-D'Amaré, A. R. Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Acc. Chem. Res.* **44**, 1329–38 (2011).
166. Lünse, C. E., Schüller, A. & Mayer, G. The promise of riboswitches as potential antibacterial drug targets. *Int. J. Med. Microbiol.* **304**, 79–92 (2014).
167. Mehdizadeh Aghdam, E., Hejazi, M. S. & Barzegar, A. Riboswitches: From living biosensors to novel targets of antibiotics. *Gene* (2016). doi:10.1016/j.gene.2016.07.035
168. Nelson, J. W., Atilho, R. M., Sherlock, M. E., Stockbridge, R. B. & Breaker, R. R. Metabolism of Free Guanidine in Bacteria Is Regulated by a Widespread Riboswitch Class. *Mol. Cell* (2016). doi:10.1016/j.molcel.2016.11.019

169. Sherlock, M. E., Malkowski, S. N. & Breaker, R. R. Biochemical Validation of a Second Guanidine Riboswitch Class in Bacteria. *Biochemistry* [acs.biochem.6b01270](https://doi.org/10.1021/acs.biochem.6b01270) (2017). doi:10.1021/acs.biochem.6b01270
170. Sherlock, M. E. & Breaker, R. R. Biochemical Validation of a Third Guanidine Riboswitch Class in Bacteria. *Biochemistry* [acs.biochem.6b01271](https://doi.org/10.1021/acs.biochem.6b01271) (2017). doi:10.1021/acs.biochem.6b01271
171. Weinberg, Z. *et al.* Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* **11**, R31 (2010).
172. Weinberg, Z. *et al.* Detection of 224 candidate structured RNAs by comparative analysis of specific subsets of intergenic regions. *Nucleic Acids Res.* **45**, 10811–10823 (2017).
173. Kalvari, I. *et al.* Non-Coding RNA Analysis Using the Rfam Database. *Curr. Protoc. Bioinforma.* **62**, e51 (2018).
174. Kalvari, I. *et al.* Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res.* **46**, D335–D342 (2018).
175. Gollop, N., Damri, B., Barak, Z. & Chipman, D. M. Kinetics and mechanism of acetohydroxy acid synthase isozyme III from *Escherichia coli*. *Biochemistry* **28**, 6310–6317 (1989).
176. Barak, Z., Chipman, D. M. & Gollop, N. Physiological implications of the specificity of acetohydroxy acid synthase isozymes of enteric bacteria. *J. Bacteriol.* **169**, 3750–3756 (1987).
177. Vyazmensky, M. *et al.* Interactions between Large and Small Subunits of Different Acetohydroxyacid Synthase Isozymes of *Escherichia coli*. *Biochemistry* **48**, 8731–8737 (2009).
178. LaRossa, R. A., Van Dyk, T. K. & Smulski, D. R. Toxic accumulation of alpha-ketobutyrate caused by inhibition of the branched-chain amino acid biosynthetic enzyme acetolactate synthase in *Salmonella typhimurium*. *J. Bacteriol.* **169**, 1372–8 (1987).
179. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
180. Fang, M. & Bauer, C. E. Regulation of stringent factor by branched-chain amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 6446–6451 (2018).
181. Chan, C. W., Chetnani, B. & Mondragón, A. Structure and function of the T-loop structural motif in noncoding RNAs. *Wiley Interdiscip. Rev. RNA* **4**, 507–522

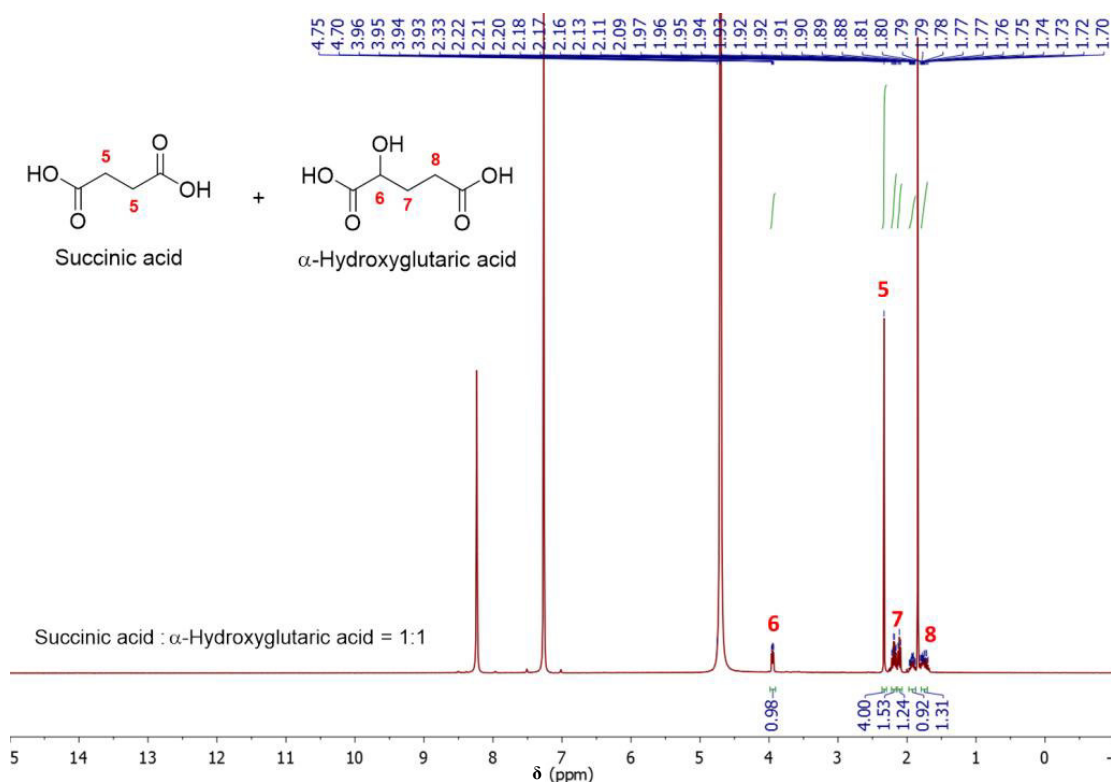
- (2013).
182. Stav, S. *et al.* Genome-wide discovery of structured noncoding RNAs in bacteria. doi:10.1186/s12866-019-1433-7
 183. Kabsch, W. XDS. *Acta Crystallogr. Sect. D* **66**, 125–132 (2010).
 184. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
 185. Afonine, P. V *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**, 352–367 (2012).
 186. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501 (2010).
 187. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. Sect. D* **65**, 1074–1080 (2009).
 188. O’Leary, N. A. *et al.* Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745 (2016).
 189. Okonechnikov, K., Golosova, O. & Fursov, M. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).
 190. Carlone, G. M. & Anet, F. A. Detection of menaquinone-6 and a novel methyl-substituted menaquinone-6 in *Campylobacter jejuni* and *Campylobacter fetus* subsp. *fetus*. *J Gen Microbiol* **129**, 3385–3393 (1983).
 191. Poderoso, J. J. *et al.* The reaction of nitric oxide with ubiquinol: kinetic properties and biological significance. *Free Radic. Biol. Med.* **26**, 925–935 (1999).
 192. BAUM, R. H. & DOLIN, M. I. ISOLATION OF 2-SOLANESYL-1,4-NAPHTHOQUINONE FROM STREPTOCOCCUS FAECALIS, 10 CL. *J. Biol. Chem.* **240**, 3425–33 (1965).
 193. Bennett, B. D., Yuan, J., Kimball, E. H. & Rabinowitz, J. D. Absolute quantitation of intracellular metabolite concentrations by an isotope ratio-based approach. *Nat Protoc* **3**, 1299–1311 (2008).
 194. Patrick, W. M., Quandt, E. M., Swartzlander, D. B. & Matsumura, I. Multicopy Suppression Underpins Metabolic Evolvability. *Mol. Biol. Evol.* **24**, 2716 (2007).

Supplementary Information

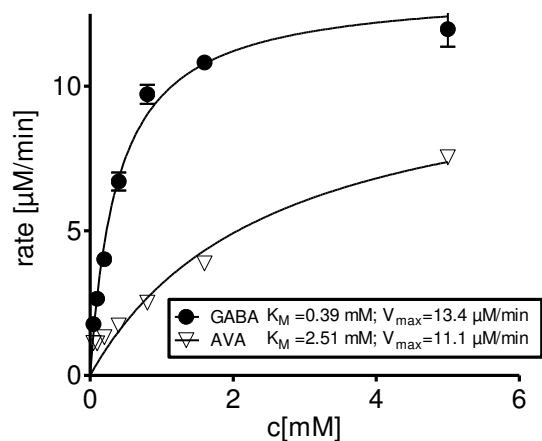
Supplementary Figures



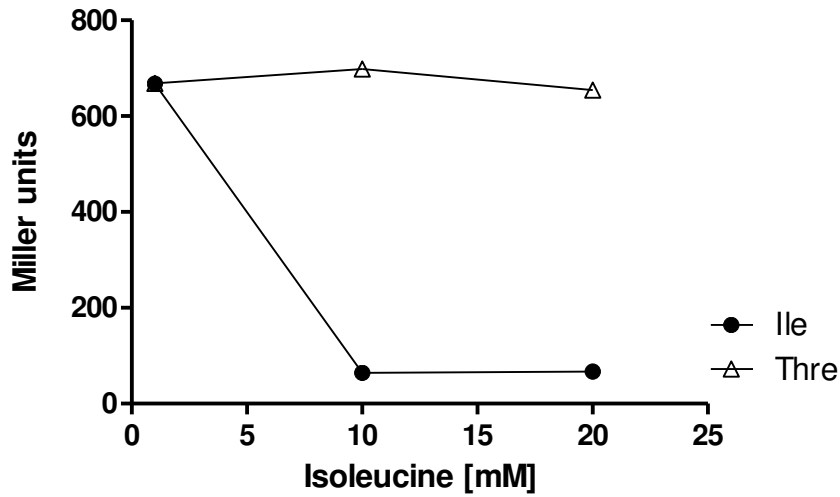
Supplementary Figure 1: ^1H NMR spectra of α -ketoglutarate and glutarate in ammonium acetate buffer. ^1H NMR (400 MHz, D_2O , α -ketoglutaric acid : glutaric acid = 1:1) 2.92 (t, $J = 6.8$ Hz, 2H, **1-H**), 2.48 (t, $J = 6.9$ Hz, 2H, **2-H**), 2.22 (t, $J = 7.6$ Hz, 4H, **3-H**), 1.76 (p, $J = 7.6$ Hz, 2H, **4-H**). Peak at 1.91 ppm arises from acetate.



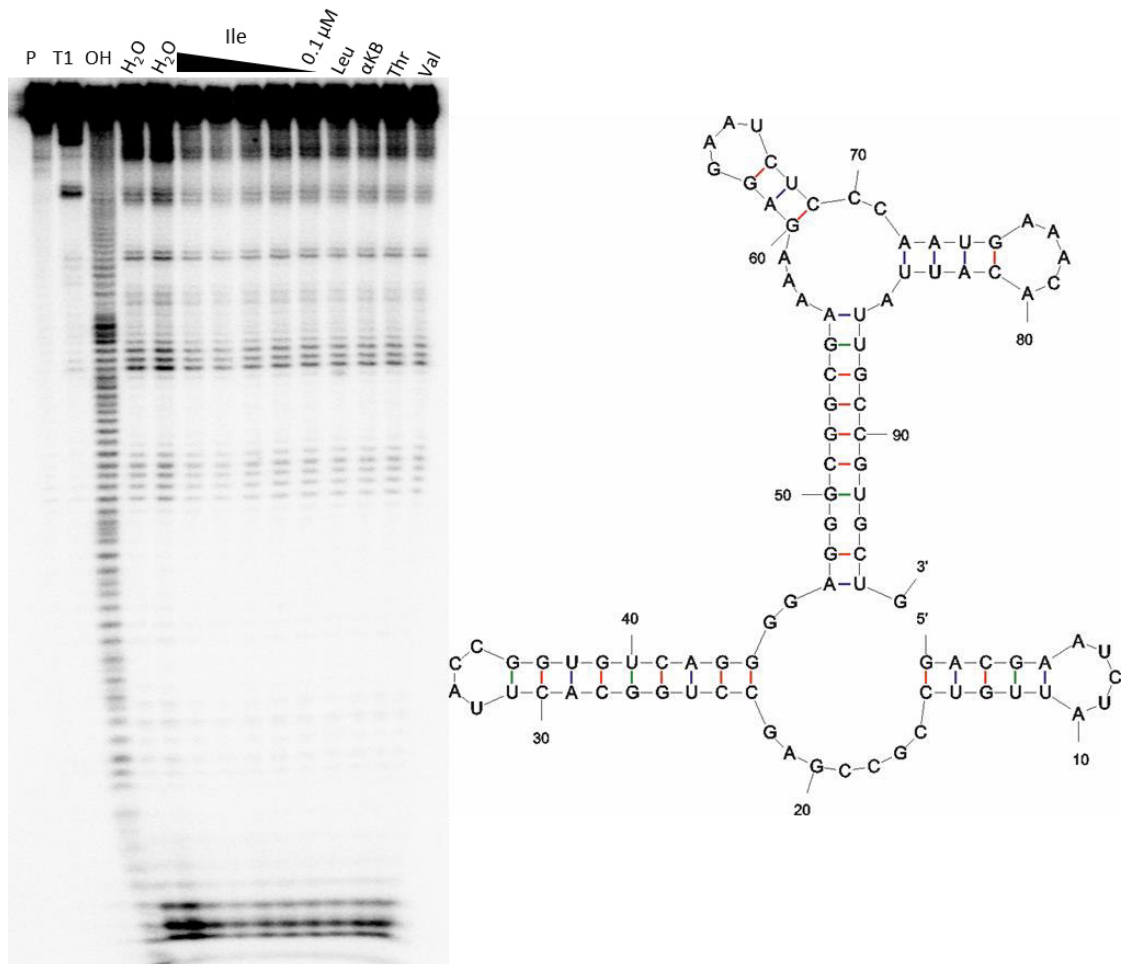
Supplementary Figure 2: ^1H NMR spectra of the products of the CsiD reaction in ammonium acetate buffer. ^1H NMR (400 MHz, D_2O , succinic acid : *L*-2-hydroxyglutarate = 1:1) δ 3.95 (dd, $J = 7.6, 4.1$ Hz, 1H, 6-H), 2.33 (s, 4H, 5-H), 2.19 (td, $J = 9.6, 5.9$ Hz, 1H, 7-H), 2.11 (t, $J = 7.7$ Hz, 1H, 7-H), 1.97 – 1.87 (m, 1H, 8-H), 1.80 – 1.70 (m, 1H, 8-H). Peak at 1.89 ppm arises from acetate. Peaks between 7-9 ppm arise from the enzyme.



Supplementary Figure 3: Coupled transamination and dehydrogenation by PuuE/Sad. Coupled enzyme kinetics of AVA (triangles) by the putrescine transaminase PuuE and the succinate semialdehyde dehydrogenase Sad in comparison to the known substrate GABA (circles). Data is shown as mean of triplicate measurements with error bars representing standard deviations. Experimental procedure was conducted as described in materials and methods.



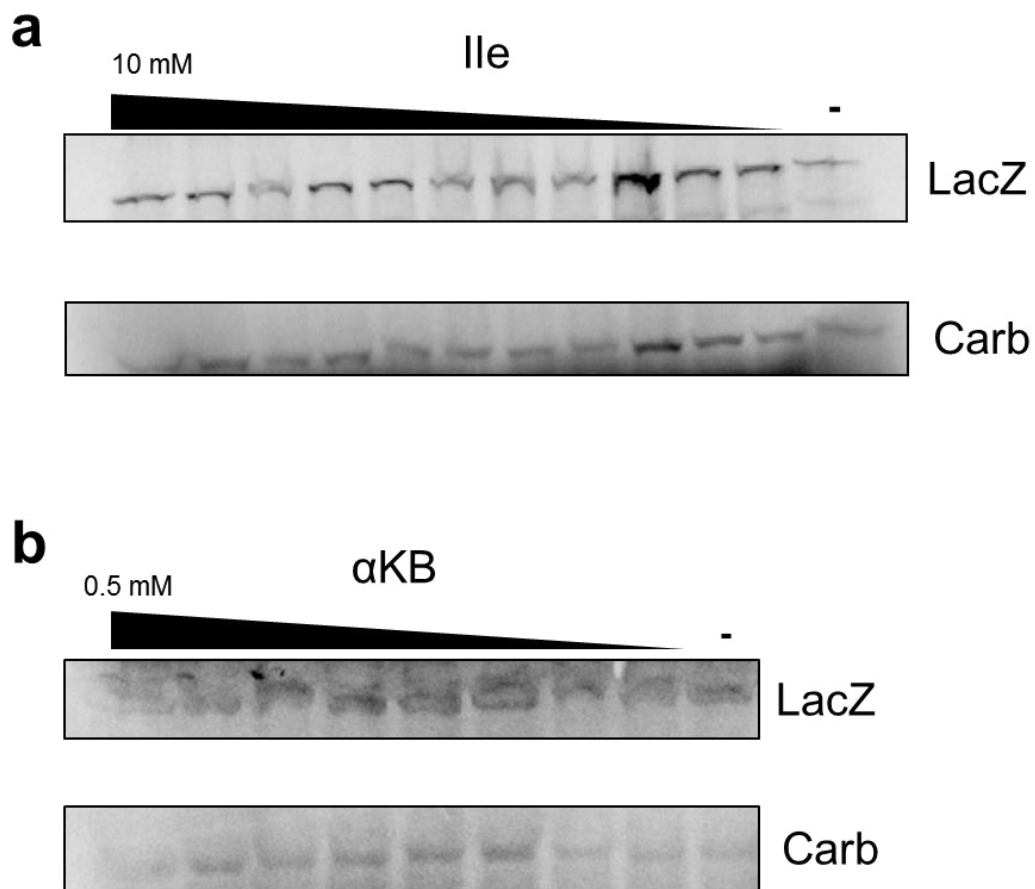
Supplementary Figure 4: Reporter activity of *ilvH*-RNA:lacZ reporter strain in response to differing Ile and Thre concentrations.



Supplementary Figure 5: In line probing of the *ilvH*-motif RNA. **Left panel,** PAGE analysis of in-line probing reactions of a 96-nucleotide construct of *Comamonas testosteroni* KF-1 *ilvH*-motif RNA. Reactions were conducted in the absence (-) or presence of potential ligands (Ile=isoleucine, αKB=α-ketobutyrate, Val=valine, Leu=leucine, Thr=Threonine). P describes the precursor RNA without reaction, T1 stands for the partially digested RNA by RNase T1 and OH stands for a complete digest due to alkaline conditions. Val, Leu, Thr and αKB were inserted in the reactions in a concentration of 100 μM. Isoleucine concentrations tested were 100 μM, 50 μM, 10 μM, 1 μM and 0.1 μM. **Right**

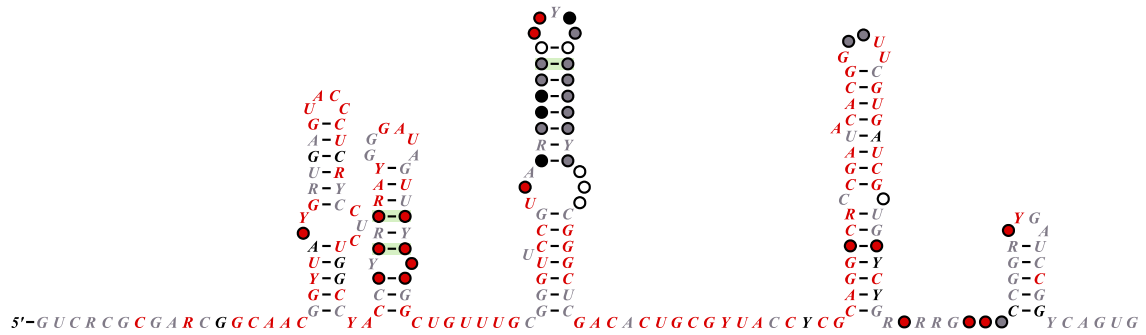
Supplementary Information

panel, Secondary structure of *Comamonas testosteroni* KF-1 96-nucleotide *ilvH*-motif RNA predicted by free energy calculations performed by the mfold algorithm.

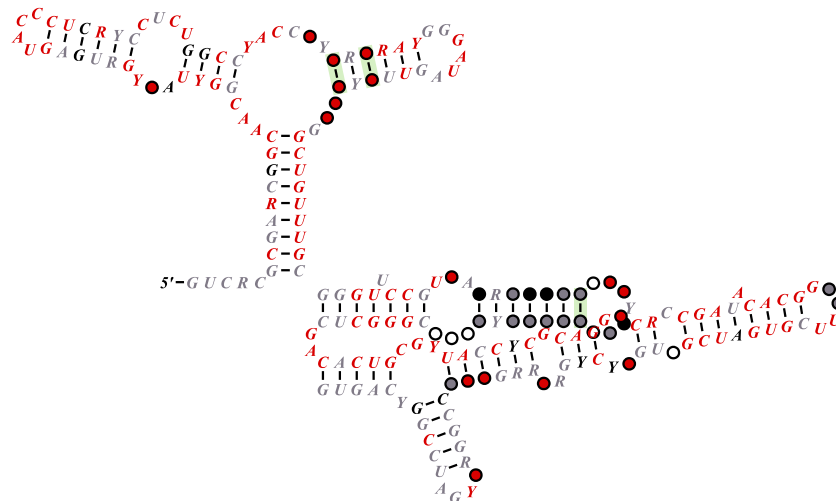


Supplementary Figure 6: Coupled Transcription-Translation assay for an *ilvH*-motif:*lacZ* construct. Levels of Met-labeled LacZ can be observed in the upper bands whereas the Carb-resistance protein (expressed from the pQE vector) serves as internal control band **a**, Titration of Ile up to 10 mM. **b**, Titration of α KB up to 0.5 mM.

ilvB-OMG-RNA motif



Alternative fold:



Supplementary Figure 7: Consensus sequence with secondary structure and genomic neighbourhood of the *ilvB*-OMG RNA motif. Consensus sequence and structure image derived from the rfam (ref) database including 34 sequences.

Supplementary Tables

Supplementary Table 1: Ions used for identification and quantification of compounds by LC-MS.

Compound	Retention time	Ions	<i>m/z</i>
Glutaric semialdehyde	4.3	[M-H] ⁻	115.0401
Glutaric acid	4.9	[M-H] ⁻	131.0350
Succinic acid	5.4	[M-H] ⁻	117.0193
L-2-Hydroxyglutaric acid	10.0	[M-H] ⁻	147.0299
Piperideine dimer ¹	10.5	[M+H] ⁺	167.1543
5-Aminopentanal ¹	11.6	[M+H] ⁺ ; [M-H ₂ O+H] ⁺	102.0913; 84.0808
α-Ketoglutarate	12.0	[M-H] ⁻	145.0425
5-Aminovalerate	12.5	[M+H] ⁺	118.0862
Glutamic acid	13.7	[M+H] ⁺	148.0604
Piperideine monomer ¹	15.3	[M+H] ⁺ ; [2M+H] ⁺	84.0808; 167.1543
Cadaverine	16.1	[M+H] ⁺	103.1230

High resolution ESI-MS (direct injection) and LC-MS measurements on an Orbitrap MS were used for compound identification. Quantification was performed in SIM mode on a quadrupole LC-MS (see Materials and Methods for details).

¹: Signal intensities of ions at *m/z* 84.1 (RT=11.6 and 15.3 min), 102.1 (RT=11.6 min) and 167.2 (RT=10.5 min) corresponding to piperideine monomer, aminopentanal and piperideine dimer, respectively, were summed up for 5-aminopentanal quantification

Supplementary Table 2: Labelling pattern and total quantification of metabolites measured via HPLC/MS of *E. coli* WT, Δ *csiD* and Δ *gabT*.

Compound		Concentration ³ μM	Intermediate distribution (%) for labeling atoms incorporated (mass shift [M+X] listed)									
			X=0	X=1	X=2	X=3	X=4	X=5	X=6	X=7	X=8	
Cadaverine	<i>WT</i>	1829±56	2.0±0.6	-	-	-	-	-	-	4.8±0.1	93.2±0.5	-
	Δ <i>csiD</i>	1233±54	1.9±0.4	-	-	-	-	-	-	4.7±0.1	93.4±0.4	-
	Δ <i>gabT</i>	768±53	4.9±0.4	-	-	-	-	-	-	4.8±0.2	90.3±0.4	-
Piperidine / Aminopentanal ¹	<i>WT</i>	753±112	-	-	-	-	-	-	4.9±0.6	95.1±0.6	-	-
	Δ <i>csiD</i>	708±90	-	-	-	-	-	-	5.2±0.9	94.8±0.9	-	-
	Δ <i>gabT</i>	535±76	-	-	-	-	-	-	4.6±0.7	95.4±0.7	-	-
Aminovalerate ²	<i>WT</i>	n.d.	n.d.	-	-	-	-	-	-	-	-	-
	Δ <i>csiD</i>	n.d.	n.d.	-	-	-	-	-	-	-	-	-
	Δ <i>gabT</i>	637±68	n.d.	-	-	-	-	-	4.8±0.4	95.2±0.4	-	-
Glutaric acid	<i>WT</i>	180±16	9.1±1.5	-	-	-	-	4.6±0.3	86.3±1.4	-	-	-
	Δ <i>csiD</i>	4994±372	2.7±0.2	-	-	-	-	4.7±0.1	92.5±0.2	-	-	-
	Δ <i>gabT</i>	357±36	9.5±1.5	-	-	-	-	4.6±0.3	85.9±1.6	-	-	-
Succinic acid	<i>WT</i>	598±101	95.4±1.5	-	3.4±0.9	-	-	1.8±0.5	-	-	-	-
	Δ <i>csiD</i>	498±31	100	-	-	-	-	-	-	-	-	-
	Δ <i>gabT</i>	782±48	93.6±1.0	-	3.8±0.6	-	-	2.7±0.6	-	-	-	-

(-): not detectable. Grey squares marked the mass shifts expected due to the described pathway. Mass shifts of [M+X-1] found at an abundance of ~ 5% can be attributed to isotopic impurity of labeled lysine used (composed of 95% [M+8] and 5% [M+7]), in accordance with the manufacturer's reported purity of the isotope-labelled lysine. Succinic acid semialdehyde was not detectable.

¹Due to the spontaneous conversion of 5-Aminopentanal to Piperidine, monomer and dimer under cell cultivation⁵, sample preparation, and measurement conditions, we used summarized concentrations of these compounds.

²Aminovalerate (RT=12.33 min) cannot be completely separated from valine (RT=12.52 min). Due to the same sum formula and its low abundance the concentration, the amount of non-labeled aminovalerate and hence the total amount of aminovalerate could not be determined (n.d.). ³ Mean ± s.d.; n ≥ 3

List of Abbreviations

(p)ppGpp	guanosine (penta)tetraphosphate
α KG	α -ketoglutarate
APA	5-aminopentanal
ATP/ADP/AMP	adenosine tri/di/monophosphate
AVA	5-aminovalerate
BCAA	branched chain amino acid
Cad	cadaverine
cAMP	cyclic AMP
Crp	cAMP receptor protein
C-source	carbon source
DCPIP	2,6-chloroindophenol
ETC	electron transport chain
GA	glutarate
GABA	γ -aminobutyrate
GDH	glutamate dehydrogenase
Gln	Glutamine
Glu	glutamate
GOGAT	glutamate synthase
GS	glutamine synthase
GSA	glutarate semialdehyde
GTP/GDP	guanosine tri/diphosphate
Hfq	phosphohistidine carrier protein
HQNO	2-heptyl-4-hydroxyquinoline n-oxide
L2HG	L-2-hydroxylglutarate
lysine	L-lysine
mRNA	messenger RNA
ncRNA	non-coding RNA
N-source	nitrogen-source
Ntr	nitrogen-regulated response
ORF	open reading frame

Ppi	pyrophosphate
PTS	phosphotransferase system
RBP	RNA binding protein
RNA	ribonucleic acid
RNAP	RNA polymerase
SA	succinate
SD	Shine-Dalgarno sequence
sRNA	small-RNA
TCA	tricarboxylic acid cycle
TRAP	<i>trp</i> RNA-binding attenuation protein
tRNA	transfer RNA
UTR	untranslated region

Danksagung

Ich möchte mich besonders bei Prof. Dr. Jörg Hartig dafür bedanken, dass er mir die Mitarbeit in seiner Forschungsgruppe ermöglichte. Die Vielseitigkeit der von ihm offerierten Themen sowie die Freiheiten, die er mir bei der Bearbeitung erlaubte, bereicherten mein wissenschaftliches Verständnis, inspirierten aber auch die eigene Kreativität ungemein. Es war für mich eine einmalige Zeit mit vielen großartigen Herausforderungen, in der er sowohl für wissenschaftliche als auch private Sorgen immer ein offenes Ohr hatte. Dieses gegenseitige Verständnis machte meine Promotionszeit zu einer besonderen Erfahrung.

Außerdem möchte ich noch den Mitgliedern meines Prüfungskomitees, Prof. Dr. Andreas Marx sowie Prof. Dr. David Schleheck, für die Bewertung und zusätzliche Aufarbeitung meiner Arbeit danken. Zudem danke ich Prof. Dr. Bernhard Schink und noch einmal Prof. Dr. David Schleheck für die Betreuung meiner Doktorarbeit als Mitglieder des Graduiertenschulen-Komitees der KoRS-CB. Ebenso danke ich der KoRS-CB selbst für die Unterstützungen während meiner Promotion sowie zahlreichen Möglichkeiten für Zusatzqualifikationen.

Ein herzlicher Dank gilt ebenso der gesamten AG-Hartig. Die gemeinsame Zeit mit allen aus der Gruppe wird für mich unvergesslich bleiben. Es war immer möglich mit euch fachlich bereichernde Diskussionen zu führen. Aber natürlich auch die Grill- und Freizeitabende sind Teil unserer gemeinsamen Geschichte. Dabei denke ich vor allem an die Bewohner von L949, Monika Sack, Julia Stifel, Michele Felletti und Dennis Kläge. Obwohl meist jeder von uns an verschiedenen Themenschwerpunkten arbeitete, hattet ihr immer einen sinnvollen fachlichen Rat zur Stelle. Zu jeder Zeit hatten wir eine Laborgemeinschaft, die sich nie zu ernst nahm und in erster Linie sehr viel Spaß verbreitete. So haben sich für mich im Laufe der Promotion einmalige Freundschaften gebildet, die hoffentlich noch lange fortbestehen werden.

Ein besonderer Dank gilt außerdem Malte Sinn. Bei unseren gemeinsamen Projekten trotzten wir allen Schwierigkeiten, hatten dabei aber immer jede Menge Spaß sowohl

inner- als auch außerhalb des Labors und hatten am Ende auch wegen dieser Art der Zusammenarbeit Erfolg.

Von Herzen danken möchte ich außerdem Astrid Joachimi. Dafür, dass sie immer den Überblick behielt, alles zu organisieren wusste und einen selbst manchmal auf Linie brachte, wenn etwas unstrukturiert ablief.

Zu guter Letzt gilt mein Dank meiner Familie. Meinen Eltern danke ich dafür, dass sie mich auf dem langen Ausbildungsweg, der mit einem naturwissenschaftlichen Studium mit anschließender Promotion einhergeht, immer tatkräftig unterstützt haben, Meiner Frau, Jessica, danke ich besonders, da sie mich immer wieder im privaten Leben auffing und mir den Fokus für das gab was am wichtigsten im Leben ist: Liebe, Glück und eine gemeinsame Familie.