Philipps-Universität Marburg
Fachbereich Biologie

Understanding and Engineering Metabolic Feedback Regulation of Amino Acid Metabolism in *Escherichia coli*

Dissertation zur Erlangung
des Doktorgrades der Naturwissenschaften
(Dr. rer. nat)

Dem Fachbereich Biologie
der Philipps-Universität Marburg vorgelegt von

Timur Sander
Master of Science, Philipps-Universität Marburg
Geboren in Volkmarsen 26. Juni 1991

Marburg, Oktober 2019

Die vorliegende Arbeit wurde in der Zeit von Oktober 2016 bis Oktober 2019 unter der Betreuung von Dr. Hannes Link am Max-Planck Institut für terrestrische Mikrobiologie in Marburg angefertigt.

Erstgutachter: Dr. Hannes Link
Zweitgutachter: Prof. Dr. Erhard Bremer

Weitere Mitglieder der Prüfungskommission: Prof. Dr. Knut Drescher
                                               Prof. Dr. Michael Bölker

Vom Fachbereich für Biologie der Philipps-Universität Marburg als Dissertation angenommen am

_____

Mündliche Prüfung am

_____

Teile dieser Arbeit sind veröffentlicht in:

Sander, T.; Farke, N.; Diehl, C.; Kuntz, M.; Glatter, T.; Link, H. Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in *E. coli* by Enforcing Enzyme Overabundance. **Cell Systems** 2019, *8* (1), 66–75.e8, doi: 10.1016/j.cels.2018.12.005

Sander, T.; Wang, CH.; Glatter, T.; Link, H. CRISPRi-Based Downregulation of Transcriptional Feedback Improves Growth and Metabolism of Arginine Overproducing *E. coli*. **ACS Synthetic Biology** 2019, *8* (9), 1983-1990, doi: 10.1021/acssynbio.9b00183

Content

Chapter 3

Transcriptional Compensation Counteracts Genetic Perturbations and Shows the Benefit of Enzyme Overabundance .... **- 91 -**

Chapter 4

Amino-Acid-Overflow is an Additional Regulatory Mechanism for the Maintenance of End-Product Homeostasis .... **- 108 -**

Closing Remarks.. **- 121 -**

Resource Table .... **- 126 -**

Acknowledgements... **- 130 -**

Curriculum Vitae .... **- 131 -**

Publications... **- 132 -**

Summary of the thesis

Metabolism is the core of what we consider to be a living cell. It covers all chemical reactions that are necessary to break down nutrients and convert them into energy and cellular building blocks for growth. These chemical reactions comprise a large metabolic network that is subject to tight feedback-regulation of enzyme activities or abundances. However, even in intensively studied model organisms like *Escherichia coli*, the knowledge about the function of feedback-regulatory mechanisms and how they interact to control metabolism is still sparse. Therefore, the first goal of this study was to understand the function and relevance of metabolic feedback regulation using amino acid metabolism in *E. coli* as a case study. The second goal was to use the knowledge about metabolic feedback regulation to engineer microbial cell factories for the production of amino acids like L-arginine.

In **Chapter 1** we constructed a panel of 7 mutants with allosterically dysregulated amino acid pathways to uncover the relevance and function of allosteric feedback inhibition *in vivo*, which was so far only demonstrated by theoretical studies. By combining metabolomics, proteomics and flux profiling we could show that allosteric feedback inhibition is crucial to adjust a reserve of biosynthetic enzymes. Such enzyme overabundance originates from a sensitive interaction between control of enzyme activity (allosteric feedback inhibition) and enzyme abundance (transcriptional regulation). Furthermore, we used a metabolic model and CRISPR interference experiments to show that enzyme overabundance renders cells more robust against genetic perturbations.

In **Chapter 2** we increased fitness of a rationally engineered arginine overproduction strain by leaving a certain level of transcriptional regulation. Therefore, we titrated the transcription factor ArgR by CRISPRi and compared this different level of transcriptional regulation with an ArgR knockout strain. Using the CRISPRi approach we elevated the growth rates of an overproduction strain by two-fold compared to the knockout strain, without impairing arginine production rates and titer. Metabolomics and proteomics experiments revealed that slow growth of the knockout strain derives from limitations in pyrimidine nucleotide metabolism and that these limitations are caused by imbalances of enzyme level at critical branching points. Thus, we demonstrated the importance of balancing enzymes in an overproduction pathway and that CRISPRi is a suitable tool for this purpose

In **Chapter 3** we show how cells respond to genetic perturbation on the molecular scale. Therefore, we perturbed amino acid biosynthesis genes with CRISPRi and analyzed the transcriptional response with GFP-reporter plasmids and proteomics. These experiments revealed that cells elevate the expression of genes in a perturbed pathway to counteract a genetic perturbation (We will refer to this mechanism as transcriptional compensation). Metabolomics and flow cytometry data of the wild-type

and the allosteric mutant demonstrated the benefit of enzyme overabundance in response to genetic perturbations: Cells without overabundance showed a heterogenic transcriptional compensation even to mild perturbations, whereas in wild-type cells such mild perturbations were buffered by enzyme overabundance.

In **Chapter 4** we consider amino acid degradation pathways as an additional regulatory mechanism for the maintenance of end-product homeostasis Nutritional downshift experiments revealed increased robustness of allosteric mutants in which the respective degradation pathway was up-regulated. By dynamic metabolite measurements we showed that *E. coli* channels an excess of arginine into the degradation pathway. This overflow mechanism might be the reason for the robustness of allosteric mutants under dynamic conditions.

Zusammenfassung der Arbeit

Stoffwechsel spielt eine zentrale Rolle in allen lebenden Organismen. Die Gesamtheit aller biochemischen Reaktionen gewährleistet den Abbau und Umwandlung von Substraten in Energie und Bausteine für zelluläres Wachstum. Solch biochemische Reaktionen sind in netzwerkartigen Strukturen angeordnet und unterliegen regulatorischen Mechanismen wie der Kontrolle von Enzym Aktivitäten (Allosterische Regulation) und -Abundanzen (Transkriptionelle Regulation). Dennoch ist selbst in Modellorganismen wie *Escherichia coli* wenig über die globalen Wirkweisen und Interaktionen dieser Regulationsmechanismen bekannt. Das Hauptziel dieser Thesis bestand darin, am Beispiel des Aminosäure Stoffwechsels von *E. coli*, die Funktion und Relevanz von Regulationsmechanismen für die lebende Zelle zu verstehen. Darüber hinaus sollte das gewonnene Wissen darauf verwendet werden, bakterielle Zellen für die biotechnologische Produktion von wertvollen Aminosäuren wie L-Arginin zu modifizieren.

In **Kapitel 1** wurde eine Zusammenstellung von sieben *E. coli* Punktmutanten generiert, in welchen die allosterische Inhibierung von jeweils einem Aminosäure Biosyntheseweg entfernt wurde. Damit sollte die bisher unklare Funktion von allosterischer End-Produkt Inhibierung *in vivo* demonstriert werden. Mit Hilfe von globalen Metabolit- und Proteomdaten, sowie Messungen des biosynthetischen Flusses konnte gezeigt werden, dass Zellen durch allosterische Inhibierung Enzymreserven generieren. Solche Enzymreserven werden durch eine sensitive Interaktion von allosterischer und transkriptioneller Regulation eingestellt. Weiterhin konnte durch eine mathematische Modellierung in Kombination mit CRISPRi-Experimenten nachgewiesen werden, dass Enzymreserven Zellen robuster gegen genetische Störungen machen.

In **Kapitel 2** wurde untersucht, ob das Wachstum eines biotechnologischen Arginin Produzenten durch verschiedene Level transkriptioneller Regulation verbessert werden kann. Dafür wurde der Transkriptionsfaktor ArgR mittels CRISPRi auf verschiedene Level titriert und mit einer ArgR Deletion verglichen. Durch die Titration von ArgR konnte die Wachstumsrate eines Überproduktionsstammes verdoppelt werden ohne dabei die Arginin Produktionsrate und den Titer negativ zu beeinflussen. Metabolit und Protein Messungen ergaben, dass der Wachstumsdefizit der Deletions-Mutante durch Limitierungen in der Pyrimidin Biosynthese entsteht. Diese Limitierungen entstehen wiederum durch Ungleichgewichte von Enzymen an metabolischen Schnittpunkten zwischen Arginin und Pyrimidin Biosyntheseweg.

In **Kapitel 3** wurden die molekularen Mechanismen in Reaktion auf genetische Perturbationen charakterisiert. Dafür wurden die finalen Experimente aus Kapitel 1 aufgegriffen und drei

verschiedene Aminosäure Produktionswege durch CRISPRi perturbiert und die transkriptionelle Reaktion mittels Fluoreszenz-Reporterplasmiden und Protein Messungen untersucht. Die Experimente ergaben, dass Zellen auf Störungen der Gen-Expression reagieren, indem sie die gesamte Expression des jeweiligen Biosyntheseweges anheben um einer Enzymlimitierung entgegenzuwirken. Anhand von Durchflusszytometry und dynamischen Metabolit Messungen konnte beobachtet werden, dass allosterische Mutanten ohne Enzymreserven eine heterogene und sensitivere transkriptionelle Reaktion auf genetische Störungen zeigen.

In **Kapitel 4** sollte untersucht werden, ob Zellen Aminosäure Abbauwege als Überlaufventil nutzen um End-produkt Homöostase zu gewährleisten. In dynamischen Experimenten mit Substratwechseln von Glukose zu Galaktose, zeigten Mutanten mit hochregulierten Abbauwegen einen Wachstumsvorteil. Mit Hilfe von dynamischen Metabolit Messungen konnte gezeigt werden, dass ein Überfluss an Arginin in den jeweiligen Abbauweg umgeleitet wird. Ein solcher Mechanismus könnte die Robustheit der allosterischen Mutanten in dynamischen Experimenten erklären.

General Introduction

Regulation of bacterial metabolism

Bacterial metabolism is a large biochemical network with thousands of different metabolites and enzymes (Guo *et al.*, 2013). This metabolic network converts nutrients into energy and building blocks for cellular growth. Metabolism is organized into so-called metabolic pathways that enable step-wise synthesis or degradation of metabolites, by either spontaneous or enzyme-driven biochemical reactions. Such pathways can be distinguished into catabolic and anabolic reactions (Chubukov *et al.*, 2014). Catabolic pathways degrade nutrients to produce energy and essential pre-cursors. A typical catabolic pathway is the glycolysis which converts glucose to pyruvate under production of the high energy products ATP and NADH. Anabolic pathways instead are used to form essential cellular building blocks such as amino acids or nucleotides. Since environmental conditions are constantly changing, bacteria evolved regulatory mechanism to ensure a need-based and efficient usage of the metabolic capacities. Such regulatory mechanism can take place on different layers, such as control of gene expression (i.e. by transcription factors or attenuation), regulation on the translational (i.e. by small RNA's or mRNA stability) and post-translational level (i.e. by adenylation or phosphorylation) or by direct enzyme-metabolite interactions that control enzymatic activity (i.e. allosteric regulation) (**Figure 1**). With these regulatory mechanisms, cells are able to maintain an equilibrium of intracellular conditions, the so-called homeostasis.

The idea of homeostasis was already postulated in 1929 by Walter B. Cannon, who described the maintenance of acceptable ranges of physiological parameters such as blood glucose and core temperature (Cannon, 1929). He realized that a homeostatic system requires regulatory mechanisms that sense perturbations and force the system back to steady state conditions. Thereby Walter Cannon described the mechanism of feedback regulation for the first time in a biological context and postulated it as a requirement for physiological homeostasis. The next milestone was set in 1956 by Edwin Umbarger, who first observed the mechanism of feedback inhibition of enzymes by pathway end-products. He could show that the first enzyme of isoleucine biosynthesis of *E. coli* is inhibited by isoleucine, which was the first example of allosteric feedback inhibition. Only five years later Jaques Monod discovered the principles of transcriptional regulation in the context of lactose up-take and hydrolysis (Jacob and Monod, 1961). In the absence of lactose, the *lacI* repressor blocks transcription of lactose degradation enzymes by binding to the promotor region of the lac-operon. If lactose is available, a degradation product (allolactose) allosterically binds to the lacI-repressor and disrupts

binding to the lac promotor, which causes expression of lactose degradation enzymes. Up to this day, the lac operon is one of the best studied examples for transcriptional feedback regulation of gene expression.
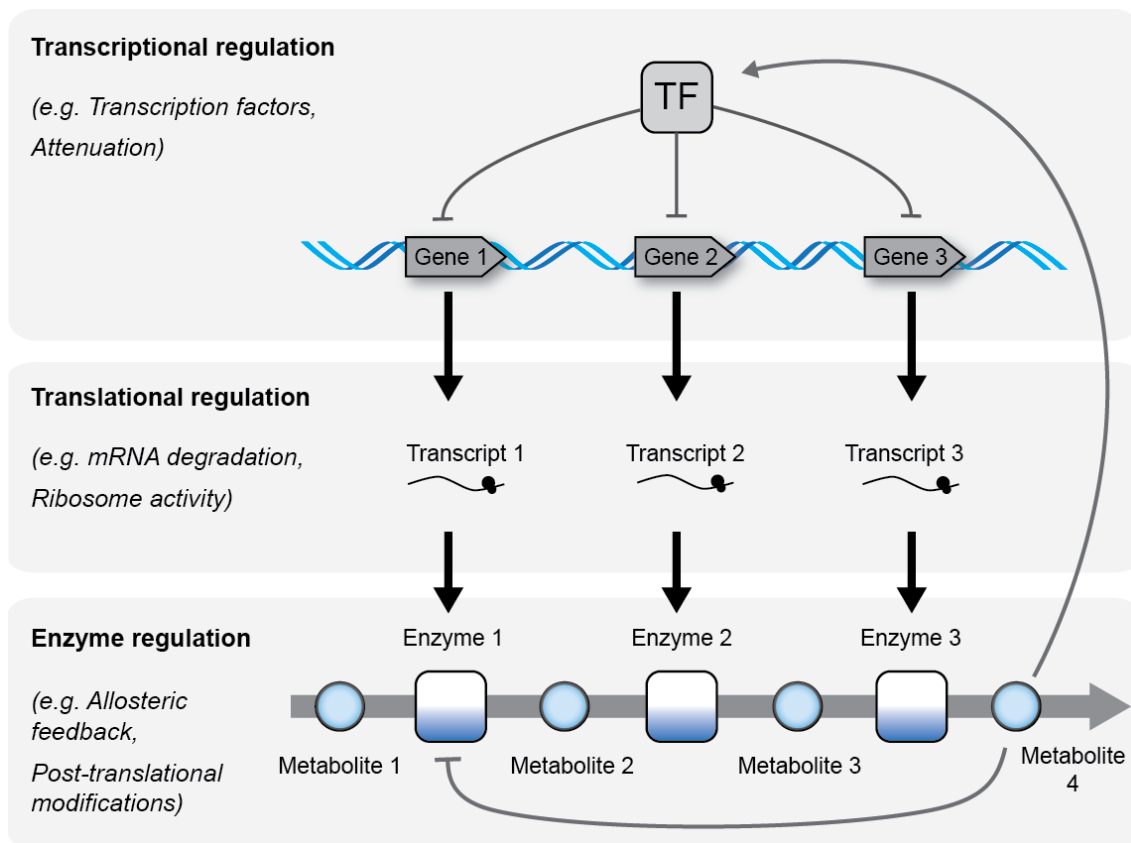

Figure 1. Metabolic pathways are regulated on different layers. The scheme shows a typical pathway regulation from amino acid biosynthesis including negative allosteric and transcriptional feedback regulation. The partial fill of the enzymes indicates that they are not working at full capacity (overabundance). TF indicates a transcription factor. Grey arrow indicates activation, grey bar arrow indicates inhibition.

The both described mechanisms represent the first examples of two of the most common regulatory mechanism controlling metabolism: Allosteric regulation of enzyme activity and transcriptional control of enzyme abundance. These feedback mechanisms are thought to be crucial in providing metabolic robustness, the ability to maintain performance in the face of perturbations and uncertainty (Stelling *et al.*, 2004). However, a direct experimental link between metabolic robustness and feedback mechanisms was only shown by theoretical analyses (Kitano, 2007). A kinetic modelling approach for instance showed that allosteric enzyme regulation significantly increases the stability of

a network (Grimbs *et al.*, 2007). In the scope of metabolic robustness, recent theoretic studies observed that enzymes reserves might be key in the ability to adapt to changing environments (O'Brien, Utrilla and Palsson, 2016; Davidi and Milo, 2017). Such enzyme overabundance was then analyzed in more specific studies, that showed a benefit in quick activation of the pentose phosphate pathway upon oxidative stress (Christodoulou *et al.*, 2018), and similar observations were made for overabundant ribosomes (Mori *et al.*, 2017) and coenzymes (Hartl *et al.*, 2017). However, beside this few studies, less is known about the function and the regulatory mechanisms that adjust enzyme overabundance. An open question is, to which extent do cells express enzymes in excess, in other words: How much enzyme is enough for a cell? This question indicates a critical trade-off between cost efficient (low enzyme level) and robust (high enzyme level) metabolic operations, which was also discussed by others (Kitano, 2007; Donati, Sander and Link, 2018). It needs to be clarified, which regulatory mechanisms allow to solve the efficiency-robustness trade-off.

Metabolite level contain information for the cell

The metabolome is a signaling platform that carries information about the current status of the cell in order to fine-tune anabolic and catabolic reactions as well as gene expression. A direct evidence that metabolites carry information was shown by Zamboni and colleagues, that measured >7000 intracellular metabolite ions in 3800 *E. coli* single gene deletion strains (Fuhrer *et al.*, 2017). Each of the single gene deletion caused a specific metabolic fingerprint that allowed tracking back the function of the deleted gene. With this approach they predicted metabolism-related function of 72 so far not annotated genes. Beside the impact of metabolites on transcription which will be reviewed later in more detail, information from metabolites can be processed on different layers. Many of these regulatory layers are based on metabolite-protein interaction, which involves binding of metabolites to the active site of enzymes as substrates or co-factors, binding to the allosteric site of enzymes or transcription factors. Metabolites were also shown to be crucial in the assembly and function of many protein complexes (Milroy *et al.*, 2014). However, the global impact of metabolite-protein interaction is still ambiguous, due to the lack of systematic approaches for detection of these interactions, such as presented for protein-DNA (Mirzaei *et al.*, 2013) or protein-protein interactions (Sokolina *et al.*, 2017). Picotti and colleagues recently presented an approach to systematically identify metabolite-protein interactions, which is based on proteolysis and mass spectrometry (Piazza *et al.*, 2018). The researchers treated a cell lysate with the metabolite of interest followed by digestion trough a broad-specificity protease. Binding of the metabolite to a protein can block cleavage by the protease which then might result in different peptide fragments. Changes in the peptide patterns due to metabolite

treatment were then detected by label-free quantitative mass-spectrometry. This approach identified 1678 metabolite-protein interactions of which more than 80% were not known. Taking into account that this study was performed with the best studied model-organism *E. coli*, suggests how less is known about the regulatory role of the metabolome.

Regulatory Metabolite-Enzyme interactions

Different multi-omics studies in different organisms showed that metabolic fluxes do not correlate with transcript level (Chubukov *et al.*, 2013; Kerkhoven *et al.*, 2016; Monk *et al.*, 2016). A conclusion of this observation is that metabolic fluxes are mainly regulated by mechanisms on the post-translational level, which highlights the role of direct enzyme-metabolite interactions such as allosteric feedback regulation. Allosteric regulation is classically defined by three characteristics: (i) the allosteric effector chemically differs from the substrate, (ii) binding of the effector causes a functional change of the protein and (iii) the effector binds at a site that topographically differs from the active site (Monod, Changeux and Jacob, 1963; Fenton, 2008).

Due to the direct binding of an effector metabolite to an enzyme and the immediate effect on enzymatic activity, allosteric regulation is known to allow fast adaption to environmental changes. Thereby cells ensure that pathway flux is not higher than actually needed. In many cases the enzyme catalyzing the committed step of a pathway is feedback regulated by the end-product. For example, the first step of anabolic pathways branching from glycolysis or TCA, such as amino acid or nucleotide biosynthesis is regulated by an allosteric feedback from a downstream product of the pathway (**Figure 2**). Thereby pathway influx for *de novo* synthesis is limited if not needed. In case of pyrimidine biosynthesis of *E. coli* it was shown, that disrupting allosteric feedback regulation leads to excretion of a pathway intermediate in order to maintain end-product homeostasis (Reaves *et al.*, 2013). This study emphasized the role of allosteric regulation as a metabolic flux control, since such overflow mechanisms are generally an indicator for flux imbalances (Paczia *et al.*, 2012). A recent study showed that *E. coli* controls flux through the pentose phosphate pathway (PPP) in response to oxidative stress by a negative allosteric feedback loop from NADPH to the G6P dehydrogenase (Christodoulou *et al.*, 2018). Upon oxidative stress, intracellular NADPH level decrease, which in turn causes a de-repression of the G6P dehydrogenase. The de-repression of the G6P dehydrogenase rapidly increases PPP-flux to replenish the NADPH pool. This study shows that allosteric feedback regulation can function like a valve that controls the influx into a metabolic pathway. Beside this flux control mechanisms of committed steps of metabolic pathways, it was shown that *E. coli* can also actively sense metabolic fluxes by allosteric enzyme-metabolite interactions. Heinemann and colleagues showed that the metabolite fructose-1,6-bisphosphate (FBP) functions as a glycolytic flux sensor by allosterically

activating the downstream enzymes pyruvate kinase and PEP carboxylase (Kochanowski *et al.*, 2013). FBP accumulates in response to increasing glycolytic flux and allosterically elevates activity of both downstream enzymes until a threshold, at which lower glycolytic flux matches upper glycolytic flux. Furthermore, this system was also shown to act on the gene regulatory layer, by allosterically inhibiting activity of the transcription factor Cra, which represses glycolytic enzymes and activate gluconeogenesis enzymes. The allosteric inhibition of Cra is mediated by binding to fructose-1-phosphate (F1P), which is a derivate of FBP (Bley Folly *et al.*, 2018). For more than 25 years it was thought that FBP is the allosteric inhibitor of Cra (Ramseier *et al.*, 1993), which was corrected only recently (Bley Folly *et al.*, 2018) and shows the complexity of detecting and validating allosteric interactions.

Although recent studies achieved fundamental progress in detecting metabolite-protein interactions (Piazza *et al.*, 2018), it is still challenging to identify allosteric interactions that are functionally relevant under particular conditions (i.e. different conditional shifts) *in vivo*. A study addressing this problem switched *E. coli* between growth on pyruvate and 13C labeled glucose or fructose and dynamically measured fluxes and metabolite level. Combining these data with a kinetic model allowed systematic identification of allosteric interactions that govern the switch between gluconeogenesis and glycolysis (Link, Kochanowski and Sauer, 2013). Detecting and understanding metabolite-enzyme interactions is a requirement for the construction of biotechnological relevant cell factories and will thereby get more and more important in a variety of different organisms.

Metabolites as signals for transcription

Unnecessary protein production is a burden for the cell, which can have drastic influence on physiological parameters like division time or cell size (Kafri *et al.*, 2016). Therefore, protein synthesis is subject to different regulatory mechanisms. Many of these mechanisms use metabolites as signals to ensure that proteins are only expressed when needed. Prominent mechanisms that use metabolite level to adjust gene expression to changing conditions are transcription factors, transcriptional attenuation, sigma factors and nucleoid proteins. Transcription factors for instance can have signal-sensing domains that allosterically bind to metabolites, which causes an activity change, which either activates or represses the expression of a gene. Activation or repression of a gene is initiated by binding to the DNA and either promoting or blocking RNA-polymerase recruitment (Latchman, 1997).

Transcriptional regulatory mechanisms are organized in so-called transcriptional regulatory networks (TRNs), which describe the relationship of thousands of genes and their transcriptional regulators. The TNR of *E. coli* is probably the most complete one and covers 210 out of the ~300 predicted

transcription factors and 4451 interactions between transcription factors and genes (Gama-Castro *et al.*, 2016). Palsson and collegues showed that they can use TNRs to predict expression for 86 % of 1364 transcriptional units of *E. coli* (Fang *et al.*, 2017). TRN have hierarchical structures with different layers of regulation in which global transcriptional regulators can control the expression of other transcriptional regulators. A recent study explaining this hierarchical structure measured metabolite concentrations and transcription rates during growth on 26 different environmental conditions in *E. coli* (Christodoulou *et al.*, 2017). The data revealed that 30% of the transcriptional changes of genes in central carbon metabolism across the different environments were caused by the two transcription factors Crp and Cra. Correlating the expression rates with the metabolome identified the three effector metabolites cyclic adenosine monophosphate (cAMP) and fructose-1-phosphate/fructose-1,6-bisphosphate, respectively, that explained the activity of Crp and Cra. Another recent study used multiomics data sets and bioinformatics to analyze *E. coli* knockout (KO) strains during an adaptive laboratory evolution (ALE) experiment (McCloskey, S. Xu, *et al.*, 2018). The results showed that gene knockouts caused imbalances of metabolite level which caused several transcriptional alterations that were re-balanced by mutations during evolution. These mutations allowed to adapt gene expression and metabolic fluxes in order to prevent the imbalances of metabolites. The authors claim that metabolite-transcription factor interactions have a dominant role in sensing and re-optimizing such perturbations during evolution.

Despite these examples showing the relevance of metabolite-transcription factor interactions, our knowledge is still scarce. Even in the arguably best-studied model microbe *E. coli*, direct interactions with metabolites have only been shown for 47 out of the 210 transcription factors. Although scalable methods for detection of metabolite-protein interactions were previously described, we are still lacking systematic approaches to identify metabolite-transcription factor interactions that are relevant under certain conditions. A promising approach could be to combine computational with experimental analysis as recently shown for cancer cells by Zampieri and colleagues (Ortmayr, Dubuis and Zampieri, 2019). The researchers integrated transcriptomic and proteomic data with metabolic profiles of different cancer cell lines to unravel interactions between transcriptional regulators and metabolism. They discovered a global regulatory signature coordinating glucose and one-carbon metabolism. A similar approach could also be used to get a deeper view into interaction of transcriptional regulation and metabolism in bacterial cells.

Regulation of amino acid metabolism in *E. coli*

Biosynthesis of amino acids is a central cellular process, which provides building blocks for protein synthesis and growth. Therefore, especially fast-growing cells without external supply of amino acids have to ensure sufficient biosynthetic capacity of the essential building blocks. *E. coli* for instance invests more than 15% of the total proteome into amino acid metabolism during growth on minimal medium. Similarly, during growth on methionine limiting media, *E. coli* was shown to devote 8% of the total protein synthesize capacity into the production of a single methionine enzyme (MetE) (Li *et al.*, 2014). According to this large investment of cellular resources, amino acid biosynthesis of *E. coli* is subject to a tight transcriptional feedback network in order to minimize unnecessary protein burden (**Figure 2**). A crucial role was shown for the four transcription factors ArgR (Arginine repressor), Lrp (Leucine-responsive regulatory protein), TrpR (Tryptophan repressor) and TyrR (Tyrosine repressor) which directly or indirectly control biosynthesis of all twenty amino acid biosynthesis pathways (Cho *et al.*, 2012). ArgR and Lrp were shown to act on a global scale by activating or repressing the expression of 140 and 283 target genes. These transcription factors are organized in a hierarchical structure with ArgR also regulating the expression of Lrp and therefore having a total regulon size of 423 target genes. ArgR is activated in response to increased arginine level and represses the expression of all enzymes in arginine and histidine and single enzymes from lysine (*dapE*), glutamate (*gltBD*) and aromatic amino acid biosynthesis (*aroB*, *aroK*).

Beside the regulation of biosynthetic genes, these transcription factors also regulate need-based dependent expression of amino acid import and degradation. In response to high intracellular arginine level, ArgR for instance represses genes of the arginine import system (*artJ*, *artMQIP*) and activates genes involved in arginine degradation (AST-pathway) (Caldara, Charlier and Cunin, 2006). Beside transcription factors, transcriptional regulation in amino acid metabolism can also occur by transcriptional attenuation. Transcriptional attenuation is known for tryptophan, threonine, valine, histidine, leucine and isoleucine biosynthesis. The tryptophan (trp) operon in *E. coli* is a well described example of transcriptional attenuation: Transcription of the trp-operon is regulated by the so-called leader peptide (trpL) which is located up-stream of the genes coding for the trp enzymes and consists of a trp-specific attenuator region. If trp-laden tRNA`s are available in excess, the transcribed trpL mRNA forms a specific hairpin structure, which will cause dissociation of the RNA-polymerase from the DNA and accordingly terminate transcription. The combined mechanisms of transcriptional repression by TrpR and transcriptional termination by attenuation allow regulation of tryptophan biosynthetic enzymes over a range of 500- to 600-fold (Yanofsky, Kelley and Horn, 1984).

Beside the regulation by transcription factors and attenuation, expression of amino acid biosynthesis enzymes is additionally regulated by global metabolic feedback mechanism through cycling AMP and ppGpp signaling (Paul, Berkmen and Gourse, 2005; You *et al.*, 2013). For example, in response to amino acid starvation, the alarmone ppGpp accumulates and triggers a stringent response that blocks activity of rRNA and tRNA promotors. Further ppGpp is an allosteric effector of the transcription factor DksA which binds and re-programs the RNA-polymerase to activate expression amino acid enzymes (Paul, Berkmen and Gourse, 2005). Although all the described mechanism (transcription factors, attenuation, global regulation) allow need-based regulation of enzyme abundance, it is not clear if enzyme levels are tuned for optimal efficiency or if enzyme overabundance (as described for ribosomes or co-enzymes) might also have a relevance in amino acid metabolism.

In addition to the control of enzyme abundance, amino acid metabolism of *E. coli* is also tightly regulated on the level of enzyme activity. 16 out of 20 pathways are regulated by allosteric feedback inhibition, in which the amino acid end-product inhibits activity of the enzyme catalyzing the committed step of the pathway. Several studies showed removal of allosteric feedback inhibition from amino acid pathways by genetic point mutations in the allosteric sites of the enzymes. However, the majority of these studies had a strong focus in engineering strains for increased amino acid production. So far, to our knowledge no study systematically analyzed the consequences of allosteric feedback disruption to show the physiological relevance of allostery in amino acid metabolism.
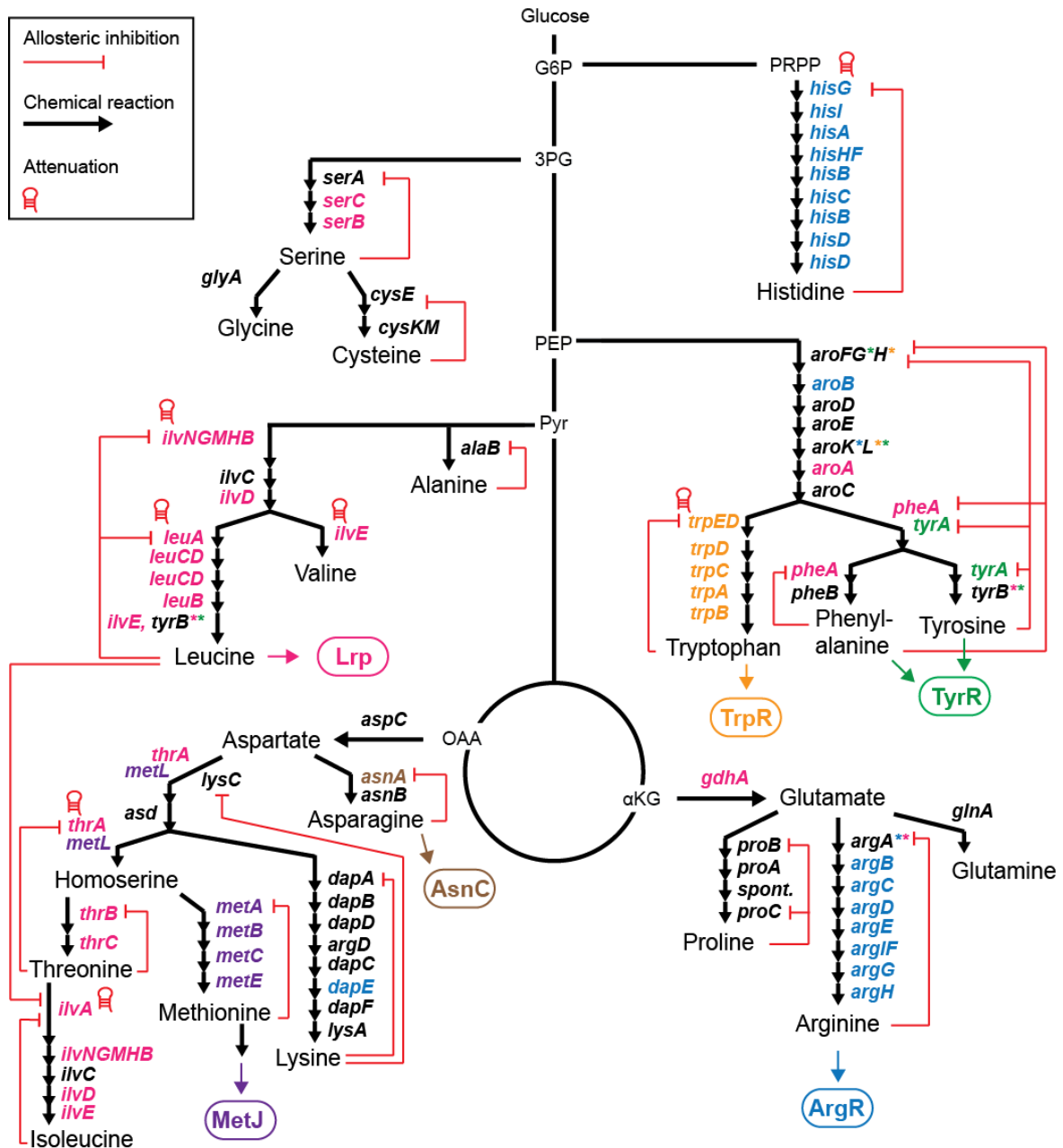
Figure 2. Amino acid metabolism of *E. coli.* Regulatory mechanisms involve allosteric feedback regulation, transcriptional attenuation and transcription factors (indicated with the rounded boxes). Gene targets of the transcription factors are indicated with the corresponding color. Targets that are regulated by multiple transcription factors are indicated with affiliations (*) in the respective colors. Sources were Ecocyc, RegulonDB and *Cho et al. 2012*.

Methods to understand and engineer microbial metabolism

The 'omic' technologies

The most direct way to understand how microbial metabolism functions is to detect and quantify the components that make up a living cell. Those components are for example DNA (genes), transcripts, proteins and metabolites. Methods to globally measure these components, the so-called 'omics' technologies involving genomics, transcriptomics, proteomics or metabolomics were essentially improved during the last decades (Franzosa *et al.*, 2015; Hasin, Seldin and Lusis, 2017). In systems biology these different omics technologies were integrated to get a holistic understanding about how cellular mechanisms act together on a global scale.

Understanding metabolism requires reliable measurement of metabolite levels, such as intermediates and products from glycolysis, TCA cycle, PPP, nucleotide-, amino acid- and co-factor metabolism. Quantification of intra cellular metabolites is challenging due to high turnover rates, chemical instabilities and diversity. Therefore, sampling plays a crucial role for reproducible metabolomics data and relies on fast quenching and avoidance of cellular stress factors (Temperature shifts, mechanical stress). Mass spectrometry was established as a gold standard for rapid identification of a large number of different metabolites. Direct sample injection into a high-resolving mass spectrometer (Flow injection electrospray-time-of-flight mass spectrometry) allows detection of up to 300 metabolites in short measurements of about 1 minute per sample (Fuhrer *et al.*, 2011). Due to the short measurement times, flow injection was shown to be suitable for dynamic metabolome profiling in example during switches between starvation and growth (Link *et al.*, 2015). However, such non-targeted approaches do not allow absolute quantification of metabolites, since each signal is affected by the sample matrix and ionization efficiency. A strategy to address this problem was to add isotopically labeled internal standards to the samples. Fast LC-MS/MS methods enable to quantify more than 200 metabolites in absolute terms by measuring the ratio between their natural 12C and 13C labeled form in a time optimized set up (Guder *et al.*, 2017). By using short chromatographic columns, this method enabled to quantitatively measure primary metabolites in 2 minutes runs. Such approaches allow high-throughput screening of cellular metabolism, which has a high relevance in clinical or biotechnological research. For example, a similar study used optimized chromatographic parameters and internal standards to quantitatively measure metabolism of 7 industrial relevant *E. coli* strains by LC-MS/MS (McCloskey, J. Xu, *et al.*, 2018). Despite short measurement times of 5 minutes, this approach showed quantification of more than 100 metabolites and revealed significant differences in glycolytic, pentose phosphate, amino acid, energy and cofactor metabolism in between the 7 *E. coli* strains.

Beside the detection of metabolites, understanding a biological system also requires detection of gene expression. The development of high-throughput sequencing approaches (Next generation sequencing), enabled global measurement of gene expression profiles by sequencing of the single mRNAs. Current transcriptomics techniques allow detection of 10^9 mRNA sequences which would cover the entire human transcriptome (Mortazavi *et al.*, 2008; Lowe *et al.*, 2017). Quantification of transcript profiles under different conditions or perturbations can give insight into gene regulatory principles of biological system. For example, a comprehensive study in yeast measured gene expression signatures of 1484 knockout mutants to understand the regulatory logics of this organism (Kemmeren *et al.*, 2014). The data were assembled into a genetic perturbation network that revealed a high number of metabolic feedbacks, with many metabolic genes assigned as key nodes in incoherent network motifs. However, transcriptomic analyses are not able to display the complete expression state of a cell, since different post-transcriptional mechanisms can affect protein synthesis (i.e. protein degradation, translation rates or protein export) (Liu, Beyer and Aebersold, 2016). Thereby, only direct measurement of protein level provides reliable information about the cellular expression state.

Just like metabolomics, current proteomics techniques are also based on liquid chromatography coupled to mass spectrometry. In the so-called "Shotgun proteomics" approach, protein samples are first treated with a proteolytic digest and the resulting peptides are then fractionated and analyzed by LC-MS/MS (Wolters, Washburn and Yates, 2001; Zhang *et al.*, 2013). The measured tandem mass spectra of a peptide sample were then compared to theoretical tandem mass spectra obtained from protein data bases. Thereby proteins can be identified due to the specific peptide spectra, which enables global measurement of a cellular proteome. Heinemann and colleagues used a shotgun proteomics approach to quantitatively measure the proteome of *E. coli* across 22 experimental conditions (Schmidt *et al.*, 2016). The researchers determined level of more than 2300 proteins which covers approximately 55% of the predicted *E. coli* genes. The data was then systematically used to unravel growth rate dependent proteome allocations and to successfully identify novel types of post-translational modifications. Moreover, such comprehensive data sets represent a valuable resource for research related to systems biology and other fields. A computational study by Palsson and co-workers for instance, combined these proteome data with a genome scale model of *E. coli* which predicts costs and utilization of the proteome under different conditions (O'Brien, Utrilla and Palsson, 2016). This work showed that almost the half of the total proteome mass of *E. coli* is unused in different environments and that this unused fraction explains >95% of the growth rate changes among the tested conditions. Reduction of this unused fraction increases growth rates of *E. coli*, for example during laboratory evolution experiments. Such protein reserves, which were already addressed

earlier, were suggested to provide robustness and readiness under varying environmental conditions within the scope of this solely computational analysis. Thus, integration of different omics techniques in addition to computational analysis enables a deep insight into biological systems, which has a high relevance in the field of synthetic biology or clinical research. Therefore, beside methods to understand biological systems it also requires methods to efficiently modify biological system which also drastically improved within the last decade.

The era of the scissor: CRISPR/Cas

Targeted modification of DNA is the basis of synthetic biology and enables the construction of organisms with novel properties. Furthermore, characterization of a gene function requires deletion of the respective loci in order to analyze consequences for the organism. Methods to modify DNA in a controlled fashion were already introduced in the 90`s with the discovery of zink finger nucleases (ZFN) (Kim, Cha and Chandrasegaran, 1996) or the use of lambda recombination for gene replacements (Murphy, 1998). These engineering techniques were further improved in the following years for example with the so-called transcription activator-like effector nucleases (TALENs) (Boch *et al.*, 2009) which are similar to ZFNs based on restriction enzyme coupled DNA-recognition proteins. However, such methods are complicated to perform due to engineering of new enzymes or time-consuming replacements of chromosomal markers. In the last decade, discovery and advancement of clustered, regularly interspaced, short palindromic repeats (CRISPR) associated to the endonuclease Cas9 brought gene editing to a new level (Doudna and Charpentier, 2014).

The CRISPR/Cas system originates from an adaptive antiviral defense system of prokaryotes, that uses specific CRIPSR RNAs (crRNAs) to guide Cas proteins to interfere and cleave virus DNA (Barrangou *et al.*, 2007; Karginov and Hannon, 2010). In 2012, Charpentier and colleagues showed that CRISPR/Cas is programmable and suggested the system for the first time as a gene editing tool (Jinek *et al.*, 2012). The researchers used a Cas protein variant from *Streptococcus pyogenes* (Cas9) and programmed the endonuclease with engineered guide RNAs (gRNAs) to introduce specific double-strand breaks in the target DNA. A crucial finding was, that beside the homologous pairing sequence, target recognition of Cas9 requires a protospacer adjacent motif (PAM) which is the short consensus sequence NGG. From this point on, laboratories started to apply CRISPR/Cas9 in many different organisms with different specified protocols, all based on the same strategy (Doudna and Charpentier, 2014; Hsu, Lander and Zhang, 2014). The no-SCAR (Scarless Cas9 Assisted Recombineering) system for instance is a fast and efficient method for genome editing in *E. coli* (Reisch and Prather, 2015). This strategy combines lambda red recombination with CRISPR/Cas9-mediated counter selection. It is based on two plasmids containing all required components: The Cas9 gene, the specific single guide RNA (sgRNA) and three

genes that compose the lambda-red system. The host cells are transformed with homologous oligonucleotides that are integrated into the desired genomic loci by lambda red recombination. These oligonucleotides can be constructed for introducing point mutations (oligonucleotide comprises the desired base-pair exchange), gene deletions (oligonucleotide consists of 3' and 5' flanking regions of the target gene) or gene insertions (oligonucleotide consists of desired gene flanked by 3' and 5' regions of genomic context). After recombination, expression of Cas9 and sgRNA is induced and will cause double-strand breaks at the specific loci where no recombination took place. In mutants where successful recombination occurred, insertion of oligonucleotides caused altering of the PAM motif and thereby prevents binding of the Cas9/sgRNA complex. The Cas9 mediated counter selection is highly efficient and the procedure can be performed within 4 days. Conventional counter selection methods are based on antibiotic resistance cassettes which need to be removed by Flp recombinase leaving "scars" on the DNA. This constitutes a great advantage of the Cas9-mediated counter selection, since such scars can cause chromosomal instability and genomic rearrangements.

Beside genome editing, CRISPR/Cas was also established as an efficient tool for control of gene expression. The CRISPR interference (CRISPRi) system is based on a catalytically dead Cas9 (dCas9) which binds to the DNA and sterically blocks transcription. Just like Cas9, dCas9 can be guided by sgRNA's allowing specific control of target gene expression. A great advantage of this knockdown strategy is, that it allows control of multiple target genes in parallel (Qi *et al.*, 2013). Further, combining CRISPRi with pooled sgRNA libraries enable genome-scale screening of gene functions, which has a high relevance in system biology and clinical research (Gilbert *et al.*, 2014; Wang *et al.*, 2018). Since CRISPRi can be applied in a variety of different organisms including human cells, such library approach can help to identify new targets that are associated to genetic diseases (Gilbert *et al.*, 2014).

Engineering regulation for synthetic biology

Engineering microbes for increased production of chemicals has become a key strategy in different industrial branches. The limited supply of fossil fuels and increasing concerns for sustainability raised the demand for biological production. However, in order to compete with classical chemical production, biotechnology has to face different economic aspects. Critical aspects of engineering microbial cell factories are a high product *titer* (product concentration at the end of a bioprocess), which should be produced with a high *rate* (product formation over time) and with an efficient substrate *yield* (product molecule formed per substrate molecule) (Nielsen and Keasling, 2016). To efficiently produce specific molecules an important first decision is the choice of the host organism, since different organism have different properties that might be beneficial for certain demands.

Relatively well characterized organisms and therefore often used host organisms are for example *E. coli*, *S. cerevisiae*, *Aspergillus niger*, *Bacillus subtilis*, *Corynebacterium glutamicum*, or Chinese hamster ovary (CHO) cells (Nielsen and Keasling, 2016). However, as already mentioned before, our knowledge about regulation of metabolism in these organisms is still scarce, although understanding the metabolic principles of a potential production host is crucial for metabolic engineering approaches. Therefore, systems biology and omics technologies go hand in hand with synthetic biology. Integration of different omics data can help to unravel the regulatory network of an organism for example on the level of allosteric (Link, Kochanowski and Sauer, 2013; Piazza *et al.*, 2018) and transcriptional (Fang *et al.*, 2017) regulation. Such information can then be included into genome scale metabolic models to make *in silico* predictions about which genetic modifications are most likely to improve production. Flux balance analysis (FBA) for instance is based on genome scale models and can help to identify knock-out targets for optimization of metabolic fluxes towards a desired product. Beside gene knock-outs, also self-inhibitory allosteric interactions between metabolites and enzymes might be crucial targets to optimize product flux.

Removing allosteric feedback inhibition and thereby increasing the enzyme activity is a long-known strategy to enhance product formation (Csonka *et al.*, 1988; Rajagopal *et al.*, 1998). Potential targets to produce allosterically insensitive enzymes (allosteric binding site, active site) can be identified with rational protein design and then verified by mutagenesis followed by *in vitro* assays (Chen *et al.*, 2014). Feedback resistance of the host strain is then either introduced by genomic point mutations (Sander *et al.*, 2019) or by expressing the dysregulated enzyme from a plasmid (Ginesy *et al.*, 2015). Beside increasing enzyme activity by removal of feedback inhibition, increasing enzyme abundance in a biosynthetic pathway is a further strategy to enhance product flux. Therefore, expression of specific enzymes can be increased by plasmid-based systems or genomic promotor modifications (Nielsen and Keasling, 2016). Those enzymes can also be expressed heterologously, as shown for the mevalonate pathway from *S. cerevisiae* which was introduced into *E. coli* for the production of isoprenoid precursors (Martin *et al.*, 2003). A large number of biosynthetic pathways are subject to transcriptional feedback control by transcription factors or attenuation. Deletion of a transcriptional repressor causes up-regulation of the corresponding pathway without the need of overexpression-plasmids or promotor modifications. Removal of transcriptional control by attenuation is performed by deletion of the leader peptide as shown for the tryptophan pathway (Tröndle *et al.*, 2018). This study achieved a tryptophan titer of 12.5 g L^{-1} by combining disruption of allosteric feedback inhibition with deletion of the transcriptional repressor TrpR and removal of the attenuator leader peptide trpL (Tröndle *et al.*, 2018). Leaving either allosteric or transcriptional control might be sufficient to control biosynthetic

flux and therefore inhibits overproduction of the desired product as shown for arginine metabolism (Caldara *et al.*, 2008).

The final titer in a production vessel not only dependents on the biosynthesis itself, but also on the ability to export the products from the intra- to the extracellular environment. Excessive accumulation of intracellular metabolites can lead to toxicity and can thereby influence physiology and productivity of a cell. Therefore, identification and overexpression of exporter systems is a classical metabolic engineering approach that was used in many studies and can elevate the product titer by more the 3-fold (Simic, Sahm and Eggeling, 2001; Ginesy *et al.*, 2015; Lubitz *et al.*, 2016). Beside modifying the pathway of interested itself, a further strategy is to re-directing the metabolic flux to a desired product by disrupting competing pathways. Different computational approaches based on FBA combined with genome-scale models can help to identify knock-out targets for optimized flux (Kim and Reed, 2010; Ranganathan, Suthers and Maranas, 2010; Cardoso *et al.*, 2018). Gene knock-outs that were previously identified by computational predictions were shown to efficiently enhance production of valine (Park *et al.*, 2007) and threonine (Lee *et al.*, 2007) in *E. coli* or sesquiterpenes in *S. cerevisiae* (Asadollahi *et al.*, 2009). The previously introduced CRISPRi was used as a tool for multiplexed knockdown of competing targets without the need of gene deletions as shown for isopentenol producing *E. coli* (Tian *et al.*, 2019). Another approach engineered *E. coli* for enhanced production of peonidin 3-*O*-glucoside (P3G) by CRISPRi-mediated down-regulation of the methionine repressor MetJ. Blocking MetJ expression led to deregulation of the methionine pathway and increased availability of S-adenosyl-L-methionine (SAM). SAM is utilized for P3G formation and thereby CRISPRi against MetJ increased P3G titer more then 2-fold compared to an overproduction base strain (Cress *et al.*, 2017).

The previously described approaches are focused on increasing production of a certain compound by disruption of regulatory mechanisms. However, dysregulating cellular metabolisms often reduces fitness and productivity of a host. Thus, engineering regulation not only focusses on increasing production, but also on controlling and autoregulating an overproduction process. A promising approach to control production aims for balancing overproduction with overall metabolism by engineered feedback mechanisms. One of the first examples was presented by Farmer and Liao, who improved lycopene production in *E. coli* by a synthetic control system (Farmer and Liao, 2000). The researches engineered a regulatory control circuit that directed flux to lycopene production in response to the metabolic state of the cells. Increased glycolytic flux led to activation of the response regulator NtrC by sensing the level of the intracellular metabolite acetyl-phosphate. NtrC led to expression of the two enzymes phosphoenolpyruvate synthase (pps) and isopentenyl diphosphate

isomerase (idi) that were fused to a NtrC regulated promotor. These two key enzymes were shown to redirect flux into an artificially introduced lycopene overproduction pathway. This synthetic feedback circuit efficiently balanced production with the glycolytic state and thereby improved productivity by 3-fold compared to the base strain for lycopene production. This study was the basis for engineering feedback circuits, as shown for controlling fuel production using a dynamic sensor-regulator system (Zhang, Carothers and Keasling, 2012), or constructing a genetically encoded metabolic switch to balance fatty acid production with the malonyl CoA pathway (Xu *et al.*, 2014). These approaches have in common, that they use transcription factor-based systems that sense a metabolic input and directly regulate expression (activation or repression) of a metabolic gene.

Voigt and colleagues showed that such transcription factor-based sensors can be coupled to CRISPRi and targeted proteolysis to control gene expression (Moser *et al.*, 2018). Such sensors were programmed to respond to generic signals during a bioproduction process (consumption of feedstock, oxygen level and by-product formation) and activating expression of dCas9 and a targeted protease. dCas9 and Proteolysis were then used to inhibit expression of target genes on the transcriptional and translational level. Targets were for example genes during early and late stages of growth that normally cause by-product formation (acetate). A further strategy to dynamically control production in a bioprocess was to use the natural quorum sensing (QS) system, which relies on the accumulation of small-molecules in a cell population (Miller and Bassler, 2001; Gupta *et al.*, 2017). Sensing of these small molecules allows density-dependent activation of control circuits to activate or repress target genes. Prather and co-workers used this approach for dynamic control of gene expression to balance 'growth mode' and 'production mode'. By placing the glycolytic enzyme Pfk1 under control of a QS regulated promotor, they were able to re-direct glycolytic flux in a cell-density-dependent manner and use it for the production of myo-inositol and glucaric acid. Thereby, specific titers were increased by ten- and fivefold respectively, compared to the strains without dynamic flux control.

Summarizing the reviewed literature reveals a strategical change of metabolic engineering in the last decades: While previous work was mainly focused on increasing production, recent studies apply system level understanding and novel methods in order to coordinate microbial metabolism with production. This allows construction of stable microbial cell factories with higher robustness and productivity.

References

Asadollahi, M. A. *et al.* (2009) 'Enhancing sesquiterpene production in Saccharomyces cerevisiae through in silico driven metabolic engineering', *Metabolic Engineering*. doi: 10.1016/j.ymben.2009.07.001.

Barrangou, R. *et al.* (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science*. doi: 10.1126/science.1138140.

Bley Folly, B. *et al.* (2018) 'Assessment of the interaction between the flux-signaling metabolite fructose-1,6-bisphosphate and the bacterial transcription factors CggR and Cra', *Molecular Microbiology*, 109(3), pp. 278–290. doi: 10.1111/mmi.14008.

Boch, J. *et al.* (2009) 'Breaking the code of DNA binding specificity of TAL-type III effectors', *Science*. doi: 10.1126/science.1178811.

Caldara, M. *et al.* (2008) 'Arginine biosynthesis in Escherichia coli: Experimental perturbation and mathematical modeling', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M705884200.

Caldara, M., Charlier, D. and Cunin, R. (2006) 'The arginine regulon of Escherichia coli: Whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation', *Microbiology*. doi: 10.1099/mic.0.29088-0.

Cannon, W. B. (1929) 'ORGANIZATION FOR PHYSIOLOGICAL HOMEOSTASIS', *Physiological Reviews*, 9(3), pp. 399–431. doi: 10.1152/physrev.1929.9.3.399.

Cardoso, J. G. R. *et al.* (2018) 'Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories', *ACS Synthetic Biology*. doi: 10.1021/acssynbio.7b00423.

Chen, Z. *et al.* (2014) 'Deregulation of feedback inhibition of phosphoenolpyruvate carboxylase for improved lysine production in Corynebacterium glutamicum', *Applied and Environmental Microbiology*. doi: 10.1128/AEM.03535-13.

Cho, B. K. *et al.* (2012) 'Deciphering the transcriptional regulatory logic of amino acid metabolism', *Nature Chemical Biology*, 8(1), pp. 65–71. doi: 10.1038/nchembio.710.

Christodoulou, D. *et al.* (2017) ' Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli ', *Molecular Systems Biology*, 13(1), p. 903. doi: 10.15252/msb.20167402.

Christodoulou, D. *et al.* (2018) 'Reserve Flux Capacity in the Pentose Phosphate Pathway Enables Escherichia coli's Rapid Response to Oxidative Stress', *Cell Systems*, 6(5), pp. 569-578.e7. doi: 10.1016/j.cels.2018.04.009.

Chubukov, V. *et al.* (2013) 'Transcriptional regulation is insufficient to explain substrate-induced flux changes in Bacillus subtilis', *Molecular Systems Biology*, 9. doi: 10.1038/msb.2013.66.

Chubukov, V. *et al.* (2014) 'Coordination of microbial metabolism', *Nature Reviews Microbiology*, pp. 327–340. doi: 10.1038/nrmicro3238.

Cress, B. F. *et al.* (2017) 'CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production', *Microbial Cell Factories*. doi: 10.1186/s12934-016-0623-3.

Csonka, L. N. *et al.* (1988) 'Nucleotide sequence of a mutation in the proB gene of Escherichia coli that confers proline overproduction and enhanced tolerance to osmotic stress', *Gene*. doi: 10.1016/0378-1119(88)90335-6.

Davidi, D. and Milo, R. (2017) 'Lessons on enzyme kinetics from quantitative proteomics', *Current Opinion in Biotechnology*, pp. 81–89. doi: 10.1016/j.copbio.2017.02.007.

Donati, S., Sander, T. and Link, H. (2018) 'Crosstalk between transcription and metabolism: how much enzyme is enough for a cell?', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. doi: 10.1002/wsbm.1396.

Doudna, J. A. and Charpentier, E. (2014) 'The new frontier of genome engineering with CRISPR-Cas9', *Science*. doi: 10.1126/science.1258096.

Fang, X. *et al.* (2017) ' Global transcriptional regulatory network for Escherichia coli robustly connects gene expression to transcription factor activities ', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1702581114.

Farmer, W. R. and Liao, J. C. (2000) 'Improving lycopene production in Escherichia coli by engineering metabolic control', *Nature Biotechnology*. doi: 10.1038/75398.

Fenton, A. W. (2008) 'Allostery: an illustrated definition for the "second secret of life"', *Trends in Biochemical Sciences*, 33(9), pp. 420–425. doi: 10.1016/j.tibs.2008.05.009.

Franzosa, E. A. *et al.* (2015) 'Sequencing and beyond: Integrating molecular "omics" for microbial community profiling', *Nature Reviews Microbiology*. doi: 10.1038/nrmicro3451.

Fuhrer, T. *et al.* (2011) 'High-throughput, accurate mass metabolome profiling of cellular extracts by flow injection-time-of-flight mass spectrometry', *Analytical Chemistry*. doi: 10.1021/ac201267k.

Fuhrer, T. *et al.* (2017) ' Genomewide landscape of gene–metabolome associations in Escherichia coli ', *Molecular Systems Biology*, 13(1), p. 907. doi: 10.15252/msb.20167150.

Gama-Castro, S. *et al.* (2016) 'RegulonDB version 9.0: High-level integration of gene regulation, coexpression, motif clustering and beyond', *Nucleic Acids Research*. doi: 10.1093/nar/gkv1156.

Gilbert, L. A. *et al.* (2014) 'Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation', *Cell*. doi: 10.1016/j.cell.2014.09.029.

Ginesy, M. *et al.* (2015) 'Metabolic engineering of Escherichia coli for enhanced arginine biosynthesis', *Microbial Cell Factories*, 14(1). doi: 10.1186/s12934-015-0211-y.

Grimbs, S. *et al.* (2007) 'The stability and robustness of metabolic states: Identifying stabilizing sites in metabolic networks', *Molecular Systems Biology*, 3. doi: 10.1038/msb4100186.

Guder, J. C. *et al.* (2017) 'Time-Optimized Isotope Ratio LC-MS/MS for High-Throughput Quantification of Primary Metabolites', *Analytical Chemistry*, 89(3). doi: 10.1021/acs.analchem.6b03731.

Guo, A. C. *et al.* (2013) 'ECMDB: The E. coli Metabolome Database', *Nucleic Acids Research*, 41(D1). doi: 10.1093/nar/gks992.

Gupta, A. *et al.* (2017) 'Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit', *Nature Biotechnology*. doi: 10.1038/nbt.3796.

Hartl, J. *et al.* (2017) 'Longevity of major coenzymes allows minimal de novo synthesis in microorganisms', *Nature Microbiology*, 2. doi: 10.1038/nmicrobiol.2017.73.

Hasin, Y., Seldin, M. and Lusis, A. (2017) 'Multi-omics approaches to disease', *Genome Biology*. doi: 10.1186/s13059-017-1215-1.

Hsu, P. D., Lander, E. S. and Zhang, F. (2014) 'Development and applications of CRISPR-Cas9 for genome engineering', *Cell*. doi: 10.1016/j.cell.2014.05.010.

Jacob, F. and Monod, J. (1961) 'Genetic regulatory mechanisms in the synthesis of proteins', *Journal of Molecular Biology*, pp. 318–356. doi: 10.1016/S0022-2836(61)80072-7.

Jinek, M. *et al.* (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science*. doi: 10.1126/science.1225829.

Kafri, M. *et al.* (2016) 'The Cost of Protein Production', *Cell Reports*, 14(1), pp. 22–31. doi: 10.1016/j.celrep.2015.12.015.

Karginov, F. V. and Hannon, G. J. (2010) 'The CRISPR System: Small RNA-Guided Defense in Bacteria and Archaea', *Molecular Cell*. doi: 10.1016/j.molcel.2009.12.033.

Kemmeren, P. *et al.* (2014) 'Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors', *Cell*. doi: 10.1016/j.cell.2014.02.054.

Kerkhoven, E. J. *et al.* (2016) 'Regulation of amino-acid metabolism controls flux to lipid accumulation in yarrowia lipolytica', *npj Systems Biology and Applications*, 2. doi: 10.1038/npjsba.2016.5.

Kim, J. and Reed, J. L. (2010) 'OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains', *BMC Systems Biology*. doi: 10.1186/1752-0509-4-53.

Kim, Y. G., Cha, J. and Chandrasegaran, S. (1996) 'Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain.', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.93.3.1156.

Kitano, H. (2007) 'Towards a theory of biological robustness', *Molecular Systems Biology*, 3. doi: 10.1038/msb4100179.

Kochanowski, K. *et al.* (2013) 'Functioning of a metabolic flux sensor in Escherichia coli', *Proceedings of the National Academy of Sciences*, 110(3), pp. 1130–1135. doi: 10.1073/pnas.1202582110.

Latchman, D. S. (1997) 'Transcription factors: An overview', *International Journal of Biochemistry and Cell Biology*. doi: 10.1016/S1357-2725(97)00085-X.

Lee, K. H. *et al.* (2007) 'Systems metabolic engineering of Escherichia coli for L-threonine production', *Molecular Systems Biology*. doi: 10.1038/msb4100196.

Li, G. W. *et al.* (2014) 'Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources', *Cell*. doi: 10.1016/j.cell.2014.02.033.

Link, H. *et al.* (2015) 'Real-time metabolome profiling of the metabolic switch between starvation and growth', *Nature Methods*. doi: 10.1038/nmeth.3584.

Link, H., Kochanowski, K. and Sauer, U. (2013) 'Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo', *Nature Biotechnology*, 31(4), pp. 357–361. doi: 10.1038/nbt.2489.

Liu, Y., Beyer, A. and Aebersold, R. (2016) 'On the Dependency of Cellular Protein Levels on mRNA Abundance', *Cell*. doi: 10.1016/j.cell.2016.03.014.

Lowe, R. *et al.* (2017) 'Transcriptomics technologies', *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1005457.

Lubitz, D. *et al.* (2016) 'Roles of export genes cgmA and lysE for the production of l-arginine and l-

citrulline by Corynebacterium glutamicum', *Applied Microbiology and Biotechnology*. doi: 10.1007/s00253-016-7695-1.

Martin, V. J. J. *et al.* (2003) 'Engineering a mevalonate pathway in Escherichia coli for production of terpenoids', *Nature Biotechnology*. doi: 10.1038/nbt833.

McCloskey, D., Xu, S., *et al.* (2018) 'Evolution of gene knockout strains of E. coli reveal regulatory architectures governed by metabolism', *Nature Communications*, 9(1). doi: 10.1038/s41467-018-06219-9.

McCloskey, D., Xu, J., *et al.* (2018) 'RapidRIP quantifies the intracellular metabolome of 7 industrial strains of E. coli', *Metabolic Engineering*. doi: 10.1016/j.ymben.2018.04.009.

Miller, M. B. and Bassler, B. L. (2001) 'Quorum Sensing in Bacteria', *Annual Review of Microbiology*. doi: 10.1146/annurev.micro.55.1.165.

Milroy, L.-G. *et al.* (2014) 'Modulators of Protein–Protein Interactions', *Chemical Reviews*, 114(9), pp. 4695–4748. doi: 10.1021/cr400698c.

Mirzaei, H. *et al.* (2013) 'Systematic measurement of transcription factor-DNA interactions by targeted mass spectrometry identifies candidate gene regulatory proteins', *Proceedings of the National Academy of Sciences*, 110(9), pp. 3645–3650. doi: 10.1073/pnas.1216918110.

Monk, J. M. *et al.* (2016) 'Multi-omics Quantification of Species Variation of Escherichia coli Links Molecular Features with Strain Phenotypes', *Cell Systems*, 3(3), pp. 238-251.e12. doi: 10.1016/j.cels.2016.08.013.

Monod, J., Changeux, J. P. and Jacob, F. (1963) 'Allosteric proteins and cellular control systems', *Journal of Molecular Biology*, 6(4), pp. 306–329. doi: 10.1016/S0022-2836(63)80091-1.

Mori, M. *et al.* (2017) 'Quantifying the benefit of a proteome reserve in fluctuating environments', *Nature Communications*, 8(1). doi: 10.1038/s41467-017-01242-8.

Mortazavi, A. *et al.* (2008) 'Mapping and quantifying mammalian transcriptomes by RNA-Seq.', *Nature methods*. doi: 10.1038/nmeth.1226.

Moser, F. *et al.* (2018) 'Dynamic control of endogenous metabolism with combinatorial logic circuits', *Molecular Systems Biology*. doi: 10.15252/msb.20188605.

Murphy, K. C. (1998) 'Use of bacteriophage λ recombination functions to promote gene replacement in Escherichia coli', *Journal of Bacteriology*.

Nielsen, J. and Keasling, J. D. (2016) 'Engineering Cellular Metabolism', *Cell*, pp. 1185–1197. doi: 10.1016/j.cell.2016.02.004.

O'Brien, E. J., Utrilla, J. and Palsson, B. O. (2016) 'Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments', *PLoS Computational Biology*, 12(6). doi: 10.1371/journal.pcbi.1004998.

Ortmayr, K., Dubuis, S. and Zampieri, M. (2019) 'Metabolic profiling of cancer cells reveals genome-wide crosstalk between transcriptional regulators and metabolism', *Nature Communications*. doi: 10.1038/s41467-019-09695-9.

Paczia, N. *et al.* (2012) 'Extensive exometabolome analysis reveals extended overflow metabolism in various microorganisms', *Microbial Cell Factories*, 11. doi: 10.1186/1475-2859-11-122.

Park, J. H. *et al.* (2007) 'Metabolic engineering of Escherichia coli for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0702609104.

Paul, B. J., Berkmen, M. B. and Gourse, R. L. (2005) 'DksA potentiates direct activation of amino acid promoters by ppGpp.', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0501170102.

Piazza, I. *et al.* (2018) 'A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication', *Cell*, 172(1–2), pp. 358-372.e23. doi: 10.1016/j.cell.2017.12.006.

Qi, L. S. *et al.* (2013) 'Repurposing CRISPR as an RNA-yuided platform for sequence-specific control of gene expression', *Cell*. doi: 10.1016/j.cell.2013.02.022.

Rajagopal, B. S. *et al.* (1998) 'Use of inducible feedback-resistant N-acetylglutamate synthetase (argA) genes for enhanced arginine biosynthesis by genetically engineered Escherichia coli K-12 strains', *Applied and Environmental Microbiology*, 64(5), pp. 1805–1811. doi: ADD1523 [pii]10.1111/j.1360-0443.2006.01523.x.

Ramseier, T. M. *et al.* (1993) 'In vitro binding of the pleiotropic transcriptional regulatory protein, FruR, to the fru, pps, ace, pts and icd operons of Escherichia coli and Salmonella typhimurium', *Journal of Molecular Biology*, 234(1), pp. 28–44. doi: 10.1006/jmbi.1993.1561.

Ranganathan, S., Suthers, P. F. and Maranas, C. D. (2010) 'OptForce: An optimization procedure for identifying all genetic manipulations leading to targeted overproductions', *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1000744.

Reaves, M. L. *et al.* (2013) 'Pyrimidine homeostasis is accomplished by directed overflow metabolism', *Nature*. Nature Publishing Group, 500(7461), pp. 237–241. doi: 10.1038/nature12445.

Reisch, C. R. and Prather, K. L. J. (2015) 'The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli', *Scientific Reports*. doi: 10.1038/srep15096.

Sander, T. *et al.* (2019) 'Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance', *Cell Systems*, 8(1), pp. 66-75.e8. doi: 10.1016/j.cels.2018.12.005.

Schmidt, A. *et al.* (2016) 'The quantitative and condition-dependent Escherichia coli proteome', *Nature Biotechnology*, 34(1), pp. 104–110. doi: 10.1038/nbt.3418.

Simic, P., Sahm, H. and Eggeling, L. (2001) 'L-threonine export: Use of peptides to identify a new translocator from corynebactedum glutamicum', *Journal of Bacteriology*. doi: 10.1128/JB.183.18.5317-5324.2001.

Sokolina, K. *et al.* (2017) 'Systematic protein-protein interaction mapping for clinically relevant human GPCRs.', *Molecular systems biology*, 13(3), p. 918. doi: 10.15252/msb.20167430.

Stelling, J. *et al.* (2004) 'Robustness of cellular functions', *Cell*, pp. 675–685. doi: 10.1016/j.cell.2004.09.008.

Tian, T. *et al.* (2019) ' Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in Escherichia coli ', *ACS Synthetic Biology*. doi: 10.1021/acssynbio.8b00429.

Tröndle, J. *et al.* (2018) 'Fed-batch production of l-tryptophan from glycerol using recombinant

Escherichia coli', *Biotechnology and Bioengineering*. doi: 10.1002/bit.26834.

Wang, T. *et al.* (2018) 'Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance-net', *Nature Communications*. doi: 10.1038/s41467-018-04899-x.

Wolters, D. A., Washburn, M. P. and Yates, J. R. (2001) 'An automated multidimensional protein identification technology for shotgun proteomics', *Analytical Chemistry*. doi: 10.1021/ac010617e.

Xu, P. *et al.* (2014) 'Improving fatty acids production by engineering dynamic pathway regulation and metabolic control', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1406401111.

Yanofsky, C., Kelley, R. L. and Horn, V. (1984) 'Repression is relieved before attenuation in the trp operon of Escherichia coli as tryptophan starvation becomes increasingly severe', *Journal of Bacteriology*.

You, C. *et al.* (2013) 'Coordination of bacterial proteome with metabolism by cyclic AMP signalling.', *Nature*. Nature Publishing Group, 500(7462), pp. 301–6. doi: 10.1038/nature12446.

Zhang, F., Carothers, J. M. and Keasling, J. D. (2012) 'Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids', *Nature Biotechnology*. doi: 10.1038/nbt.2149.

Zhang, Y. *et al.* (2013) 'Protein analysis by shotgun/bottom-up proteomics', *Chemical Reviews*. doi: 10.1021/cr3003533

Chapter 1

Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in *E. coli* by Enforcing Enzyme Overabundance

Timur Sander, Niklas Farke, Christoph Diehl, Michelle Kuntz, Timo Glatter, Hannes Link

TS designed the study, performed experiments, performed data analysis and co-wrote the manuscript.

NF performed kinetic modeling and co-wrote the manuscript

CD and MK performed experiments

TG performed proteome measurements

HL supervised the study and co-wrote the manuscript

Summary

Microbes must ensure robust amino acid metabolism in the face of external and internal perturbations. This robustness is thought to emerge from regulatory interactions in metabolic and genetic networks. Here, we explored the consequences of removing allosteric feedback inhibition in seven amino acid biosynthesis pathways in *Escherichia coli* (arginine, histidine, tryptophan, leucine, isoleucine, threonine and proline). Proteome data revealed that enzyme levels decreased in five of the seven dysregulated pathways. Despite lower enzyme levels flux through the dysregulated pathways was not limited, indicating that enzyme levels are higher than absolutely needed in wild-type cells. We show that such enzyme overabundance renders the arginine, histidine and tryptophan pathway robust against perturbations of gene expression, using a metabolic model and CRISPR interference experiments. The results suggest a sensitive interaction between allosteric feedback inhibition and enzyme level regulation that ensures robust yet efficient biosynthesis of histidine, arginine and tryptophan in *E. coli*.

Graphical Abstract
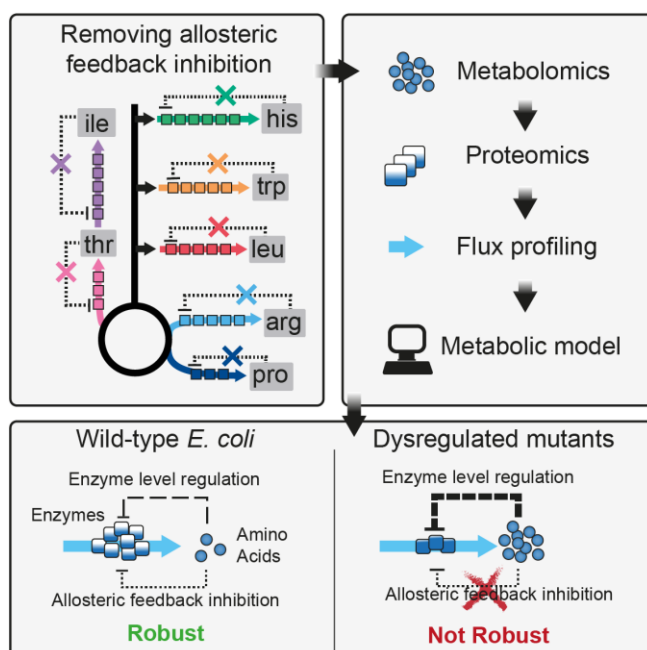
Introduction

Regulation of microbial metabolism involves a wide range of mechanisms that act on different cellular layers and together control the abundance and activity of enzymes (Chubukov et al., 2014). An example is end-product inhibition of amino acid biosynthesis in *Escherichia coli*, which can act on enzyme abundance through transcriptional regulatory cues, and enzyme activities through allosteric feedback inhibition. However, since metabolic reaction rates are determined by both enzyme abundance and enzyme activity it has been difficult to disentangle the specific roles of the two regulatory layers, and to understand how they interact to control metabolism (Chubukov et al., 2013; Daran-Lapujade et al., 2007; ter Kuile and Westerhoff, 2001).

Allosteric feedback inhibition of the committed step in biosynthetic pathways is thought to maintain homeostasis of end-products (Umbarger, 1956), and 16 out of 20 amino acids in *E. coli* feedback inhibit enzymes of their own biosynthesis pathway (Reznik et al., 2017). The consequences of dysregulating these enzymes were mainly studied *in vitro* (Schomburg et al., 2013), or in the context of biotechnological overproduction strains (Hirasawa and Shimizu, 2016). For the case of nucleotide biosynthesis in *E. coli*, a detailed *in vivo* study showed that removing allosteric feedback inhibition did not perturb nucleotide homeostasis (Reaves et al., 2013). In the absence of allosteric feedback inhibition, additional regulatory mechanisms accomplished proper control of the pathway by channeling the excess of nucleotides into degradation pathways (so-called directed overflow). Theoretical analyses, in contrast, suggest a key role of allosteric feedback inhibition in achieving end-product homeostasis (Hofmeyr and Cornish-Bowden, 2000), metabolic robustness (Grimbs et al., 2007), flux control (Kacser and Burns, 1973; Schuster and Heinrich, 1987) and optimal growth (Goyal et al., 2010).

The abundance of enzymes in *E. coli* amino acid metabolism is mainly regulated at the layer of transcription, either by transcriptional attenuation (Yanofsky, 1981) or transcription factors (Cho et al., 2008, 2012). For example, a set of four transcription factors (ArgR, TrpR, TryR and Lrp) control expression of 19 out of 20 amino acid pathways, by sensing the availability of amino acids via allosteric binding (Cho et al., 2012). This regulation ensures that enzymes in amino acid pathways are only made when they are needed (Schmidt et al., 2016; Zaslaver et al., 2004). As a consequence of such need-based enzyme level regulation, one would expect that enzyme levels are not higher than absolutely needed for amino acid biosynthesis. However, recent data suggest that cells express the majority of enzymes at higher levels than necessary to fulfill biosynthetic demands, and that such enzyme overabundance provides a benefit in changing environments (Davidi and Milo, 2017; O'Brien et al., 2016). For example, enzyme overabundance enables a quick activation of the pentose phosphate

pathway upon stresses (Christodoulou et al., 2018), and similar benefits were attributed to overabundant ribosomes (Mori et al., 2017) and coenzymes (Hartl et al., 2017).

Here we constructed seven *E. coli* mutants, each with a different feedback-dysregulated amino acid biosynthesis pathway (arginine, histidine, tryptophan, leucine, isoleucine, threonine and proline), and measured proteins, metabolites, fluxes and growth of the mutants. In all seven feedback-dysregulated pathways the concentration of amino acid end-products increased, and in five pathways we measured lower enzyme levels. Despite the lower enzyme levels biosynthetic flux was not limited, indicating that these enzymes are not operating at maximal capacity in wild-type cells. By combining theoretical and experimental analysis, we show that this enzyme overabundance provides a robustness benefit against genetic perturbations, in case of the arginine, tryptophan and histidine pathway.

Results

Dysregulating Allosteric Enzymes Changes Levels of Specific Amino Acids in *E. coli*

To explore the function of allosteric feedback inhibition in the arginine, histidine, tryptophan, leucine, isoleucine, threonine and proline biosynthesis pathways, we first created a panel of seven allosterically dysregulated *E. coli* mutants (Figure 1A and Table S1). Using a scarless CRISPR method (Reisch and Prather, 2015), we introduced point mutations into genes encoding the allosteric enzyme that catalyzes the committed reaction in each pathway (*argA*, *hisG*, *trpE*, *leuA*, *ilvA*, *thrA* and *proB*). These mutations have been shown previously to abolish the allosteric interaction while not affecting enzyme activity, thereby allowing us to study regulation of the pathway in the absence of allosteric feedback (Caligiuri and Bauerle, 1991; Csonka et al., 1988; Doroshenko et al., 2013; Gusyatiner et al., 2005; LaRossa et al., 1987; Lee et al., 2003; Rajagopal et al., 1998). For N-acetylglutamate synthase (ArgA), we confirmed with *in vitro* assays that the mutation does not affect enzymatic activity, and abolishes inhibition by arginine (Figure S1). To analyze metabolism of the mutants we quantified intracellular metabolites during exponential growth on glucose by LC-MS/MS (Guder et al., 2017). Stronger metabolic changes were restricted to amino acid biosynthesis, with specific increases between 2- and 16-fold of only the amino acid products of the dysregulated pathways (Figure 1B). Despite these changes within the dysregulated pathways, the remaining amino acid concentrations as well as the global metabolite profile remained relatively stable (Figure 1B and S2). Thus, dysregulating allosteric enzymes in *E. coli* amino acid biosynthesis elevated the intracellular concentration of the corresponding amino acid product.
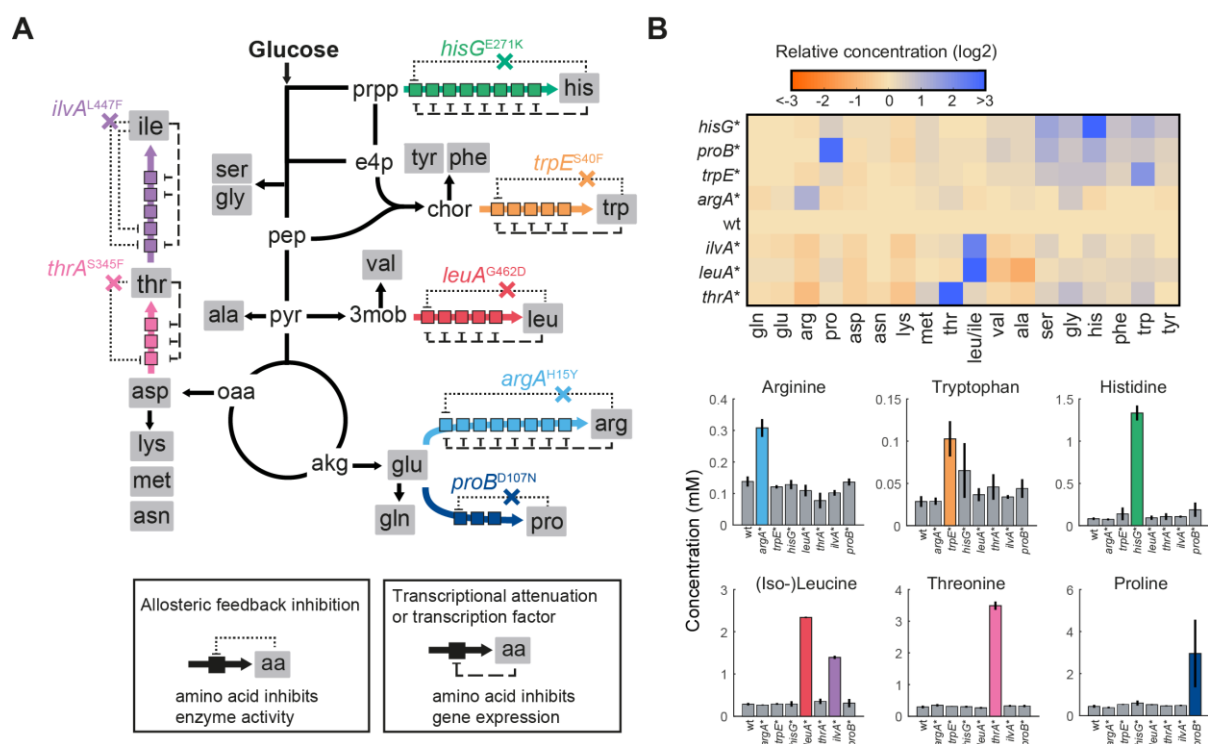
Figure 1. Amino Acid Profile of Feedback-dysregulated *E. coli* Mutants

(A) Seven amino acid pathways were dysregulated by genomic point mutations in the indicated genes. See also Table S1. Negative allosteric feedbacks of amino acids on enzymes in the biosynthetic pathways are shown as dotted lines. Negative transcriptional feedbacks of amino acids are shown as dashed lines. Boxes indicate enzymes in the biosynthesis pathways.

(B) Relative concentrations of intracellular amino acids in wild-type *E. coli* and the seven dysregulated mutants. Bar plots show absolute concentrations of the amino acid in the dysregulated pathways. See also Figure S2. Data are represented as mean, and error bars are ± SD (n = 3).

Lower Expression of Enzymes in Feedback-Dysregulated Pathways

With the exception of proline biosynthesis, all of the dysregulated pathways are additionally controlled at the layer of enzyme abundance, either via transcription factors or transcriptional attenuation. To probe if elevated amino acid concentrations in our mutants affected enzyme levels in the corresponding pathways, we measured their proteomes (Figure 2A). The data covered relative abundances of 173 out of the 204 enzymes annotated to amino acid metabolism in the latest *E. coli* metabolic model (Monk et al., 2017). Enzyme expression was indeed lower in five of the seven dysregulated pathways (*argA**, *trpE**, *hisG**, *leuA**, *thrA**), indicating that the elevated amino acid concentrations caused a compensatory downregulation of their associated pathway (Figure 2A). Enzyme levels did not change in the *proB** and *ilvA** mutant, which is expected because proline

biosynthesis lacks enzyme level regulation and isoleucine biosynthesis is subject to a second allosteric feedback that was not removed (Figure 1A and 2A). The *leuA** mutant showed more global changes in enzyme levels than the other mutants. The high leucine concentration in this strain likely activates the leucine responsive transcription factor Lrp, which acts on many genes in amino acid metabolism (Cho et al., 2008). In the *argA** mutant we observed an expected accompanying decrease in histidine biosynthesis enzymes, which are additional targets of the transcription factor ArgR (Gama-Castro et al., 2016). Apart from the compensatory downregulation of biosynthetic enzymes, enzymes in dedicated amino acid degradation pathways were upregulated in three mutants (AstC in the arginine mutant, TnaA in the tryptophan mutant and PutA in the proline mutant, Figure 2A). This likely constitutes an additional compensatory mechanism similar to the directed overflow reported for nucleotides (Reaves et al., 2013).

To obtain additional evidence for lower enzyme levels in the dysregulated pathways, we used GFP-promoter fusions and measured fluorescence in single cells (Figure 2B). GFP expression from an ArgR-regulated promoter was indeed ~3-fold lower in the *argA** mutant compared to the wild-type. Similarly, a TrpR-regulated promoter was ~3-fold stronger repressed in the *trpE** mutant. The cell-to-cell variation in GFP content was similar in wild-type cells and the mutants, thus indicating that all cells in the population of allosteric feedback mutants have lower enzyme levels in the dysregulated pathway. A GFP reporter with the *thrL* leader peptide was only 17% repressed in the *thrA** mutant compared to the wild-type, which is consistent with the small decrease of enzymes levels in the dysregulated threonine pathway (Figure 2A and 2B). We also fused GFP to the *hisL* and *leuL* leader peptides, but they did not report repression by amino acids even when they were added to the medium (Figure S3). Probably transcriptional attenuation by *hisL* and *leuL* requires the genomic context, and cannot function on plasmids. In summary, proteome data revealed a lower expression of enzymes for five of the seven dysregulated pathways (*argA**, *trpE**, *hisG**, *leuA** and *thrA**). GFP-promoter fusions confirm this enzyme level regulation at the single cell level, and indicate that downregulation of enzymes in the *argA**, *trpE** and *thrA** mutant occurs at the transcriptional layer.
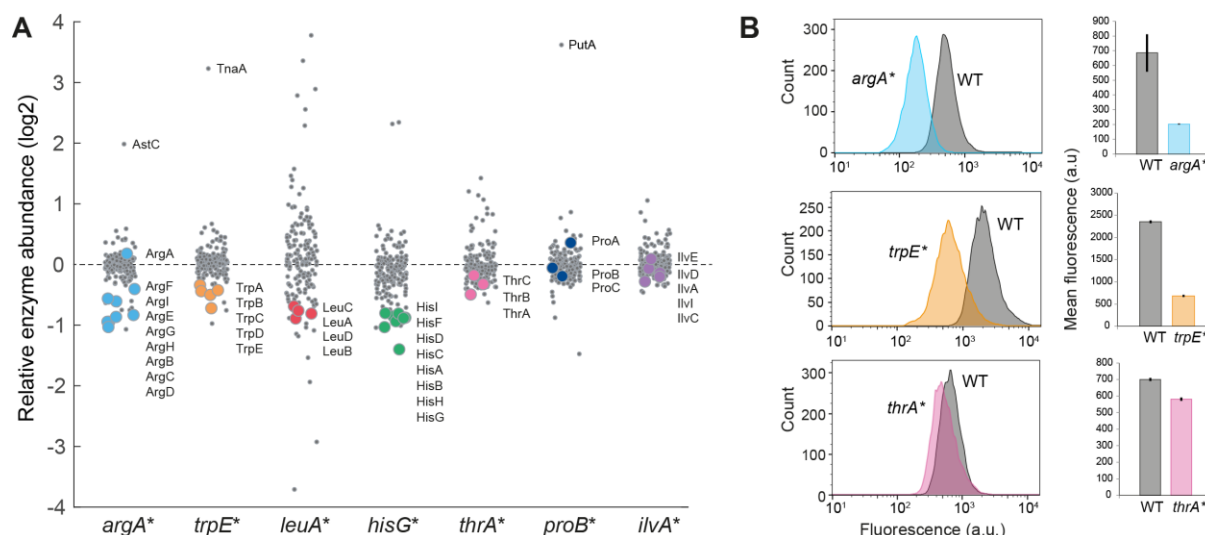
Figure 2. Expression of Enzymes in Feedback-dysregulated Pathways

(A) Abundance of 173 enzymes in amino acid metabolism (out of 204 enzymes in total), relative to the level in the wild-type. Data are represented as mean (n = 3). For each strain the enzymes in the dysregulated pathway are shown as colored dots. Enzymes in degradation pathways of arginine, tryptophan and proline are indicated by their names.

(B) GFP-fluorescence measured by flow cytometry. GFP-promoter fusions were transformed in wild-type cells and the indicated mutant. Upper panel: pP*argA-gfp*; middle panel: pP*trpL-gfp;* lower panel: pP*thrL-gfp*. Histograms represent fluorescence of 10,000 single cells. Mean fluorescence was calculated from 10,000 single cells of n = 3 independent cultures. See also Figure S3.

Allosteric Feedback Inhibition Enforces Enzyme Overabundance

Next, we wondered if lower expression of enzymes limits the biosynthetic capacity of the mutants. First, we tested steady state growth on glucose minimal medium and seven other carbon sources (Figure S4). All mutants showed wild-type like growth, except the *leuA** mutant, which grew in average 10% slower than the wild-type. To test if lower enzyme levels affect biosynthetic capacity in dynamics shifts, we starved cells for carbon and measured growth resumption on glucose minimal medium (Figure 3A). During the initial phase of growth resumption all mutants had the same growth rate as the wild-type. Only the *leuA**, *ilvA** and *thrA** mutants reached lower growth rates than the wild-type during the subsequent 4 hours. The three strains had also lower ODs after 20 hours starvation. Similarly, nutritional upshifts from galactose to glucose did not affect growth of the mutants (Figure S5). The three strains with highest reduction in enzyme levels (*argA**, *trpE** and *hisG**) grew like the wild-type in all tested conditions, indicating that biosynthetic capacity is not limited by lower enzyme

level. The advantage of lower protein costs in these pathways was either too subtle to be detected by growth assays, or counterbalanced by negative effects of feedback-dysregulation.
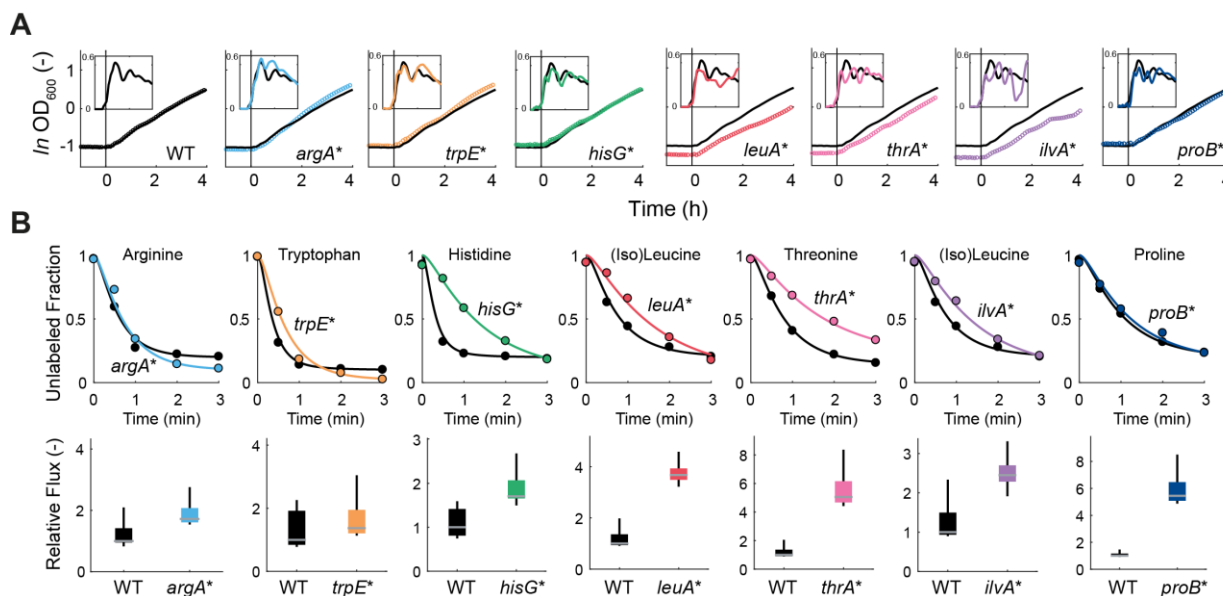

Figure 3. Growth and Biosynthetic Flux of Feedback-dysregulated *E. coli* Mutants

(A) Growth resumption after 20 hours carbon starvation of wild-type *E. coli* and the seven dysregulated mutants. Cells were starved in minimal medium and glucose was added at t = 0 h. OD was measured in 5 minute intervals in a plate reader. Shown are means of n = 3 cultures. Inserts show the specific growth rate in h^{-1} during the same time period. Growth rates were estimated by linear regression over a moving 30 minute window. The same wild-type growth curve and growth rate is shown in each graph in black as a reference. See also Figure S4 and S5.

(B) Decay of unlabeled amino acids in the wild-type *E. coli* (black) and the seven dysregulated mutants (color). The measured amino acid is indicated above each graph. Cells were loaded from shake flasks onto filters and perfused with ^{15}N-medium for different lengths of time (0, 30, 60, 120 and 180 seconds). Dots are means of n = 2 samples for each time point. Lines are means of 1000 fits of decay rates based on equations for kinetic flux profiling. Box plots show fluxes based on the 1000 fits, relative to the median flux estimate in the wild-type. Boxes contain 50% and whiskers 99% of the flux estimates.

To directly probe biosynthetic capacity, we traced intracellular fluxes of amino acids with ^{15}N labeling experiments (Figure 3B). Labeling of arginine, tryptophan and proline was similar in the respective mutant and the wild-type, whereas histidine, (iso)-leucine and threonine labeled slower in the mutants. However, it is important to consider that labeling rates depend on fluxes, and also on absolute pool sizes of amino acids. Because amino acid pools were higher in the mutants we used a

method for quantitative analysis of the labeling profiles to estimate fluxes (Yuan et al., 2008). To account for unknown labeling profiles of upstream nitrogen precursors, we calculated fluxes for a wide range of precursor labeling rates in the literature (Yuan et al., 2006). The flux estimates show that none of the mutants had lower flux through the dysregulated pathways than the wild-type (Figure 3B). In most cases biosynthetic flux was even higher, indicating that downregulation of enzyme levels could not fully compensate the loss of allosteric feedback inhibition in some of the mutants. This might be the reason for the growth-phenotype of the *leuA**, *ilvA** and *thrA** mutants in dynamic growth experiments (Figure 3A).

In conclusion, the feedback-dysregulated mutants showed the same or higher flux through the dysregulated amino acid pathways than wild-type cells, although in five mutants (*argA**, *trpE**, *hisG**, *leuA** and *thrA**) enzyme levels in the dysregulated pathway were lower. Especially, the *argA**, *trpE** and *hisG** mutant had ~2-fold lower enzyme levels in the dysregulated pathways compared to the wild-type, while fluxes were 1-2 fold higher and growth was unaffected. This indicates that these enzymes are not operating at maximal capacity in wild-type *E. coli* during growth on glucose. We then hypothesized that this enzyme overabundance emerges from allosteric feedback inhibition by maintaining low concentration of end-products, which in turn increases production of enzymes (e.g. by de-repression of transcription). Next, we explored this interplay between control of enzyme activity and enzyme abundance and its relevance for cellular metabolism.

Interdependence of Allosteric Feedback Inhibition and Enzyme Level Regulation

To obtain a better mechanistic understanding of the interplay between allosteric feedback inhibition and enzyme level regulation, we developed a kinetic model of metabolism and enzyme expression (Figure 4A). Briefly, the model includes two enzymes e_1 and e_2, and two metabolites m_1 and m_2 in a two-step pathway. The end-product m_2 represents an amino acid, which is consumed in the last reaction for protein synthesis and growth. The end-product m_2 feedback inhibits the expression of both enzymes, as well as the activity of the first enzyme. The first reaction and the expression of both enzymes follow simple inhibition kinetics, whereas the second reaction follows Michaelis-Menten kinetics (Figure 4A). As such this model is a simplified representation of an amino acid biosynthesis pathway that is controlled at two layers (Figure 1A).
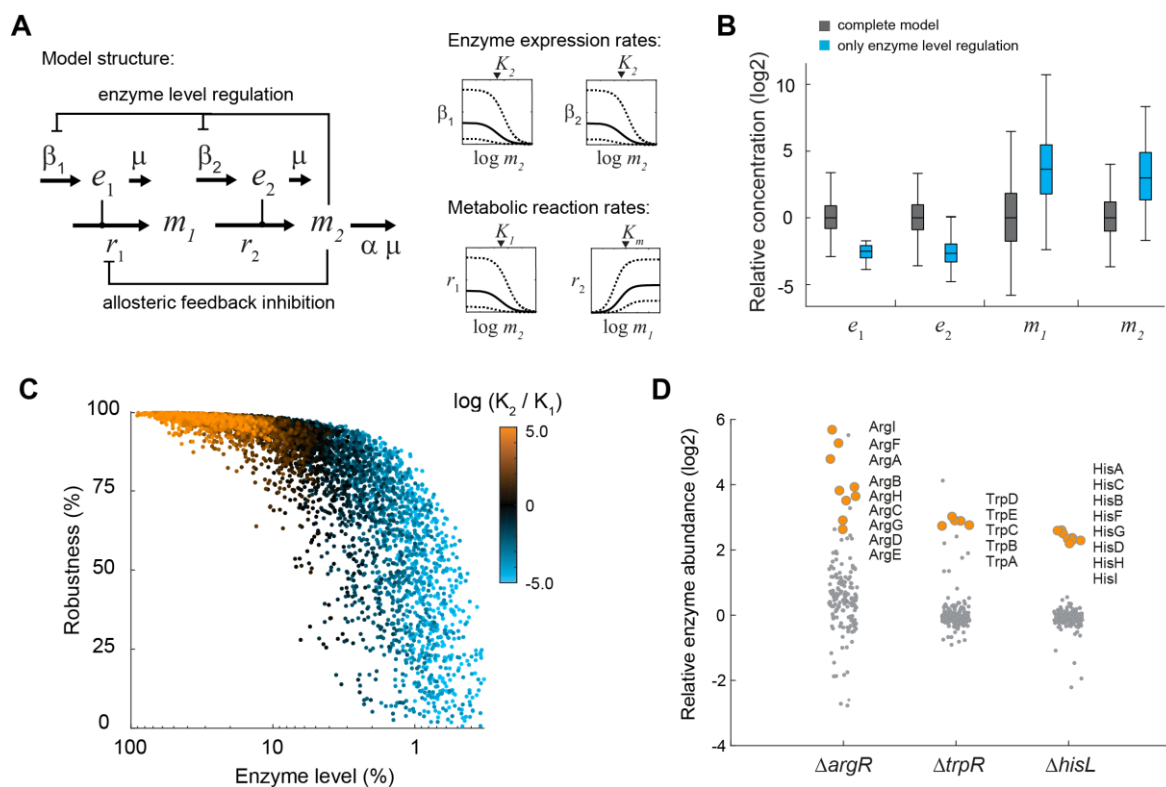
Figure 4. A Kinetic Model Predicts a Robustness-Efficiency Tradeoff

(A) Stoichiometry and structure of the kinetic model. m_1 and m_2 are metabolites, e_1 and e_2 are enzymes. Kinetics of the enzyme catalyzed reactions r_1 and r_2, as well as kinetics of enzyme expression rates β_1 and β_2 are sampled in the indicated intervals.

(B) Steady state concentrations of e_1, e_2, m_1 and m_2 calculated with 5000 random parameter sets for the complete model (grey), and the model with only enzyme level regulation (blue). Boxes contain 50% and whiskers 99% of the simulated concentrations. All concentrations are normalized to the median concentrations of the complete model. See also Figure S6 and S7.

(C) Enzyme levels (sum of e_1 and e_2) and robustness against perturbations of $\beta_{2,max}$, for 5000 simulations of the complete model (dots). The color of each dot shows the ratio of inhibition constants for allosteric feedback inhibition (K_1) and enzyme level regulation (K_2) in the respective model. Robustness corresponds to the percentage downregulation of $\beta_{2,max}$ that was tolerated by each model. 100% enzyme abundance corresponds to the maximum theoretical enzyme concentration in the model.

As a starting point for the model analysis, we fixed the flux in the pathway to the amino acid requirement given by the growth rate of *E. coli* on glucose. We randomly sampled seven model parameters (maximal rates and binding constants) 5000 times from physiologically meaningful ranges based on literature values (Davidi and Milo, 2017; Li et al., 2014; Milo et al., 2010). For each of the thus derived 5000 parameter sets we calculated concentrations of e_1, e_2, m_1 and m_2, for a model

including feedback on enzyme activity and enzyme abundance (complete model, grey in Figure 4B), and also for a model including only feedback on enzyme abundance (single feedback model, blue in Figure 4B). The simulated concentrations of e_1, e_2, m_1 and m_2 matched qualitatively the measured protein and metabolite data: the two enzymes decreased in the single feedback model (Figure 2A), whereas the end-product m_2 increased (Figure 1B). Also the simulated concentration of the intermediate m_1 matched the measured increase of intermediates in amino acid pathways (Figure S6). Thus a simple model confirms our hypothesis that allosteric feedback inhibition enforces enzyme overabundance. In theory, other types of enzyme inhibition could cause a similar increase in enzyme expression. To test this, we replaced the allosteric feedback in the model with competitive product inhibition of the second reaction (Figure S7). However, removing competitive product inhibition was compensated by lower substrate concentrations (m_1), and not by lower enzyme levels. This model result indicates that enzyme overabundance does not emerge from all types of enzyme inhibition.

The Interplay of two Feedbacks solves a Robustness-Efficiency Tradeoff

Next, we set out to investigate the function that emerges from the interplay between feedback on enzyme activity and enzyme abundance. While low enzyme levels are obviously advantageous due to lowering protein cost, high enzyme levels could provide a cellular benefit by improving robustness against perturbations in enzyme expression. To test this with the model, we made use of a numerical parameter continuation method to quantify robustness (Lee et al., 2014). This method iteratively decreases a model parameter until instabilities occur in the model. Robustness can then be defined as the percentage change of this parameter that was tolerated. Using this method we calculated robustness against perturbations of the maximal expression rate of the second enzyme ($\beta_{2,max}$) in the complete model with 5000 randomly sampled parameter sets (Figure 4C). Changing $\beta_{2,max}$ reflects genetic or environmental perturbations of gene expression that can lead to a bottleneck in the pathway. Consistent with our expectations, models with high enzyme levels showed increased robustness, while models with lower enzyme levels were more sensitive to perturbations of enzyme expression (Figure 4C). However, robustness was not proportional to the enzyme level: a relatively small increase of enzyme levels already conferred a large robustness benefit. Very high enzyme levels, in comparison, did not increase robustness substantially over more subtle changes in enzyme abundance. Our model thus reveals a tradeoff between protein costs and robustness, which can be solved by sensitively balancing enzyme levels.

The optimal balance of enzyme levels occurs in models occupying the middle of the tradeoff frontier, those models with equally strong feedback on enzyme activity and enzyme abundance (indicated by similar inhibition constants Ki, black dots in Figure 4C). We then wondered if amino acid biosynthesis

in *E. coli* operates in the middle of the tradeoff frontier, meaning that both feedbacks are simultaneously active. In particular enzyme levels in the *argA*, trpE* and hisG** mutant demonstrated that wild-type *E. coli* does not operate with minimal enzyme levels in these pathways (blue dots in Figure 4C). To test if enzymes in these pathways are maximally expressed (orange dots in Figure 4C), we removed their transcriptional regulation, which functions by different mechanisms: a transcription factor (arginine), transcriptional attenuation (histidine), or both (tryptophan). In the arginine and tryptophan pathway we deleted the respective transcription factor (Δ*argR* and Δ*trpR*), and in histidine biosynthesis we removed the leader peptide *hisL*. Removing transcriptional regulation of all three pathways resulted in higher expression of enzymes in the respective pathway (Figure 4D): arginine enzymes increased between 5 and 60 fold, histidine enzymes about 6 fold, and tryptophan enzymes about 8 fold. This shows that *E. coli* does not operate at maximal expression of arginine, tryptophan and histidine enzymes, but rather in the middle of the tradeoff frontier. Previous studies that support this observation showed that ArgR binds to promoters of arginine genes more than 80% of the time when *E. coli* grows on glucose (Gerosa et al., 2013). Deletion of ArgR caused more global changes of amino acid enzymes than removing TrpR or HisL. This reflects the potential of ArgR to control metabolism of almost all amino acid pathways (Cho et al., 2012).

Taken together, both model and dysregulated mutants indicate a regulatory interplay in the arginine, tryptophan and histidine pathway: removing transcriptional regulation increased enzyme levels (Figure 4D), whereas removing allosteric regulation decreased enzyme levels (Figure 2A). The model shows that if feedback on enzyme activity and enzyme abundance are simultaneously active, inhibition constant of the two feedbacks must have similar values (black dots in Figure 4C). Inhibition constants and binding affinities in the literature show that feedbacks on enzyme activity and enzyme abundance are indeed equally strong for many amino acids (Table S5), corroborating the existence of a two-pronged regulation strategy.

Enzyme Overabundance Provides Robustness against Genetic Perturbations

To test if arginine, tryptophan and histidine biosynthesis are more robust against perturbations of gene expression in wild-type cells than in the feedback-dysregulated mutants, we used CRISPR interference (CRISPRi) (Larson et al., 2013). We designed single guide RNAs (sgRNA) targeting the genes *argE* in arginine biosynthesis, *hisB* in histidine biosynthesis and *trpA* in tryptophan biosynthesis. The sgRNAs were cloned on a plasmid, which harbors an inducible dCas9 and the constitutively expressed sgRNA. The three CRISPRi plasmids and a control without target sequence were transformed into the wild-type, and also into the *argA**, *trpE** and *hisG** mutants. This resulted in 16 strains with all combinations of genetic perturbations and dysregulation of the three pathways (Figure

5A). All strains expressing the control sgRNA without target sequence grew almost identically and also induction of dCas9 did not affect growth (Figure 5B).

Induction of dCas9 in strains with sgRNAs targeting *argE*, *hisB* and *trp*A reduced growth of all strains by more than 50% (Figure 5C). However, we observed the strongest growth defect when perturbing a gene in a dysregulated pathway. For example, CRISPRi of *argE* reduced growth of the *argA** mutant more than twice as much as the other strains. Similarly the *hisG** and *trpE** mutant were most sensitive to perturbations of expression of *hisB* and *trpA*, respectively. The *argA**mutant was also sensitive to a perturbation of *hisB*, which matches the lower expression of histidine enzymes in this mutant (Figure 2A). These data confirm that feedback-dysregulated mutants are indeed more sensitive to a perturbation of gene expression. Notably, the mutants were only more sensitive to a perturbation within pathways that had lower enzyme levels, and they did not lack a general robustness.

While these data support the hypothesis that high enzyme levels render arginine, histidine and tryptophan biosynthesis more robust against perturbations of gene expression, bacteria would hardly face such strong perturbations in nature. Therefore, we designed the sgRNAs in such a way that the wild-type showed only a small growth defect without induction of dCas9 (Figure 5D). The mild perturbations in un-induced cultures still affected the respective mutants stronger than the other strains, causing instable growth and lower growth rates (Figure 5D). Thus, feedback-dysregulation renders the arginine, tryptophan and histidine pathways more sensitive against perturbations of gene expression, which may arise in nature due to the stochasticity of gene expression.
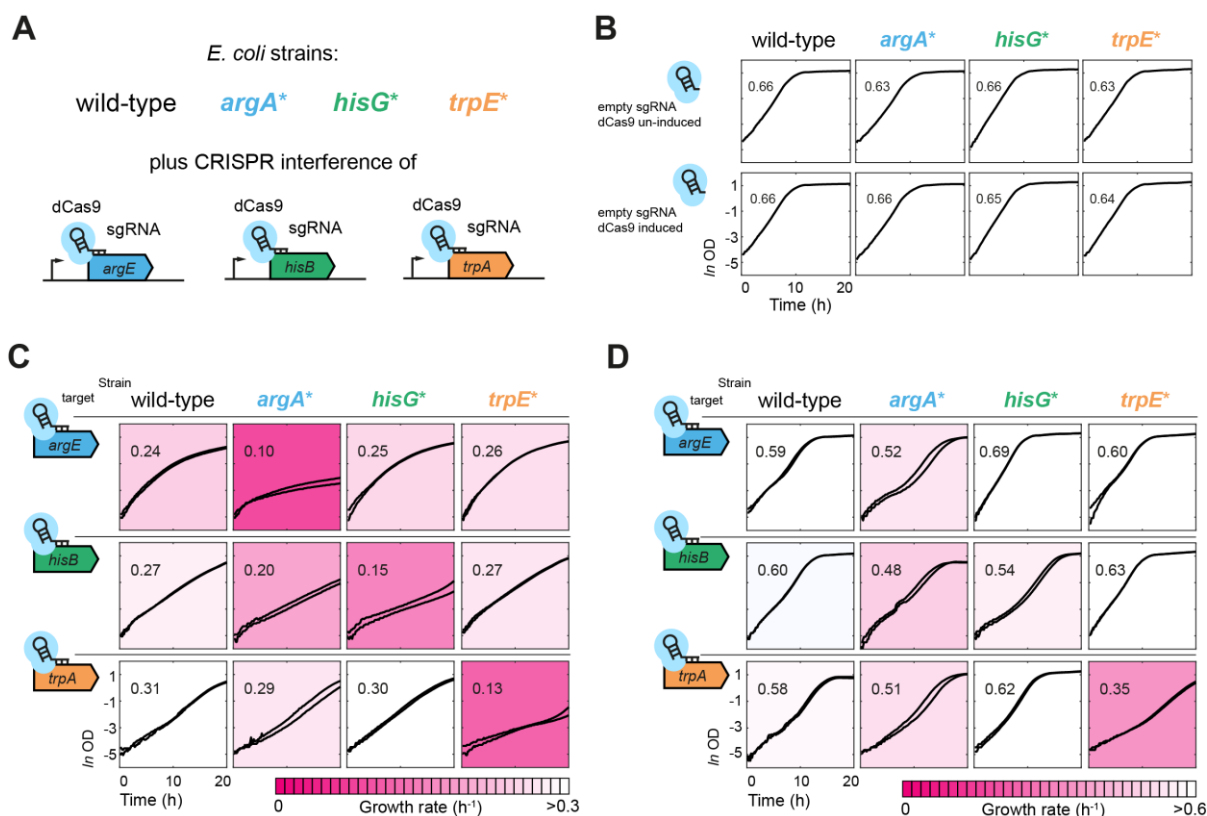
Figure 5. Enzyme Overabundance Achieves Robustness Against Perturbations of Gene Expression by CRISPR Interference.

(A) CRISPR interference in wild-type cells and the allosteric feedback mutants *argA**, *hisG** and *trpE**. Strains were transformed with single guide RNAs targeting genes of the arginine (*argE*), histidine (*hisB*) and tryptophan (*trpA*) pathway, as well as an empty sgRNA without target.

(B) Growth of wild-type, *argA**, *hisG** and *trpE** with the empty control sgRNA. Upper panels show un-induced cultures and lower panel induced cultures (100 μM IPTG). Growth curves show means from n=3 cultures cultivated in minimal glucose medium in a plate reader. Numbers are specific growth rates (in h^{-1}), and were estimated by linear regression between OD 0.2 and 0.6. (c) Growth of wild-type, *argA**, *hisG** and *trpE** with sgRNAs targeting *argE*, *hisB* and *trpA*. dCas9 expression was induced with 100 μM IPTG. Growth curves are means of n=3 cultures; two curves per graph show experiments that were performed at different days. Numbers and colors indicate specific growth rates (in h^{-1}), which were estimated by linear regression between 5 and 15 hours. All axes have ranges shown in the lower left graph.

(D) Same as C) but without induction of dCas9. Growth rates were estimated by linear regression between OD 0.2 and 0.6. All axes have ranges shown in the lower left graph.

Discussion

In this study we explored the consequences of missing allosteric feedback inhibition in seven *E. coli* mutants with dysregulated amino acid biosynthesis pathways: arginine (*argA**), histidine, (*hisG**), tryptophan (*trpE**), leucine (*leuA**), threonine (*thrA**), isoleucine (*ilvA**), and proline (*proB**). In all mutants, the amino acid product of the feedback-dysregulated pathway increased, showing that allosteric feedback inhibition is relevant to maintain end-products at a desired level. In five mutants (*argA**, *trpE**, *hisG**, *thrA**, and *leuA**), we observed a downregulation of enzymes in the dysregulated pathways, presumably because high end-products caused stronger inhibition of enzyme expression. However, these low enzyme levels did not limit biosynthetic flux, thus indicating that wild-type cells maintain higher enzyme levels than would be necessary to ensure sufficient biosynthetic flux (enzyme overabundance). These results are consistent with enzyme overabundance in other pathways (Davidi and Milo, 2017; O'Brien et al., 2016), and the observation that enzymes are rarely operating at maximal capacity (Fendt et al., 2010; Hackett et al., 2016).

Both model analysis and dysregulated mutants indicate that enzyme overabundance is enforced by allosteric feedback inhibition, which maintains low end-product levels and thereby increases production of enzymes. In case of amino acid biosynthesis it is likely that low end-products de-repress transcription, because amino acid levels are known signals for transcription factors and transcriptional attenuation (Cho et al., 2012). Additionally, GFP-promoter fusions indicated regulation at the transcriptional layer in the *argA**, *trpE** and *thrA**mutant. It will be important to clarify if enzyme overabundance emerges also from other inhibitory interactions, which are abundant in metabolic networks (Alam et al., 2017). Besides inhibition of enzymes by metabolites, other sources for enzyme overabundance might be post-translational modifications. For example, it was recently shown that deleting kinases in yeast has a strong effect on enzyme levels (Zelezniak et al., 2018), pointing towards a similar interplay between post-translational modifications of enzymes and enzyme level regulation.

The strongest and most localized decrease of enzyme levels occurred when we removed allosteric feedback inhibition in the arginine, tryptophan and histidine pathway. Removing transcriptional regulation in the same pathways caused higher expression of enzymes, which is in agreement with previous reports of a role for transcriptional regulation in minimizing protein costs in metabolic pathways (Chubukov et al., 2012; You et al., 2013). This antagonistic regulation by allosteric feedback inhibition and transcriptional regulation enables an optimal balance of enzyme levels in amino acid metabolism of wild-type cells. Optimization of enzyme levels has been shown for the global *E. coli* proteome (Scott et al., 2010; You et al., 2013), for the lac system (Dekel and Alon, 2005), and for a single enzyme in the methionine pathway (Li et al., 2014). Here we provided first indication that

enzyme abundance is optimized in the arginine, histidine and tryptophan pathway, to meet multiple, conflicting objectives – robustness and efficiency. Using a simplified model of amino acid metabolism, we show that cells can solve this tradeoff between protein costs and robustness through the interplay of allosteric feedback inhibition and enzyme level regulation. CRISPRi of metabolic enzymes in the dysregulated arginine, tryptophan and histidine pathways showed that allosteric feedback inhibition provides a substantial robustness benefit against perturbations of gene expression. While such robustness effects were attributed to allosteric feedback by previous modelling approaches (Chandra et al., 2011; Grimbs et al., 2007), we quantified it *in vivo* by studying mutants lacking allosteric control. During the lifetime of a cell, perturbations of gene expression could result from stochastic effects at the level of transcription or in response to fluctuating environments.

In conclusion, our case study of *E. coli* amino acid metabolism demonstrated that regulation of enzyme activity and enzyme abundance are not isolated from each other, but interact to control metabolism. Allosteric feedback inhibition sets amino acid concentrations, which are signals for enzyme level regulation. Considering the high precision of metabolite concentrations (Fuhrer et al., 2017; Mülleder et al., 2016), it seems possible that the proposed regulatory principle goes beyond *E. coli* amino acid metabolism.

Material & Methods

Experimental Model and Subject Details

Strains and Culture

E. coli MG1655 (DSMZ No. 18039) was the wild-type strain. Chemically competent *E. coli* TOP10 (One Shot™ TOP10, Invitrogen) were used for cloning. All mutants created in this study derive from the MG1655 strain and are listed in Key Resources Table. Genomic point mutations were created by scarless Cas9 Assisted recombineering (Reisch and Prather, 2015). Therefore, we constructed 7 specific sgRNA-plasmids, derived from the backbone plasmid pKDsgRNA-ack (Addgene #62654). The sgRNAs consist of a gene specific 20 base pair region (*argA*: ggtcgagggattccgccatt; *trpE*: acacaactggtgaaaaagcg; *hisG*: tggaaaaactgaaagcgctg; *thrA*: tggtgctgattacgcaatca; *leuA*: cggtaaagatgcgctgggtc; *ilvA*: caacacgctgggtacgtact; *proB*: cgacaccctgcgagcgttgc), which pairs adjacent to a NGG PAM site. Each sgRNA-plasmid was transformed together with pCas9-CR4 (Addgene #62655) into MG1655 wild-type cells. The resulting strains were grown at 30°C (pKDsgRNA-ack is temperature sensitive at 37°C) and supplemented with arabinose (final concentration 1.2 %) to induce the λ-Red recombinase genes which are located on the sgRNA-plasmid. The induced strains were then transformed with the 70-80 bp homologous oligonucleotides (Table S2), which contain the desired base pair exchanges of PAM site and the point mutation disrupting allosteric feedback (*argA*^{H15Y}, *trpE*^{S40F}, *hisG*^{E271K}, *thrA*^{S345F}, *leuA*^{G462D}, *ilvA*^{L447F}, *proB*^{D107N}). Cells were plated on LB agar containing 100 ng ml^{-1} anhydrotetracycline (aTc) to induce Cas9 expression, which recognizes the sgRNA adjacent to the PAM sequence and cleaves the chromosomal DNA. Only cells that successfully integrated the homologous oligonucleotides will survive due to the modified PAM sequence which prevents Cas9 recognition. Thereby we counter selected for clones harboring the desired amino acid exchanges, which were verified by sequencing. The transcriptional knockout mutants Δ*argR* and Δ*trpR* were constructed with the same cloning procedure according to the noSCAR protocol, while Δ*hisL* was constructed by P1 Phage transduction with the donor strain JW2000-1 (*ΔhisL*) from the Keio collection (Baba et al., 2006).

All cultivations were performed using M9 minimal medium with 5 g L^{-1} glucose (or the respective carbon source in Figure S4). The M9 medium consisted of the following components (per liter): 7.52 g Na$_2$HPO$_4$ 2 H$_2$O, 5 g KH$_2$PO$_4$, 1.5 g (NH$_4$)$_2$SO$_4$, 0.5 g NaCl. The following components were sterilized separately and then added (per liter of final medium): 1 ml 0.1 M CaCl$_2$, 1 ml 1 M MgSO$_4$, 0.6 ml 0.1 M FeCl$_3$, 2 ml 1.4 mM thiamine-HCL and 10 ml trace salts solution. The trace salts solution contained (per liter): 180 mg ZnSO$_4$ 7 H$_2$O, 120 mg CuCl$_2$ 2 H$_2$O, 120 mg MnSO$_4$ H$_2$O, 180 mg CoCl$_2$ 6 H$_2$O. Where appropriate, 50 μg mL^{-1} kanamycin, 34 μg mL^{-1} chloramphenicol, 15 μg mL^{-1} gentamycin, 50 μg mL^{-1}

spectinomycin or 100 µg mL^{-1} ampicillin was added. For cultivations in microtiter plates, LB pre-culture in 96-deep-well format plates were inoculated from glycerol stocks and grown to an exponential stage. From this first pre-culture a second M9 pre-culture in 96-deep-well plates was inoculated 1:100 and incubated overnight at 37 °C under shaking. Finally, 96-well flat transparent plates (Greiner Bio-One International) containing 150 µl M9 minimal medium were inoculated 1:150 from the overnight culture. Online measurements of optical density at 600 nm (OD$_{600}$) were performed at 37°C with shaking in a plate reader (Epoch, BioTek Instruments Inc, USA; Spark 10M, Tecan Trading AG, Switzerland). For induction of CRISPRi, IPTG was added to the main culture to a final concentration of 100 µM. Growth rates were calculated as dln(OD)/dt by linear regression over the indicated time windows. For cultivations in shake flask, 5 ml LB pre-culture in cultivation tubes were inoculated from glycerol stocks and grown to an exponential stage. From this first pre-culture, 5 ml of a second M9 glucose pre-culture in cultivation tubes was inoculated 1:100 and incubated overnight at 37°C in a rotary shaker. For the main culture, a 500 ml shake flask containing 35 ml M9 minimal medium (5 g L^{-1} glucose) was inoculated 1:150 from the overnight culture, and incubated at 37 °C under shaking at 220 rpm.

Method Details

CRISPR Interference

CRISPR interference experiments were performed with a single plasmid (pNUT1533) expressing the sgRNA from a constitutive and the dCas9 protein from an IPTG inducible Ptac promotor. For construction of this plasmid, the sgRNA and its constitutive promotor were amplified from the pgRNA plasmid (Addgene #44251) and the dCas9 gene was amplified from the pdCas9 plasmid (Addgene #44249). The promotor of dCas9 was replaced by an IPTG inducible Ptac promotor. To assure tight regulation of dCas9 expression, the gene coding for the lacIQ1 repressor (Glascock and Weickert, 1998) was added to the vector. The two single fragments were joined together by PCR and the resulting fragment was inserted into pNUT542 with PacI and NotI restriction enzymes (New England Biolabs, USA). This plasmid was used as a backbone for cloning of the specific plasmids targeting the arginine (pNUT1533-*argE*), histidine (pNUT1533-*hisB*) and tryptophan pathway (pNUT1533-*trpA*). Therefore, sgRNAs guide sequences were customized by site-directed mutagenesis using the primer listed in Table S6.

Metabolite Measurements

Shake flask cultivations on M9 glucose were performed as described above. Cells were grown to an optical density (OD$_{600}$) of 0.5 and 2 mL culture aliquots were vacuum-filtered on a 0.45 µm pore size filter (HVLP02500, Merck Millipore). Filters were immediately transferred into 40:40:20 (v-%)

acetonitrile/methanol/water at -20°C for extraction. Extracts were centrifuged for 15 minutes at 13,000 rpm at -9 °C. Centrifuged extracts were mixed with ^{13}C-labeled internal standard and analyzed by LC-MS/MS, with an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies) as described previously (Guder et al., 2017). An Agilent 1290 Infinity II UHPLC system (Agilent Technologies) was used for liquid chromatography. Temperature of the column oven was 30 °C, and the injection volume was 3 µL. LC solvents A were water with 10 mM ammonium formate and 0.1% formic acid (v/v) (for acidic conditions); and water with 10 mM ammonium carbonate and 0.2% ammonium hydroxide (for basic conditions). LC solvents B were acetonitrile with 0.1% formic acid (v/v) for acidic conditions and acetonitrile without additive for basic conditions. LC columns were an Acquity BEH Amide (30 x 2.1 mm, 1.7 µm) for acidic conditions, and an iHILIC-Fusion(P) (50 x 2.1 mm, 5 µm) for basic conditions. The gradient for basic and acidic conditions was: 0 min 90% B; 1.3 min 40 % B; 1.5 min 40 % B; 1.7 min 90 % B; 2 min 90 % B. Absolute concentrations of amino acids in the ^{13}C-labeled internal standard were determined with authentic standards. Quantification of intracellular metabolite concentrations was based on the ratio of ^{12}C and ^{13}C peak heights, and a specific cell volume of 2 µL mg^{-1} was used to calculate the cell volume.

Proteomics

Shake flask cultivations on M9 glucose were performed as described above. Cells were grown to an optical density (OD$_{600}$) of 0.5 and 2 mL culture aliquots were transferred into 2 ml reaction tubes and washed two times with PBS buffer (0.14 mM NaCl, 2.7 mM KCL, 1.5 KH$_2$PO$_4$, 8.1 Na$_2$HPO$_4$). Cell pellets were resuspended in 300 µl lysis buffer containing 100 mM ammonium bicarbonate, 0.5 % sodium laroyl sarcosinate (SLS) and 5 mM Tris(2-carboxyethyl)phosphine (TCEP). Cells were lysed by 5 minutes incubation at 95 °C and ultra-sonication for 10 seconds (Vial Tweeter, Hielscher). Cells were again incubated for 30 minutes at 90 °C followed by alkylation with 10 mM iodoacetamide for 30 minutes at 25 °C. To clear the cell lysate, samples were centrifuged for 10 minutes at 15,000 rpm and the supernatant transferred into a new tube. Proteins in the cell lysates were digested with 1 µg trypsin (Promega) overnight at 30 °C. To remove the SLS by precipitation, trifluoroacetic acid (TFA) was added to a final concentration of 1.5 % and rested at room temperature for 10 minutes. Samples were centrifuged for 10 minutes at 10,000 rpm and the supernatant used for C18 purification. The peptide purification was performed using the C18 microspin columns according to the manufactors instructions (Harvard Apparatus). Eluted peptide solutions were dried and resuspended in 0.1 % TFA. The concentration of peptides in the samples was measured with a colorimetric peptide assay (Pierce™ Quantitative Colorimetric Peptide Assay, Thermo Fischer Scientific). Analysis of peptides was performed by liquid chromatography-mass spectrometry. Analysis of peptides was performed by liquid chromatography-mass spectrometry, carried out on a Q-Exactive Plus instrument connected to

an Ultimate 3000 RSLC nano with a Prowflow upgrade and a nanospray flex ion source (Thermo Scientific). Peptide separation was performed on a reverse-phase HPLC column (75 µm x 42 cm) packed in-house with C18 resin (2.4 µm, Dr. Maisch GmbH, Germany). The following separating gradient was used: 98 % solvent A (0.15% formic acid) and 2 % solvent B (99.85 acetonitrile, 0.15 % formic acid) to 25 % solvent B over 105 minutes and to 35 % solvent B for additional 35 minutes at a flow rate of 300 nl/min. The data acquisition mode was set to obtain one high resolution MS scan at a resolution of 70,000 full width at half maximum (at m/z 200) followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS attempts, the charged state screening modus was enabled to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 seconds. The ion accumulation time was set to 50 ms for MS and 50 ms at 17,500 resolution for MS/MS. The automatic gain control was set to 3x106 for MS survey scans and 1x105 for MS/MS scans. Label-free quantification (LFQ) of the data was performed using Progenesis QIP (Waters), and for MS/MS searches of aligned peptide features MASCOT (v2.5, Matrix Science) was used. The following search parameters were used: full tryptic search with two missed cleavage sites, 10ppm MS1 and 0.02 Da fragment ion tolerance. Carbamidomethylation (C) as fixed, oxidation (M) and deamidation (N,Q) as variable modification. Progenesis outputs were further processed with SafeQuant.

Kinetic Flux Profiling

Incorporation of ^{15}N label into amino acids was measured with a filter cultivation method (Link et al., 2013). Briefly, cells were cultured on M9 glucose medium, which contains unlabeled ammonium sulfate as sole nitrogen source. At mid-exponential phase when cells reached ODs between 0.4 and 0.6, 2 mL of the culture was vacuum-filtered, and cell-loaded filters were continuously perfused with M9 glucose medium containing labeled ammonium-^{15}N sulfate. Filters were repeatedly loaded and perfused with ^{15}N-medium for different lengths of time: 0, 30, 60, 120 and 180 seconds. Subsequently, filters were immediately transferred into 40:40:20 (v-%) acetonitrile/methanol/water kept at -20 °C. Extracts were centrifuged for 15 minutes at 13,000 r.p.m. at -9 °C and the supernatant was directly used for LC-MS/MS. For LC separation of tryptophan, proline, threonine and (iso)leucine a ZIC-pHILIC column (150 x 2.1 mm, 5 µm, Merck) was used, and an Acquity BEH Amide (100 x 2.1 mm, 1.7 µm, Waters) for LC separation of histidine and arginine. Buffers were as described for metabolite measurements and gradients were for Acquity BEH Amide: 0 min 90% B; 2.6 min 40 % B; 3 min 40 % B; 3.4 min 90 % B; 5 min 90 % B. For ZIC-pHILIC: 0 min 90% B; 4.5 min 40 % B; 5 min 40 % B; 6 min 90 % B; 8 min 90 % B. Transitions for all isotopologues per amino acid were measured by LC-MS/MS and the amount of each isotopologue was used to calculate the fraction of unlabeled amino acid F^U as:

$$F^U = \frac{M^0}{\sum_0^N M^{+i}} = \frac{Peak\ Area\ (unlabeled\ AA)}{Sum\ of\ Peak\ Area\ (all\ AA\ \text{isotopologues})}$$

Where M^0 is the amount of the unlabeled amino acid, M^{+1} is the amount of all isotopologues with one ^{15}N atom, etc. N is the number of ^{15}N atoms in the amino acid: arginine ($N = 4$ from 2x glutamate, 1x glutamine, 1x aspartate), tryptophan ($N = 2$ from 1x glutamine, 1x serine), histidine ($N = 3$ from ATP, 1x glutamate), threonine ($N = 1$ from glutamate), proline ($N = 1$ from glutamate), iso-/leucine ($N = 1$ from glutamate). Fluxes were estimated based on equations for kinetic flux profiling (Yuan et al., 2008), which considers the decay of the unlabeled fraction F^U:

$$F^U = \left[\frac{(1-a)(1-b)}{k_{pc} - k_{aa}}\right]\left[k_{pc}\ e^{-k_{aa}t} - k_{aa}\ e^{-k_{pc}t}\right] + [1 - (1-a)(1-b)]$$

The rate constant k_{aa} is the flux into the amino acid ($flux_{aa}$) divided by their absolute concentration: $k_{aa} = flux_{aa}\ /\ c_{aa}$. The rate constant k_{aa} was obtained by fitting the equation to the measured unlabeled fraction F^U. The rate constant k_{pc} describes labeling of upstream nitrogen precursor. Because amino acids like arginine receive ^{15}N label from several sources, the rate constant of precursor labeling k_{pc} was unknown. To account for this uncertainty the parameter k_{pc} was randomly sampled between boundaries of 0.8 min^{-1} and 14.2 min^{-1}, which are the highest and lowest first order rate constants measured for nitrogen assimilation in *E. coli* (Yuan et al., 2006). a and b consider amino acid production from degradation of protein and other macromolecules and they were estimable parameters within bounds of 0 and 0.2.

GFP-promoter Fusions

GFP reporter plasmids for detection of promotor activity of *argA*, *trpL*, *hisL* and *leuL* were obtained from a library of fluorescent transcriptional reporters for *E. coli* (Zaslaver et al., 2006). Since the original plasmids pUA66-P*hisL-gfp* and pUA66-P*leuL-gfp* lacked parts of the attenuator region, we modified the respective promotor resulting in the plasmids pP*hisL-gfp* and pP*leuL-gfp*. Therefore we amplified leader sequence including the rho-independent terminator of *hisL* and *leuL* from chromosomal DNA of *E. coli* MG1655 (P*hisL*: hisL_fwd_gfp ccgctcgaggctttcatcattgttgccg, hisL_rev_gfp ccgggatcccgcagaatatcaatcggc; P*leuL*: leuL_fwd_gfp ccgctcgagttgtccccttttttcctcg, leuL_rev_gfp ccgggatccgatggtttgcaccgattc). The resulting two single fragments were introduced into an empty pUA66 backbone with the restriction enzymes XhoI and BamHI. The threonine reporter plasmid which was not available in the library was constructed with the same strategy. The attenuator region of *thrL* was amplified with the primer pair thrA_fwd_gfp (ccgctcgagactgcaacgggcaatatg) and thrA_rev_gfp

(ccgggatcctcggcatcgctgatattg) and the single fragment was introduced into pUA66 (XhoI and BamHI) resulting in pP*thrL-gfp*.

Flow Cytometry

Activity of the *argA*, *trpL* and *thrL* promoter was assayed using plasmid-based GFP reporters that were described in the previous section. Strains for flow cytometry were cultivated in three independent shake flasks (100 ml) containing 10 ml M9 minimal medium (5 g L^{-1} glucose; 50 µg mL^{-1} kanamycin) as described in Strains and Culture. After reaching an OD between 0.5 and 0.8 cells were diluted 1:2000 in tethering buffer (10 mM KH_2PO_4, 100 µM EDTA, 1 µM L-methionine and 10 mM lactic acid, pH=7.0) and fluorescence was measured with BD LSRFortessa SORP cell analyser (BD Biosciences, Germany). 488-nm lasers, 600 long pass and a 520/30 band pass filters were used for detection of green fluorescence. Per sample, fluorescence of 10,000 single cells was measured. Before the measurements, cell aggregates were dispersed by vigorous mixing. BD FACSDiva software version 8.0 (BD Biosciences, NJ, USA) and FlowJo v10.4.1 (FlowJo LLC, Ashland, OR, USA) were used for analysis of the acquired data.

Purification and *In Vitro* Activity Assays of N-Acteylglutamate Synthase

E. coli BL21 cells harboring the overexpression vector pET28a(+)-*argA* respectively pET28a(+)-*argA*(H15Y) were cultivated at 37 °C (220 rpm) in 500 ml of LB medium (5 L shake flasks) containing 30 µg ml^{-1} kanamycine. When cells reached OD_{600} 0.6, the culture was shifted to 16 °C to cool down the cell broth. To induce protein expression, 10 µl of IPTG stock solution (final concentration is 10 µM) were added. The culture was incubated overnight at 16 °C (220 rpm). The cells were harvested by centrifugation at 6000 x *g* for 10 minutes at 4 °C. The supernatant was completely removed. The cell pellet was resuspended in Lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazol) (2-5 ml per gram wet weight). 50 µl protease inhibitor cocktail and 5 mg of DNAse I powder were added. Lysis of cells was performed by french press (1100 bar). The lysate was centrifuged at 4,000 x *g* for 45 minutes at 4 °C to pellet the cellular debris. The supernatant was filtered using a 0.2-µm-pore-size syringe filter and transferred into a new collection tube. Purification was performed with columns purchased from GE Healthcare Life Science (His GraviTrap; 11-0033-99). 10 ml of equilibration buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM Imidazol) was added to the column. As soon as equilibration buffer flowed through, up to 35 ml of filtered supernatant were added to the column. The column was washed twice with 10 ml washing buffer (same as equilibration buffer). Elution of the protein was performed 3 times with 3 ml elution buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM Imidazol). Protein concentration of all fractions was determined (660 nm protein assay, life technologies PIERCE^{TM} #22660). Activity of

purified N-acetylglutamate-synthase (ArgA) as well as for the feedback-resistant version ArgA (H15Y) was assayed in 30 mM TRIS buffer (with 40 mM L-glutamate, 0.65 mM N-acetyl-CoA and 10 mM $MgCl_2$). To start the enzymatic reaction 10 µL of enzyme stock solution (0.15 mg/ml) was transferred to 90 µL assay buffer and mixed by pipetting up and down. To stop the reaction, 10 µL were transferred into 40 µl of 50:50 (v-%) acetonitrile/methanol at -20°C. Samples were taken every minute in a total time interval of 8 minutes. The reaction product N-acetylglutamate was measured by LC-MS and calibrated with authentic standards.

Kinetic Model

The stoichiometry of the model is shown in Figure 4A. Mass balancing results in the system of ordinary differential equations (ODEs), F, that is a temporal function of the state variables x and the kinetic parameters p:

$$F(x,p) = \frac{dx}{dt} = \begin{cases} \dfrac{dm_1}{dt} = r_1 - r_2 \\ \dfrac{dm_2}{dt} = r_2 - \alpha\,\mu \\ \dfrac{de_1}{dt} = \beta_1 - e_1\,\mu \\ \dfrac{de_2}{dt} = \beta_2 - e_2\,\mu \end{cases}$$

(Equation 1)

The five reactions (r_1, r_2, β_1, β_2, μ) are described by the following kinetic equations:

Reaction 1 is feedback inhibited by m_2 according to normal inhibition kinetics:

$$r_1 = k_{cat,1}\,e_1\,\frac{K_1}{K_1 + m_2}$$

(Equation 2)

In the model without allosteric regulation the equation reduces to:

$$r_1 = k_{cat,1}\,e_1$$

(Equation 3)

Reaction 2 follows Michaelis-Menten kinetics:

$$r_2 = k_{cat,2} \, e_2 \, \frac{m_1}{m_1 + K_m} \qquad \text{(Equation 4)}$$

Expression rates of enzyme 1 and enzyme 2 follow inhibition kinetics

$$\beta_1 = \beta_{1,\text{max}} \, \frac{K_2}{K_2 + m_2} \qquad \text{(Equation 5)}$$

$$\beta_2 = \beta_{2,\text{max}} \, \frac{K_2}{K_2 + m_2} \qquad \text{(Equation 6)}$$

The growth rate depends on availability of the end-product m2:

$$\mu = \mu_{max} \, \frac{m_2}{m_2 + K_\mu} \qquad \text{(Equation 7)}$$

Dilution of metabolites by growth was not considered, due to large difference in time scales between growth dilution and metabolic flux. Dilution of enzymes by growth is included in equation 1, because the time scales of enzyme synthesis and growth dilution are closer.

Together, the kinetic equations include eight kinetic parameters k_{cat1}, k_{cat2}, $\beta_{1,max}$, $\beta_{2,max}$, K_1, K_2, K_m and α. The physiological ranges for these parameters were derived from literature values. The boundaries of enzyme turnover number ($k_{cat,1}$ and $k_{cat,2}$) are based on *in vitro* measured k_{cat} values of enzymes in amino acid biosynthesis (Table S3) and have values between 930 min^{-1} and 4140 min^{-1}. The maximal enzyme expression rates ($\beta_{1,max}$ and $\beta_{2,max}$) are defined by the translation rate of ribosomes according to equation 8. The equation considers the following parameters that were derived from the Bionumbers Database (Milo et al., 2010): average translation rate (r_T = 8.4 amino acids s^{-1}), the median and abundance weighted protein length (L = 209 amino acids), the fraction of active ribosomes (f_R = 0.8), the cellular volume ($V_{c,0.6}$ = 3 x 10^{-15} L) at a growth rate of μ = 0.6 h^{-1}, the Avogadro number (N_A = 6.02 x 10^{23} mol^{-1}), the amount of ribosomes per cell at that growth rate ($R_{0.6}$ = 8000 ribosomes $cell^{-1}$) and the fraction of ribosomes (p) that synthesize the enzyme:

$$\beta_{k,max} = \frac{r_t \cdot R_{0.6} \cdot f_R}{L \cdot N_A \cdot V_c} \cdot p \qquad \text{(Equation 8)}$$

The limits of $\beta_{k,max}$ are then derived by varying the fraction of ribosomes (p) that synthesize the enzymes in the pathway. According to the literature the maximal number for a single amino acid biosynthesis enzyme in *E. coli* is 7% (Li et al., 2014), therefore we set the boundaries to 1% and 10% (p = 0.01 - 0.1). The parameter limits for the Ki and Km values were set to 0.01 mM and 1 mM. The amino acid requirement (α = 86.6 mM) was a fixed parameter based on the average amino acid requirement of an *E. coli* cell (Table S4). We assumed that the amino acid limits the growth rate reaction only at very low concentrations. This reflects the low Km values of tRNA ligases. Therefor we fixed K_μ at a low value of 10^{-5} mM and set μ_{max} to the measured growth rate on glucose of 0.6 h^{-1}.

Steady State and Robustness Analysis

For steady state analysis a parameter set was randomly sampled from the intervals given above. With a specific parameter set the steady state concentrations of e_1, e_2, m_1 and m_2 were calculated numerically for each of the two models (complete model and single feedback model). Starting values of the numerical solver were 0.01 mM for m_1 and m_2, and 10^{-5} mM for e_1 and e_2. The convergence criterion was defined as $<10^{-8}$ change in all variables. To test stability of the steady state we calculated eigenvalues of the Jacobian matrix, and tested if all eigenvalues are negative ($\lambda < -10^{-5}$). This procedure was repeated until 5000 steady states (with different parameter sets) were achieved. Note that both models share the same parameter sets and reach the same steady state flux. In order to estimate robustness of the model against perturbations of the maximal enzyme expression rate $\beta_{2,max}$, we used a numerical parameter continuation method (Lee et al., 2014). The method is based on finding a connected path of steady state concentrations (x_{ss}: steady state concentration vector containing $e_{1,ss}$, $e_{2,ss}$, $m_{1,ss}$, $m_{2,ss}$), as a parameter, p, is varied. As the system is in steady state it follows that:

$$\frac{dx}{dt} = F(x_{SS}, p) = 0 \qquad \text{(Equation 9)}$$

The derivative of $F(x_{SS}, p)$ with respect to the parameters is also zero:

$$\frac{dF(x_{SS}, p)}{dp} = \frac{\delta F}{\delta x_{SS}} \cdot \frac{dx_{SS}}{dp} + \frac{\delta F}{\delta p} = 0 \qquad \text{(Equation 10)}$$

After rearranging Equation 10, Equation 11 is obtained:

$$\frac{dx_{SS}}{dp} = -\left(\frac{\delta F}{\delta x_{SS}}\right)^{-1} \cdot \frac{\delta F}{\delta p} \qquad \text{(Equation 11)}$$

which describes the changes in the steady-state concentrations as a kinetic parameter is varied iteratively. The iteration stops when one of the following three stability criteria is no longer fulfilled.

1st criterion: all real parts of the eigenvalues of the system's Jacobian need to be negative. This implies stability of a steady state. Furthermore, in equation 11 the inverse of the Jacobian Matrix ($\delta F / \delta x_{SS}$) is required. The inversion is only possible as long as the matrix is regular. Once an eigenvalue reaches zero, the Jacobian becomes singular and matrix inversion is no longer possible. This bifurcation point defines the boundary between the stable and unstable parameter space. In other words: after this point is passed, the system no longer returns to a stable steady state. By checking the eigenvalues of the Jacobian at each step, we make sure that the iteration is terminated when one eigenvalue becomes bigger than $\lambda = -10^{-5}$. 2nd criterion: all variables are required to be positive. 3rd criterion: a model is considered unstable when a certain time limit (t > 1 s) is exceeded, which can be the case when numerical errors occur during the numerical integration process. The maximum theoretical enzyme amount in the model was calculated as:

$$0 = \beta_{i,max} - e_{i,max}\,\mu \qquad \text{(Equation 12)}$$

After rearranging equation 12 and substituting the upper parameter bound of the maximum protein translation rate ($\beta_{i,max}^{ub}$), the maximum theoretical enzyme amount of each enzyme is:

$$e_{i,max} = \frac{\beta_{i,max}^{ub}}{\mu} = \frac{8.5 \cdot 10^{-4}\ mM\ min^{-1}}{0.01\ min^{-1}} = 0.085\ mM \qquad \text{(Equation 13)}$$

Considering that the model includes two enzymes, the maximum amount of total enzyme is 0.17 mM, which was defined as the maximal enzyme level (100%).

Quantification and Statistical Analysis

Statistical analysis was done with MATLAB. The statistical details of each experiment can be found in the respective figure caption. For proteomics and metabolomics n represents the number of independent shake flask cultures. In growth assays, n represents the number of independent microtiter plate cultures. For *in vitro* assays, n represents the number of independent reaction vessels.

Software

All codes for model analysis are available in the Github repository: https://github.com/nfarke/Sander_et_al.

References

Alam, M.T., Olin-Sandoval, V., Stincone, A., Keller, M.A., Zelezniak, A., Luisi, B.F., and Ralser, M. (2017). The self-inhibitory nature of metabolic networks and its alleviation through compartmentalization. Nat. Commun. *8*, 16018.

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

Caligiuri, M.G., and Bauerle, R. (1991). Subunit communication in the anthranilate synthase complex from *Salmonella typhimurium*. Science *252*, 1845–1848.

Chandra, F.A., Buzi, G., and Doyle, J.C. (2011). Glycolytic Oscillations and Limits on Robust Efficiency. Science *333*, 187.

Cho, B.-K., Barrett, C.L., Knight, E.M., Park, Y.S., and Palsson, B.Ø. (2008). Genome-scale reconstruction of the Lrp regulatory network in *Escherichia coli*. Proc. Natl. Acad. Sci. *105*, 19462–19467.

Cho, B.-K., Federowicz, S., Park, Y.-S., Zengler, K., and Palsson, B.Ø. (2012). Deciphering the transcriptional regulatory logic of amino acid metabolism. Nat. Chem. Biol. *8*, 65.

Christodoulou, D., Link, H., Fuhrer, T., Kochanowski, K., Gerosa, L., and Sauer, U. (2018). Reserve Flux Capacity in the Pentose Phosphate Pathway Enables *Escherichia coli*'s Rapid Response to Oxidative Stress. Cell Syst. *6*, 569-578.e7.

Chubukov, V., Zuleta, I.A., and Li, H. (2012). Regulatory architecture determines optimal regulation of gene expression in metabolic pathways. Proc. Natl. Acad. Sci. *109*, 5127–5132.

Chubukov, V., Uhr, M., Le Chat, L., Kleijn, R.J., Jules, M., Link, H., Aymerich, S., Stelling, J., and Sauer, U. (2013). Transcriptional regulation is insufficient to explain substrate-induced flux changes *in Bacillus subtilis.* Mol. Syst. Biol. *9*, 709.

Chubukov, V., Gerosa, L., Kochanowski, K., and Sauer, U. (2014). Coordination of microbial metabolism. Nat. Rev. Microbiol. *12*, 327–340.

Csonka, L.N., Gelvin, S.B., Goodner, B.W., Orser, C.S., Siemieniak, D., and Slightom, J.L. (1988). Nucleotide sequence of a mutation in the proB gene of *Escherichia coli* that confers proline overproduction and enhanced tolerance to osmotic stress. Gene *64*, 199–205.

Daran-Lapujade, P., Rossell, S., van Gulik, W.M., Luttik, M.A.H., de Groot, M.J.L., Slijper, M., Heck, A.J.R., Daran, J.-M., de Winde, J.H., Westerhoff, H.V., et al. (2007). The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. Proc. Natl. Acad. Sci. *104*, 15753–15758.

Davidi, D., and Milo, R. (2017). Lessons on enzyme kinetics from quantitative proteomics. Curr. Opin. Biotechnol. *46*, 81–89.

Dekel, E., and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. Nature *436*, 588–592.

Doroshenko, V.G., Lobanov, A.O., and Fedorina, E.A. (2013). The directed modification of *Escherichia coli* MG1655 to obtain histidine-producing mutants. Appl. Biochem. Microbiol. *49*, 130–135.

Fendt, S.-M., Buescher, J.M., Rudroff, F., Picotti, P., Zamboni, N., and Sauer, U. (2010). Tradeoff between enzyme and metabolite efficiency maintains metabolic homeostasis upon perturbations in enzyme capacity. Mol. Syst. Biol. *6*, 356.

Fuhrer, T., Zampieri, M., Sévin, D.C., Sauer, U., and Zamboni, N. (2017). Genomewide landscape of gene–metabolome associations in *Escherichia coli*. Mol. Syst. Biol. *13*, 907.

Gama-Castro, S., Salgado, H., Santos-Zavaleta, A., Ledezma-Tejeida, D., Muñiz-Rascado, L., García-Sotelo, J.S., Alquicira-Hernández, K., Martínez-Flores, I., Pannier, L., Castro-Mondragón, J.A., et al. (2016). RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. Nucleic Acids Res. *44*, D133-143.

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

Glascock, C.B., and Weickert, M.J. (1998). Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in *Escherichia coli*. Gene *223*, 221–231.

Goyal, S., Yuan, J., Chen, T., Rabinowitz, J.D., and Wingreen, N.S. (2010). Achieving Optimal Growth through Product Feedback Inhibition in Metabolism. PLoS Comput. Biol. *6*, e1000802.

Grimbs, S., Selbig, J., Bulik, S., Holzhütter, H.-G., and Steuer, R. (2007). The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. Mol. Syst. Biol. *3*, 146.

Guder, J.C., Schramm, T., Sander, T., and Link, H. (2017). Time-Optimized Isotope Ratio LC-MS/MS for High-Throughput Quantification of Primary Metabolites. Anal. Chem. *89*, 1624–1631.

Gusyatiner, M.M., Ivanovskaya, L.V., Kozlov, Y.I., Lunts, M.G., and Voroshilova, E.B. (2005). DNA coding for mutant isopropylmalate synthase, l-leucine-producing microorganism and method for producing l-leucine.

Hackett, S.R., Zanotelli, V.R.T., Xu, W., Goya, J., Park, J.O., Perlman, D.H., Gibney, P.A., Botstein, D., Storey, J.D., and Rabinowitz, J.D. (2016). Systems-level analysis of mechanisms regulating yeast metabolic flux. Science *354*.

Hartl, J., Kiefer, P., Meyer, F., and Vorholt, J.A. (2017). Longevity of major coenzymes allows minimal de novo synthesis in microorganisms. Nat. Microbiol. *2*, 17073.

Hirasawa, T., and Shimizu, H. (2016). Recent advances in amino acid production by microbial cells. Curr. Opin. Biotechnol. *42*, 133–146.

Hofmeyr, J.-H.S., and Cornish-Bowden, A. (2000). Regulating the cellular economy of supply and demand. FEBS Lett. *476*, 47–51.

Kacser, H., and Burns, J.A. (1973). The control of flux. Symp. Soc. Exp. Biol. *27*, 65–104.

ter Kuile, B.H., and Westerhoff, H.V. (2001). Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. FEBS Lett. *500*, 169–171.

LaRossa, R.A., Dyk, T.K.V., and Smulski, D.R. (1987). Toxic accumulation of alpha-ketobutyrate caused by inhibition of the branched-chain amino acid biosynthetic enzyme acetolactate synthase in *Salmonella typhimurium*. J. Bacteriol. *169*, 1372–1378.

Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat. Protoc. *8*, 2180–2196.

Lee, J.-H., Lee, D.-E., Lee, B.-U., and Kim, H.-S. (2003). Global analyses of transcriptomes and proteomes of a parent strain and an L-threonine-overproducing mutant strain. J. Bacteriol. *185*, 5442–5451.

Lee, Y., Lafontaine Rivera, J.G., and Liao, J.C. (2014). Ensemble Modeling for Robustness Analysis in engineering non-native metabolic pathways. Metab. Eng. *25*, 63–71.

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

Link, H., Kochanowski, K., and Sauer, U. (2013). Systematic identification of allosteric protein-metabolite interactions that control enzyme activity *in vivo*. Nat. Biotechnol. *31*, 357–361.

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

Monk, J.M., Lloyd, C.J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., et al. (2017). iML1515, a knowledgebase that computes *Escherichia coli* traits. Nat. Biotechnol. *35*, 904–908.

Mori, M., Schink, S., Erickson, D.W., Gerland, U., and Hwa, T. (2017). Quantifying the benefit of a proteome reserve in fluctuating environments. Nat. Commun. *8*.

Mülleder, M., Calvani, E., Alam, M.T., Wang, R.K., Eckerstorfer, F., Zelezniak, A., and Ralser, M. (2016). Functional Metabolomics Describes the Yeast Biosynthetic Regulome. Cell *167*, 553-565.e12.

O'Brien, E.J., Utrilla, J., and Palsson, B.O. (2016). Quantification and Classification of *E. coli* Proteome Utilization and Unused Protein Costs across Environments. PLoS Comput. Biol. *12*, e1004998.

Rajagopal, B.S., DePonte, J., Tuchman, M., and Malamy, M.H. (1998). Use of inducible feedback-resistant N-acetylglutamate synthetase (argA) genes for enhanced arginine biosynthesis by genetically engineered *Escherichia coli* K-12 strains. Appl. Environ. Microbiol. *64*, 1805–1811.

Reaves, M.L., Young, B.D., Hosios, A.M., Xu, Y.-F., and Rabinowitz, J.D. (2013). Pyrimidine homeostasis is accomplished by directed overflow metabolism. Nature *500*, 237–241.

Reisch, C.R., and Prather, K.L.J. (2015). The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in *Escherichia coli*. Sci. Rep. *5*, 15096.

Reznik, E., Christodoulou, D., Goldford, J.E., Briars, E., Sauer, U., Segrè, D., and Noor, E. (2017). Genome-Scale Architecture of Small Molecule Regulatory Networks and the Fundamental Trade-Off between Regulation and Enzymatic Activity. Cell Rep. *20*, 2666–2677.

Schmidt, A., Kochanowski, K., Vedelaar, S., Ahrné, E., Volkmer, B., Callipo, L., Knoops, K., Bauer, M., Aebersold, R., and Heinemann, M. (2016). The quantitative and condition-dependent *Escherichia coli* proteome. Nat. Biotechnol. *34*, 104–110.

Schomburg, I., Chang, A., Placzek, S., Söhngen, C., Rother, M., Lang, M., Munaretto, C., Ulas, S., Stelzer, M., Grote, A., et al. (2013). BRENDA in 2013: integrated reactions, kinetic data, enzyme function data, improved disease classification: new options and contents in BRENDA. Nucleic Acids Res. *41*, D764-72.

Schuster, S., and Heinrich, R. (1987). Time hierarchy in enzymatic reaction chains resulting from optimality principles. J. Theor. Biol. *129*, 189–209.

Scott, M., Gunderson, C.W., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010). Interdependence of cell growth and gene expression: origins and consequences. Science *330*, 1099–1102.

Umbarger, H.E. (1956). Evidence for a Negative-Feedback Mechanism in the Biosynthesis of Isoleucine. Science *123*, 848–848.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. Nature *289*, 751–758.

You, C., Okano, H., Hui, S., Zhang, Z., Kim, M., Gunderson, C.W., Wang, Y.-P., Lenz, P., Yan, D., and Hwa, T. (2013). Coordination of bacterial proteome with metabolism by cyclic AMP signalling. Nature *500*, 301–306.

Yuan, J., Fowler, W.U., Kimball, E., Lu, W., and Rabinowitz, J.D. (2006). Kinetic flux profiling of nitrogen assimilation in *Escherichia coli*. Nat. Chem. Biol. *2*, 529–530.

Yuan, J., Bennett, B.D., and Rabinowitz, J.D. (2008). Kinetic flux profiling for quantitation of cellular metabolic fluxes. Nat. Protoc. *3*, 1328–1340.

Zaslaver, A., Mayo, A.E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M.G., and Alon, U. (2004). Just-in-time transcription program in metabolic pathways. Nat. Genet. *36*, 486–491.

Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M.G., and Alon, U. (2006). A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat. Methods *3*, 623–628.

Zelezniak, A., Vowinckel, J., Capuano, F., Messner, C.B., Demichev, V., Polowsky, N., Mülleder, M., Kamrad, S., Klaus, B., Keller, M.A., et al. (2018). Machine Learning Predicts the Yeast Metabolome from the Quantitative Proteome of Kinase Knockouts. Cell Syst. *7*, 269-283.e6.
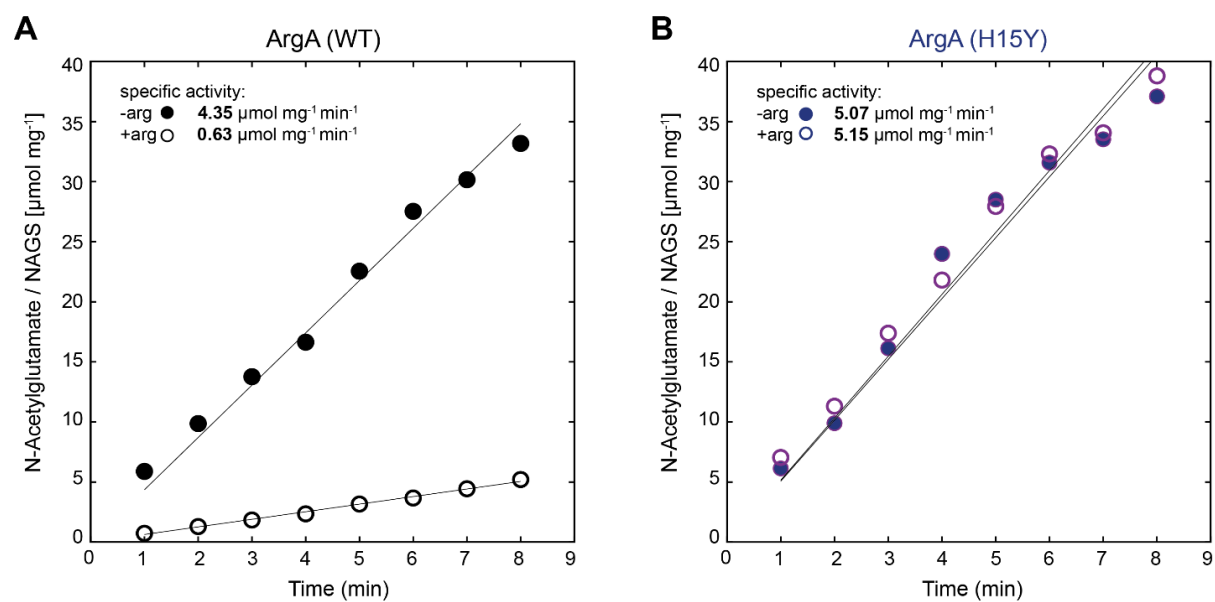
Supplementary Material

Figure S1. Related to Figure 1; *In vitro* kinetics of N-acetylglutamate-synthase (NAGS) from *E. coli* (ArgA) in the **A** native and **B** the allosteric feedback resistant version ArgA (H15Y). Dots represent means from n=2 independent assays (filled = no arginine; empty = 1 mM arginine). Activity of His-tagged purified enzymes was assayed in 30 mM TRIS buffer (40 mM L-glutamate, 0.65 mM Acetyl-CoA and 10 mM MgCl$_2$). For sampling 10 µL of reaction solution was transferred into 40 µL of 50:50 (v-%) acetonitrile/methanol at -20°C. The reaction product N-acteylglutamate was measured by LC-MS/MS. Specific activity in [µmol mg^{-1} min^{-1}] was calculated from linear regression through the 8 time points.

Figure S2. Related to Figure 1; Relative concentrations of 110 intracellular metabolites in wild-type *E. coli* and seven dysregulated mutants (n = 3).
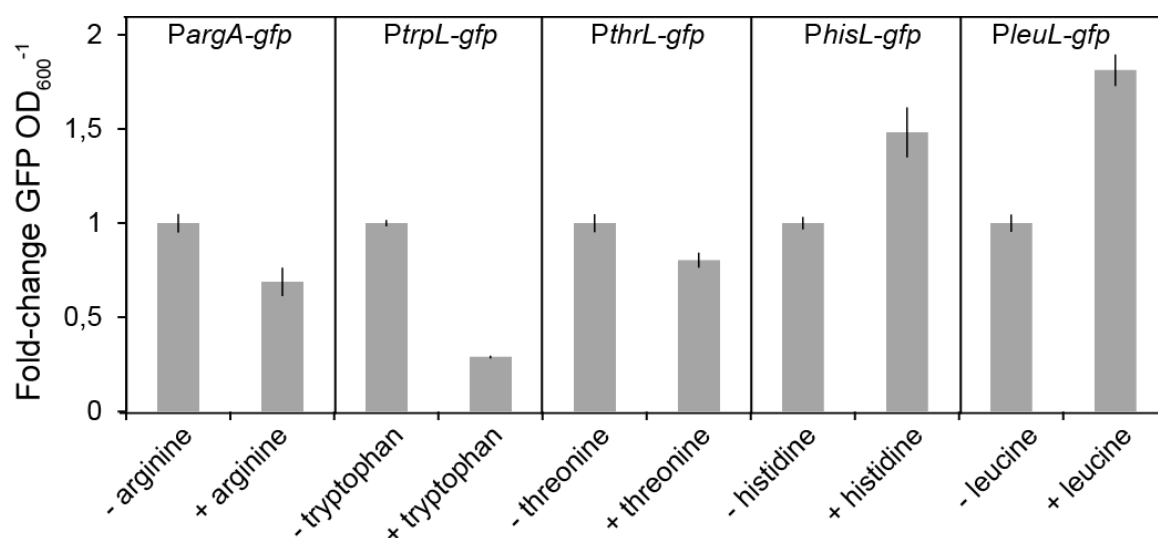
Figure S3. Related to Figure 2; GFP expression of promotor fusions P*argA-gfp*, P*trpL-gfp*, P*thrL-gfp*, P*hisL-gfp* and P*leuL-gfp* in *E. coli* wild-type with and without addition of external amino acids. Bar plots show fold-changes of GFP per OD_{600} relative to the condition without external amino acids (n=3). Cells were grown in M9 minimal medium (5 g L^{-1} glucose) and GFP expression was measured in mid-exponential phase at OD_{600} ~0.5 with a plate reader. Amino acids were supplemented to a final concentration of 2 mM.
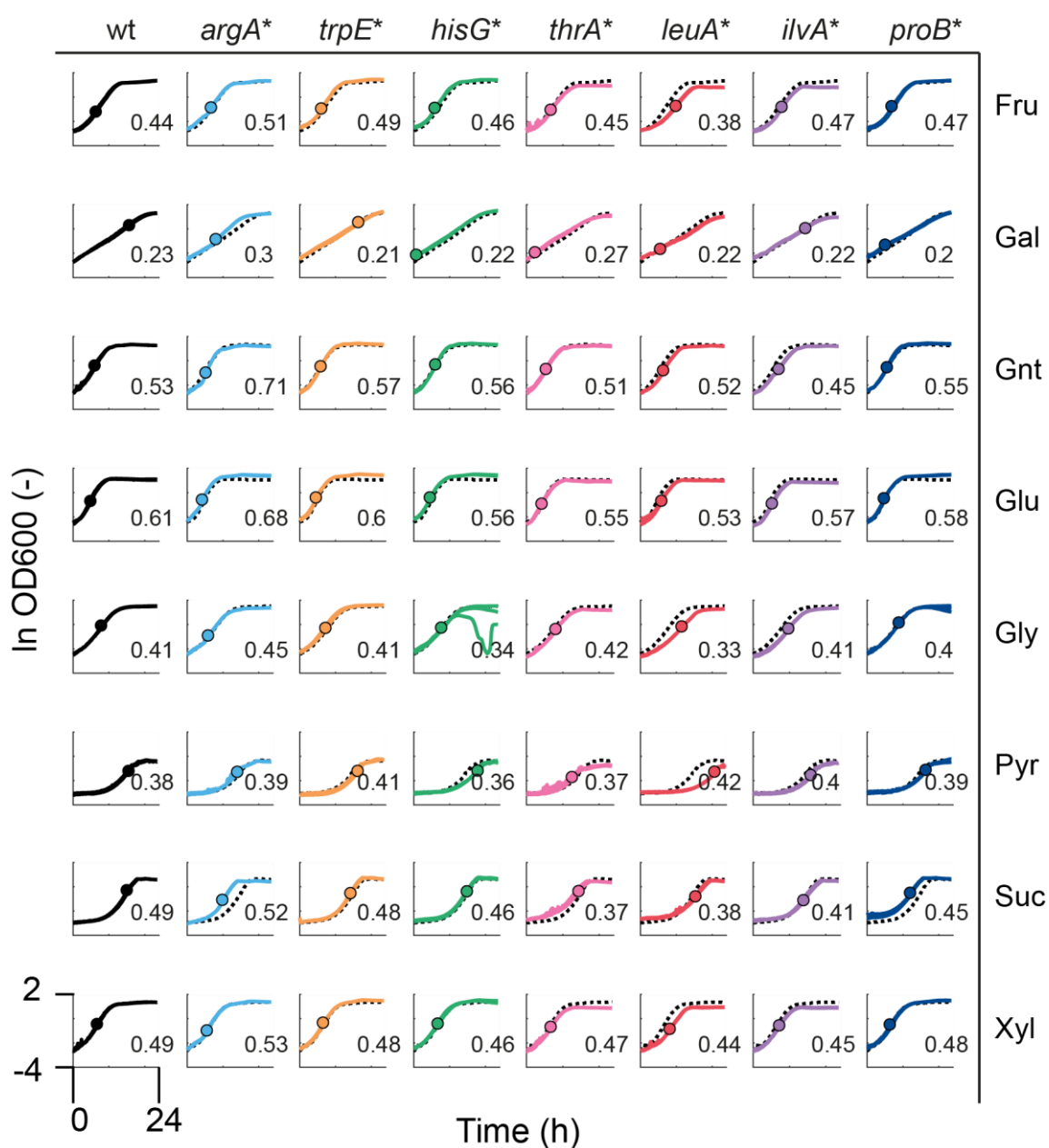
Figure S4. Related to Figure 3; Growth of wild-type *E. coli* and 7 mutants (see also Figure 1A) on fructose (Fru), galactose (Gal), gluconate (Gnt), glucose (Glu), glycerol (Gly), pyruvate (Pyr), succinate (Suc), and xylose (Xyl). Shown are three cultivations in microtiter plates. The dashed line is the mean of the wild-type in the particular condition (n = 3). Numbers are the maximal growth rates in h^{-1}, which is reached at the time indicated by dots. All x-axes range from 0 to 24 hours. All y-axes range from -4 to 2 ($\ln OD_{600}$).
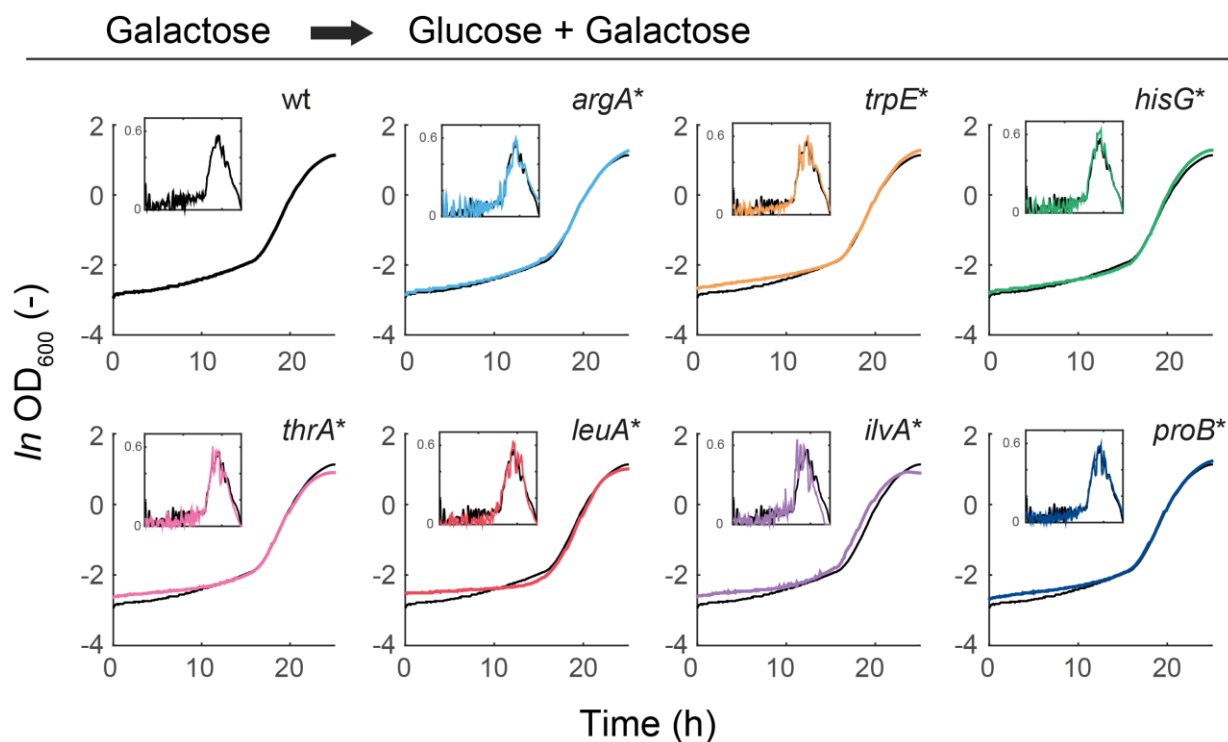
Figure S5. Related to Figure 3; Growth of wild-type *E. coli* and the seven dysregulated mutants in shifts from galactose to glucose. For up-shifts from galactose to glucose, cells were grown in M9 minimal medium with 5 g L^{-1} galactose and glucose was added to a final concentration of 5 g L^{-1} at an OD of 0.1. Shown are means of n = 3 cultures. Inserts show the growth rate during the same time period. Growth rates were estimated by linear regression over a moving 30-minute window. The same wild-type growth curve is shown in each graph in black as a reference.

Figure S6. Related to Figure 4; Intermediates in dysregulated pathways measured by LC-MS in wild-type *E. coli* and seven dysregulated mutants (n = 3).
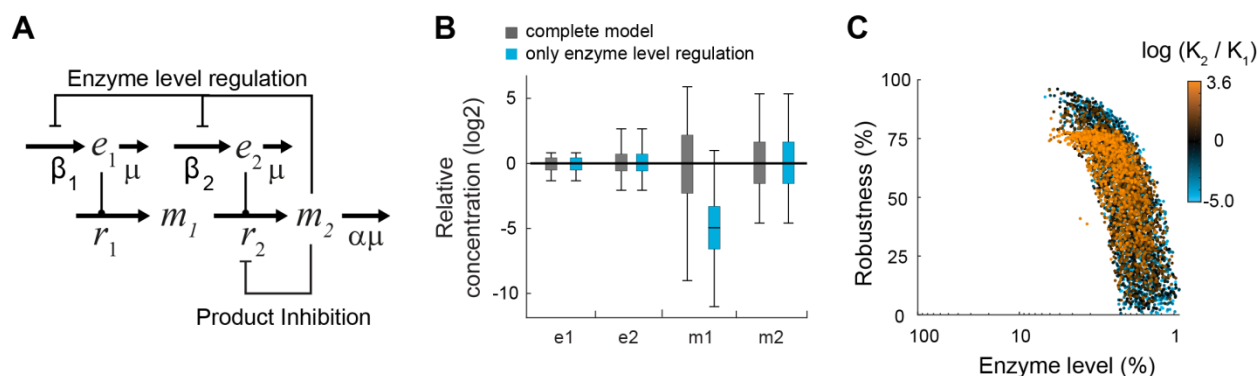
Figure S7. Related to Figure 4;

(A) Model with product inhibition, instead of allosteric feedback inhibition. Metabolite 2 inhibits reaction 2 by competitive product inhibition, which was modelled using the following equation:
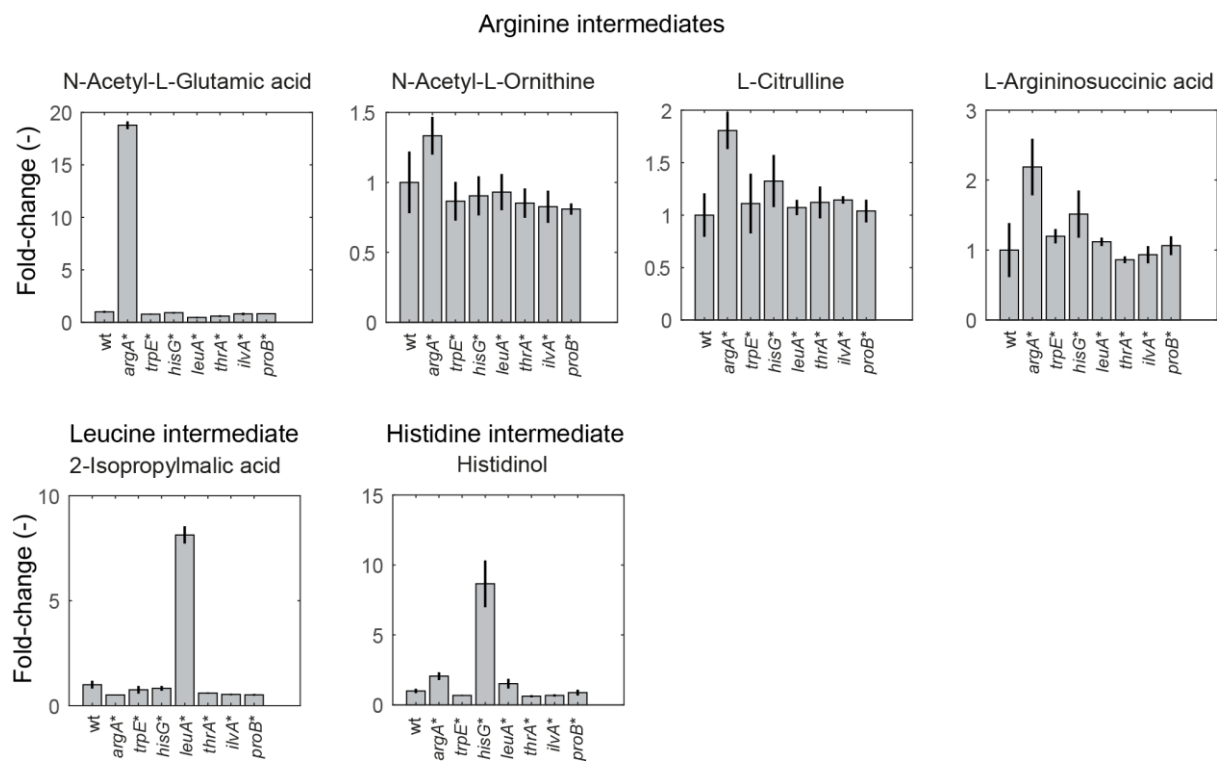
$$r_2 = k_{cat,2} \cdot e_2 \cdot \frac{m1}{m1 + Km \cdot \left(1 + \frac{m2}{K1}\right)}$$

(B) Steady state concentrations of e_1, e_2 m_1 and m_2 calculated with 5000 simulations for the complete model (grey), and the model with only enzyme level regulation (blue). Boxes contain 50% and whiskers 99% of the simulated concentrations. All concentrations are normalized to the median concentrations of the complete model.

(C) Enzyme levels (sum of e_1 and e_2) and robustness against perturbations of $\beta_{2,max}$ for 5000 simulations of the complete model (dots). The color of each dot shows the ratio of inhibition constants for allosteric feedback inhibition (K1) and enzyme level regulation (K2) in the respective model. Robustness corresponds to the percentage downregulation of $\beta_{2,max}$ that was tolerated by each model. 100% enzyme abundance corresponds to the maximum theoretical enzyme concentration in the model.

Table S1. Related to Figure 1; Mutations in allosteric enzymes that were investigated in this study.

Pathway	Gene	Enzyme	Mutation	Reference
L-arginine biosynthesis	*argA*	N-acetylglutamate synthase	H15Y	Rajagopal et al., 1998
L-isoleucine biosynthesis	*ilvA*	Threonine deaminase	L447F	LaRossa et al., 1987
L-histidine biosynthesis	*hisG*	ATP phosphoribosyl transferase	E271K	Doroshenko et al., 2013
L-leucine biosynthesis	*leuA*	2-isopropylmalate synthase	G462D	Gusyatiner et al., 2002
L-proline biosynthesis	*proB*	Glutamate-5-kinase	D107N	Csonka et al., 1988
L-threonine biosynthesis	*thrA*	Aspartate kinase	S345F	Lee et al., 2003
L-tryptophan biosynthesis	*trpE*	Anthranilate synthase	S40F	Caligiuri and Bauerle, 1991

Table S2. Related to Figure 1; Oligonucleotides for recombineering

Gene	Oligonucleotides for recombineering (5'-3')	Protospacer sequence (5'-3')
argA	GTGGTAAAGGAACGTAAAACCGAGTTGGTCGAGGGAT TCCGCTATTCAGTTCCCTATATCAATACCCACCGGGGAA	GGTCGAGGGATT CCGCCATT
ilvA	GGAATCACCGGGCGCGTTCCTGCGCTTTCTCAACACG CTGGGTACGTACTGGAACATTTCTTTGTTCCACTATCG	CAACACGCTGG GTACGTACT
hisG	GTCAGCAGCAAAACCCTGTTCTGGGAAACTATGGAAA AACTGAAAGCGCTGGGGCCAGTTCAATTCTGGTCCTG	TGGAAAAACTGA AAGCGCTG
leuA	CTGGTGAAATACAGCCTGACCGCCAAAGGACACGGTA AAGATGCGCTGGATCAGGTGGATATCGTCGCTAACTAC	CGGTAAAGATGC GCTGGGTC
proB	ACCCGTGCTAATATGGAAGACCGTGAACGCTTCCTGAACGCTCGCGACAC CCTGCGAGCGTTGCTCGATAACAATATC	CGACACCCTGCG AGCGTTGC
thrA	GCGCGCGTCTTTGCAGCGATGTCACGCGCCCGTATTT TCGTGGTGCTGATTACGCAATCATCTTCCGAATACAGC	TGGTGCTGATTA CGCAATCA
trpE	CTTATCGCGACAATCCCACTGCGCTTTTTCACCAGTTGTGTGGGGATCGTC CGGCAACGCTGCTGCTGGAATTCGCAGAT	CGCTTTTTCACC AGTTGTGT

Table S3. Related to Figure 4; Literature k$_{cat}$ values for enzymes in amino acid biosynthesis. The values were collected from the BRENDA database, and from Davidi and Milo, 2017. - indicates that no value could be found in both sources. The 25th and 75th quartiles of these k$_{cat}$ values are 930 min^{-1} and 4140 min^{-1}, respectively.

Name	k$_{cat}$, s^{-1}	Name	k$_{cat}$, s^{-1}	Name	k$_{cat}$, s^{-1}
argA	654.00	cysK	378.50	ilvN	40.00
argB	-	cysM	24.00	leuA	-
argC	-	cysN	-	leuB	69.00
argD	-	cysQ	11.00	leuC	-
argE	1800.00	dadX	33.66	leuD	-
argF	-	dapA	104.00	lysA	33.00
argG	-	dapB	382.00	lysC	22.13
argH	-	dapD	36.00	metA	22.00
argI	-	dapE	-	metB	121.00
aroA	32.00	dapF	84.00	metC	34.10
aroB	14.00	gdhA	37.00	metE	3.50
aroC	39.00	glnA	33.00	metH	-
aroD	75.00	gltB	-	metL	-
aroE	237.00	gltD	-	pheA	32.00
aroF	-	glyA	10.00	proA	10.00
aroG	4.20	hisA	7.20	proB	53.00
aroH	-	hisB	-	proC	717.00
aroK	-	hisC	-	prs	-
aroL	-	hisD	12.00	serA	29.00
asd	-	hisF	-	serB	-
asnA	-	hisG	-	serC	1.80
asnB	4.50	hisH	-	thrA	-
aspC	-	hisI	-	thrB	17.00
avtA	-	ilvA	-	thrC	-
cysC	50.00	ilvB	38.50	trpA	-
cysD	-	ilvC	0.30	trpB	-
cysE	772.00	ilvD	69.00	trpC	18.77
cysH	-	ilvE	-	trpE	-
cysI	47.00	ilvH	-	tyrA	71.00
cysJ	-	ilvI	-	tyrB	-

Table S4. Related to Figure 4; Amino acid requirements of *E. coli* (Monk et al., 2017). The mean of 86.6 mM was used as parameter α in the model.

Amino Acid	Coefficients, mmol g_{dw}^{-1}	alpha, mM
ala-L	0.499	166.4
arg-L	0.287	95.8
asn-L	0.234	78.1
asp-L	0.234	78.1
cys-L	0.089	29.7
gln-L	0.256	85.2
glu-L	0.256	85.2
gly	0.595	198.4
his-L	0.092	30.7
ile-L	0.282	94.1
leu-L	0.438	145.9
lys-L	0.333	111.1
met-L	0.149	49.8
phe-L	0.180	60.0
pro-L	0.215	71.6
ser-L	0.210	69.9
thr-L	0.247	82.2
trp-L	0.055	18.4
tyr-L	0.134	44.7
val-L	0.411	137.1
Mean	**0.260**	**86.6**

Table S5. Related to Figure 4; Inhibition constants of allosteric enzymes (K$_i$-value), transcriptional attenuation (tRNA-ligase K$_m$-value) and metabolite-transcription factor interactions (K$_d$-value). Values were obtained from EcoCyc (Keseler et al., 2017), Brenda (Schomburg et al., 2002) or RegulonDB (Gama-Castro et al., 2016). When more than one value was available, an upper and a lower bound are given. The grey background indicates the seven pathways that were investigated during this work. The Ki of ArgA was measured in this work with *in vitro* assays.

Biosynthesis pathway	Allosteric Feedback		K$_i$ mM		Transcriptional Feedback			K$_{m/d}$ mM	
	Enzyme	Metabolite	LB	UB	Mechanism	Protein	Metabolite	LB	UB
Arginine	ArgA	arg	0.15		Repressor	ArgR	arg	0.28	
Asparagine	AsnA	asn	0.12		Repressor	AsnC	asn	1	
Cysteine	CysE	cys	0.001						
Histidine	HisG	his	0.012	0.1	Attenuation	his-tRNA ligase	his	0.008	0.03
Isoleucine	IlvA	ile	0.06		Attenuation	ile-tRNA ligase	ile	0.0036	1.3
Leucine	LeuA	leu	0.28		Attenuation	leu-tRNA ligase	leu	0.0015	0.05
Lysine	DapA	lys	0.21	3.9					
Methionine	MetA	met	0.1	4	Repressor	MetJ	sam	0.01	0.05
Phenylalanine	PheA	phe	0.1	0.6		TyrR	phe	>0.18	
Proline	ProB	pro	0.02						
Serine	SerA	ser	0.005	0.37					
Threonine	ThrA	thr	0.097	0.167	Attenuation	thr-tRNA ligase	thr	0.11	0.2
Tryptophan	TrpE	trp	0.17		Repressor	TrpR	trp	0.16	
Tryptophan	TrpE	trp	0.17		Attenuation	trp-tRNA ligase	trp	0.017	
Tyrosine	TyrA	tyr	0.1		Repressor	TyrR	tyr	0.18	
Valine	IlvB	val	0.078	0.1	Attenuation	val-tRNA ligase	val	0.0043	0.1

Table S6. Oligonucleotides used in this study.

Oligonucleotide	Sequence (5'-3')	Description
argA_Forward	GGTCGAGGGATTCCGCCATTG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against argA
argA_Reverse	AATGGCGGAATCCCTCGACCG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against argA
ilvA_Forward	AGTACGTACCCAGCGTGTTGG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against ilvA
ilvA_Reverse	CAACACGCTGGGTACGTACTG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against ilvA
hisG_Forward	CAGCGCTTTCAGTTTTTCCAGT TTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against hisG
hisG_Reverse	TGGAAAAACTGAAAGCGCTGG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against hisG
leuA_Forward	GACCCAGCGCATCTTTACCGG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against leuA
leuA_Reverse	CGGTAAAGATGCGCTGGGTCG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against leuA
proB_Forward	GCAACGCTCGCAGGGTGTCGG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against proB
proB_Reverse	CGACACCCTGCGAGCGTTGCG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against proB
thrA_Forward	TGATTGCGTAATCAGCACCAG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against thrA
thrA_Reverse	TGGTGCTGATTACGCAATCAG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against thrA
trpE_Forward	ACACAACTGGTGAAAAAGCGG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against trpE
trpE_Reverse	CGCTTTTTCACCAGTTGTGTG TGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against trpE
argR_Forward	ATTCTTCAATGGACTGGAGGG TTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA targeted against argR
argR_Reverse	CCTCCAGTCCATTGAAGAATGT GCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA targeted against argR
CPEC001	TTTATAACCTCCTTAGAGCTCGA	Reverse primer for amplification of fragment 1 for pKDsgRNA
CPEC002	CCAATTGTCCATATTGCATCA	Forward primer for amplification of fragment 2 for pKDsgRNA
Ec-F	GTTTTAGAGCTAGAAATAGCAAGTTAA AATAAGGC	Foward primer used with **guide_Rev** for ampflification of customized pNUT1533-ctrl
Ec-F-argE-mm5	TTTTTCATTGTTGACACCCCTCGTTTTAG AGCTAGAAATAGCAAGTTAAAATAAGG C	Foward primer used with **guide_Rev** for ampflification of customized pNUT1533-argE
Ec-F-trpA	TTCTTTGCGCTCCTTCAACTGTTTTAGA GCTAGAAATAGCAAGTTAAAATAAGGC	Foward primer used with **guide_Rev** for ampflification of customized pNUT1533-trpA
Ec-F-hisB	TCACTCGGCGGTTCGCTAATCAGTTTTA GAGCTAGAAATAGCAAGTTAAAATAAG GC	Foward primer used with **guide_Rev** for ampflification of customized pNUT1533-hisB
Ec-R	ACTAGTATTATACCTAGGACTGAGCTA GC	Reverse primer for amplification of customized pNUT1533 plasmids
ArgA_fwd_NdeI	TGACCATATGATGGTAAAGGAACGTAA AAC	Amplification of genomic argA
ArgA_rev_BamHI	TGACGGATCCTTACCCTAAATCCGCCAT CA	Amplification of genomic argA
ArgA_H15Y_fwd	AGGGAACCGAATAGCGGAATCCCTC	Forward primer for amplification pET28a(+)-argA
ArgA_H15Y_rev	ATATCAATACCCACCGGG	Reverse primer for amplification pET28a(+)-argA
hisL_fwd_gfp	CCGCTCGAGGCTTTCATCATTGTTGCCG	Forward primer for amplification of hisL attenuator region
hisL_rev_gfp	CCGGGATCCCGCAGAATATCAATCGGC	Reverse primer for amplification of hisL attenuator region
leuL_fwd_gfp	CCGCTCGAGTTGTCCCCTTTTTCCTCG	Forward primer for amplification of leuL attenuator region
leuL_rev_gfp	CCGGGATCCATGGTTTGCACCGATTC	Reverse primer for amplification of leuL attenuator region
thrA_fwd_gfp	CCGCTCGAGACTGCAACGGGCAATATG	Forward primer for amplification of thrL attenuator region
thrA_rev_gfp	CCGGGATCCTCGGCATCGCTGATATTG	Reverse primer for amplification of thrL attenuator region

Chapter 2

CRISPRi-Based Downregulation of Transcriptional Feedback Improves Growth and Metabolism of Arginine Overproducing *E. coli*

Timur Sander, Chun Ying Wang, Timo Glatter and Hannes Link

TS designed and performed experiments, analyzed data and co-wrote the manuscript.

CYW performed experiments.

TG performed proteome measurement

HL analyzed data, co-wrote the manuscript and directed the project.

Summary

Removing transcriptional feedback regulation of metabolic pathways is a classical approach to enhance overproduction of chemicals in microbes. However, disrupting transcriptional regulation can have broad physiological consequences that decrease cellular growth and productivity. Here, we compared downregulation and deletion of the transcriptional repressor ArgR in arginine overproducing *Escherichia coli*. Different levels of ArgR downregulation were achieved with CRISPR interference (CRISPRi) and resulted in 2-times higher growth rates compared to deletion of ArgR, while specific arginine production was similar (~2 mmol g_{DW}^{-1} h^{-1}). Metabolomics and proteomics data revealed that poor growth of the ArgR deletion strain was caused by a limitation of pyrimidine nucleotide biosynthesis, because a 17-fold overexpression of ornithine carbamoyltransferase (ArgI) perturbed the arginine-pyrimidine branch point. These results demonstrate that overexpression of enzymes in an engineered pathway can impair metabolism of the host, especially in case of branch point enzymes. Thus, balancing enzyme-levels is important to optimize industrial microbes and CRISPRi of a transcription factor is a versatile tool for this purpose.

Graphical Abstract
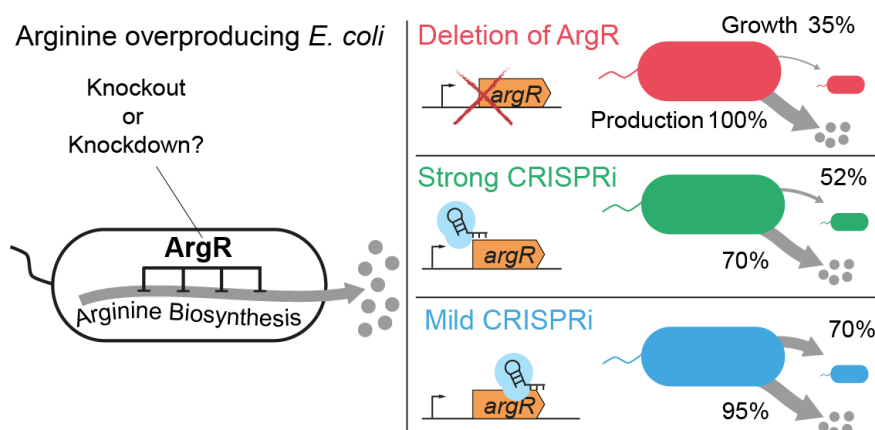
Introduction

Dysregulating cellular control mechanisms is a common approach to engineer microbes for overproduction of chemicals (Park *et al.*, 2017). Control of metabolic pathways is often mediated by end-products that feedback inhibit transcription or enzymatic activity, and removing this layer of control can increase flux into overproduction pathways (Nielsen and Keasling, 2016; Lee and Wendisch, 2017). However, although removing feedback inhibition enhances overproduction, the missing regulation can have broad physiological consequences for the host that reduce cellular growth and robustness (He, Murabito and Westerhoff, 2016). This problem has been addressed by creating overproduction pathways with new control mechanisms, which sense for example the availability of intracellular metabolites (Farmer and Liao, 2000), or the concentration of pathway intermediates by synthetic feedback mechanisms (Zhang, Carothers and Keasling, 2012). Another approach is to separate growth and production phases by dynamically switching between these states using metabolic valves (Burg *et al.*, 2016; Gupta *et al.*, 2017). Instead of engineering new regulation, an alternative strategy could be to remove feedback regulation just partially, such that flux into the production pathway is high enough, but the host retains a certain level of regulation. Such partial removal of regulation could be achieved by downregulating a transcription factor, for instance by CRISPR interference.

The CRISPR interference system consists of a catalytically deactivated Cas9 protein (dCas9) and a single guide RNA (sgRNA) that form a complex and block transcription (Lim *et al.*, 2013). CRISPRi has been used to engineer new regulation by placing expression of dCas9 or the sgRNA under control of stress promoters (Ceroni *et al.*, 2018), or to create synthetic feedbacks that control gene expression in response to glucose or acetate levels in the host (Moser *et al.*, 2018). Other metabolic engineering approaches used CRISPRi to downregulate genes that are involved in competing pathways, which redirected flux towards desired products (Kim *et al.*, 2017; Tian *et al.*, 2019). CRISPRi was also used to downregulate the transcription factor MetJ in *E. coli*, in order to improve overproduction of O-methylated anthocyanin (Cress *et al.*, 2017). However, it is not known whether downregulation or complete deletion of a transcription factor is better for overproduction.

Removing transcriptional feedback is especially important to overproduce amino acids and their intermediates, because transcription of genes in amino acid metabolism is tightly regulated in most organisms (Becker and Wittmann, 2012). Arginine, for instance, is a high demand product in the pharmaceutical and food-industry (Álvares *et al.*, 2011; Nascimento, 2018), and has been produced with engineered *E. coli* strains that lacked allosteric feedback inhibition and transcriptional regulation in the arginine pathway (Ginesy *et al.*, 2015). One of these *E. coli* strains achieved titers of 11.6 g L^{-1}

arginine, but this strain grew slowly (Ginesy *et al.*, 2015). One reason for the slow growth could be the deletion of the transcriptional repressor ArgR (**Figure 1a**), which has a regulatory function beyond arginine metabolism and controls 423 direct and indirect targets (Cho *et al.*, 2012). However, because ArgR strongly represses genes in the arginine pathway even during growth on glucose minimal medium (Gerosa *et al.*, 2013; Sander *et al.*, 2019), it needs to be clarified if ArgR can be reduced to a level at which overproduction is sufficiently high, while cellular physiology is not affected by the missing regulation.

Here we tested if downregulation of ArgR improves arginine overproduction compared to deletion of ArgR. Therefore, we first rationally engineered *E. coli* for arginine overproduction by dysregulating the first enzyme in the pathway (ArgA^{H15Y}) and expressing an arginine transporter (ArgO). In this base strain we compared deletion of the arginine repressor ArgR (complete dysregulation) with CRISPRi mediated downregulation of ArgR to different levels (partial dysregulation). Both approaches achieved a specific arginine production of about 2 mmol g_{DW}^{-1} h^{-1}, but the partial dysregulated strain grew 2-times faster. Metabolomics data revealed that these differences in cellular growth originate from pyrimidine nucleotide biosynthesis, which competes with arginine biosynthesis at the branch point of ornithine carbamoyltransferase (ArgI) and aspartate carbamoyltransferase (PyrB).

Results and Discussion

To understand how feedback dysregulation affects physiology and overproduction of arginine in *E. coli,* we measured growth on minimal glucose medium and the concentration of arginine in the culture supernatant of wild-type *E. coli* (MG1655), and three dysregulated *E. coli* strains that derive from MG1655. The first dysregulated strain (*argA**) lacked allosteric feedback inhibition of the N-acetylglutamate synthase (ArgA) by arginine due to a genomic point mutation (H15Y) in the *argA* gene, which does not affect enzymatic activity of ArgA (Sander *et al.*, 2019) (**Figure 1a**). The second strain (Δ*argR*) lacked transcriptional feedback by arginine, due to deletion of ArgR (Sander *et al.*, 2019). The third strain (*argA**Δ*argR*) was doubly dysregulated, because both feedback mechanisms were removed in this strain. Deletion of either allosteric or transcriptional feedback resulted in low arginine concentrations in the culture supernatant (<0.01 mmol g_{DW}^{-1}), and only the doubly feedback dysregulated *argA**Δ*argR* strain produced higher amounts of arginine (2.23 mmol g_{DW}^{-1}). The observation that overproduction of arginine requires dysregulation of both allosteric and transcriptional feedback regulation is consistent with previous reports about the interplay of the two

regulatory layers (Caldara *et al.*, 2008; Sander *et al.*, 2019). Thus, the loss of one feedback mechanism is compensated by the other, in order to avoid overproduction of arginine.
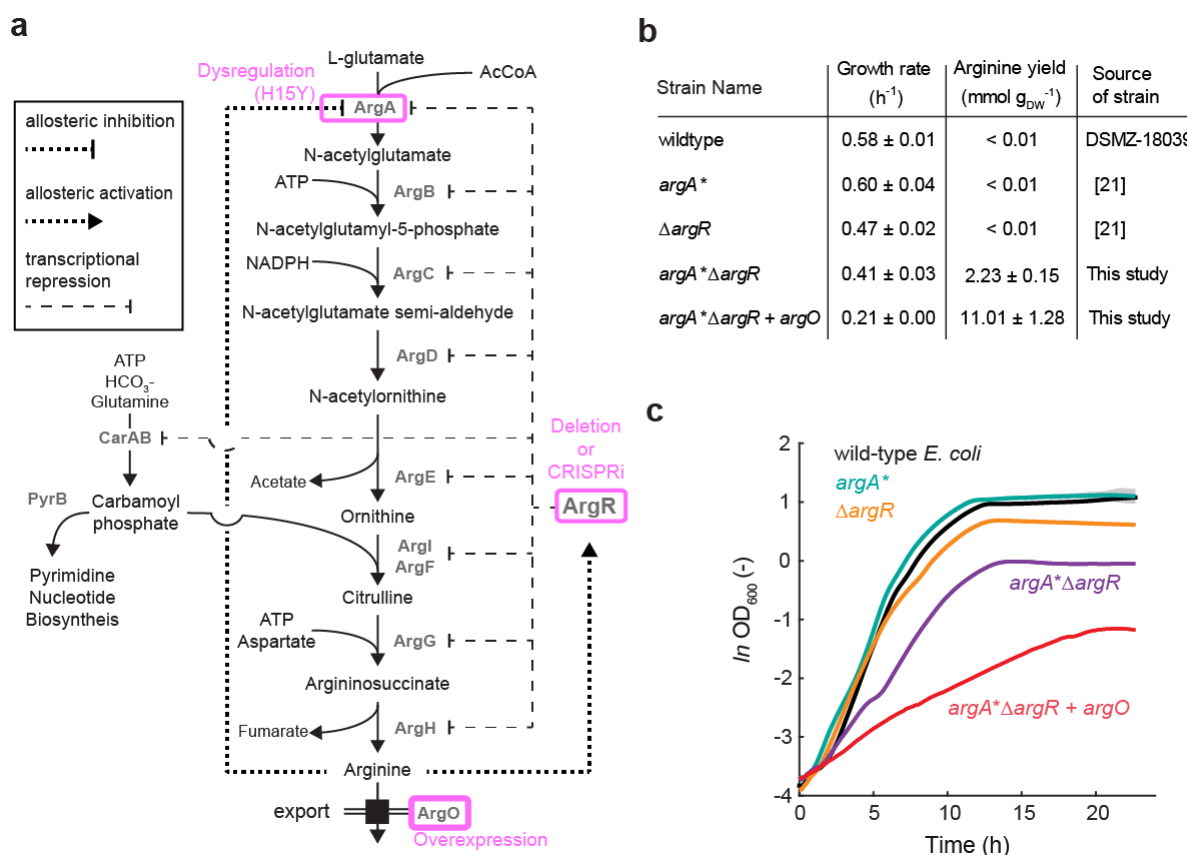
a

b

Strain Name	Growth rate (h^{-1})	Arginine yield (mmol g$_{DW}^{-1}$)	Source of strain
wildtype	0.58 ± 0.01	< 0.01	DSMZ-18039
argA*	0.60 ± 0.04	< 0.01	[21]
ΔargR	0.47 ± 0.02	< 0.01	[21]
argA*ΔargR	0.41 ± 0.03	2.23 ± 0.15	This study
argA*ΔargR + argO	0.21 ± 0.00	11.01 ± 1.28	This study

c

Figure 1. **Overproduction of arginine in *E. coli*.**

(a) Biosynthesis of arginine in *E. coli* is regulated by allosteric feedback inhibition (dotted line) and transcriptional regulation by ArgR (dashed line). Targets for metabolic engineering are indicated with pink boxes.

(b) Growth rates (h^{-1}) and biomass specific arginine yield (mmol g$_{DW}^{-1}$) of *E. coli* wild-type, the allosteric mutant *argA**, the transcriptional mutant Δ*argR*, the double mutant *argA*ΔargR* and the double mutant expressing the arginine exporter *argO* (*argA*ΔargR + pargO*). Extracellular arginine was measured at ODs of 0.4 and 0.7 to calculate the biomass specific arginine yield by linear regression. Standard deviations are based on n = 3 independent shake flask cultures.

(c) Growth curves in micro-titer plates of wild-type *E. coli*, the allosteric mutant *argA**, the transcriptional mutant Δ*argR*, the double mutant *argA*ΔargR* and the double mutant expressing the *argO* exporter (*argA*ΔargR + pargO*).

Expressing the arginine exporter ArgO (Nandineni and Gowrishankar, 2004) in the doubly dysregulated strain further increased arginine levels to 11 mmol g$_{DW}^{-1}$ (**Figure 1b**), and we will refer to this *argA*ΔargR + pargO* strain as the KO-strain (KO for knockout of ArgR). A problem was that the KO-

strain grew about 70% slower than the wild-type, which could be caused by: i) arginine overproduction (flux burden), ii) the missing transcriptional feedback regulation (regulatory burden) or iii) expression of the arginine exporter (intracellular arginine limitation) (**Figure 1b, c**). To exclude that expression of the arginine exporter ArgO depleted intracellular arginine and thereby reduced growth, we expressed ArgO in the wild-type, in the *argA** strain and in the Δ*argR* strain. In none of the strains expression of ArgO caused a growth defect, showing that slow growth of the KO-strain did not result from expression of the arginine exporter (**Figure S2**). To understand the impact of a regulatory burden by missing ArgR regulation, we next sought to tune transcriptional feedback regulation by gradual downregulation of ArgR (instead of ArgR deletion).

To gradually tune transcriptional feedback regulation in the arginine pathway, we downregulated ArgR with CRISPRi. Targeting different loci of a gene results in different levels of downregulation by CRISPRi(Larson *et al.*, 2013), and therefore we designed 10 sgRNAs targeting different positions of the promoter and the coding region of the *argR* gene on the template strand (sgRNA 1-5) and on the non-template strand (sgRNA 6-10) (**Figure 2a**). Plasmids with the 10 sgRNAs were individually transformed into the wild-type and the allosterically dysregulated *argA** strain. dCas9 was expressed from a second plasmid pdCas9(Lim *et al.*, 2013). This resulted in 20 strains, which were cultivated in 96-well plates and screened for arginine production by measuring arginine in the whole culture broth after 20 hours (**Figure 2b**). As expected on the basis of low arginine production in the Δ*argR* strain, arginine levels did not change in the wild-type with any of the 10 sgRNA's targeting *argR* (**Figure 2c**). This is presumably caused by the remaining allosteric feedback regulation that compensates transcriptional dysregulation. In the *argA** mutant, in contrast, downregulating of *argR* resulted in three different arginine levels: low <5 fold increase (sgRNA 1, 3, 4 and 5), medium 213-243 fold increase (sgRNA 2, 6, 8, 9 and 10) and a strong increase of 613 fold in case of sgRNA 7 that targets the ATG region on the non-template strand (**Figure 2c**). These results are consistent with previous studies showing that repression efficiency is higher when targeting the non-template strand, than when targeting the template strand (Larson *et al.*, 2013). An exception is sgRNA 2, which led to high arginine levels although it is targeting the template strand, most likely because sgRNA 2 binds to the -10 promotor region. Thus, by targeting different loci of the *argR* gene with CRIPSRi we could tune arginine overproduction in an allosterically dysregulated *argA** strain. We then decided to characterize CRISPRi strains with sgRNA 7 (613-fold higher arginine levels) and sgRNA 10 (231-fold higher arginine levels) in more detail and compared them to deletion of ArgR.
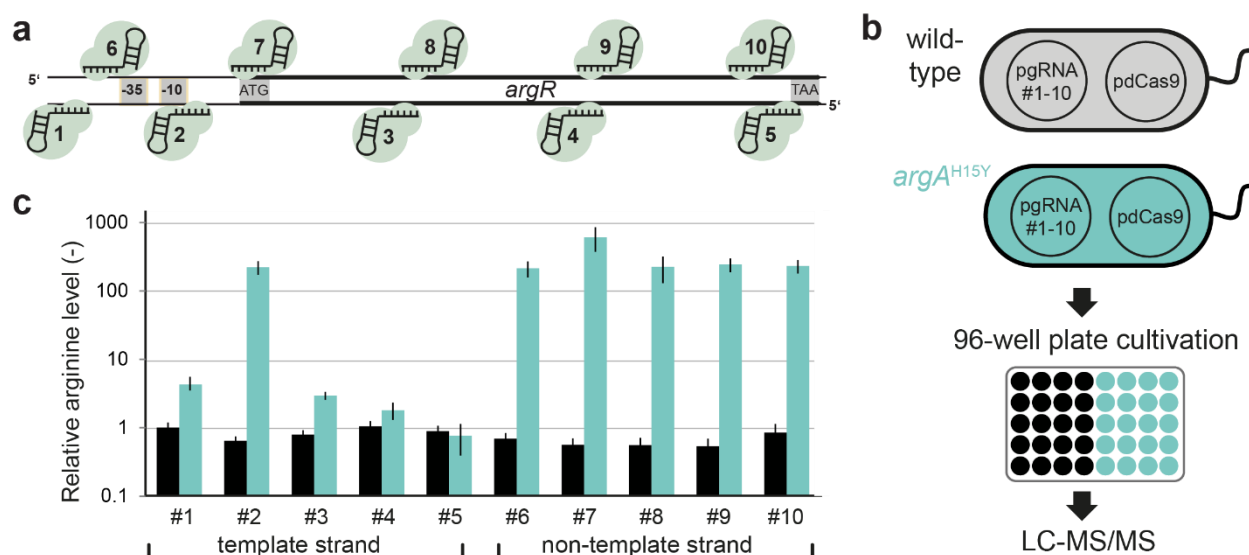
Figure 2. Different levels of ArgR downregulation with CRISPRi results in different arginine production.

(a) CRISPR interference of the *argR* gene at 10 different positions. sgRNA #1-5 target the template strand and sgRNA #6-10 the non-template strand.

(b) Experimental set-up to screen arginine production in 20 strains with CRISPR interference. Wild-type *E. coli* and the allosteric mutant *argA** were each transformed with a plasmid that harbors dCas9 and one of the 10 sgRNA plasmids (pgRNA#1-10). The resulting 20 strains were cultivated in 96-well plate cultures on glucose minimal medium (n = 4) and arginine was measured after 20 hours.

(c) Relative arginine levels in the wild-type (black) and the allosteric mutant *argA** (cyan) with 10 different sgRNAs. Arginine levels were normalized to biomass and are shown relative to wild-type with sgRNA#1. Error bars are the standard deviation of n = 4 cultures.

To compare deletion of ArgR with the two levels of ArgR downregulation, we used the *argR*-deletion strain *argA*ΔargR* + p*argO* (KO-strain) as the reference strain. For the two knockdown strains we targeted *argR* in the allosteric mutant *argA** that expressed the arginine exporter (*argO*). Therefore, we first cloned sgRNA 7 and sgRNA 10 (see **Figure 2a**) on a single CRISPRi plasmid that harbors both, dCas9 and the sgRNA (based on pNUT1527(Beuter *et al.*, 2018)). This resulted in the two strains *argA** + pCRISPRi-*argR*#7 + p*argO* and *argA** + pCRISPRi-*argR*#10 + p*argO*, and we will refer to these strains as CRISPRi#7 and CRISPRi#10, respectively.

First, we tested how CRISPRi and deletion of ArgR affects the expression of enzymes in the arginine biosynthesis pathway and measured the proteome of the wild-type, the KO-strain, as well as the

strains CRISPRi#7 and CRISPRi#10. The abundance of all nine enzymes in the arginine pathway showed the expected pattern, which was: KO-strain > CRISPRi#7 > CRISPRi#10 > wild-type (**Figure 3a**). Thus, all dysregulated strains overexpressed arginine biosynthesis enzymes, with stronger overexpression in the strain with ArgR deletion than in the two CRISPRi strains. The degree of overexpression varied among the different enzymes, and N-acetylglutamate synthase (ArgA) as well as ornithine carbamoyltransferase (ArgI) were the two most strongly overexpressed enzymes in all strains. Notably, ArgA catalyzes the committed step of the arginine pathway and ArgI catalyzes a key reaction at the branch point between arginine and pyrimidine nucleotide biosynthesis (**Figure 1a**).
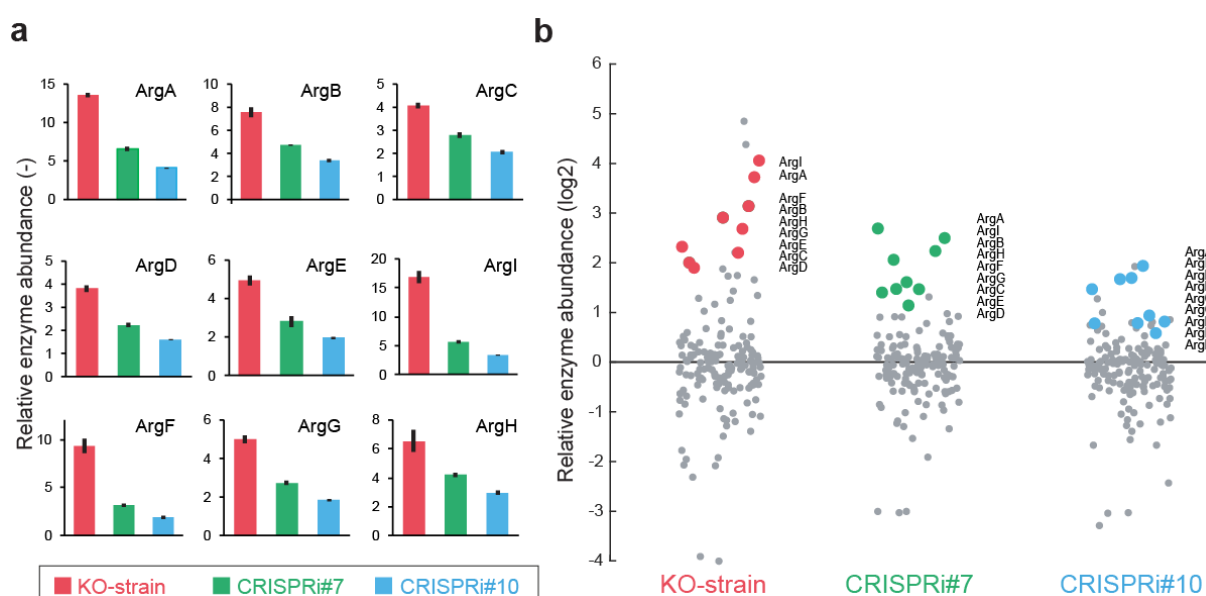

Figure 3. Expression of enzymes in strains with CRISPRi of ArgR and deletion of ArgR.

(a) Enzyme levels in the arginine biosynthesis pathway (relative to the wild-type). KO-strain (red), strain CRISPRi#7 (green) and strain CRISPRi#10 (blue). Error bars show standard deviations of proteomics samples from n = 3 independent shake flask culture.

(b) Abundance of 156 enzymes in amino acid metabolism of *E. coli* (relative to wild-type, log2 scale). The colored dots are arginine biosynthesis enzymes.

Apart from arginine enzymes, we inspected changes of the remaining 156 enzymes in amino acid metabolism that were covered by the proteomics data (**Figure 3b**). In the KO-strain, 40 enzymes changed more than 2-fold, while only 21 and 25 enzymes changed in CRISPRi#7 and CRISPRi#10, respectively. This shows that deletion of ArgR causes stronger and more global changes of amino acid enzymes than downregulation of ArgR.

Having established that the three arginine overproduction strains (KO, CRISPRi#7, CRISPRi#10) overexpress arginine biosynthesis enzymes differently, we wondered how this affects arginine production. Therefore, the KO-strain, CRISPRi#7, CRISPRi#10 and *E. coli* wild-type were cultivated in shake flasks on glucose minimal medium to measure growth and arginine concentration in the supernatant (three independent cultures per strain). Both CRISPRi#7 and CRISPRi#10 grew much better than the complete dysregulated KO-strain, and growth rates were consistent between independent experiments that were performed on different days (**Figure 4a, b**). CRISPRi#10 (lowest overexpression of arginine enzymes) grew almost like the wildtype, while CRISPRi#7 (medium overexpression of arginine enzymes) had a lower growth rate. Yet, CRISPRi#7 still grew ~50% faster than the KO-strain (strongest overexpression of arginine enzymes).

To determine the specific arginine production rate, we measured extracellular arginine during the exponential growth phase and collected samples of the supernatant when cells reached optical densities of 0.2, 0.5 and 0.8 (**Figure 4a**). From these data we calculated the specific arginine production rate, which is directly related to the flux of arginine per cell (**Figure 4a, S1**). Surprisingly, specific arginine production was in a similar range in all three strains. CRISPRi#10 had almost the same specific rate as the KO-strain (2.2 and 2.3 mmol g_{DW}^{-1} h^{-1}). Specific arginine production of the CRISPRi#7 strain was a bit lower (1.6 mmol g_{DW}^{-1} h^{-1}). Similarly, arginine titer at the end of the exponential growth phase were comparable between the KO strain (1.03 g L^{-1}), CRISPRi#7 (0.96 g L^{-1}) and CRISPRi#10 (0.90 g L^{-1}) (**Figure 4b**).

Thus, the three strains express different amounts of arginine enzymes but flux through the arginine pathway is comparable. This indicates that other factors than enzyme abundance limit arginine overproduction (e.g. metabolic precursors and cofactors). Moreover, it seems that arginine overproduction *per se* is not burdensome for *E. coli*, because the KO-strain and CRISPRi#10 had the same arginine flux (2.2 mmol g_{DW}^{-1} h^{-1}) but large differences in growth (**Figure 4c**). Therefore, we hypothesized that other factors than arginine flux influence growth of the production strains, and they may be caused by the different enzyme-levels in the arginine pathway.
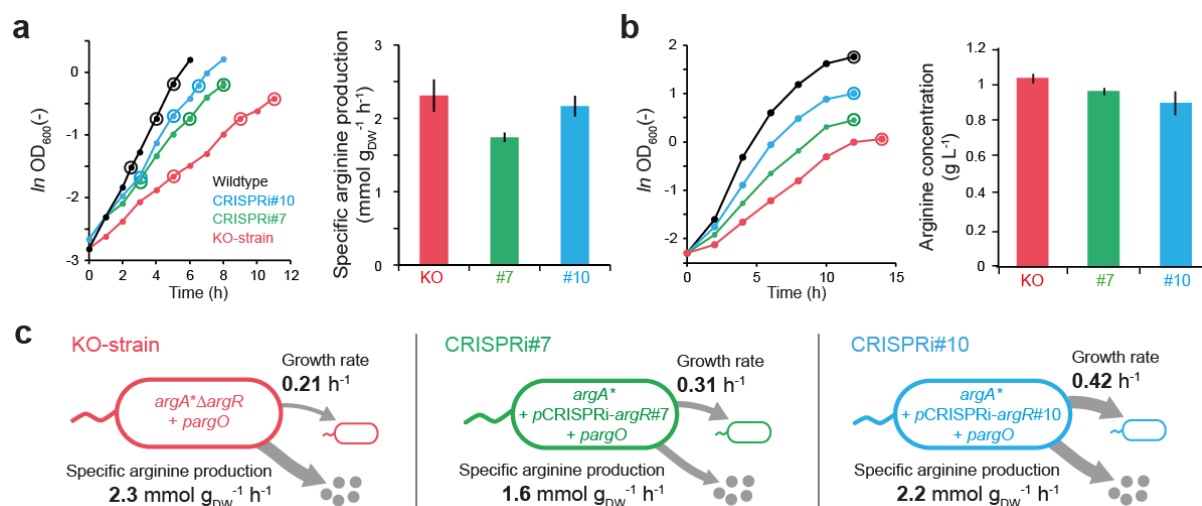
Figure 4. Growth and specific arginine production in strains with CRISPRi of ArgR and deletion of ArgR.

(a) Specific arginine production during exponential growth of wild-type *E. coli* (black), strain *argA*ΔargR + pargO* (KO-strain, red), strain *argA** pCRISPRi-*argR*#7 + *pargO* (CRISPRi#7-strain, green) and strain *argA** pCRISPRi-*argR*#10 + *pargO* (CRISPRi#10-strain, blue). Specific arginine production was calculated by linear regression at the three time points, which are indicated in the growth curves with circles. Shown are means of n = 3 independent shake flask cultures.

(b) Absolute arginine concentrations at the end of exponential growth phase of the same strains as in (a). Arginine was measured at the last time point, which is indicated in the growth curves with a circle. Shown are means of n = 3 independent shake flask cultures.

(c) Summary of growth and specific arginine production of the KO-strain (red), CRISPRi#7 (green) and CRISPRi#10 (blue). Growth rates are calculated from the three time points indicated with circles in (a). The thickness of the arrows illustrates the relative amount of growth and production.

To understand what caused the differences in growth rates of the KO-strain and the two CRISPRi strains we measured intracellular metabolites in these strains. Therefore, we collected samples of the wildtype, the KO-strain, as well as the CRISPRi#7 and CRISPRi#10 strains during exponential growth in shaking flasks and measured metabolites by LC-MS/MS (Guder *et al.*, 2017). The metabolite data revealed changes in different parts of cellular metabolism, including central-, nucleotide-, cofactor- and amino acid metabolism (**Figure 5**). Across all 104 measured metabolites, the strongest decreases occurred in pyrimidine nucleotide metabolism of the KO-strain, and the first pyrimidine intermediate - carbamoyl-aspartate - was by far the most strongly decreased metabolite (**Figure 5**). However, while carbamoyl-aspartate decreased >100-fold in the KO-strain, the concentration of this metabolite was

much higher in the two CRISPRi strains (**Figure 5**). In the CRISPRi#10 strain for example, carbamoyl-aspartate was only about 3-fold lower than in the wild-type. Similarly, other metabolites in pyrimidine nucleotide biosynthesis, in particular the end-products UTP and CTP were closest to wild-type levels in the CRISPRi#10 strain (**Figure 5**).
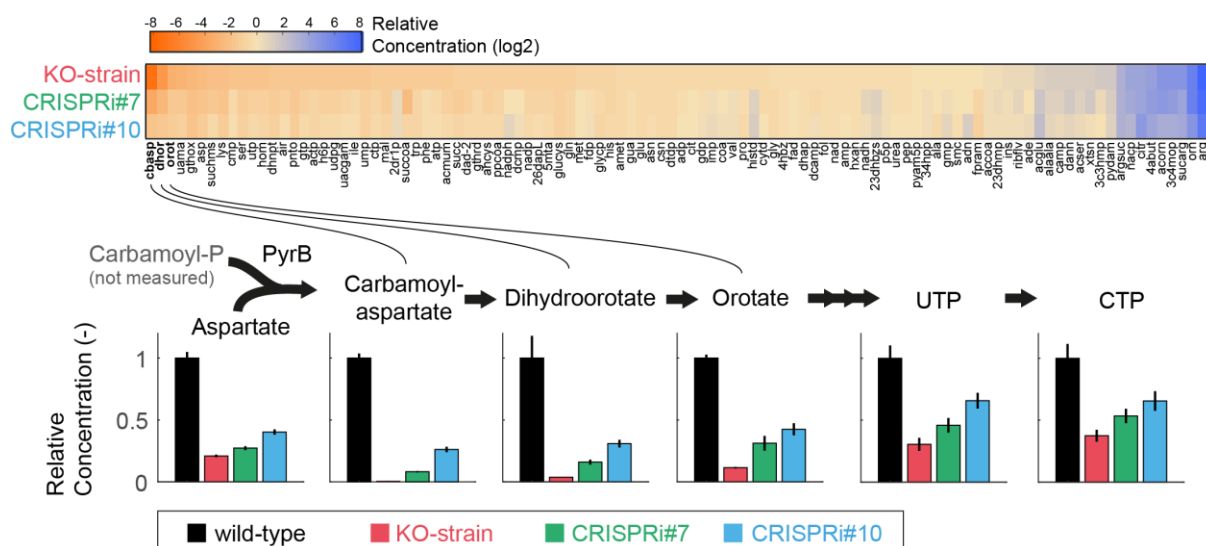

Figure 5. Intracellular metabolite concentrations in strains with CRISPRi of ArgR and deletion of ArgR.

Relative concentrations of 104 intracellular metabolites in the KO-strain, the CRISPRi#7-strain and the CRISPRi#10-strain. The bar plots show concentrations of pyrimidine nucleotide intermediates and end-products. Concentrations are relative to the wild-type. Bars show means of n = 3 samples from independent shake flasks, and error bars are standard deviations.

In summary, metabolomics data revealed a pyrimidine nucleotide limitation especially in the KO-strain. The low carbamoyl-aspartate concentration in the KO-strain indicates that the limitation occurs at the first step of the pathway, which is catalyzed by aspartate carbamoyltransferase (PyrB). A hypothesis is that the 17-fold overexpression of ornithine carbamoyltransferase (ArgI) in the KO-strain (**Figure 3a**) perturbs the ArgI-PyrB branch point, e.g. by leaving less substrate for PyrB. This would also explain why the pyrimidine limitation is less severe in the CRSIPRi#10 strain, because ArgI was only 3-fold overexpressed in this strain. Thus, it seems that the overcapacity of ArgI impairs growth of the KO-strain, and not flux through the arginine pathway (which is comparable in the three strains, **Figure**

4a). Aspartate levels were also lower in the three overproduction strains than in the wild-type and this could additionally limit pyrimidine biosynthesis (**Figure 5**).

Finally, we tested if pyrimidine nucleotides are indeed growth limiting in the arginine production strains, and supplemented the minimal glucose medium with the pyrimidine intermediate orotate, the pyrimidine precursor aspartate, the nucleoside cytidine and the nucleobase uracil. The KO-strain grew indeed 2.3-fold faster in the presence of the supplements, confirming that pyrimidine nucleotides are limiting growth of this strain (**Figure 6**). The growth rate of CRISPRi#7 increased 1.5-fold with supplements, while growth rates of the wild-type and CRISPRi#10 were almost not affected (1.1-fold increase, **Figure 6**).
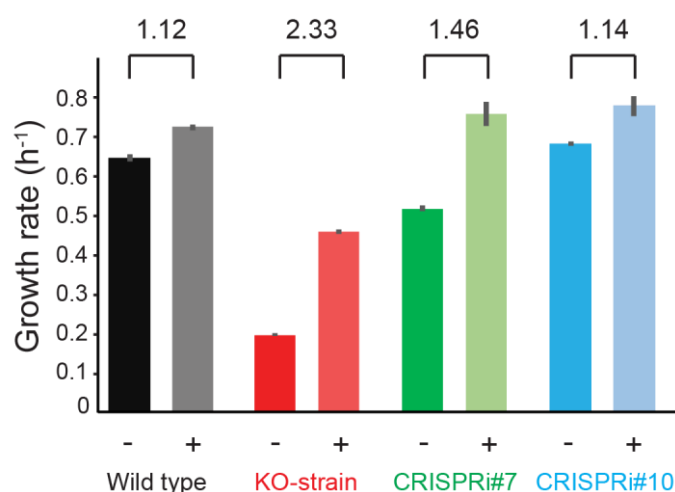

Figure 6. Growth rates with (+) and without (-) supplementation of orotate, aspartate, cytidine and uracil (each 0.5 mM) for the wild-type, KO-strain, CRISPRi#7 and CRISPRi#10. Number above brackets show the fold change between growth rates with and without supplements. Growth rates (h^{-1}) were calculated by linear regression over a 60 minutes window at OD 0.3. Shown are means of n = 3 microtiter-plate cultures.

Therefore, the limited supply of pyrimidine nucleotides seems responsible for slow growth of the KO-strain, and ArgR-downregulation instead of ArgR-deletion almost abolishes this limitation. However, with supplements the growth rate of the KO-strain (0.46 h^{-1}) was still below the wild-type growth rate (0.65 h^{-1}), but similar to the growth rate of the strain with only ArgR deletion ($\Delta argR$, 0.47 h^{-1}, **Figure 1b**). This suggests that in addition to the pyrimidine limitation the global changes of amino acid enzymes (**Figure 3b**) impair growth of the KO-strain.

In conclusion, we could show that growth of an arginine overproduction strain can be improved by carefully balancing enzyme levels in the arginine pathway. The KO-strain with deletion of ArgR overexpressed arginine biosynthesis enzymes much stronger than strains with CRISPRi of ArgR (CRISPRi#7 and CRISPRi#10). Despite this, the specific arginine production of the three strains was in a similar range, showing that it was not arginine overproduction *per se* that caused differences in growth between the three strains. One reason for these differences was that deletion of ArgR caused stronger and more global changes of amino acid enzymes than downregulation of the transcription factors. These global changes probably cause protein burden if enzymes are upregulated or limit biosynthesis of other amino acids if enzymes are downregulated. Apart from global effects on amino acid metabolism, we concluded that overexpression of arginine biosynthesis enzymes was responsible for the slow growth of the ArgR deletion KO-strain. In particular the very high levels of ArgI caused a limitation in pyrimidine nucleotide biosynthesis, probably by perturbing enzyme ratios at the ArgI-PyrB branch point. A solution to this problem could be additional overexpression of PyrB, to restore the ratio between ArgI and PyrB levels. These results show that if enzyme-levels are higher than needed for overproduction, they can impair cellular growth and thereby reduce productivity. It is especially important to balance the expression of branch point enzymes, because their abundance can affect other metabolic pathways of the host. Balancing expression of enzymes in overproduction pathways is possible with various methods (Jones, Toparlak and Koffas, 2015), and here we show that downregulating transcription factors with CRISPRi is an alternative approach. Since deletion of transcription factors is a common approach to engineer industrial microbes, the proposed CRISPRi strategy could be applicable to other pathways and products in order to design fitter and more robust overproduction strains.

Material and Methods

Strains and Culture

E. coli MG1655 (DSMZ No. 18039) was used as the wild-type. Chemically competent *E. coli* TOP10 (One Shot^{TM} TOP10, Invitrogen Cat#C404003) were used for cloning. All mutants created in this study derive from the MG1655 strain and are listed in **Table S3**. For heterologous expression of the arginine exporter *argO* (KEGG T00007: b2923), the encoding gene was cloned into a p15A backbone (Addgene #44249; pdCas9-bacteria) by Gibson assembly (**Figure S3**). *argO* was expressed from a P$_{tet}$ promotor which was not induced due to sufficient basal expression. The genomic point mutation that removes allosteric feedback of *argA* (H15Y) and deletion of *argR* were created by scarless Cas9 assisted recombineering (Reisch and Prather, 2015).

All cultivations were performed in M9 minimal medium with 5 g L^{-1} glucose. The M9 medium consisted of the following components (per liter): 7.52 g Na$_2$HPO$_4$ 2 H$_2$O, 5 g KH$_2$PO$_4$, 1.5 g (NH$_4$)$_2$SO$_4$, 0.5 g NaCl. The following components were sterilized separately and then added (per liter of final medium): 1 mL 0.1 M CaCl$_2$, 1 mL 1 M MgSO$_4$, 0.6 mL 0.1 M FeCl$_3$, 2 mL 1.4 mM thiamine-HCL and 10 mL trace salts solution. The trace salts solution contained (per liter): 180 mg ZnSO$_4$ 7 H$_2$O, 120 mg CuCl$_2$ 2 H$_2$O, 120 mg MnSO$_4$ H$_2$O, 180 mg CoCl$_2$ 6 H$_2$O. Where appropriate, 34 µg mL^{-1} chloramphenicol, 15 µg mL^{-1} gentamycine, 50 µg mL^{-1} spectinomycin or 100 µg mL^{-1} ampicillin was added. For cultivations in microtiter plates, LB pre-cultures in 96-deep-well format plates were inoculated from glycerol stocks and incubated until the cells grew exponentially. From this first pre-culture a second M9 pre-culture in 96-deep-well plates was inoculated 1:100 and incubated overnight at 37 °C under shaking. Finally, 96-well flat transparent plates (Greiner Bio-One International) containing 150 µL M9 minimal medium were inoculated 1:150 from the overnight culture. Online measurements of optical density at 600 nm (OD$_{600}$) were performed at 37°C with shaking in a plate reader (Tecan Trading AG, Switzerland). Growth rates were calculated as dln(OD)/dt by linear regression during exponential growth. For cultivations in shake flasks, 5 mL LB pre-culture in cultivation tubes were inoculated from glycerol stocks and incubated until the cells grew exponentially. From this first pre-culture, 5 mL of a second M9 batch pre-culture (5 g L^{-1} glucose) in cultivation tubes was inoculated 1:100 and incubated overnight at 37°C in a rotary shaker. Finally, a 500 mL shake flask containing 25 mL M9 minimal medium (5 g L^{-1} glucose) were inoculated 1:150 from the overnight culture, and incubated at 37 °C under shaking at 220 rpm.

CRISPR interference

CRISPR interference of *argR* was performed according to Larson et al. 2013(Larson *et al.*, 2013) with the plasmids pdCas9 (Addgene #44249) and pgRNA (Addgene #44251) (**Table S4**). The 10 different

gRNA plasmids were customized by site-directed mutagenesis with the forward primers EcF_*argR*#1-10, carrying the 20-nt base-pairing sequence (protospacer) and the reverse primer EcR (**Table S5**). For arginine overproduction strains CRISPRi#7 and #10 we used a single plasmid (pNUT1533) expressing the respective sgRNA from a constitutive and the dCas9 protein from an IPTG inducible Ptac promotor (Sander *et al.*, 2019). Therefore, the protospacer regions were customized by site-directed mutagenesis using the primer EcF_#7, respectively EcF_#10 and the reverse primer EcR. All CRISPRi experiments were performed without addition of inducer.

Sampling for measurement of intra- and extra cellular metabolites

Shake flask cultivations on M9 glucose were performed as described above. For measurement of intracellular metabolites, cells were grown to an optical density (OD_{600}) of 0.5 and 2 mL culture aliquots were vacuum-filtered on a 0.45 µm pore size filter (HVLP02500, Merck Millipore). Filters were immediately transferred into 40:40:20 (v-%) acetonitrile/methanol/water at -20°C for extraction. Extracts were centrifuged for 15 minutes at 13,000 rpm at -9 °C. Centrifuged extracts were mixed with ^{13}C-labeled internal standard and analyzed by LC-MS/MS (Guder *et al.*, 2017).

For measurement of extracellular arginine of the overproduction strains **(Figure 3)**, 500 µL of cell culture were sampled at OD 0.2, 0.5 and 0.8 and transferred into a reaction tube. The extracellular arginine yield in Figure 1b derived from samples at OD 0.4 and 0.7. After 30 s centrifugation at 13,000 rpm, 100 µL of the cell free supernatant were transferred into 400 µL 50:50 (v-%) acetonitrile/methanol at -20°C. Extracts were again centrifuged for 15 minutes at 13,000 rpm at -9 °C. Centrifuged extracts were mixed with ^{13}C-labeled internal standard and analyzed by LC-MS/MS. Samples were diluted 1:10 with 40:40:20 acetonitrile/methanol/water if necessary.

LC-MS/MS for metabolite measurements

LC-MS/MS analysis was performed with an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies) as described previously (Guder *et al.*, 2017). An Agilent 1290 Infinity II UHPLC system (Agilent Technologies) was used for liquid chromatography. Temperature of the column oven was 30°C, and the injection volume was 3 µL. LC solvents A were water with 10 mM ammonium formate and 0.1% formic acid (v/v) (for acidic conditions); and water with 10 mM ammonium carbonate and 0.2% ammonium hydroxide (for basic conditions). LC solvents B were acetonitrile with 0.1% formic acid (v/v) for acidic conditions and acetonitrile without additive for basic conditions. LC columns were an Acquity BEH Amide (30 x 2.1 mm, 1.7 µm) for acidic conditions, and an iHILIC-Fusion(P) (50 x 2.1 mm, 5 µm) for basic conditions. The gradient for basic and acidic conditions was: 0 min 90% B; 1.3 min 40 % B; 1.5 min 40 % B; 1.7 min 90 % B; 2 min 90 % B. Relative quantification of intracellular metabolite

concentrations was based on the ratio of ^{12}C and ^{13}C peak heights and normalized to the OD. Absolute concentrations of arginine were determined with an authentic standard.

Proteomics

Shake flask cultivations on M9 glucose were performed as described above. Cells were grown to an optical density (OD$_{600}$) of 0.5 and 2 mL culture aliquots were transferred into 2 ml reaction tubes and washed two times with PBS buffer (0.14 mM NaCl, 2.7 mM KCL, 1.5 KH$_2$PO$_4$, 8.1 Na$_2$HPO$_4$). Cell pellets were resuspended in 300 µL lysis buffer containing 100 mM ammonium bicarbonate, 0.5 % sodium laroyl sarcosinate (SLS) and 5 mM Tris(2-carboxyethyl)phosphine (TCEP). Cells were lysed by 5 minutes incubation at 95 °C and ultra-sonication for 10 seconds (Vial Tweeter, Hielscher). Cells were again incubated for 30 minutes at 90 °C followed by alkylation with 10 mM iodoacetamide for 30 minutes at 25 °C. To clear the cell lysate, samples were centrifuged for 10 minutes at 15,000 rpm and the supernatant transferred into a new tube. Proteins in the cell lysates were digested with 1 µg trypsin (Promega) overnight at 30 °C. Analysis of peptides was performed by liquid chromatography-mass spectrometry, carried out on a Q-Exactive Plus instrument connected to an Ultimate 3000 RSLC nano with a Prowflow upgrade and a nanospray flex ion source (Thermo Scientific). Peptide separation was performed on a reverse-phase HPLC column (75 µm x 42 cm) packed in-house with C18 resin (2.4 µm, Dr. Maisch GmbH, Germany). The following separating gradient was used: 98 % solvent A (0.15% formic acid) and 2 % solvent B (99.85 acetonitrile, 0.15 % formic acid) to 25 % solvent B over 105 minutes and to 35 % solvent B for additional 35 minutes at a flow rate of 300 nl min^{-1}. The data acquisition mode was set to obtain one high resolution MS scan at a resolution of 70,000 full width at half maximum (at m/z 200) followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS attempts, the charged state screening modus was enabled to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 seconds. The ion accumulation time was set to 50 ms for MS and 50 ms at 17,500 resolution for MS/MS. The automatic gain control was set to 3x106 for MS survey scans and 1x105 for MS/MS scans. Label-free quantification (LFQ) of the data was performed using Progenesis QIP (Waters), and for MS/MS searches of aligned peptide features MASCOT (v2.5, Matrix Science) was used. The following search parameters were used: full tryptic search with two missed cleavage sites, 10ppm MS1 and 0.02 Da fragment ion tolerance. Carbamidomethylation (C) as fixed, oxidation (M) and deamidation (N,Q) as variable modification. Progenesis outputs were further processed with SafeQuant.

References

Álvares, T. S. *et al.* (2011) 'L-arginine as a potential ergogenic aid in healthy subjects', *Sports Medicine*, pp. 233–248. doi: 10.2165/11538590-000000000-00000.

Becker, J. and Wittmann, C. (2012) 'Systems and synthetic metabolic engineering for amino acid production - the heartbeat of industrial strain development', *Current Opinion in Biotechnology*, pp. 718–726. doi: 10.1016/j.copbio.2011.12.025.

Beuter, D. *et al.* (2018) 'Selective Enrichment of Slow-Growing Bacteria in a Metabolism-Wide CRISPRi Library with a TIMER Protein', *ACS Synthetic Biology*. doi: 10.1021/acssynbio.8b00379.

Burg, J. M. *et al.* (2016) 'Large-scale bioprocess competitiveness: the potential of dynamic metabolic control in two-stage fermentations', *Current Opinion in Chemical Engineering*, pp. 121–136. doi: 10.1016/j.coche.2016.09.008.

Caldara, M. *et al.* (2008) 'Arginine biosynthesis in Escherichia coli: Experimental perturbation and mathematical modeling', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M705884200.

Ceroni, F. *et al.* (2018) 'Burden-driven feedback control of gene expression', *Nature Methods*. doi: 10.1038/nmeth.4635.

Cho, B. K. *et al.* (2012) 'Deciphering the transcriptional regulatory logic of amino acid metabolism', *Nature Chemical Biology*, 8(1), pp. 65–71. doi: 10.1038/nchembio.710.

Cress, B. F. *et al.* (2017) 'CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production', *Microbial Cell Factories*, 16(1). doi: 10.1186/s12934-016-0623-3.

Farmer, W. R. and Liao, J. C. (2000) 'Improving lycopene production in Escherichia coli by engineering metabolic control', *Nature Biotechnology*, 18(5), pp. 533–537. doi: 10.1038/75398.

Gerosa, L. *et al.* (2013) 'Dissecting specific and global transcriptional regulation of bacterial gene expression', *Molecular Systems Biology*, 9. doi: 10.1038/msb.2013.14.

Ginesy, M. *et al.* (2015) 'Metabolic engineering of Escherichia coli for enhanced arginine biosynthesis', *Microbial Cell Factories*, 14(1). doi: 10.1186/s12934-015-0211-y.

Guder, J. C. *et al.* (2017) 'Time-Optimized Isotope Ratio LC-MS/MS for High-Throughput Quantification of Primary Metabolites', *Analytical Chemistry*, 89(3). doi: 10.1021/acs.analchem.6b03731.

Gupta, A. *et al.* (2017) 'Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit', *Nature Biotechnology*. doi: 10.1038/nbt.3796.

He, F., Murabito, E. and Westerhoff, H. V. (2016) 'Synthetic biology and regulatory networks: Where metabolic systems biology meets control engineering', *Journal of the Royal Society Interface*. doi: 10.1098/rsif.2015.1046.

Jones, J. A., Toparlak, T. D. and Koffas, M. A. G. (2015) 'Metabolic pathway balancing and its role in the production of biofuels and chemicals', *Current Opinion in Biotechnology*. doi: 10.1016/j.copbio.2014.11.013.

Kim, S. K. *et al.* (2017) 'CRISPR interference-guided multiplex repression of endogenous competing pathway genes for redirecting metabolic flux in Escherichia coli', *Microbial Cell Factories*, 16(1). doi: 10.1186/s12934-017-0802-x.

Larson, M. H. *et al.* (2013) 'CRISPR interference (CRISPRi) for sequence-specific control of gene expression', *Nature Protocols*, 8(11), pp. 2180–2196. doi: 10.1038/nprot.2013.132.

Lee, J. H. and Wendisch, V. F. (2017) 'Production of amino acids – Genetic and metabolic engineering approaches', *Bioresource Technology*, pp. 1575–1587. doi: 10.1016/j.biortech.2017.05.065.

Lim, W. A. *et al.* (2013) 'Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.', *Cell*, 152(5), pp. 1173–1183. doi: 10.1016/j.cell.2013.02.022.

Lutz, R. and Bujard, H. (1997) 'Independent and tight regulation of transcriptional units in escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements', *Nucleic Acids Research*. doi: 10.1093/nar/25.6.1203.

Moser, F. *et al.* (2018) 'Dynamic control of endogenous metabolism with combinatorial logic circuits', *Molecular Systems Biology*. doi: 10.15252/msb.20188605.

Nandineni, M. R. and Gowrishankar, J. (2004) 'Evidence for an arginine exporter encoded by yggA (argO) that is regulated by the LysR-type transcriptional regulator argP in Escherichia coli', *Journal of Bacteriology*. doi: 10.1128/JB.186.11.3539-3546.2004.

Nascimento, M. M. (2018) 'Potential Uses of Arginine in Dentistry', *Advances in dental research*, pp. 98–103. doi: 10.1177/0022034517735294.

Nielsen, J. and Keasling, J. D. (2016) 'Engineering Cellular Metabolism', *Cell*, pp. 1185–1197. doi: 10.1016/j.cell.2016.02.004.

Park, S. Y. *et al.* (2017) 'Metabolic Engineering of Microorganisms for the Production of Natural Compounds', *Advanced Biosystems*, 2(1), p. 1700190. doi: 10.1002/adbi.201700190.

Qi, L. S. *et al.* (2013) 'Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression', *Cell*. doi: 10.1016/j.cell.2013.02.022.

Reisch, C. R. and Prather, K. L. J. (2015) 'The no-SCAR (Scarless Cas9 Assisted Recombineering) system for genome editing in Escherichia coli.', *Scientific reports*. Nature Publishing Group, 5(October), p. 15096. doi: 10.1038/srep15096.

Sander, T. *et al.* (2019) 'Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance', *Cell Systems*, 8(1), pp. 66-75.e8. doi: 10.1016/j.cels.2018.12.005.

Tian, T. *et al.* (2019) ' Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in Escherichia coli ', *ACS Synthetic Biology*. doi: 10.1021/acssynbio.8b00429.

Zhang, F., Carothers, J. M. and Keasling, J. D. (2012) 'Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids', *Nature Biotechnology*, 30(4), pp. 354–359. doi: 10.1038/nbt.2149.

Supplementary Material

a Biomass specific arginine yield ($Y_{arg,x}$)
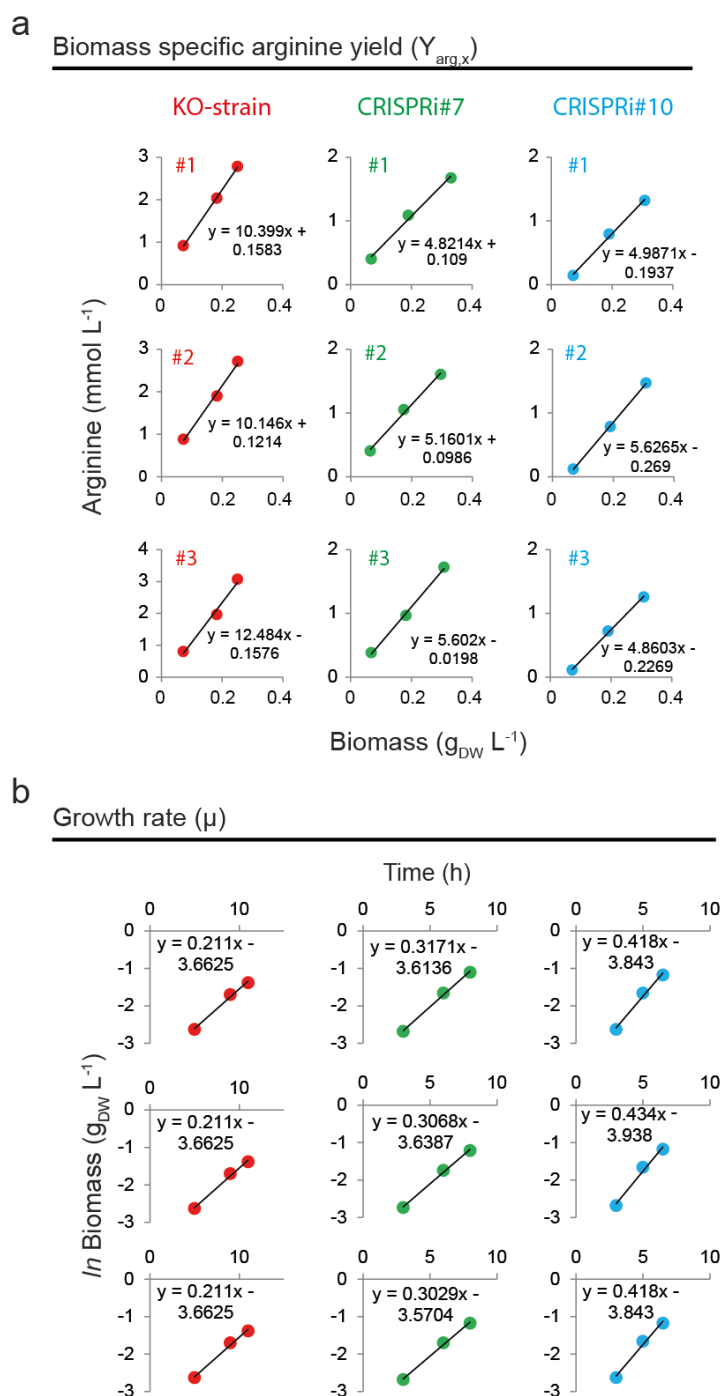
b Growth rate (μ)

Figure S1. Biomass specific arginine yields and growth rates that were used to calculate the specific arginine production rates shown in **Figure 4.** (a) Biomass specific arginine yield ($Y_{arg,x}$) of the KO-strain, the CRISPRi#7-strain and the CRISPRi#10-strain was calculated by linear regression of arginine concentration (mmol L^{-1}) and the biomass (g$_{DW}$ L^{-1}) for 3 shake flask cultures per strain. (b) Growth rates (μ) were calculated at the same time points by linear regression of *In* biomass (gDW L^{-1}) and the time (h). Specific arginine production rates follow as $q_{arg} = Y_{arg,x} \cdot \mu$.
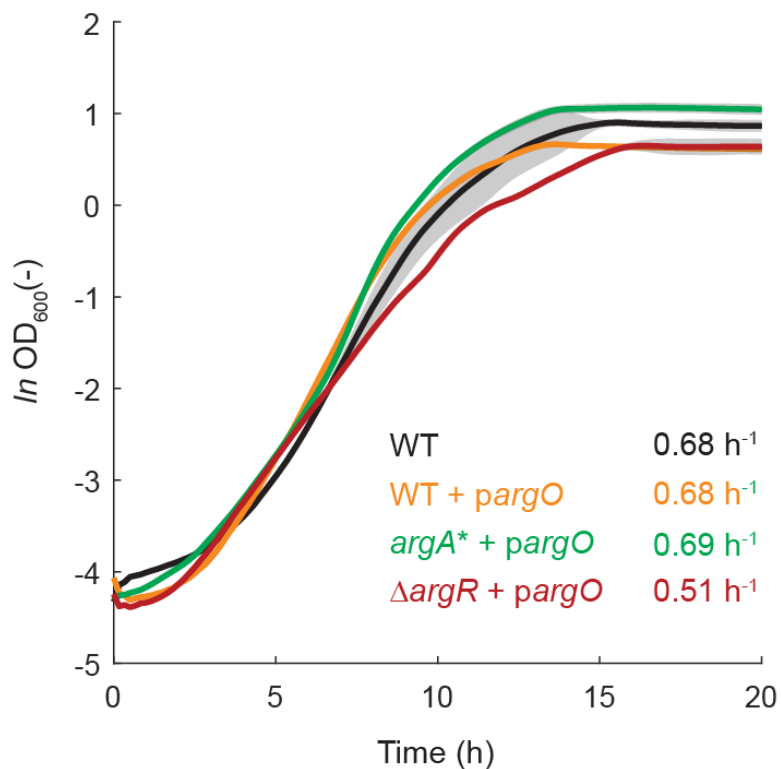
Figure S2. Growth of *E. coli* wildtype (WT), WT + p*argO*, *argA** + p*argO* and Δ*argR* + p*argO*. Cells were cultivated in glucose minimal medium (5 g L^{-1}) and optical density (OD600) was measured in 10 minute intervals in a plate reader. Growth rates (h^{-1}) were estimated by linear regression through 8 time points during exponential growth. Shown are means of n = 3 independent cultures.
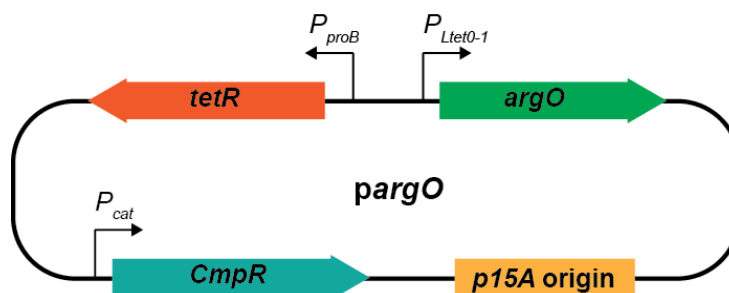
Figure S3: Plasmid map of p*argO*. The plasmid was constructed from a pdCas9 backbone (Addgene #44249) and *dCas9* was replaced by *argO* by gibson cloning. *argO* was expressed from the $P_{LtetO-1}$ promotor[1].

Table S1: Substrate specific arginine yield (g arginine/g glucose) and arginine concentration (g arginine/L) from wildtype, the KO-strain and the CRISPRi strains #7 and #10. Samples were taken from three independent shake flask cultures grown in glucose minimal medium (5 g L^{-1}).

	g arginine/g glucose	g arginine/L
WT	<0.01	<0.01
KO	0.16 ± 0.02	1.03 ± 0.06
CRISPRi#7	0.12 ± 0.02	0.96 ± 0.02
CRISPRi#10	0.11± 0.02	0.9 ± 0.02

Table S2: Growth rates of the wildtype, KO-strain, CRISPRi#7 and CRISPRi#10 after addition of orotate, aspartate, cytidine and uracil (each 0.5 mM). Growth rates (h^{-1}) were calculated by linear regression over an 60 minutes window at OD 0.3. Cells were cultivated in glucose minimal medium (5 g L^{-1}) and optical density (OD600) was measured in 10 minutes intervals in a plate reader. Shown are means of n = 3 cultures.

	- Oro/Asp/Cyt/Ura	+ Oro/Asp/Cyt/Ura
WT	0.65 ± 0.01	0.72 ± 0,02
KO	0.20 ± 0.00	0.46 ± 0.01
CRISPRi#7	0.51 ± 0.01	0.76 ± 0.03
CRISPRi#10	0.68 ± 0.01	0.78 ± 0.03

Table S3: Oligonucleotides used in this study

Oligonucleotide	Sequence (5'-3')	Description
*argA*_Forward	GGTCGAGGGATTCCGCCATTGTTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customized pKDsgRNA-*argA*(H15Y)
*argA*_Reverse	AATGGCGGAATCCCTCGACCGTGCTCAGTATCTCTATCACTGA	Forward primer used with **CPEC002** for amplification of fragment 1 for customized pKDsgRNA-*argA*(H15Y)
*argR*_Forward	ATTCTTCAATGGACTGGAGGGTTTTAGAGCTAGAAATAGCAAG	Forward primer used with **CPEC001** for amplification of fragment 1 for customzied pKDsgRNA-Δ*argR*
*argR*_Reverse	CCTCCAGTCCATTGAAGAATGTGCTCAGTATCTCTATCACTGA	Reverse primer used with **CPEC002** for amplification of fragment 2 for customzied pKDsgRNA-Δ*argR*
CPEC001	TTTATAACCTCCTTAGAGCTCGA	Reverse primer for amplification of fragment 1 for pKDsgRNA
CPEC002	CCAATTGTCCATATTGCATCA	Forward primer for amplification of fragment 2 for pKDsgRNA
argA(H15Y)_recomb	GTGGTAAAGGAACGTAAAACCGAGTTGGTCGAGGGATTCCGCTATTCAGTTCCCTATATCAATACCCACCGGGGAA	Recombination fragment for introduction of H15Y point mutation
Δ*argR*_recomb	ACAATAATGTTGTATCAACCACCATATCGGGTGACTTTCTCTGCCCCGTCGTTTCTGACGGCGGGGAAAATGT	Recombination fragment for *argR* deletion
EcR	ACTAGTATTATACCTAGGACTGAGCTAGC	Reverse primer for apmlification of customized pgRNA plasmids
EcF_*argR*#1	TAACATTATTTCAGCCTTCTTCgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#2	AATGTTGTATCAACCACCATATgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#3	AAAGAAGAGAAATTTAGCTCCCgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#4	ACCTCCAGTCCATTGAAGAATCgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#5	GCGATTTTAGAGCTGTTCGACCgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#6	TGCAAACAGTCAGCCCTGAAGAgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#7	GCTTCGcatAAGTCACCCGATAgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#8	TCAAAGCCTTGCTCCTGCAACGgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#9	TATCCAGCACCAGATTCTTCAAgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids
EcF_*argR*#10	CAGCTCTAAAATCGCTTCGTACgttttagagctagaaatagcaagttaaaataaggc	Forward primer for amplification of customized pgRNA plasmids

Chapter 3

Transcriptional Compensation Counteracts Genetic Perturbations and Shows the Benefit of Enzyme Overabundance

Timur Sander, Hannes Link

TS designed the study, performed experiments, performed data analysis and wrote the manuscript.

HL supervised the study.

Summary

Cellular metabolism is constantly faced with genetic perturbation that might arise from mutations or stochastic effects. Such perturbations can disturb the expression of a gene and cause bottlenecks in biosynthetic pathways. Here, we analyze how cells respond to decreases of gene expression. Therefore, we targeted genes from arginine, tryptophan and histidine biosynthesis with CRISPR interference and recorded the transcriptional response with *gfp*-reporter plasmids (arginine and tryptophan) and proteomics (histidine). Perturbing all three amino acid pathways caused up-regulation of the respective pathway enzymes in order to counteract the genetic bottleneck. Dynamic metabolite measurements revealed that the transcriptional compensation is initiated by a 30% decrease of the amino acid end-product which caused a transcriptional de-repression of the pathway enzymes. Previous experiments revealed significant instabilities of allosteric feedback mutants as a result of genetic perturbations (Chapter 1) and here we show that such perturbations cause oscillatory arginine level in the allosteric arginine mutant. Most likely, such oscillations cause a heterogenic transcriptional compensation in the allosteric mutant which was measured by flow cytometry. Moreover, flow cytometry data could show that the allosteric mutant with low enzyme level showed a transcriptional response even to mild perturbations, which are normally buffered by enzyme overabundance.

Introduction

Cells have to maintain robustness of metabolic functions in the face of variations in gene expression which might cause bottlenecks in biosynthetic pathways. Changes in gene expression can arise from internal perturbations like mutations, that directly affect expression or activity of an enzyme, or perturb regulatory circuits that usually control enzyme level (Stelling *et al.*, 2004). Further, genetic variation can derive from stochastic effects of gene expression which can cause a high cell-to-cell variability in a growing population (Elowitz *et al.*, 2002; Ozbudak *et al.*, 2002). Such fluctuations of gene expression can be an advantage under different conditions as shown for an increased fitness of *B. subtilis* under fluctuating environments (Çağatay *et al.*, 2009) or in the resistance towards drugs as shown for *E. coli* (Balaban *et al.*, 2004). However, beside this advantages cells obviously have to maintain performance of metabolic processes in a given condition. A third source for decreases in gene expression are artificial modifications in the scope of metabolic engineering. Nowadays different techniques like CRISPR interference are established tools to tune gene expression (Qi *et al.*, 2013), for example to knockdown competing pathways (Kim *et al.*, 2016) or disturb feedback regulation (Cress *et al.*, 2017). It is of particular interest how cells respond to such artificial bottlenecks on the molecular scale. Therefore, we ask which mechanisms provide robustness towards genetic perturbations in order to maintain metabolic reactions?

A simple protective strategy against genetic perturbations is redundancy of metabolic enzymes which can be described as genetic compensation (El-Brolosy and Stainier, 2017). After loss of function of an enzyme, alternative enzymes can get upregulated that are capable to catalyze the same chemical reaction. A metabolic network analysis in yeast revealed that up to 28 % of all metabolic gene knockouts that do not cause a phenotype, are compensated by alternative enzymes (Papp, Pál and Hurst, 2004). Beside expressing alternative enzymes, rerouting flux through alternative pathways is also a strategy to circumvent genetic perturbations. It is known that many organisms evolved alternative routes for catabolic and anabolic reactions. A prominent example for alternative metabolic routes, are the acetyl-CoA oxidizing pathways tricarboxylic acid cycle and glyoxylate shunt (Ahn *et al.*, 2016). A third strategy to face genetic fluctuations are enzyme reserves, so-called "enzyme overabundance". Recently we demonstrated how *E. coli* adjusts overabundant enzyme level in amino acid metabolism by an interplay between metabolic feedbacks (Sander *et al.*, 2019). A CRISPR interference mediated genetic perturbation of three amino acid pathways revealed instabilities of mutants lacking overabundance compared to wild-type cells. Different studies in different organisms suggest that expressing enzymes at higher level than actually needed might be a general strategy to make metabolism more robust against internal and external perturbations (Basan *et al.*, 2015; O'Brien,

Utrilla and Palsson, 2016; Mori *et al.*, 2017a). These studies demonstrated a phenotype for cells lacking enzyme overabundance either at the level of growth or theoretically. However, so far, no study was able to show the benefit of enzyme overabundance on the molecular scale.

In this study we use *gfp*-promotor fusions to analyze the response to perturbations of gene expression. Perturbing arginine, tryptophan and histidine biosynthesis revealed a compensatory up-regulation of the respective pathway enzymes. For the arginine pathway we showed that transcriptional compensation is initiated by a slight decrease of the amino acid end-product. Systematic comparison between the wild-type and the allosteric mutant *argA**, which was previously shown to be more sensitive towards genetic perturbations revealed a critical role of allosteric regulation for transcriptional compensation. The allosteric mutant failed to adjust arginine homeostasis which is the basis for a homogenic transcriptional response and the missing enzyme overabundance in this strain caused an increased sensitivity towards the genetic perturbations.

Results & Discussion

Transcriptional Compensation Counteracts Genetic Perturbations

To analyze the mechanistic response to a genetic perturbation we transformed wildtype strains with CRISPR interference plasmids targeting genes from arginine (*argE*) and tryptophan (*trpC*) biosynthesis. The plasmids pCRISPRi-*argE* and pCRISPRi-*trpC* are based on the pNUT1533 plasmid (Resource table), expressing the respective sgRNA constitutively and the dCas9 protein from an IPTG inducible promotor. In order to measure changes of gene expression in the perturbed pathways we additionally transformed the strains with fluorescent transcriptional reporter plasmids. These reporter plasmids express *gfp* from an arginine (pP*argA-gfp*) and a tryptophan promotor (pP*trpL-gfp*). This resulted in the strains WT + pCRISPRi-argE + pP*argA-gfp* and WT + pCRISPRi-trpC + pP*trpL-gfp* (Figure 1a) to which we will further refer to as WT-arg and WT-trp. Both strains were cultivated in a plate reader and CRISPRi was induced after 2 hours. After induction of CRISPRi in the WT-arg strain, we observed a growth phenotype after 1 hour (Figure 1b). The delay between induction of CRISPRi and the reduced growth is presumably caused by the fact that the targeted enzymes need to be diluted by growth in order to observe a phenotype. Measuring the transcriptional response during this perturbation revealed strong changes in gene expression of the arginine promotor (Figure 1c, d). The gfp expression (GFP OD^{-1}) from an *argA* promotor was 4-fold higher after induction of the genetic bottleneck (Figure 1c) and the promotor activity (GFP OD^{-1} h^{-1}) changed in the same range (Figure 1d). Interestingly, the

promotor activity increased already ~20 minutes after induction which is significantly faster compared to the response observed for cellular growth (Figure 1b)
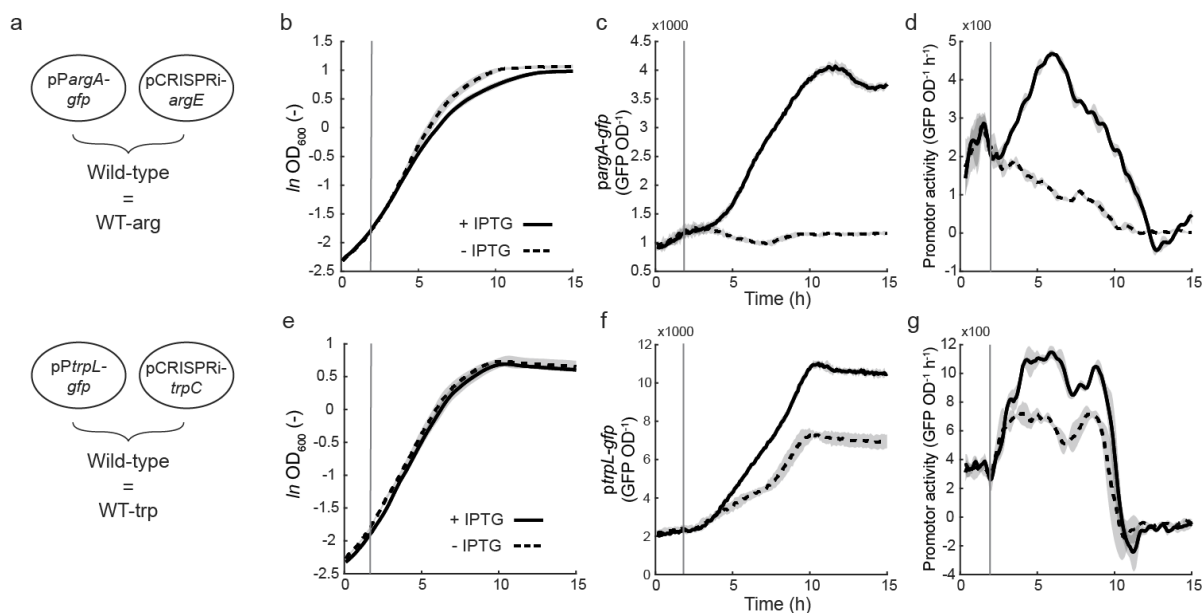

Figure 1. Transcriptional compensation in arginine and tryptophan biosynthesis pathways.

(a) Wild-type was transformed with the *gfp*-reporter plasmid pP*argA-gfp* and IPTG-inducible pCRISPRi-*argE* plasmid to assay the arginine pathway, resulting in WT-arg. The plasmids pP*trpL-gfp* and IPTG-inducible pCRISPRi-*trpC* were used to assay the tryptophan pathway resulting in WT-trp.

(b) Growth of WT-arg with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Cells were grown in a plate reader and curves show mean of n = 3 replicates. Grey line indicates supplementation of IPTG. Standard deviation is indicated with the grey shadow.

(c) *gfp* expression (GFP OD^{-1}) from an *argA* promoter in WT-arg with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Data derive from the experiment shown in panel b.

(d) Promotor activity of an *argA* promoter (GFP OD^{-1} h^{-1}) in WT-arg with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Data derive from the experiment shown in panel b.

(e), (f) and (g) same as (a), (b), (c) but with the strain WT-trp.

The increase of gene expression after the genetic perturbation was also observed for WT-trp (Figure 1f, g). Perturbing *trpC* elevated expression of *gfp* from a *trpL* promotor and promotor activity increased by 60%. This is a less severe increase compared to WT-arg which is presumably due the fact that induction of CRISPRi against *trpC* had no effect on growth (Figure 1e). However, the enhanced transcription in the WT-trp strain indicates that the observed response is not caused by decreased

cellular growth. Arginine and tryptophan biosynthesis are both regulated by transcription factors (ArgR and TrpR) which repress expression of all pathway genes by sensing amino acid end-product level. Therefore, we wanted to test if this strong transcriptional response is limited to pathways that are directly controlled by specific transcription factors.

Transcriptional Compensation in the Histidine Pathway

Transcriptional regulation in amino acid metabolism is mediated by transcription factors or by transcriptional attenuation. A typical example of transcriptional attenuation in amino acid metabolism is the histidine pathway. The eight biosynthetic histidine enzymes are arranged in a single operon and we choose *hisB* as a target for CRISPRi perturbation since it is located in the middle of the operon (Figure 2a). Unfortunately, transcriptional *gfp* reporter plasmids that are sensitive to attenuation were not available and therefore we decided to directly measure protein level by LC-MS/MS in a wild type strain expressing CRISPRi against *hisB* (WT + gRNA-hisB) and compare them to a CRISPRi-control strain expressing a non-targeted sgRNA (WT + gRNA-ctrl). Protein data of WT + gRNA-hisB revealed 3-fold overexpression of the enzymes HisG and HisD which are coded upstream of the perturbed *hisB* gene. Interestingly, level of the third enzyme which is coded upstream of *hisB*, are decreased by 5-fold. This is presumably caused by polar effects of dCas9, which not only blocks expression of the target gene, but also of the adjacent genes. The CRISPRi target *hisB* and also the down-stream genes were strongly repressed, which provides important information about the general nature of CRISPRi. dCas9 not only blocks transcription of the target gene but also of the down-stream genes in an operon structure as already observed by others (Wu *et al.*, 2015; Kim *et al.*, 2017). However, we could confirm the observations made on the transcriptional level for WT-arg and WT-trp also on the protein level for WT + gRNA-hisB: Perturbing the expression of a gene in a biosynthetic pathway elevates the expression of the surrounding genes (If not located down-stream of the target in an operon). We hypothesize that this might be a general strategy to compensate bottlenecks in a biosynthetic pathway, which can be caused by decreases in enzyme activity and abundance. Next we wanted to analyze the molecular mechanism and origin of this transcriptional compensation and decided to continue with WT-arg as case study since perturbation of arginine biosynthesis revealed the strongest effects.
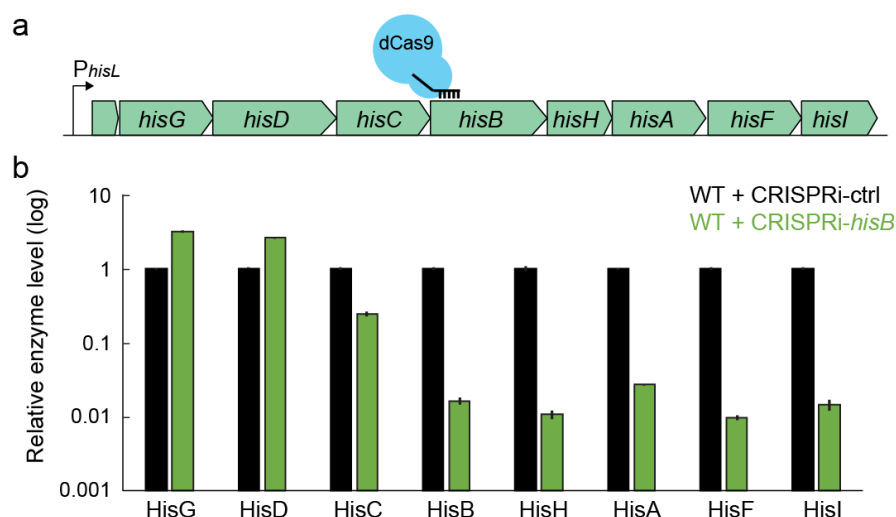
Figure 2. Level of histidine enzymes in response to perturbation of *hisB*

(a) Genes coding for enzymes of histidine *de novo* biosynthesis are organized in an operon. The schematic dCas9 illustration shows the position of binding on the *hisB* gene.

(b) Enzyme level of HisG, HisD, HisC, HisB, HisH, HisA, HisF and HisI in wild-type with pdCas9 and pgRNA-ctrl (WT + gRNA-ctrl; black), respectively pgRNA-hisB (WT + gRNA-hisB, green). Samples were taken at an OD_{600} of 0.5 from glucose minimal medium shake flask cultures and were measured by liquid chromatography – mass spectrometry. Enzyme level shown mean of n = 3 independent replicates.

Transcriptional Compensation Differs in the Wild-type and the Allosteric Mutant *argA**

We decided to include the allosterically dysregulated strain *argA** in the further analysis, since we could show in the previous chapter, that feedback resistant mutants were more sensitive towards genetic perturbations. To get a better insight into the mechanistic response to genetic perturbations we compared WT-arg to the *argA** mutant, which was also transformed with pCRISPRi-*argE* and pP*argA-gfp* resulting in *argA**-arg. Expression of GFP and promotor activity were reduced by ~50% in the uninduced *argA**-arg strain compared to the WT-arg, which corresponds to 50% decreased arginine enzymes in *argA** reported in chapter 1 (Figure 3b and Figure 2a in Chapter 1). After induction of CRISPRi, WT-arg showed a similar response as shown in figure 1 with slightly decreased growth and an ~4-fold increased promotor activity (Figure 3a, b, c). The allosteric mutant *argA**-arg showed a stronger growth defect in response to CRISPRi-*argE* induction which matches previous observations in chapter 1 (Figure 3a).
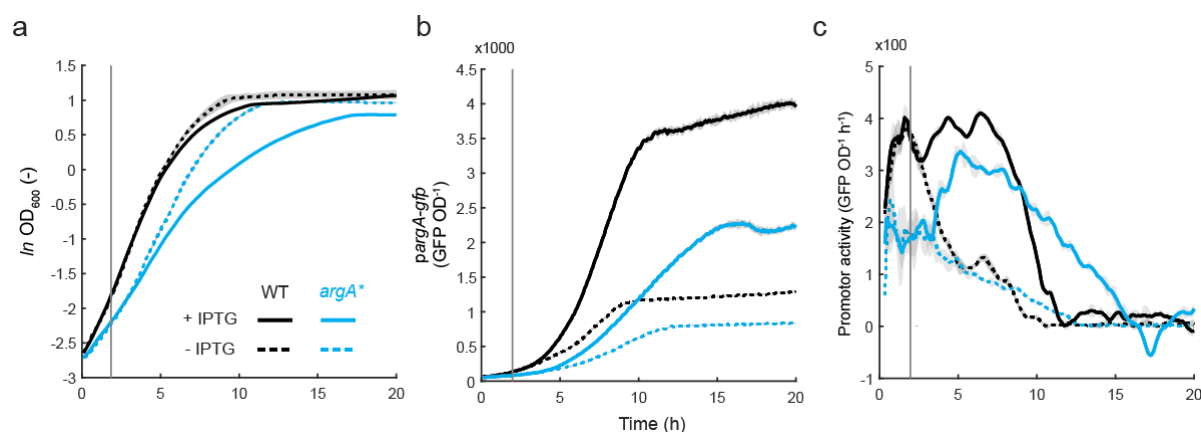
Figure 3. Transcriptional compensation of the wild-type compared to the allosteric mutant *argA* in response to perturbation of *argE*.

(a) Growth of WT-arg (black) and *argA**-arg (blue) with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Cells were grown in a plate reader and curves show mean of n = 3 replicates. Grey line indicates supplementation of IPTG. Standard deviation is indicated with the grey shadow.

(b) *gfp* expression (GFP OD^{-1}) from an *argA* promotor in WT-arg (black) and *argA**-arg (blue) with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Cells were grown in a plate reader and curves show mean of n = 3. Grey line indicates supplementation of IPTG. Standard deviation is indicated with the grey shadow.

(c) Promotor activity of an *argA* promotor (GFP OD^{-1} h^{-1}) in WT-arg (black) and *argA**-arg (blue) with (solid line) and without (dashed line) induction by 0.1 mM IPTG. Cells were grown in a plate reader and curves show mean of n = 3 replicates. Grey line indicates supplementation of IPTG. Standard deviation is indicated with the grey shadow.

Transcriptional activity also increased in the *argA**-arg strain after induction, however GFP expression and promotor activity were only 3-fold higher compared to the uninduced strain (Figure 3b, c). Thus, transcriptional compensation was stronger in WT-arg (4-fold increase) compared to *argA**-arg (3-fold increase). Furthermore, the time until the transcriptional compensation occurred was delayed by approximately 1 hour in the *argA**-arg compared to the WT-arg. These findings indicate a role of allosteric feedback regulation within this mechanism of transcriptional compensation. Since allosteric feedback regulation is known to control end-product homeostasis we suggested that the compensatory response might be caused by changes in the respective amino acid level. To test this hypothesis, we dynamically measured arginine level in the wild-type and the *argA** mutant in response to CRISPRi against *argE*.

Interference of *argE* Causes Intracellular Arginine Oscillation in the Allosteric Mutant *argA**

To measure intracellular arginine level in response to a genetic perturbation, we cultivated the Wild-type + pCRISPRi-*argE* and *argA** + pCRISPRi-*argE* in shake flasks on glucose minimal medium and collected samples of intracellular metabolites during 2 hours (Figure 4a). The first sample was collected at time-point 0 before induction with IPTG. Cells were induced directly after this measurement and further samples post-induction were taken after 5, 10, 20, 30, 45, 60, 90 and 120 minutes. The arginine level shown in figure 4 were quantified by LC-MS/MS and derive from two independent experiments on two different days (Day 1: solid lines; Day 2: dashed lines). For the wildtype we observed a decrease of intracellular arginine by ~1.5-fold after 10 minutes post-induction and the arginine level were constantly kept at this lower level in the subsequent 2 hours (Figure 4b). However, the *argA** mutant showed oscillatory arginine level in the first hour after induction (Decrease of arginine by 2.5-fold followed by an increase of 3.5-fold and again decreased by 3.5-fold). Thus, the *argA** mutant failed to adjust a homeostatic arginine level in the first hour after induction. This might be the reason for the delayed transcriptional compensation, which occurred ~1 hour later in the *argA** mutant compared to the wildtype (Figure 3c).
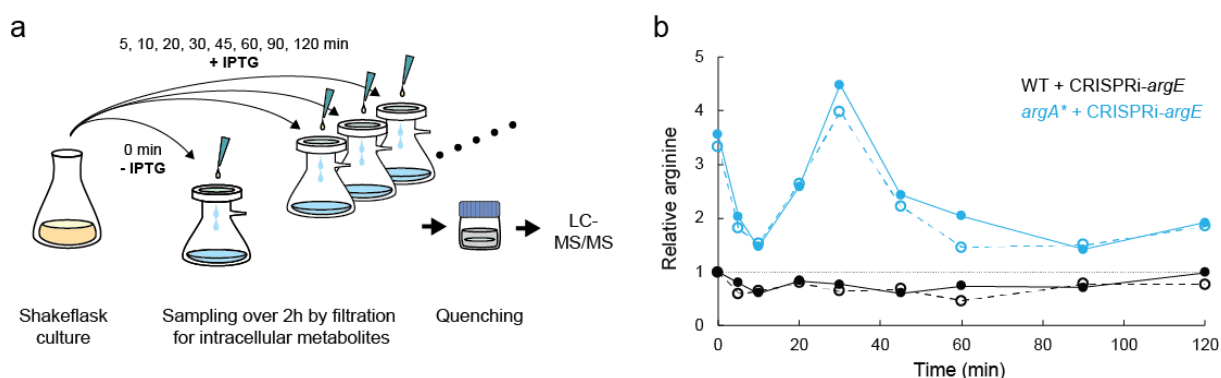

Figure 4. Dynamic arginine level of the wild type and the *argA mutant after perturbation of *argE*.**

(a) Experimental set-up for dynamic measurement of intracellular arginine level. Cells were cultivated in glucose minimal medium in shake flasks at 37°C. Filter samples were taken before (timepoint 0 min) and after (5, 10, 20, 30, 45, 60, 90 and 120 min) induction with 0.1 mM IPTG and were analyzed by LC-MS/MS.

(b) Relative arginine level of WT + CRISPRi-*argE* (black) and *argA** + CRISPRi-*argE* (blue) over 2 hours after induction with 0.1 mM IPTG. Shown are two independent experiments performed on two different days (day 1: filled circles and solid lines; day 2: empty circles and dashed lines). Dashed grey line indicates arginine level in wild-type before induction.

In wildtype cells such oscillations are prevented by allosteric feedback regulation as suggested by the homeostatic arginine level. Thus, transcriptional compensation is mediated by a decrease of intracellular arginine, which presumably decreases the activity of the transcriptional repressor ArgR. The decreased ArgR activity causes transcriptional up-regulation of the arginine genes which, at least to a certain extent, eliminates the bottleneck. Our data indicate, that in order to ensure a proper transcriptional compensation, also allosteric feedback inhibition is involved by finetuning the arginine end-product level.

Although these data have the potential to explain how wild type cells react to decreased enzyme level, it is still ambiguous why allosteric feedback mutants show such a strong growth phenotype in response to a genetic bottleneck (Chapter 1, Figure 5; Chapter 3, Figure 3a). The dynamic measurement of transcriptional activity in combination with the arginine level revealed instabilities especially in the first hour after induction of CRISPR interference against *argE*. However, the strong growth defect over 20 hours can barely be explained by this data, especially because arginine level neither dropped below a critical threshold nor showed an excessive increase that could cause flux burden. Importantly, plate reader experiments with GFP-reporter and dynamic metabolite measurements, only allow claims on the population level. The effects on transcription and arginine level might be more significant on the single cell level. This hypothesis is supported by the oscillatory arginine dynamics in the first hour after induction. Induction of the CRISPRi bottleneck by addition of IPTG might synchronize the cellular response which disappears over time due to stochasticity on the population level. If such single cell effects would be relevant during transcriptional compensation of a metabolic bottleneck, we would expect increased heterogeneity in perturbed cultures.

Genetic Perturbations Causes Cell-to-Cell Heterogeneity and Prove the Benefit of Enzyme Overabundance

To probe transcriptional compensation in single cells, we measured *gfp*-expression from an arginine promotor (pP*argA-gfp*) by flow cytometry. Therefore, we used the strains WT-arg and *argA**-arg (both harboring pCRISPRi-argE and pP*argA-gfp*) and compared them to the respective control strains WT-ctrl and *argA**-ctrl (both harboring pCRISPRi-ctrl and pP*argA-gfp*). To point out the differences between wild-type and *argA** mutant we used the uninduced CRISPRi system, because it was previously shown that basal expression of dCas9 only perturbed growth of the *argA** mutant, whereas the wild-type was almost not affected (Chapter 1). Flow cytometry data showed, that perturbing *argE*-expression caused heterogenic gene expression in wild-type and the *argA** compared to the respective control strains (Figure 5). WT-arg and *argA**-arg both showed a wider distribution than WT-ctrl and *argA**-ctrl respectively. However, the heterogeneity of *argA**-arg was more significant

compared to WT-arg and even divided in two sub-populations. We hypothesize that the heterogeneity of the *argA** mutant derives from the previously described transcriptional compensation. As indicated by the arginine dynamics (Figure 4b), perturbing gene expression in *argA** causes transcriptional compensation and subsequent oscillatory arginine level. Arginine oscillations on a single cell level are presumably stronger compared to the population level and might cause a wide range of promotor activity within a population.
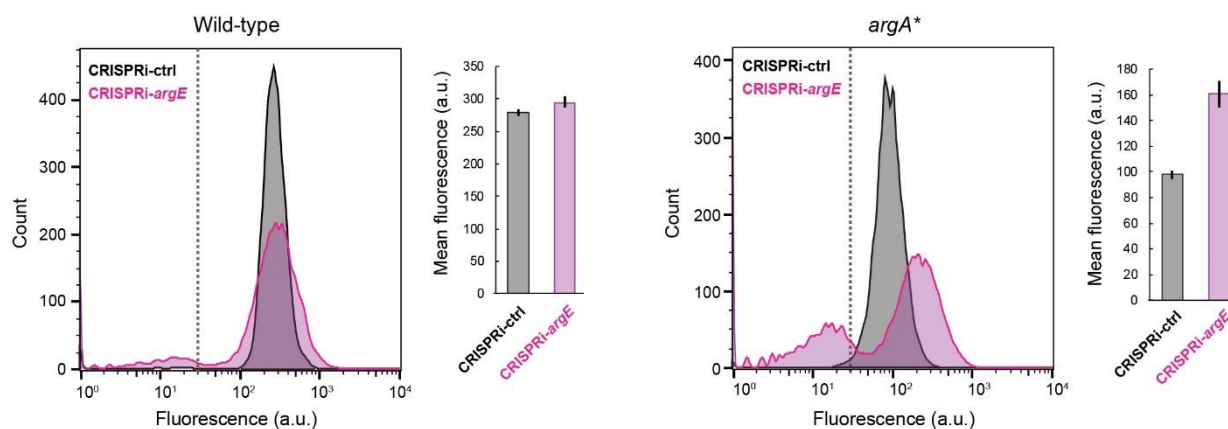

Figure 5. Single-cell fluorescence of *gfp* expressed from an *argA* promotor in the wild-type and the *argA* mutant after perturbation of *argE*.

Histograms show fluorescence of WT-arg and *argA**-arg both harboring pCRIPSPRi-*argE* and pP*argA-gfp* (purple), each compared to the respective control strain harboring pCRIPSPRi-*ctrl* + pP*argA-gfp* (grey). Bar-plots show mean-fluorescence of the same strains including all cells above the gate threshold for green fluorescence (indicated with the dashed line). Fluorescence was measured by flow-cytometry for 10,000 single cells per strain. Error bars show standard deviation of n = 3 independent cultivations.

The differences between the wid-type and *argA** are directly related to the previous reports about enzyme overabundance: The mild perturbation of *argE* is buffered in the wild-type by overabundant enzyme level and therefore mean expression of *gfp* from an arginine promotor was almost not affected (increase by 5%) (Figure 5). However, enzyme levels in the *argA** mutant are catalyzing reactions with full capacity and therefore also mild perturbations caused a transcriptional response. Consequently, mean *gfp* expression from an arginine promoter was increased by 66% due to mild *argE* perturbation in the *argA** mutant. These results demonstrate the benefit of enzyme overabundance on the molecular scale: Wildtype cells can react to genetic fluctuations by increasing activity of the

enzyme reserves, whereas allosteric mutants need to respond on the transcriptional level by increasing gene expression. Thus, in combination with the dynamic arginine level, these data can explain the significant growth phenotype of the allosteric mutants in response to genetic perturbations. The mutants are presumably constantly faced with oscillatory gene expression and fail to adjust homeostasis.

Conclusions

Here we show the mechanistic principles of transcriptional compensation in response to genetic perturbations for 3 amino acid pathways. After reduction of an enzyme below a critical threshold the cell senses a decrease of the amino acid end-product, which results in a de-repression of transcription. This de-repression leads to up-regulation of the pathway enzymes in order to counteract against the bottleneck. Importantly, wildtype cells adjust a new homeostatic state with constantly lower amino acid end-product level and higher enzyme expression. Our data indicate that this mechanism requires a sensitive interplay between allosteric and transcriptional feedback regulation and that allosteric mutants fail to adjust a homeostatic transcriptional response. Moreover, the flow cytometry data indicate that overabundant enzyme level in wild-type cells can buffer genetic perturbations, whereas the allosteric mutant, which has no enzyme reserves also responds to mild genetic perturbations by transcriptional compensation. Thus, by measuring transcriptional compensation we were able to detect a benefit of enzyme overabundance at the molecular scale. So far, the advantage of enzyme overabundance was only shown by growth reduction on the population level (Mori *et al.*, 2017; Christodoulou *et al.*, 2018; Sander *et al.*, 2019) or by computational studies (O'Brien, Utrilla and Palsson, 2016).

It needs to be clarified if the heterogeneity in the allosteric *argA*-mutant indeed derived from oscillatory gene expression or might be due to static sub-populations. Also plasmid loss might be a reason for single cells that show no or very low level of fluorescence. Therefore, we constructed a pP*argA-gfp* reporter plasmid that additionally expresses *rfp* under a constitutive promotor. Analysis of this plasmid system in future experiments with microscopic time-lapse imaging in microfluidic chambers could exclude plasmid loss or cell division effects.

Beside the general understanding of a biological systems, these data are also important for the field of metabolic engineering. CRISPRi is a widely used tool to manipulate gene expression in metabolic pathways. Here we show that cells counteract such perturbations on the transcriptional level. Such transcriptional compensation causes up-regulation of the surrounding enzymes which can have an

effect on overall fitness of a host organism due to protein burden or overflow metabolism. Especially in combination with allosteric dysregulation, which is a common strategy in metabolic engineering, gene knockdown methods can strongly disturb metabolism and physiology of a cell.

Material and Methods

Strains and Culture

All strains in this study were derived from *E. coli* MG1655 (DSMZ No. 18039) and are listed in the general resource table. Chemically competent *E. coli* TOP10 (One Shot^{TM} TOP10, Invitrogen Cat#C404003) were used for cloning.

All cultivations were performed in M9 minimal medium with 5 g L^{-1} glucose. The M9 medium consisted of the following components (per liter): 7.52 g Na$_2$HPO$_4$ 2 H$_2$O, 5 g KH$_2$PO$_4$, 1.5 g (NH$_4$)$_2$SO$_4$, 0.5 g NaCl. The following components were sterilized separately and then added (per liter of final medium): 1 ml 0.1 M CaCl$_2$, 1 ml 1 M MgSO$_4$, 0.6 ml 0.1 M FeCl$_3$, 2 ml 1.4 mM thiamine-HCL and 10 ml trace salts solution. The trace salts solution contained (per liter): 180 mg ZnSO$_4$ 7 H$_2$O, 120 mg CuCl$_2$ 2 H$_2$O, 120 mg MnSO$_4$ H$_2$O, 180 mg CoCl$_2$ 6 H$_2$O. Where appropriate, 34 µg mL^{-1} chloramphenicol, 15 µg mL^{-1} gentamycine, 50 µg mL^{-1} spectinomycin or 100 µg mL^{-1} ampicillin was added. For plate reader experiments, LB pre-culture in 96-deep-well format plates were inoculated from glycerol stocks and grown to an exponential stage. From this first pre-culture a second M9 pre-culture in 96-deep-well plates was inoculated 1:100 and incubated overnight at 37 °C under shaking. Finally, 96-well flat transparent plates (Greiner Bio-One International) containing 150 µl M9 minimal medium were inoculated 1:150 from the overnight culture. Online measurements of optical density at 600 nm (OD$_{600}$) were performed at 37°C with shaking in a plate reader (Tecan Trading AG, Switzerland). For cultivations in shake flask, 5 ml LB pre-culture in cultivation tubes were inoculated from glycerol stocks and grown to an exponential stage. From this first pre-culture, 5 ml of a second M9 batch pre-culture (5 g L^{-1} glucose) in cultivation tubes was inoculated 1:100 and incubated overnight at 37°C in a rotary shaker. Finally, a 500 ml shake flask containing 25 ml M9 minimal medium (5 g L^{-1} glucose) were inoculated 1:150 from the overnight culture, and incubated at 37 °C under shaking at 220 rpm.

CRISPR interference

CRISPR interference of *hisB* was performed according to Larson et al. 2013 (Larson *et al.*, 2013) with the plasmids pdCas9 (Addgene #44249) and pgRNA (Addgene #44251). The gRNA plasmid was

customized by site-directed mutagenesis with the forward primers EcF_*hisB*, carrying the 20-nt base-pairing sequence (protospacer) and the reverse primer EcR.

For targeting argE and trpC we used single plasmid (pNUT1533) expressing the respective sgRNA from a constitutive and the dCas9 protein from an IPTG inducible Ptac promotor (Sander *et al.*, 2019). Therefore, the protospacer regions were customized by site-directed mutagenesis using the primer EcF_argEmm5, respectively EcF_trpC and the reverse primer EcR. For induction of dCas9 expression we supplemented 0.1 mM IPTG to the medium.

Sampling for measurement of intra cellular arginine

Shake flask cultivations on M9 glucose were performed as described above. For measurement of intracellular arginine, 2 mL culture aliquots at the indicated time points 0, 5, 10, 20, 30, 45, 60, 90 and 120 min were vacuum-filtered on a 0.45 µm pore size filter (HVLP02500, Merck Millipore). Filters were immediately transferred into 40:40:20 (v-%) acetonitrile/methanol/water at -20°C for extraction. Extracts were centrifuged for 15 minutes at 13,000 rpm at -9 °C. Centrifuged extracts were mixed with ^{13}C-labeled internal standard and analyzed by LC-MS/MS (Guder *et al.*, 2017).

Measurement of intra-cellular arginine by LC-MS/MS

LC-MS/MS analysis was performed with an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies) as described previously (Guder *et al.*, 2017). An Agilent 1290 Infinity II UHPLC system (Agilent Technologies) was used for liquid chromatography. Temperature of the column oven was 30°C, and the injection volume was 3 µL. LC solvents A were water with 10 mM ammonium formate and 0.1% formic acid (v/v) for acidic conditions. LC solvents B were acetonitrile with 0.1% formic acid (v/v). LC columns were an Acquity BEH Amide (30 x 2.1 mm, 1.7 µm) for acidic conditions. The gradient for acidic conditions was: 0 min 90% B; 1.3 min 40 % B; 1.5 min 40 % B; 1.7 min 90 % B; 2 min 90 % B. Relative quantification of intracellular arginine concentrations was based on the ratio of ^{12}C and ^{13}C peak heights and normalized to the OD.

Proteomics

Shake flask cultivations on M9 glucose were performed as described above. Cells were grown to an optical density (OD_{600}) of 0.5 and 2 mL culture aliquots were transferred into 2 ml reaction tubes and washed two times with PBS buffer (0.14 mM NaCl, 2.7 mM KCL, 1.5 KH_2PO_4, 8.1 Na_2HPO_4). Cell pellets were resuspended in 300 µl lysis buffer containing 100 mM ammonium bicarbonate, 0.5 % sodium laroyl sarcosinate (SLS) and 5 mM Tris(2-carboxyethyl)phosphine (TCEP). Cells were lysed by 5 minutes incubation at 95 °C and ultra-sonication for 10 seconds (Vial Tweeter, Hielscher). Cells were again incubated for 30 minutes at 90 °C followed by alkylation with 10 mM iodoacetamide for 30 minutes

at 25 °C. To clear the cell lysate, samples were centrifuged for 10 minutes at 15,000 rpm and the supernatant transferred into a new tube. Proteins in the cell lysates were digested with 1 µg trypsin (Promega) overnight at 30 °C. To remove the SLS by precipitation, trifluoroacetic acid (TFA) was added to a final concentration of 1.5 % and rested at room temperature for 10 minutes. Samples were centrifuged for 10 minutes at 10,000 rpm and the supernatant used for C18 purification. The peptide purification was performed using the C18 microspin columns according to the manufactors instructions (Harvard Apparatus). Eluted peptide solutions were dried and resuspended in 0.1 % TFA. The concentration of peptides in the samples was measured with a colorimetric peptide assay (Pierce™ Quantitative Colorimetric Peptide Assay, Thermo Fischer Scientific). Analysis of peptides was performed by liquid chromatography-mass spectrometry, carried out on a Q-Exactive Plus instrument connected to an Ultimate 3000 RSLC nano with a Prowflow upgrade and a nanospray flex ion source (Thermo Scientific). Peptide separation was performed on a reverse-phase HPLC column (75 µm x 42 cm) packed in-house with C18 resin (2.4 µm, Dr. Maisch GmbH, Germany). The following separating gradient was used: 98 % solvent A (0.15% formic acid) and 2 % solvent B (99.85 acetonitrile, 0.15 % formic acid) to 25 % solvent B over 105 minutes and to 35 % solvent B for additional 35 minutes at a flow rate of 300 nl/min. The data acquisition mode was set to obtain one high resolution MS scan at a resolution of 70,000 full width at half maximum (at m/z 200) followed by MS/MS scans of the 10 most intense ions. To increase the efficiency of MS/MS attempts, the charged state screening modus was enabled to exclude unassigned and singly charged ions. The dynamic exclusion duration was set to 30 seconds. The ion accumulation time was set to 50 ms for MS and 50 ms at 17,500 resolution for MS/MS. The automatic gain control was set to 3x106 for MS survey scans and 1x105 for MS/MS scans. Label-free quantification (LFQ) of the data was performed using Progenesis QIP (Waters), and for MS/MS searches of aligned peptide features MASCOT (v2.5, Matrix Science) was used. The following search parameters were used: full tryptic search with two missed cleavage sites, 10ppm MS1 and 0.02 Da fragment ion tolerance. Carbamidomethylation (C) as fixed, oxidation (M) and deamidation (N,Q) as variable modification. Progenesis outputs were further processed with SafeQuant.

Flow Cytometry

Cell-to-cell variability of pP*argA-gfp* activity in wild-type + CRISPRi-*argE* and *argA** + CRISPRi-*argE* was assayed by flow cytometry. Strains for flow cytometry were cultivated in three independent shake flasks (100 ml) containing 10 ml M9 minimal medium (5 g L^{-1} glucose; 50 µg mL^{-1} kanamycin and 15 µg mL^{-1} gentamycine) as described in Strains and Culture. After reaching an OD between 0.5 and 0.8 cells were diluted 1:2000 in tethering buffer (10 mM KH$_2$PO$_4$, 100 µM EDTA, 1 µM L-methionine and 10 mM lactic acid, pH=7.0) and fluorescence was measured with BD LSRFortessa SORP cell analyser (BD Biosciences, Germany). 488-nm lasers, 600 long pass and a 520/30 band pass

filters were used for detection of green fluorescence. Per sample, fluorescence of 10,000 single cells was measured. Before the measurements, cell aggregates were dispersed by vigorous mixing. BD FACSDiva software version 8.0 (BD Biosciences, NJ, USA) and FlowJo v10.4.1 (FlowJo LLC, Ashland, OR, USA) were used for analysis of the acquired data.

References

Ahn, S. *et al.* (2016) 'Role of glyoxylate shunt in oxidative stress response', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M115.708149.

Balaban, N. Q. *et al.* (2004) 'Bacterial persistence as a phenotypic switch', *Science.* doi: 10.1126/science.1099390.

Basan, M. *et al.* (2015) 'Overflow metabolism in Escherichia coli results from efficient proteome allocation', *Nature*, 528(7580), pp. 99–104. doi: 10.1038/nature15765.

Burg, J. M. *et al.* (2016) 'Large-scale bioprocess competitiveness: the potential of dynamic metabolic control in two-stage fermentations', *Current Opinion in Chemical Engineering*, pp. 121–136. doi: 10.1016/j.coche.2016.09.008.

Çağatay, T. *et al.* (2009) 'Architecture-Dependent Noise Discriminates Functionally Analogous Differentiation Circuits', *Cell*. doi: 10.1016/j.cell.2009.07.046.

Christodoulou, D. *et al.* (2018) 'Reserve Flux Capacity in the Pentose Phosphate Pathway Enables Escherichia coli's Rapid Response to Oxidative Stress', *Cell Systems*. doi: 10.1016/j.cels.2018.04.009.

Cress, B. F. *et al.* (2017) 'CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production', *Microbial Cell Factories*, 16(1). doi: 10.1186/s12934-016-0623-3.

El-Brolosy, M. A. and Stainier, D. Y. R. (2017) 'Genetic compensation: A phenomenon in search of mechanisms', *PLoS Genetics*. doi: 10.1371/journal.pgen.1006780.

Elowitz, M. B. *et al.* (2002) 'Stochastic gene expression in a single cell', *Science*. doi: 10.1126/science.1070919.

Guder, J. C. *et al.* (2017) 'Time-Optimized Isotope Ratio LC-MS/MS for High-Throughput Quantification of Primary Metabolites', *Analytical Chemistry*, 89(3). doi: 10.1021/acs.analchem.6b03731.

Kim, S. K. *et al.* (2016) 'CRISPR interference-guided balancing of a biosynthetic mevalonate pathway increases terpenoid production', *Metabolic Engineering*, 38, pp. 228–240. doi: 10.1016/j.ymben.2016.08.006.

Kim, S. K. *et al.* (2017) 'CRISPR interference-guided multiplex repression of endogenous competing pathway genes for redirecting metabolic flux in Escherichia coli', *Microbial Cell Factories*. doi: 10.1186/s12934-017-0802-x.

Larson, M. H. *et al.* (2013) 'CRISPR interference (CRISPRi) for sequence-specific control of gene expression', *Nature Protocols*, 8(11), pp. 2180–2196. doi: 10.1038/nprot.2013.132.

Lee, J. H. and Wendisch, V. F. (2017) 'Production of amino acids – Genetic and metabolic engineering approaches', *Bioresource Technology*, pp. 1575–1587. doi: 10.1016/j.biortech.2017.05.065.

Mori, M. *et al.* (2017) 'Quantifying the benefit of a proteome reserve in fluctuating environments', *Nature Communications*, 8(1). doi: 10.1038/s41467-017-01242-8.

O'Brien, E. J., Utrilla, J. and Palsson, B. O. (2016) 'Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments', *PLoS Computational Biology*, 12(6), pp. 1–22. doi: 10.1371/journal.pcbi.1004998.

Ozbudak, E. M. *et al.* (2002) 'Regulation of noise in the expression of a single gene', *Nature Genetics*. doi: 10.1038/ng869.

Papp, B., Pál, C. and Hurst, L. D. (2004) 'Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast', *Nature*. doi: 10.1038/nature02636.

Qi, L. S. *et al.* (2013) 'Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression', *Cell*. doi: 10.1016/j.cell.2013.02.022.

Sander, T. *et al.* (2019) 'Allosteric feedback inhibition enables robust amino acid biosynthesis in *E. coli* by enforcing enzyme overabundance', *Cell Systems*. doi: 10.1016/j.cels.2018.12.005.

Stelling, J. *et al.* (2004) 'Robustness of cellular functions', *Cell*, pp. 675–685. doi: 10.1016/j.cell.2004.09.008.

Wu, J. *et al.* (2015) 'Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in Escherichia coli', *Scientific Reports*. doi: 10.1038/srep

Chapter 4

Amino-Acid-Overflow is an Additional Regulatory Mechanism for the Maintenance of End-Product Homeostasis

Timur Sander, Hannes Link

TS designed the study, performed experiments, performed data analysis and co-wrote the manuscript.

HL supervised the study and co-wrote the manuscript

Summary

Microbes developed regulatory mechanism in order to efficiently use nutrients for generating energy and biomass. These regulatory mechanisms, as the previously described allosteric feedback inhibition, allow fast response to changing environments for need-based production of biomass precursors such as amino acids. However, dynamic experiments of allosteric feedback mutants revealed surprisingly high tolerance towards conditional shifts. Proteomics data showed up-regulation of degradation enzymes of the dysregulated arginine, tryptophan and proline biosynthesis pathways. Dynamic metabolite measurements of the allosteric arginine mutant (*argA**) revealed that a transient excess of arginine is channeled into the arginine degradation pathway (AST-pathway). Thus *E. coli* might use degradation pathways in amino acid metabolism as a regulatory mechanism to maintain end-product homeostasis and to avoid wasteful metabolite accumulations. Targeting the AST-pathway by CRISPRi had stronger impact on growth of the *argA** mutant compared to the wildtype.

Abstract Figure
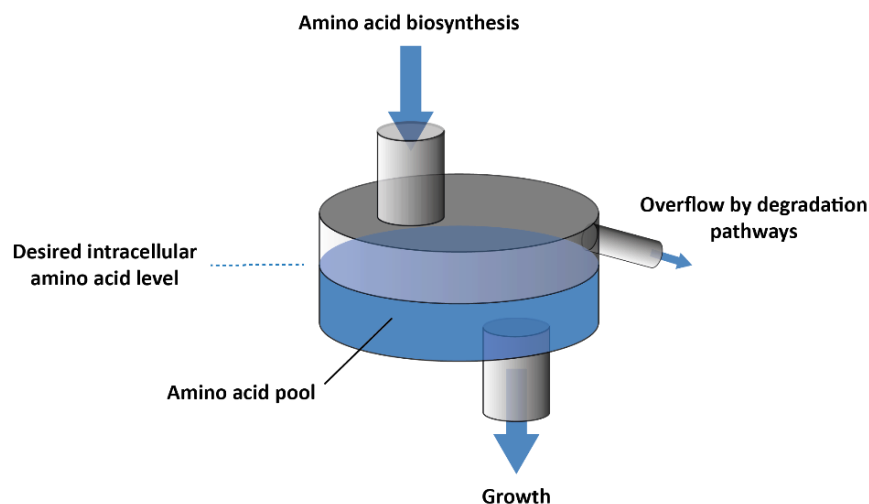
Introduction

Microbes are subject to a variety of self-regulatory mechanisms that control biosynthetic fluxes and intracellular metabolite pools. Regulation is crucial since microbes must ensure a constant set of intracellular conditions in constantly changing environments. The equilibrium of metabolic conditions, the so-called homeostasis was first postulated by Walter B. Cannon (Cannon, 1929). He realized that homeostasis requires feedback regulatory mechanisms that maintain constancy. Nowadays feedback mechanisms, such as allosteric feedback are often postulated to be crucial in maintaining end-product homeostasis (Hofmeyr and Cornish-Bowden, 2000) and fast adaption to changing conditions (Link, Kochanowski and Sauer, 2013). However, recent studies which removed allosteric feedback regulation from amino acid metabolism or pyrimidine metabolism of *E. coli*, revealed surprisingly high tolerance of such mutants towards dynamic conditions (Reaves *et al.*, 2013; Sander *et al.*, 2019).

In case of pyrimidine metabolism, it was shown that pyrimidine homeostasis is accomplished by directed overflow metabolism. Increased pyrimidine pathway flux triggers degradation of the intermediate uridine monophosphate and allows homeostasis of pyrimidine triphosphates at the expense of uracil excretion (Reaves *et al.*, 2013). It is unclear if cells use a similar strategy for maintenance of amino acid homeostasis. It is known that *E. coli* is subject to degradation pathways for 12 of the 20 amino acids, which might be potential routes for a mechanism similar to pyrimidine overflow metabolism (Biocyc, RegulonDB). Generally amino acid degradation pathways are considered as sources for nitrogen, carbon and energy (Reitzer, 2005). A well described example is the arginine degradation pathway (AST-pathway), which converts one molecule arginine to two molecules of ammonia and glutamate and can provide *E. coli's* total nitrogen requirement (Schneider, Kiupakis and Reitzer, 1998). The AST-pathway consists of five enzymes whose genes are organized in an operon under positive control of the arginine repressor ArgR (Kiupakis and Reitzer, 2002). ArgR activates expression of the AST-operon in response to increased arginine level. Deletion of ast-genes (*astB* and *astC*) only caused a growth phenotype when arginine was used as the only nitrogen source, indicating that the prior role of the AST-pathway consists in nitrogen supply. However, it is unclear if degradation pathways like the AST-pathway could also be important in end-product maintenance, especially under such conditions of intra-cellular arginine excess.

Here we show up-regulation of degradation pathways in allosteric dysregulated arginine, tryptophan and proline pathways of *E. coli*. Allosteric feedback mutants that are subject to a degradation pathway (arginine, tryptophan, proline) were more robust in conditional down shifts from glucose to galactose minimal medium compared to mutants in which the degradation pathway was not up-regulated (threonine), respectively to those where no degradation pathway is known (leucine, isoleucine).

Metabolite measurements during conditional down-shifts with the wildtype and the dysregulated arginine mutant (*argA**) showed that a transient excess of arginine is channeled into the degradation pathway (AST-pathway). This arginine overflow mechanism was observed for the wildtype and *argA**, however the AST-pathway showed a stronger activation in the allosteric mutant. Down-regulation of the AST-pathway by CRISPR interference caused decreased growth rates of the allosteric mutant *argA**.

Results

Degradation Pathways are Up-Regulated in Robust Allosteric Mutants

In the previous chapter we could show that enzyme level in allosteric dysregulated amino acid pathways decrease. However, decreased enzyme level did not affect the ability of the mutants to switch from starvation or slow growth to fast growth (readiness). To test if allosteric feedback regulation serves as a mechanism that quickly reduces biosynthetic flux in nutritional down shifts, we tested a switch from fast growth (glucose) to slow growth (galactose). Therefore, the seven allosteric feedback mutants were grown on glucose depletion medium (0.5 g L^{-1} glucose; 5 g L^{-1} galactose). After glucose depletion, we measured reduced growth rates for the *leuA** (leucine), *ilvA** (isoleucine) and *thrA** (threonine) mutant, revealing a relevance of allosteric feedback inhibition in response to conditional shifts (Figure 1). However, the *argA** (arginine), *trpE** (tryptophan), *hisG** (histidine) and *proB** (proline) mutant grew like the wildtype (Figure 1). This indicates that either adaption of biosynthetic flux has no relevance for the respective pathways, or other mechanisms fulfill this function. A possible compensating mechanism might be mediated by amino acid degradation mechanisms. Importantly, proteomics data revealed that enzymes from degradation pathways were up-regulated only in the mutants that showed stable growth under dynamic conditions. We measured a strong increase of AstC (arginine degradation) in *argA**, TnaA (tryptophan degradation) in *trpE** and PutA (proline degradation) in *proB** (Figure 2). For the mutants that showed decreased growth rates under dynamic conditions, no degradation pathway is known (*leuA**, *ilvA**) or degradation pathway showed no up-regulation respectively (*thrA**).
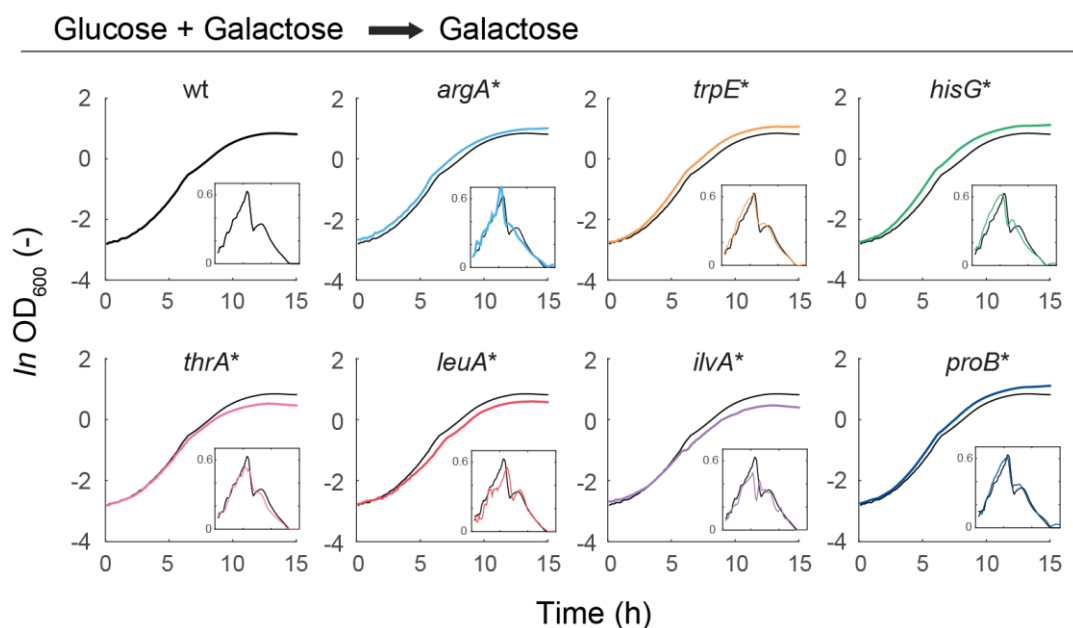
Figure 1. Growth of wild-type *E. coli* and the seven dysregulated mutants in a shift from glucose to galactose.

Cells were grown in M9 minimal medium with 0.5 g L^{-1} glucose and 5 g L^{-1} galactose. Shown are means of n = 3 cultures. Inserts show the growth rate during the same time period. Growth rates were estimated by linear regression over a moving 30 minute window. The same wild-type growth curve is shown in each graph in black as a reference. Published in Sander *et al*. 2019.

In case of histidine no degradation pathway is known in *E. coli*, however we observed a strong up-regulation of the glutaminase GlsA in the *hisG** mutant, which could have a promiscuous activity for deamination of histidine (Figure 2). We hypothesize that the degradation pathways might have an additional role in mediating robustness and end product homeostasis by recycling an excess of amino acids. A similar overflow mechanism was recently discovered in pyrimidine biosynthesis. The strong activation of degradation pathways in our mutants indicates that directed overflow is also relevant for amino acid pathways.
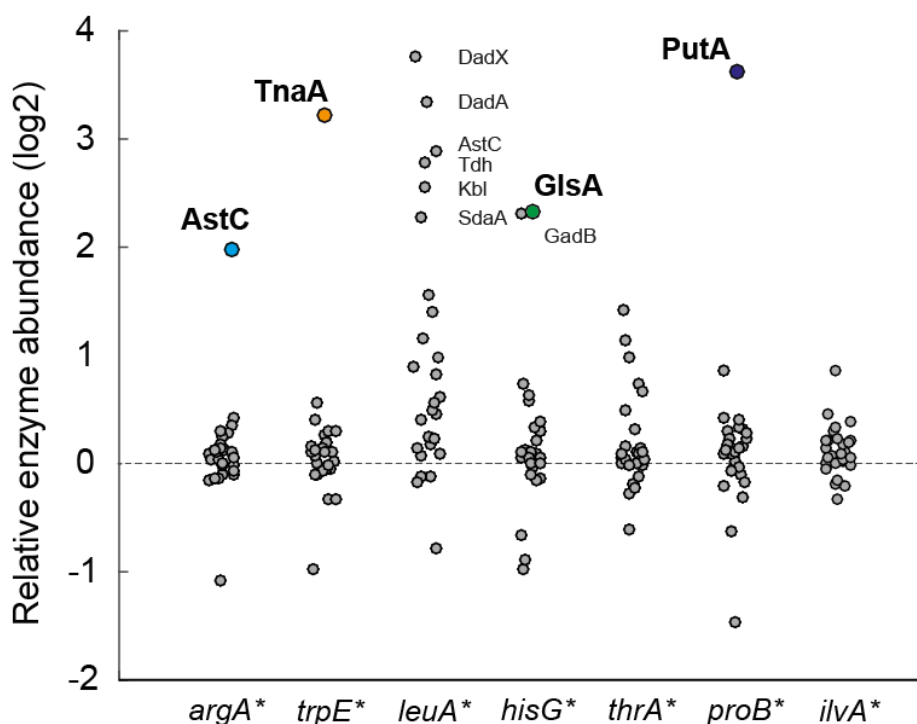
Figure 2. Enzymes of amino acid degradation pathways are up-regulated in respective allosteric mutants.

Enzymes in amino acid degradation pathways (27 out of all 50 enzymes) of seven allosteric dysregulated mutants relative to the wildtype (log2). Enzymes with fold-change (log2) ≥ 2 are indicated with their names. Published in Sander *et al.* 2019.

AST Degradation Pathway Buffers Arginine Excess

To test functioning of directed overflow we used arginine as a case study and dynamically measured intracellular metabolites during a transient excess of arginine. Therefore, we limited protein synthesis by switching the wildtype and the *argA** mutant from external supply of methionine to *de novo* biosynthesis of methionine (Figure 3a). This switch should cause a transient methionine limitation and therefore an excess of other amino acids. If directed overflow buffers this excess in case of arginine, we expected an increase of arginine degradation during the methionine switch. Indeed, immediately after methionine depletion intracellular arginine increased in the dysregulated mutant (17-fold), and rapidly leveled off after 5 minutes (Figure 3b). After the acute increase of arginine, the first intermediate succinyl-arginine of the arginine degradation pathway (AST-pathway) starts accumulating (13-fold). The wildtype showed similar dynamics of arginine and succinyl-arginine pattern but smaller (3-fold) increases of both metabolites. This transient activation of arginine degradation suggests that amino acid overflow functions in dynamic fashion: When arginine exceeds

a certain threshold, it is recycled via degradation pathways. The high *Km* value of the first enzyme in the degradation pathway (AstA) in related species supports the hypothesis of a threshold-based activation of amino acid overflow (Tricot *et al.*, 1994). Thus, we conclude that degradation pathways in amino acid metabolism might serve as an additional regulatory layer to provide end-product homeostasis. To test the physiological role of arginine degradation, we tested how downregulation of the AST pathway affects growth of the wildtype and the *argA** mutant.
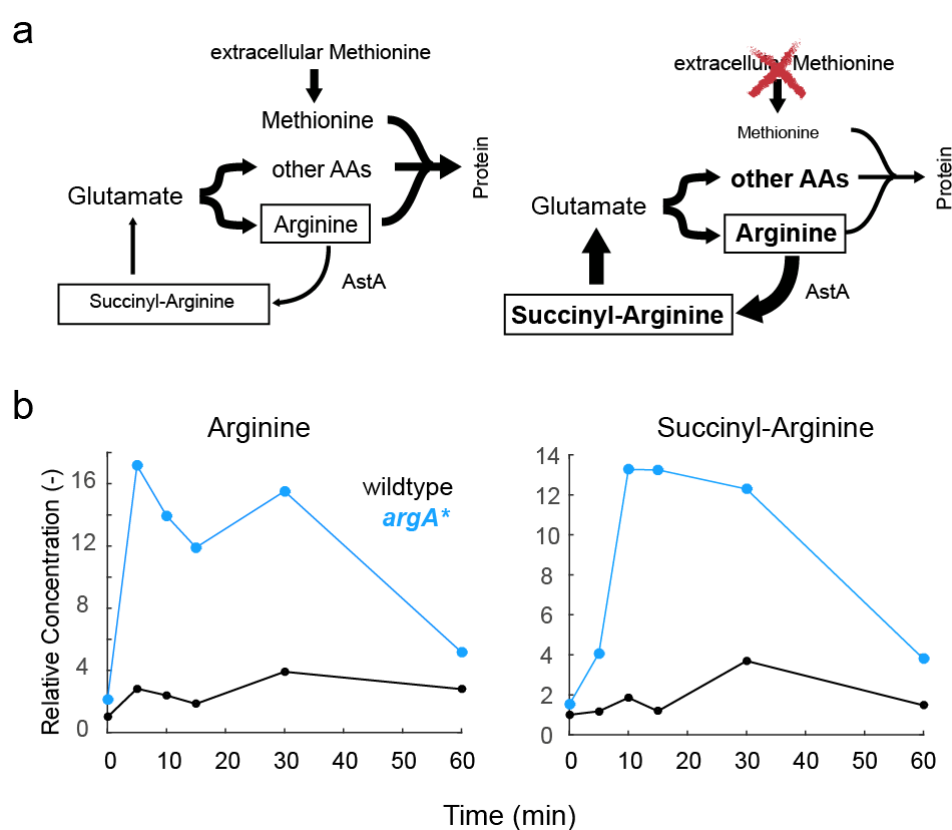
Figure 3. Methionine depletion causes arginine overflow metabolism

(a) Schematic methionine depletion. Switch from external supply of methionine to *de novo* synthesis causes a transient halt of protein synthesis which accordingly leads to an accumulation of the other amino acids including arginine.

(b) Intracellular concentrations of arginine and the first degradation product succinyl-arginine during 60 minutes of transient arginine excess. Arginine excess was created by limiting protein synthesis due to methionine depletion.

CRISPRi Knockdown of AST Pathway Reduces Growth and Gene Expression in Arginine Biosynthesis

Previous studies observed a phenotypical relevance of the AST pathway in providing ammonia during growth on nitrogen limiting conditions. If the AST pathway also functions as a regulatory overflow mechanism, we would expect a growth phenotype of the *argA** mutant when expression of the AST pathway is disturbed. To test this hypothesis, we targeted *astC* by CRISPRi. All genes coding for enzymes of the AST-pathway are organized in a single operon, with *astC* at the first position downstream of the promotor. Thus, blocking *astC* by CRISPRi should have an inhibitory impact on transcription of all genes downstream in the operon. The gRNA-plasmid (pgRNA; Addgene plasmid # 44251) targeting *astC* and the plasmid harboring dCas9 (pdCas9; Addgene plasmid # 44249) were both transformed into the wildtype and the *argA** mutant. In addition, we transformed a reporter plasmid with *gfp* fused to a promotor of an arginine pathway enzyme (pP*argA-gfp*) into both strains (Zaslaver *et al.*, 2006), in order to analyze the effect on gene expression.

Cultivating the strains in a plate reader revealed similar growth rates of 0.65 (h^{-1}) for the wildtype and 0.64 (h^{-1}) for the *argA** mutant without induction of the CRISPRi system (Figure 4a). Induction of CRISPRi against *astC* with anhydrotetracycline (aTc) resulted in reduced growth rates for both, the wildtype and *argA**. However, growth rates of the *argA** mutant decreased stronger by ~50% compared to the un-induced *argA** mutant after initial 2 hours growth. Growth rates of the wildtype decreased only by 20 % after induction. The initial 2 hours are presumably the time until the total copy number of AST enzymes drops below a critical threshold due to transcriptional halt. The expression of P*argA-gfp* decreased by 30 % in both strains after induction, indicating reduced enzyme expression of the arginine pathway due to AST-knockdown (Figure 4b). P*argA-gfp* expression was always lower in the *argA** mutant, matching previous results reporting decreased enzyme expression in allosteric mutants (Sander *et al.*, 2019).

In summary, we could show that the arginine degradation pathway is also relevant under steady state growth without nitrogen limitation. Knockdown of the AST-pathway caused decreased growth rates and reduced activity of an arginine pathway promotor. Importantly, knockdown of the AST-pathway was more critical for the *argA** mutant. This indicates that the allosteric feedback inhibition can to some extent compensate the AST-pathway knockdown in the wildtype. We hypothesize that similar to pyrimidine metabolism (Reaves *et al.*, 2013), allosteric feedback inhibition and degradation pathway could both be involved in maintaining end-product homeostasis.
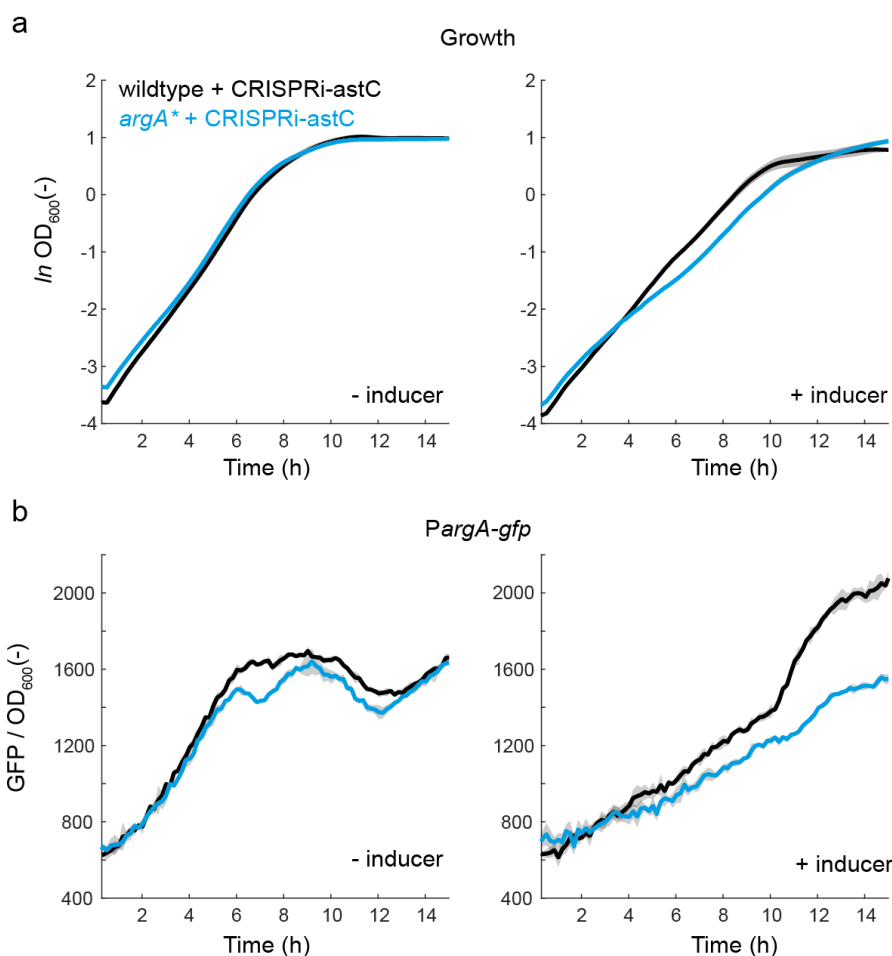
Figure 4. Growth and promotor activity of the wild-type and the allosteric mutant *argA** in response to CRISPRi against *astC*.

(a) Growth of wildtype + CRISPRi-astC and *argA** + CRISPRi-*astC* with and without addition of 0.1 mM aTc. Cells were grown in a plate reader in M9 minimal medium with 5 g L^{-1} glucose and optical density (OD600) was measured in 5 minute intervals. Shown are means of n = 3 cultures.

(b) GFP expression of pP*argA-gfp* in wildtype + CRISPRi-*astC* and *argA** + CRISPRi-*astC* with and without addition of 0.1 mM aTc. GFP was normalized to optical density and measured in 5 minute intervals.

Discussion

Here we could show that *E. coli* uses the AST degradation pathway as an overflow valve: An excess of intracellular arginine is channeled into the degradation pathway indicated by the dynamic increase of the catabolic intermediate succinyl-arginine. A CRISPRi mediated knockdown of the degradation pathway revealed decreased growth in glucose minimal medium without nitrogen limitation. Previous studies observed a growth phenotype only under nitrogen limiting conditions, indicating a sole

function of the AST pathway in providing ammonia (Schneider, Kiupakis and Reitzer, 1998). Thus, our results indicate an additional function of the AST pathway which consists in the maintenance of arginine end-product homeostasis. One reason why cells try to maintain metabolic homeostasis is that metabolites are important signals for enzyme expression (Donati, Sander and Link, 2018). Expression of arginine enzymes for instance is directly regulated by a negative transcriptional feedback in response to intracellular arginine level. Thus, a recurring arginine excess (i.e. due to stochastic effects or dynamic conditions) could imbalance enzyme expression, which might be the reason for impaired growth in the AST-knockdown strains. This hypothesis is supported by the decreased GFP expression from an arginine regulated promotor (P*argA*) after knocking-down the AST-pathway. However, the decreased promotor activity might also result from the decreased growth rates. This could be clarified in future experiments with different level of CRISPR interference, in order to find a condition in which growth rates are not affected but promotor activity is decreased.

Importantly we measured a stronger activation of the AST-pathway in the allosterically dysregulated *argA**-mutant (higher level of the intermediate succinyl-arginine and the degradation enzyme AstC). Also, the growth reduction after AST-knockdown was stronger in the *argA**-mutant compared to the wildtype. This indicates that allosteric dysregulation is to a certain extent compensated by the degradation pathway. We hypothesize that both mechanisms fulfil a role in maintaining end-product homeostasis. While allosteric feedback inhibition regulates the influx into a metabolic pathway, degradation pathways might regulate the outflux. A similar mechanism was previously observed in pyrimidine metabolism, where end-product homeostasis is accomplished by directed overflow through degradation of the intermediate uridine monophosphate (Reaves *et al.*, 2013). Also, in case of pyrimidine overflow metabolism, the mechanism was observed in an allosteric dysregulated mutant. It needs to be clarified if overflow metabolism can also be observed in other amino acid pathways and might compensate allosteric dysregulation in the other described mutants (Figure 1).

The *hisG** mutant was the only stable growing mutant in the down-shift experiments, that is not subject to a degradation mechanism in the respective dysregulated pathway (Figure 1). However, proteomic analysis of the *hisG** mutant revealed a 5-fold increase of the glutaminase GlsA that is known to catalyze the deamination of glutamine to glutamate and free ammonia (NH_4^+) (Brown *et al.*, 2008). Since deamination is a key step in all known amino acid degradation pathways of *E. coli* (Sawers, 2006), GlsA could have a promiscuous function in a so far overlooked histidine degradation mechanism. Potential histidine degradation by GlsA will be tested in future experiments, for example with LC-MS based *in vitro* assays.

Material & Methods

Strains and Culture

All strains in this study were derived from *E. coli* MG1655 (DSMZ No. 18039) and are listed in the general resource table. Chemically competent *E. coli* TOP10 (One Shot™ TOP10, Invitrogen) were used for cloning.

All steady state cultivations were performed using M9 minimal medium with 5 g L^{-1} glucose. For glucose depletion experiments, cells were cultivated in M9 medium with 0.5 g L^{-1} glucose and 5 g L^{-1} galactose. The M9 medium consisted of the following components (per liter): 7.52 g Na$_2$HPO$_4$ 2 H$_2$O, 5 g KH$_2$PO$_4$, 1.5 g (NH$_4$)$_2$SO$_4$, 0.5 g NaCl. The following components were sterilized separately and then added (per liter of final medium): 1 ml 0.1 M CaCl$_2$, 1 ml 1 M MgSO$_4$, 0.6 ml 0.1 M FeCl$_3$, 2 ml 1.4 mM thiamine-HCL and 10 ml trace salts solution. The trace salts solution contained (per liter): 180 mg ZnSO$_4$ 7 H$_2$O, 120 mg CuCl$_2$ 2 H$_2$O, 120 mg MnSO$_4$ H$_2$O, 180 mg CoCl$_2$ 6 H$_2$O. Where appropriate, 50 μg mL^{-1} kanamycin, 34 μg mL^{-1} chloramphenicol or 100 μg mL^{-1} ampicillin were added. For cultivations in microtiter plates, LB pre-culture in 96-deep-well format plates were inoculated from glycerol stocks and grown to an exponential stage. From this first pre-culture a second M9 pre-culture in 96-deep-well plates was inoculated 1:100 and incubated overnight at 37 °C under shaking. Finally, 96-well flat transparent plates (Greiner Bio-One International) containing 150 μl M9 minimal medium were inoculated 1:150 from the overnight culture. Online measurements of optical density at 600 nm (OD$_{600}$) were performed at 37°C with shaking in a plate reader (Epoch, BioTek Instruments Inc, USA; Spark 10M, Tecan Trading AG, Switzerland).

For methionine depletion experiments, cells were cultivated in 25 ml M9 minimal medium (5 g L^{-1} glucose) supplemented with 2 mM methionine. After reaching an OD600 of ~0.4, cells were vacuum filtered on a 0.45 μm pore size nitrocellulose filter (Millipore) and washed with 5 ml glucose minimal medium (pre-warmed). Cell laden filters were transferred to fresh and pre-warmed 25 ml M9 minimal medium (5 g L^{-1} glucose) in a 500 ml shake flask. Shake flask cultures were incubated at 37°C under continuous shaking at 210 rpm. Filter samples for quantification of intracellular metabolites were taken before methionine depletion and after 5, 10, 15, 30 and 60 minutes.

Metabolite Measurements

Shake flask cultivations for methionine depletion experiments were performed as described above. Immediately before and 5, 10, 15, 30 and 60 minutes after methionine depletion 2 mL culture aliquots were vacuum-filtered on a 0.45 μm pore size filter (HVLP02500, Merck Millipore). Filters were directly

transferred into 40:40:20 (v-%) acetonitrile/methanol/water at -20°C for extraction. Extracts were centrifuged for 15 minutes at 13,000 rpm at -9 °C. Centrifuged extracts were mixed with ^{13}C-labeled internal standard and analyzed by LC-MS/MS, with an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies) as described previously[14].

CRISPR interference

CRISPR interference mediated knockdown of *astC* was performed according to Larson et al. 2013[15] under usage of pdCas9 (Addgene #44249) and pgRNA (Addgene #44251). The gRNA plasmid was customized by site-directed mutagenesis with the forward primer EcF_astC (TCAAAGTTTTCACGCGTAATGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC), carrying the 20-nt base-pairing sequence (protospacer) and the reverse primer EcR (ACTAGTATTATACCTAGGACTGAGCTAGC). The protospacer region was designed to target the *astC* gene on the coding strand adjacent to the PAM motif TGG. Expression of dCas9 was induced with 0.1 mM anhydrotetracycline (aTc).

References

Brown, G. *et al.* (2008) 'Functional and structural characterization of four glutaminases from Escherichia coli and Bacillus subtilis', *Biochemistry*, 47(21), pp. 5724–5735. doi: 10.1021/bi800097h.

Cannon, W. B. (1929) 'ORGANIZATION FOR PHYSIOLOGICAL HOMEOSTASIS', *Physiological Reviews*, 9(3), pp. 399–431. doi: 10.1152/physrev.1929.9.3.399.

Donati, S., Sander, T. and Link, H. (2018) 'Crosstalk between transcription and metabolism: how much enzyme is enough for a cell?', *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*. doi: 10.1002/wsbm.1396.

Guder, J. C. *et al.* (2017) 'Time-Optimized Isotope Ratio LC-MS/MS for High-Throughput Quantification of Primary Metabolites', *Analytical Chemistry*, 89(3). doi: 10.1021/acs.analchem.6b03731.

Hofmeyr, J. H. S. and Cornish-Bowden, A. (2000) 'Regulating the cellular economy of supply and demand', *FEBS Letters*, pp. 47–51. doi: 10.1016/S0014-5793(00)01668-9.

Kiupakis, A. K. and Reitzer, L. (2002) 'ArgR-independent induction and ArgR-dependent superinduction of the astCADBE operon in Escherichia coli', *Journal of Bacteriology*, 184(11), pp. 2940–2950. doi: 10.1128/JB.184.11.2940-2950.2002.

Larson, M. H. *et al.* (2013) 'CRISPR interference (CRISPRi) for sequence-specific control of gene expression', *Nature Protocols*, 8(11), pp. 2180–2196. doi: 10.1038/nprot.2013.132.

Link, H., Kochanowski, K. and Sauer, U. (2013) 'Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo', *Nature Biotechnology*, 31(4), pp. 357–361. doi: 10.1038/nbt.2489.

Reaves, M. L. *et al.* (2013) 'Pyrimidine homeostasis is accomplished by directed overflow metabolism', *Nature*. Nature Publishing Group, 500(7461), pp. 237–241. doi: 10.1038/nature12445.

Reitzer, L. (2005) 'Catabolism of Amino Acids and Related Compounds', *EcoSal Plus*, 1(2). doi: 10.1128/ecosalplus.3.4.7.

Sander, T. *et al.* (2019) 'Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance', *Cell Systems*, 8(1), pp. 66-75.e8. doi: 10.1016/j.cels.2018.12.005.

Sawers, G. (2006) 'Amino Acid Degradation', in *Encyclopedia of Life Sciences*. doi: 10.1038/npg.els.0001388.

Schneider, B. L., Kiupakis, A. K. and Reitzer, L. J. (1998) 'Arginine catabolism and the arginine succinyltransferase pathway in Escherichia coli', *Journal of Bacteriology*, 180(16), pp. 4278–4286.

Tricot, C. *et al.* (1994) 'Purification and Properties of a Succinyltransferase from Pseudomonas aeruginosa Specific for both Arginine and Ornithine', *European Journal of Biochemistry*, 224(3), pp. 853–861. doi: 10.1111/j.1432-1033.1994.00853.x.

Zaslaver, A. *et al.* (2006) 'A comprehensive library of fluorescent transcriptional reporters for Escherichia coli', *Nature Methods*, 3(8), pp. 623–628. doi: 10.1038/nmeth8

Closing Remarks

Key Findings

Chapter 1

Enzymes in amino acid metabolism of arginine, tryptophan, leucine, histidine, threonine, proline and isoleucine biosynthesis in *E. coli* do not operate at maximal capacity. Such enzyme overabundance is adjusted by an interaction of allosteric and transcriptional feedback regulation: Allosteric regulation maintains low end-product level, which causes inactivation of the transcriptional repression.

Two-pronged metabolic feedback on allosteric and transcriptional layer solves a critical trade-off between efficiency (low enzyme level) and robustness (high enzyme level). Proteomic and computational data showed that enzymes in amino acid metabolism work in the middle of this trade-off frontier. Removal of allosteric feedback inhibition causes low and efficient enzyme expression however cells are less robust against perturbations of gene expression. Removal of transcriptional regulation causes high and robust enzyme expression however cellular growth suffers from protein burden.

CRISPR interference mediated perturbation of enzyme expression in arginine, histidine and tryptophan biosynthesis showed that wild-type cells (with overabundant enzyme level) are more robust compared to allosteric feedback mutants (without overabundant enzyme level).

Removing allosteric feedback inhibition causes increased metabolic flux in histidine, isoleucine, threonine, leucine and proline biosynthesis, demonstrating a further role of allosteric regulation to control biosynthetic flux. However, biosynthetic flux in the allosteric arginine and tryptophan mutant is compensated by enzyme level regulation of the transcription factors ArgR and TrpR. Thus, biosynthetic flux in the allosteric arginine and tryptophan mutant is not increased.

Chapter 2

Overproduction of arginine in *E. coli* requires removal of allosteric and transcriptional feedback regulation, since both mechanisms have the ability to compensate the loss of the other. Overexpression of the arginine exporter ArgO further elevates the product yield by 5-fold.

CRISPRi-mediated knockdown of the arginine repressor ArgR enables testing different level of remaining transcriptional regulation. Different level of ArgR expression allows tuning of enzyme expression in the arginine biosynthesis pathway.

A residual expression of ArgR results in two-times higher growth rates compared to the rationally engineered knockout strain. Despite this faster growth, arginine production is not lower in the CRISPRi strains (~2 mmol gDW^{-1} h^{-1}).

Metabolomics and proteomics data reveal that slow growth of the knockout strain derives from limitations in pyrimidine biosynthesis, which are caused by enzyme imbalances at the arginine-pyrimidine branching point (17-fold overexpression of the ornithine carbamoyltransferase ArgI). CRISPRi against *argR* efficiently re-balances enzyme level and thereby restores pyrimidine metabolism.

Chapter 3

E. coli reacts to genetic perturbations in amino acid metabolism by elevating the gene expression in the respective biosynthetic pathway. Such 'transcriptional compensation' emerges from a decrease of the respective amino acid end-product, which causes transcriptional de-repression of the pathway enzymes. This mechanism is relevant for pathways regulated by a transcription factor (Arginine), transcriptional attenuation (Histidine) or both (Tryptophan).

Allosteric feedback inhibition has a relevance for the mechanism of transcriptional compensation by re-adjusting end-product level in response to a genetic perturbation. Removing allosteric regulation causes metabolic imbalances in response to a genetic perturbation which causes a delayed transcriptional compensation.

FACS reveal that such metabolic imbalances result in a significant transcriptional heterogeneity of the allosteric mutant compared to the wild-type, when faced with a genetic perturbation. These observations are related to the previous reports about enzyme overabundance (Chapter 1): High enzyme level in wild-type cells buffer genetic perturbations, whereas genetic perturbations in the allosteric mutant (enzymes are working at full capacity) cause metabolic and transcriptional instabilities.

Chapter 4

Allosteric arginine, tryptophan, histidine and proline mutants show no phenotype during conditional down-shift experiments from fast (Glucose) to slow (Galactose) growth. Proteomic data of the

allosteric mutants reveal up-regulation of the respective degradation pathways of arginine, tryptophan and proline metabolism. For histidine metabolism, no degradation pathway is known, although proteomic data show up-regulation of the glutaminase GlsA which might have a promiscuous role in histidine deamination.

Dynamic metabolite measurements of the allosteric arginine mutant reveal that a transient excess of arginine during down-shift experiments is channeled into the degradation pathway. We conclude that amino acid degradation pathways are used as an overflow mechanism, providing end-product homeostasis in dynamic conditions.

CRISPRi-mediated knockdown of the degradation pathway has stronger impact on growth of the allosteric mutants compared to the wild-type, indicating a critical role for amino-acid overflow in compensating allosteric dysregulation.

Open Questions & Outlook

How general is the novel function of allosteric feedback regulation in enforcing enzyme overabundance?

In chapter 1 we showed that amino acid enzymes in *E. coli* do not work at fully capacity and that such overabundance renders the cell more robust against genetic perturbations. Mechanistically, overabundance of amino acid enzymes is adjusted by a sensitive interaction of allosteric and transcriptional feedback regulation. This arises the question, if these regulatory principles go beyond amino acid metabolism of *E. coli*. Indeed, different computational studies in *E. coli* and yeast indicate that cells express higher enzyme level than actually needed also in other parts of metabolism (Hackett *et al.*, 2016; O'Brien, Utrilla and Palsson, 2016). Furthermore, the metabolic regulation on two layers (transcriptional and allosteric) which is the mechanistic basis of overabundance in amino acid metabolism, is a common used regulatory principle. For example, glycolysis is subject to transcriptional control by transcription factors such as Crp or Cra, and in addition tightly regulated by allosteric feedbacks on key enzymes like PfkA or PykF (Chubukov *et al.*, 2014). Also, pyrimidine and purine biosynthesis pathways are regulated by combined allosteric and transcriptional feedbacks, similar to amino acid metabolism. Thus, we can conclude that at least all requirements are given to observe similar regulatory principles for the adjustment of enzyme overabundance in other parts of metabolism. A first step to test this assumption would be the construction of further allosteric feedback mutants of different metabolic pathways. Such strains could then be tested by a proteome

analysis combined with a global genetic perturbation analysis for example by CRISPRi. This could reveal changes in gene expression or genetic instabilities compared to the wild-type. Further, such strategy can also be applied for yeast or other eukaryotic cells to show that the system is not limited to prokaryotes.

Is the CRISPRi approach targeted against ArgR useful for the biotechnological production of arginine?

In chapter 2 we titrated expression of the transcription factor ArgR by CRISPRi and thereby doubled growth rates of an arginine overproduction strain without decreasing production rates. However, the presented approach represents a prove of principle and as such would not be applicable in an industrial process. The reason for this is the usage of plasmids and therefore antibiotics, but also the heterologous expression of the endonuclease Cas9 (although it lacks endonuclease activity). However, we think that our approach has a high relevance for industrial biotechnology, since production processes aim for time-efficiency and robustness. Thus, a next step in this study would be to make the engineered arginine production strains applicable for an industrial fermentation process. This would involve the genetic integration of the arginine exporter ArgO and using an alternative strategy to tune ArgR expression. Different expression level of ArgR could for instance be achieved by mutating the promotor instead of blocking transcription by CRISPRi.

Is the titration of a transcription factor also relevant for other overproduction pathways?

We could show the importance of balancing enzyme level in an overproduction pathway and that titration of a transcription factor is a suitable tool for this. Especially imbalances of enzymes at branching points can have drastic effects on the physiology of a production host. Therefore, we hypothesize that this strategy is also applicable to other overproduction pathways that are regulated by transcriptional repressors. Further case studies could be for instance the branched aromatic amino acid pathways that are regulated by the transcription factors TrpR and TyrR or the methionine pathway under control of MetJ. Our findings could have a particular relevance for global transcriptional regulators, such as Lrp which affects the expression of 38 % of the *E. coli* genes (Kroner, Wolfe and Freddolino, 2019) and is therefore often considered as a metabolic engineering target (Park and Lee, 2010). Knocking down instead of knocking out such global regulators might reduce the impact on global metabolism of the host strain.

Does allosteric dysregulation cause oscillations of metabolites and gene expression?

In chapter 3 we observed instabilities of arginine level in the allosteric mutants which might be the reason for the heterogenic gene expression showed in the FACS experiment. A potential explanation would be, that allosteric feedback mutants are not able to adjust metabolic flux in the face of internal- or external perturbations. This might cause oscillatory arginine level for example during dynamically changing environmental conditions. However, although fast adaptation during dynamic condition was often postulated to be the key function of allosteric feedback inhibition it was so far not shown experimentally (Kochanowski, Sauer and Noor, 2015). A limiting factor is, that methods to quantify metabolite level or gene expression only allow observation of the population level. Oscillations of metabolite level or gene expression might be covered by the stochasticity of the population scale. Therefore, we have a strong need for methods allowing dynamic analysis of single cells. A promising approach might be microscopic time-lapse imaging in microfluidic chambers. Combining such an experimental approach with fluorescent sensors for gene expression (GFP-promotor fusions) and/or direct fluorescence sensors for metabolites (FRET-sensors) could uncover single cell oscillations.

References

Chubukov, V. *et al.* (2014) 'Coordination of microbial metabolism', *Nature Reviews Microbiology*, pp. 327–340. doi: 10.1038/nrmicro3238.

Hackett, S. R. *et al.* (2016) 'Systems-level analysis of mechanisms regulating yeast metabolic flux', 354(6311), pp. 1–37. doi: 10.1126/science.aaf2786.Systems-level.

Kochanowski, K., Sauer, U. and Noor, E. (2015) 'Posttranslational regulation of microbial metabolism', *Current Opinion in Microbiology*. doi: 10.1016/j.mib.2015.05.007.

Kroner, G. M., Wolfe, M. B. and Freddolino, P. L. (2019) 'Escherichia coli Lrp regulates one-third of the genome via direct, cooperative, and indirect routes', *Journal of Bacteriology*. doi: 10.1128/JB.00411-18.

O'Brien, E. J., Utrilla, J. and Palsson, B. O. (2016) 'Quantification and Classification of E. coli Proteome Utilization and Unused Protein Costs across Environments', *PLoS Computational Biology*, 12(6), pp. 1–22. doi: 10.1371/journal.pcbi.1004998.

Park, J. H. and Lee, S. Y. (2010) 'Fermentative production of branched chain amino acids: A focus on metabolic engineering', *Applied Microbiology and Biotechnology*. doi: 10.1007/s00253-009-2307-y.

Resource Table

Resource Table: Listed are all strains, plasmids and other resources used in this doctoral thesis. Strains are listed as follows: **Strain background: Name of the strain used in the study: Genotype**.

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial Strains		
Escherichia coli TOP10: F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 Δ lacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG	Invitrogen, Thermo Fischer Scientific	Cat#C404003
Escherichia coli MG1655: wildtype: F-, lambda-, rph-1	DZMS-German Collection of Microorganisms and Cell Cultures	DSM-No.: 18039
MG1655: argA*: F-, lambda-, rph-1, argA(H15Y)	This study	N/A
MG1655: ilvA*: F-, lambda-, rph-1, ilvA(L447F)	This study	N/A
MG1655: hisG*: F-, lambda-, rph-1, hisG(E271K)	This study	N/A
MG1655: leuA*: F-, lambda-, rph-1, leuA(G462D)	This study	N/A
MG1655: proB*: F-, lambda-, rph-1, proB(D107N)	This study	N/A
MG1655: thrA*: F-, lambda-, rph-1, thrA(S345F)	This study	N/A
MG1655: trpE*: F-, lambda-, rph-1, trpE(S40F)	This study	N/A
MG1655: ΔargR: F-, lambda-, rph-1, ΔargR	This study	N/A
MG1655: ΔtrpR: F-, lambda-, rph-1, ΔtrpR	This study	N/A
MG1655: ΔhisL: F-, lambda-, rph-1, ΔhisL	This study	N/A
MG1655: wildtype CRISPRi-ctrl: F-, lambda-, rph-1, pNUT1533-ctrl	This study	N/A
MG1655: wildtype CRISPRi-argE: F-, lambda-, rph-1, pNUT1533-argE	This study	N/A
MG1655: wildtype CRISPRi-hisB: F-, lambda-, rph-1, pNUT1533-hisB	This study	N/A
MG1655: wildtype CRISPRi-trpA: F-, lambda-, rph-1, pNUT1533-trpA	This study	N/A
MG1655: argA* CRISPRi-ctrl: F-, lambda-, rph-1, argA(H15Y) pNUT1533-ctrl	This study	N/A
MG1655: argA* CRISPRi-argE: F-, lambda-, rph-1, argA(H15Y) pNUT1533-argE	This study	N/A
MG1655: argA* CRISPRi-hisB: F-, lambda-, rph-1, argA(H15Y) pNUT1533-hisB	This study	N/A
MG1655: argA* CRISPRi-trpA: F-, lambda-, rph-1, argA (H15Y) pNUT1533-trpA	This study	N/A
MG1655: hisG* CRISPRi-ctrl: F-, lambda-, rph-1, hisG (E271K) pNUT1533-ctrl	This study	N/A
MG1655: hisG* CRISPRi-hisB: F-, lambda-, rph-1, hisG (E271K) pNUT1533-hisB	This study	N/A
MG1655: hisG* CRISPRi-argE: F-, lambda-, rph-1, hisG (E271K) pNUT1533-argE	This study	N/A
MG1655: hisG* CRISPRi-ctrl: F-, lambda-, rph-1, hisG (E271K) pNUT1533-trpA	This study	N/A
MG1655: trpE* CRISPRi-ctrl: F-, lambda-, rph-1, trpE (S40F) pNUT1533-ctrl	This study	N/A
MG1655: trpE* CRISPRi-trpA: F-, lambda-, rph-1, trpE (S40F) pNUT1533-trpA	This study	N/A
MG1655: trpE* CRISPRi-argE: F-, lambda-, rph-1, trpE (S40F) pNUT1533-argE	This study	N/A

MG1655: trpE* CRISPRi-hisB: F-, lambda-, rph-1, trpE (S40F) pNUT1533-hisB	This study	N/A
MG1655: pPargA-gfp: F-, lambda-, rph-1,	This study	N/A
MG1655: argA* pPargA-gfp: F-, lambda-, rph-1, argA (H15Y)	This study	N/A
MG1655: pPtrpL-gfp: F-, lambda-, rph-1,	This study	N/A
MG1655: trpE* pPtrpL-gfp: F-, lambda-, rph-1, trpE (S40F)	This study	N/A
MG1655: pPthrL-gfp: F-, lambda-, rph-1,	This study	N/A
MG1655: thrA* pPthrL-gfp: F-, lambda-, rph-1, thrA (S345F)	This study	N/A
MG1655: pPhisL-gfp: F-, lambda-, rph-1,	This study	N/A
MG1655: hisG* pPhisL-gfp: F-, lambda-, rph-1, hisG (E271K)	This study	N/A
MG1655: pPleuL-gfp: F-, lambda-, rph-1,	This study	N/A
MG1655: leuA* pPleuL-gfp: F-, lambda-, rph-1, leuA (G462D)	This study	N/A
MG1655: WT-arg: F-, lambda-, rph-1, pNUT1533-argE, pPargA-gfp	This study	N/A
MG1655: WT-trpC: F-, lambda-, rph-1, pNUT1533-trpC, pPtrpL-gfp	This study	N/A
MG1655: WT + gRNA-hisB: F-, lambda-, rph-1, pgRNA-hisB, pdCas9	This study	N/A
MG1655: WT + gRNA-ctrl: F-, lambda-, rph-1, pgRNA-ctrl, pdCas9	This study	N/A
MG1655: argA*-arg: F-, lambda-, rph-1, argA (H15Y), pNUT1533-argE, pPargA-gfp	This study	N/A
MG1655: WT + CRISPRi-astC: F-, lambda-, rph-1, pgRNA-astC, pdCas9, pPargA-gfp	This study	N/A
MG1655: argA* + CRISPRi-astC: F-, lambda-, rph-1, argA (H15Y), pgRNA-astC, pdCas9, pPargA-gfp	This study	N/A
MG1655: argA*ΔargR: F-, lambda-, rph-1, argA(H15Y), ΔargR	This study	N/A
MG1655: Wildtype + pargO: F-, lambda-, rph-1, pargO	This study	N/A
MG1655: argA* + pargO F-, lambda-, rph-1, argA(H15Y), pargO	This study	N/A
MG1655: ΔargR + pargO F-, lambda-, rph-1, ΔargR, pargO	This study	N/A
MG1655: argA* ΔargR + pargO (KO-strain): F-, lambda-, rph-1, argA(H15Y), ΔargR, pargO	This study	N/A
MG1655: CRISPRi#7: F-, lambda-, rph-1, argA(H15Y), pCRISPRi-argR#7, pargO	This study	N/A
MG1655: CRISPRi#10: F-, lambda-, rph-1, argA(H15Y), pCRISPRi-argR#10, pargO	This study	N/A
MG1655: Wildtype CRISPRi-argR#1: F-, lambda-, rph-1, pdCas9, pgRNA-argR#1	This study	N/A
MG1655: Wildtype CRISPRi-argR#2: F-, lambda-, rph-1, pdCas9, pgRNA-argR#2	This study	N/A
MG1655: Wildtype CRISPRi-argR#3: F-, lambda-, rph-1, pdCas9, pgRNA-argR#3	This study	N/A
MG1655: Wildtype CRISPRi-argR#4: F-, lambda-, rph-1, pdCas9, pgRNA-argR#4	This study	N/A
MG1655: Wildtype CRISPRi-argR#5: F-, lambda-, rph-1, pdCas9, pgRNA-argR#5	This study	N/A
MG1655: Wildtype CRISPRi-argR#6: F-, lambda-, rph-1, pdCas9, pgRNA-argR#6	This study	N/A
MG1655: Wildtype CRISPRi-argR#7: F-, lambda-, rph-1, pdCas9, pgRNA-argR#7	This study	N/A
MG1655: Wildtype CRISPRi-argR#8: F-, lambda-, rph-1, pdCas9, pgRNA-argR#8	This study	N/A
MG1655: Wildtype CRISPRi-argR#9: F-, lambda-, rph-1, pdCas9, pgRNA-argR#9	This study	N/A
MG1655: Wildtype CRISPRi-argR#10: F-, lambda-, rph-1, pdCas9, pgRNA-argR#10	This study	N/A

MG1655: *argA** CRISPRi-argR#1: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#1	This study	N/A
MG1655: *argA** CRISPRi-argR#2: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#2	This study	N/A
MG1655: *argA** CRISPRi-argR#3: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#3	This study	N/A
MG1655: *argA** CRISPRi-argR#4: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#4	This study	N/A
MG1655: *argA** CRISPRi-argR#5: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#5	This study	N/A
MG1655: *argA** CRISPRi-argR#6: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#6	This study	N/A
MG1655: *argA** CRISPRi-argR#7: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#7	This study	N/A
MG1655: *argA** CRISPRi-argR#8: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#8	This study	N/A
MG1655: *argA** CRISPRi-argR#9: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#9	This study	N/A
MG1655: *argA** CRISPRi-argR#10: F-, *lambda*-, *rph*-1, *argA*(H15Y) pdCas9, pgRNA-*argR*#10	This study	N/A
Plasmids		
pKDsgRNA-ack	Reisch et al. 201539	Addgene plasmid # 62654
pCas9-CR4	Reisch et al. 201539	Addgene plasmid # 62655
pKDsgRNA-p15	Reisch et al. 201539	Addgene plasmid # 62656
pdCas9	Qi et al. 2013	Addgene plasmid # 44249
pgRNA	Qi et al. 2013	Addgene plasmid # 44251
pKDsgRNA-argA(H15Y)	This study	N/A
pKDsgRNA-ilvA(L447F)	This study	N/A
pKDsgRNA-hisG(E271K)	This study	N/A
pKDsgRNA-leuA(G462D)	This study	N/A
pKDsgRNA-proB(D107N)	This study	N/A
pKDsgRNA-thrA(S345F)	This study	N/A
pKDsgRNA-trpE(S40F)	This study	N/A
pKDsgRNA-ΔargR	This study	N/A
pKDsgRNA-ΔtrpR	This study	N/A
pgRNA-hisB	This study	N/A
pgRNA-astC	This study	N/A
pgRNA-ctrl	This study	N/A
pNUT542	Singh et al. 2017	
pNUT1533-ctrl	This study	N/A
pNUT1533-argE	This study	N/A
pNUT1533-trpA	This study	N/A
pNUT1533-trpC	This study	N/A
pNUT1533-hisA	This study	N/A
pUA66-PargA-gfp: pPargA-gfp	Zaslaver et al. 2006	N/A
pUA66-PtrpL-gfp: pPtrpL-gfp	Zaslaver et al. 2006	N/A
pUA66-PhisL-gfp: pPhisL-gfp	Zaslaver et al. 2006	N/A

pUA66-PleuL-gfp: pPleuL-gfp	Zaslaver et al. 2006	N/A
pUA66 based plasmid with pPthrA	This study	pPthrA-gfp
pgRNA-*argR*#1	This study	N/A
pgRNA-*argR*#2	This study	N/A
pgRNA-*argR*#3	This study	N/A
pgRNA-*argR*#4	This study	N/A
pgRNA-*argR*#5	This study	N/A
pgRNA-*argR*#6	This study	N/A
pgRNA-*argR*#7	This study	N/A
pgRNA-*argR*#8	This study	N/A
pgRNA-*argR*#9	This study	N/A
pgRNA-*argR*#10	This study	N/A
pCRISPRi-*argR*#7	This study	N/A
pCRISPRi-*argR*#10	This study	N/A
pargO	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
Acetonitrile	Honeywell Riedel-de Haën	Cat#14261-2L
Methanol	VWR	Cat#83638.320
Anhydrotetracycline	Sigma-Aldrich	Cat#1035708-25MG
IPTG	Roth	Cat#CN08.2
Ampicillin	Roth	Cat#K029.2
Kanamycin	Roth	Cat#T832.3
Gentamycin	Roth	Cat#0233.3
Spectinomycine	Roth	Cat# HP66.2
Critical Commercial Assays		
PierceTM Quantitative Colometric Peptide Assay	Thermo Fisher Scientific	Cat#23275
His GraviTrapTM	Merck	11-0033-99
Deposited Data		
kcat-values for enzymes in amino-acid biosynthesis (Table S3)	(Schomburg et al., 2014) (Davidi and Milo, 2017)	BRENDA doi:10.1016/j.copbio.2017.02.007
Amino acid requirement of *Escherichia coli* (Table S4)	(Monk et al., 2017)	doi: 10.1038/nbt.3956
Inhibition Constants (Table S5)	(Keseler et al. 2017) (Schomburg et al., 2014) (Gama-Castro et al., 2016)	EcoCyc BRENDA; RegulonDB
Software and Algorithms		
Matlab Version 9.3.0.713579 (R2017b) for the modelling section and analysis of experimental data	mathworks.com	

Acknowledgements

This thesis was an intense, but stimulating and important part of my life and would not have been possible without a couple of people helping and supporting me over the years. Thus, I warmly thank...

... my supervisor **Hannes Link** who picked me up as a master student and invested lots of time to teach me all the basics that are necessary to do and understand science. Thank you for the support over the last years!

... my thesis committee members **Erhard Bremer** and **Knut Drescher** for the critical input during the thesis advisory meetings.

... my friend and colleague **Christoph Diehl** who provided some of the most important data of this thesis during his time as a master student. Thanks Chris, this thesis would not have been possible in this way and in this time without your support. The same applies to **Timo Glatter**, who was involved in almost every project of this thesis. Thanks Timo, for always taking some time, not only for measuring proteins but also giving helpful input.

... my whole working group for the nice atmosphere and all the important discussions during the group meetings. Special thanks to **Niklas Farke, Michelle Kuntz** and **Chun-Ying Wang** for the collaborations in different projects and thereby helping me to survive also the stressful times. Also special thanks to my office partners **Stefano Donati** and **Martin Lempp** for the scientific and also (predominantly) non-scientific debates.

... my family and friends for the love and support and always reminding me what really matters.

… Michael Crichton for evoking my passion for science.

... **Sina** for taking care of my mental health especially during a very demanding summer `18 and also for the enormous contribution in writing this thesis.

Curriculum Vitae

Personal Data

Name:	Timur Sander
Date of birth:	26.06.1991
Nationality:	German

Education

10/2016 – present:	Ph.D. student in the group 'Dynamic control of metabolic networks' Max-Planck Institute for Terrestrial Microbiology Phillips-University Marburg
10/2014 – 10/2016:	M.Sc. Molecular & Cellular Biology Max-Planck Institute for Terrestrial Microbiology Phillips-University Marburg
10/2011 – 10/2014:	B.Sc. Biology Phillips-University Marburg
10/2010 – 05/2011:	Civil Service
08/2001 – 06/2010:	Gymnasium Marianum Warburg

Publications

The following works were published during the completion of this thesis:

Sander T, Wang CH, Glatter T, Link H; *CRISPRi-Based Downregulation of Transcriptional Feedback Improves Growth and Metabolism of Arginine Overproducing E. coli.* **ACS Synthetic Biology**, 2019. doi: 10.1021/acssynbio.9b00183

Sander T, Farke N, Diehl C, Kuntz M, Glatter T, Link H; *Allosteric Feedback Inhibition Enables Robust Amino Acid Biosynthesis in E. coli by Enforcing Enzyme Overabundance*. **Cell Systems,** 2019. doi: 10.1016/j.cels.2018.12.005

Laganenka L, **Sander T**, Lagonenko A, Chen Y, Link H, Sourjik V; Quorum Sensing and Metabolic State of the Host Control Lysogeny-Lysis Switch of Bacteriophage T1. **mBio,** 2019. Doi: 10.1128/mBio.01884-19

Donati S, **Sander T**, Link H; *Crosstalk between transcription and metabolism: how much enzyme is enough for a cell?* **Wiley Interdisciplinary Reviews Systems Biology and Medicine**, 2017. doi: 10.1002/wsbm.1396

Guder J, Schramm T, **Sander T**, Link H; Time-Optimized Isotope Ratio LC–MS/MS for High-Throughput Quantification of Primary Metabolites. **ACS Analytical Chemistry**, 2017. doi: 10.1021/acs.analchem.6b03731

Abgrenzung der Eigenleistung

Die in dieser Arbeit präsentierten Ergebnisse wurden von mir eigenständig und nur mit Hilfe der aufgeführten Personen angefertigt.

Niklas Farke hat die mathematische Modellierung in Kapitel 1 angefertigt.

Christoph Diehl war involviert in der Klonierung der allosterischen Punktmutanten, sowie deren Charakterisierung auf Metabolit- und Proteinebene in Kapitel 1.

Michelle Kuntz war involviert in die Klonierung der CRISPRi Plasmide in Kapitel 1.

Timo Glatter hat die Messungen der Protein Level in Kapitel 1, 2 und 3 angefertigt.

Chun Ying Wang hat die Proben für die Protein Messung in Kapitel 2 aufbereitet.

Hannes Link hat die präsentierten Arbeiten betreut und ist Co-Autor der Kapitel 1 und 2.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass die vorgelegte Dissertation von mir selbst und ausschließlich mit den angegebenen Hilfen verfasst, keine anderen als die angegebenen Quellen benutzt und alle übernommenen Zitate als solche gekennzeichnet wurden.

Diese Dissertation wurde in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen in- oder ausländischen Hochschule anlässlich eines Promotionsgesuchs oder zu anderen Prüfungszwecken eingereicht.

_____   _____

Ort, Datum   Timur Sander