Max-Planck-Institut für terrestrische Mikrobiologie
ENG Dynamic Control of Metabolic Networks

Philipps-Universität Marburg
Fachbereich Biologie

Construction of Enzymes with Synthetic Allosteric Regulation to Control Metabolic Pathways of *Escherichia coli*

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

dem Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Dominik Beuter
aus Marburg an der Lahn

Marburg, März 2019

Die Untersuchungen zur vorliegenden Arbeit wurden von Januar 2015 bis Februar 2019 unter der Betreuung von Herrn Dr. Hannes Link am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg durchgeführt.

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

_____

Erstgutachter: Herr Dr. Hannes Link

Zweitgutachter: Herr Prof. Dr. Victor Sourjik

Tag der mündlichen Prüfung am:

_____

Teile dieser Arbeit wurden im folgenden Artikel veröffentlicht:

Beuter D, Gomes-Filho J, Randau L, Díaz-Pascual F, Drescher K, Link H (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

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Acronyms

Δ	gene deletion
% (v/v)	percent per volume
% (w/v)	percent per volume
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	basepairs
C-terminal	carboxy-terminal
Cmp	Chloramphenicol
CRISPR	Clustered Regulary Interspaced Short Palindromic Repeats
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphate
dsDNA	Doublestranded DNA
dsRed	Discosoma red fluorescent protein
e.g.	*Exempli gratia*, for example
EDTA	Ethylene-diamine-tetraacetic acid
et al.	*et alteri*, and others
g	gramm
GFP	Green gluorescent protein
h	Hours
i.e.	*id est*, that is
IPTG	Isopropyl β-D-1-thiogalactopyranoside
Kan	Kanamycin
Kb	Kilobases
L	Liter
LB medium	Lysogeny broth medium
LB buffer	Lithium bromide buffer
M	Molar (mol/L)
min	Minutes
mRNA	Messenger RNA
µ	Micro (10^{-6})
n	Nano (10^{-9})
N-terminal	Amino-terminal
nt	Nucleotides
OD_{600nm}	Optical density at 600 nm
PCR	Polymerase chain reaction
pH	Potential of hydrogen
rpm	Revolutions per minute
s	Seconds
TAE	Tris-acetate EDTA buffer
Tris	Tris-(hydroxymethyl)-aminomethane
V	Volt
W	Watt
WT	wildtype

Zusammenfassung

In Metabolic Engineering werden Stämme generiert, die bestimmte Stoffwechselprodukte überproduzieren. Dazu werden bestehende Stoffwechselwege von jeglicher transkriptioneller, translationeller und post-transkriptioneller Regulation befreit, was dazu führt, dass beteiligte Enzyme in hoher Anzahl vorliegen und unabhängig der Konzentration des Endprodukts aktiv sind. Eine komplette Deregulation eines Stoffwechselweges ist allerdings auch problematisch: Eine Reaktion auf interne oder externe Störungen ist häufig nicht möglich und der Gesamtmetabolismus überfordert mit der Überproduktion eines bestimmten Metaboliten, was dazu führt, dass die zelluläre Fitness reduziert wird und Wachstumsraten abnehmen.

Um dieses Problem zu lösen, kann es hilfreich sein, neue Regulationsmechanismen in das metabolische Netzwerk und im Besonderen in den Überproduktions-Stoffwechselweg einzuführen. Bisher wurde dazu in der Regel vor allem die Enzymproduktion neu reguliert. Ein Beispiel dafür sind Zwei-Phasen-Bioprozesse, welche unterteilt sind in eine Wachstumsphase, in der ausreichend Biomasse akkumuliert wird, und eine Produktionsphase, in welcher die für die Überproduktion benötigten Enzyme hergestellt und gegebenenfalls Flüsse in konkurrierende Stoffwechselwege blockiert werden. Allerdings erlaubt eine Regulation der Enzymmenge keine schnelle Regulation innerhalb von Sekunden oder Minuten, was insbesondere in großen Bioreaktoren wichtig wäre, da unzureichende Vermischungen hier dazu führen, dass die Verfügbarkeit von Nährstoffen und Sauerstoff sehr stark schwanken kann. Dies kann zur Bildung von Bereichen innerhalb des Bioreaktors mit nur geringer Nährstoffversorgung, so genannter *dead zones*, führen. Zellen, in denen der Überproduktions-Stoffwechselweg dereguliert oder lediglich die Enzymmenge gesteuert wird, sind nicht in der Lage sich kurzfristig auf sich ändernde Umweltbedingungen zu reagieren, mit der Folge, dass solche Zellen in *dead zones* stärker gestresst werden und damit die Produktivität des gesamten Bioprozesses abnimmt.

Eine schnellere und dynamische Kontrolle der Stoffwechselwege, beispielsweise durch synthetische allosterische Regulation eines an der Überproduktion beteiligten Enzyms könnte daher hilfreich sein. Allerdings ist die Konstruktion und Nutzung eines solchen Enzyms kompliziert. Ein wichtiges Ziel dieser Arbeit war deren Konstruktion und Überprüfung, ob mit diesen Flüsse abhängig des gewünschten Effektors kontrolliert werden können.

Da synthetische allosterische Enzyme als eine Art „Metabolisches Ventil“ in den von ihnen katalysierten Reaktionen agieren und wir diese *in vivo* charakterisieren wollen, war unser erstes Ziel das Schalten von metabolischen Ventilen an Engpässen im Stoffwechsel (sogenannte *Metabolic Bottlenecks*) zu untersuchen (Kapitel 3). Dazu haben wir das Wachstum und metabolische Profile von Wildtypisolaten und Laborstämmen analysiert und konnten zeigen, dass in Laborstämmen ein zuvor beschriebenes *Bottleneck* in der Pyrimidin-Biosynthese (verursacht durch zu geringe *pyrE*-Expressionsraten), dafür sorgt, dass der Fluss in den Stoffwechselweg so gering ist, dass letztlich sogar die Wachstumsrate verringert ist.

Zusätzlich dazu haben wir *CRISPR interference* verwendet um 30 künstliche *Bottlenecks* in unterschiedliche Teile des metabolischen Netzwerks einzubringen. In 16 dieser Stämme mit einem solchen *Bottleneck* konnten wir erhöhte Substrat- und/oder reduzierte Produktkonzentrationen feststellen, die auf ein erfolgreich eingefügtes *Bottleneck* hindeuten. Allerdings konnten wir nur in 6 dieser 16 Stämme auch eine reduzierte Wachstumsrate feststellen, was unterstreicht, dass der Einfluss metabolischer *Bottlenecks* auf die zelluläre Fitness von der Stärke des *Bottlenecks* sowie der Reaktion, in die das Bottleneck eingeführt wurde, abhängig ist.

Im zweiten Teil haben wir dann zwei Methoden evaluiert, mit denen synthetische allosterische Enzyme konstruiert werden können und welche beide auf dem Konzept der gerichteten Evolution basieren, nämlich „Split Proteins“ (Kapitel 4) und „Domain Insertion“ (Kapitel 5). Mit Hilfe des Split Protein-Ansatzes waren wir in der Lage, zwei Fragmente der Dihydrofolatreduktase (DHFR) mit den konditionell interagierenden Proteinen FRAP und FKBP12 zu verbinden, resultierend in einem Rapamycin-abhängigen metabolischem Enzym, das zur Kontrolle des Folat-Biosynthese und letztlich der Wachstumsrate genutzt werden kann.

Mit dem Domain Insertion-Ansatzes konnten wir Chimere aus metabolischen Enzymen – 2-isopropylmalate-Synthase (LeuA) bzw. DHFR – und dem Maltose-bindenden Protein MBP als regulatorischer Domäne bilden. Wir waren in der Lage funktionelle Proteine zu identifizieren, die ausreichend katalytische Aktivität haben um Deletionsmutanten zu komplementieren. Allerdings konnten is jetzt keine Varianten gefunden werden, die sensitiv zum Effektor Maltose sind. Das Domain Insertion-Protokoll haben wir dahingehend optimiert, dass nun Stammbibliotheken, bestehend aus tausenden Stämmen, die potentiell Maltose-abhängige Enzymvarianten exprimieren können, hergestellt werden können, sodass eine Hochdurchsatzmethode zur Identifikation der interessanten Stämme benötigt wurde.

Im dritten Teil haben wir daher das Fluoreszenzprotein TIMER, das als Einzelzell-Wachstumssensor verwendet werden kann, für seine Verwendung in *E. coli* und im Besonderen zur Anreicherung langsam wachsender Zellen aus einer großen Stammbibliothek evaluiert (Kapitel 6). Die darauf aufbauende Anreicherungsmethode soll in Zukunft dafür verwendet werden, Stämme mit Enzymvarianten, die mit Domain Insertion hergestellt wurden und in welche synthetische allosterische Regulation erfolgreich implementiert wurde, zu identifizieren und anzureichern.

1 Abstract

In metabolic engineering strains are created that overproduce a certain product. For that, the production pathway is often released from any transcriptional, translational and post-translational regulation, resulting in a high abundance of enzymes in the production pathway and enzyme variants with feedback-resistance. However, complete dysregulation has several disadvantages: the pathway cannot respond to internal and external perturbations and metabolism of the host is overloaded, resulting in a lowered cellular fitness and reduced growth rates.

To circumvent this problem, it is desirable to implement new layers of regulation in the metabolic network and in particular in the overproduction pathway. So far, such regulation has usually been implemented by controlling enzyme abundance. Two-phase processes for instance are dividing a bioprocess in two phases, a growth phase in which a sufficient amount of biomass is accumulated, and a production phase. In this second phase, the expression of enzymes needed for overproduction is induced, often in combination with the introduction of metabolic bottlenecks in competing pathways.

However, the regulation of enzyme abundances does not allow fast response at the second or minute time-scale. Especially in large-scale bioreactors fast response is important, because of fluctuating availabilities of nutrients and oxygen caused by insufficient mixing which leads to the formation of microenvironments and dead-zones. Cells in which the overproduction pathway is either dysregulated or regulated only by implemented control of enzyme abundance are not able to adjust their metabolic networks according to fast changing microenvironments, leading to stressed and therefore unproductive strains which might negatively affect the stability and durability of bioprocesses.

This highlights the need for faster acting dynamic control of metabolic pathways, for example through enzymes with synthetic allosteric regulation. However, the creation and usage of such enzymes is very challenging. A major goal of this work was to create such enzymes with synthetic allosteric regulation and to test their ability to control fluxes through their pathway.

Synthetic allosteric enzymes are ‘metabolic valves’ that implement bottlenecks in the reaction they are catalyzing and we sought to characterize functioning of these valves *in vivo*. Therefore, our first goal was to examine functioning of these valves, resulting metabolic bottlenecks and their impact on the general fitness (Chapter 3). For that, we analyzed growth and metabolic profiles of wildtype isolate and laboratory strains and could show that in laboratory strains a previously reported bottleneck caused by low *pyrE* gene expression causes insufficient fluxes through the pyrimidine biosynthesis pathway and subsequently lowered growth rates.

In addition to that, we used CRISPR interference to introduce artificial bottlenecks in 30 reactions in different parts of the metabolic network. In 16 of the resulting 30 strains we were able to detect elevated substrate or lowered product concentration, indicating a metabolic

bottleneck. However, only 6 of these 16 strains also had a reduced growth rate, underlining that the impact of metabolic bottlenecks on the growth rate is generally dependent on the reaction and strength of the bottleneck.

In the second part, we then evaluated two methods to create synthetic allosteric enzymes, both of which are based on the concept of directed evolution: Split Proteins (Protein Fragment Complementation, Chapter 4) and Domain Insertion (Chapter 5). With the Split Protein approach we were able to couple two fragments of a split dihydrofolate reductase (DHFR) to the conditionally interacting proteins FRAP and FKBP12, resulting in a rapamycin-dependent metabolic enzyme that can be used to control the folate biosynthesis pathway and consequently the growth rate.

With the Domain Insertion approach, we created enzyme-regulatory domain chimera consisting of 2-Isopropylmalate synthase (LeuA) and murine DHFR as enzymes and the maltose binding protein MBP as regulatory domain. We isolated functional proteins, but could so far not identify a variant that is sensitive to the effector. We optimized the protocol to an extent that libraries of thousands of strain variants expressing potentially switching enzymes can be generated and the screening for strains of interest became the limiting factor.

In a third part of this work we therefore evaluated the fluorescent single cell growth rate reporter TIMER for its utilization in *E. coli* and especially to enrich slow growing cells out of large genetic variant strain libraries in a high-throughput manner using fluorescence-activated cell sorting (Chapter 6). The herewith developed enrichment method is planned to be applied in the future to strain libraries created with the Domain Insertion library approach.

2 Introduction

2.1 Structure and regulation of metabolic networks

The metabolic network of *Escherichia coli* is very complex and consists of numerous metabolites, reactions and enzymes catalyzing these reactions. The components of this network seem to be known for the most parts: For instance, by combining genome, transcriptome, proteome and metabolome data sets, a comprehensive model of the metabolic network of *E. coli* K-12 MG1655, called *i*ML1515 [1], has recently been constructed, including 1515 in metabolism involved genes, 1192 metabolites and 2719 metabolic reactions. Interactions between the different components, i.e. metabolites and proteins [2,3], metabolites and transcription factors [4] or metabolites and RNA [5,6] on the other hand are more difficult to identify and therefore presumably still unknown, highlighting that our knowledge about the metabolic network of *E. coli* is, despite decades of research, still limited.

As a consequence of the complexity and the large metabolic capabilities as well as the need to use valuable nutrients in the most economic manner, one of the key characteristics of metabolic networks is that they are tightly regulated through controlled expression and activity of involved metabolic enzymes.

4 levels of metabolic pathway control exist, 2 levels on which enzyme abundance is regulated and 2 levels on which enzyme activity can be controlled (Figure 1): Enzyme abundance is usually controlled by transcriptional and translational regulation. For *Mycobacterium tuberculosis*, enzyme abundance is also controlled by the post-translational modification pupylation which, similar to ubiquitination in eukaryotes, tags enzymes for targeted protein degradation [7]. However, such a mechanism is not known for *E. coli* and only synthetic mechanisms for targeted protein degradation have been reported [8]. Enzymatic activity is controlled through post-translational modifications and allostery.

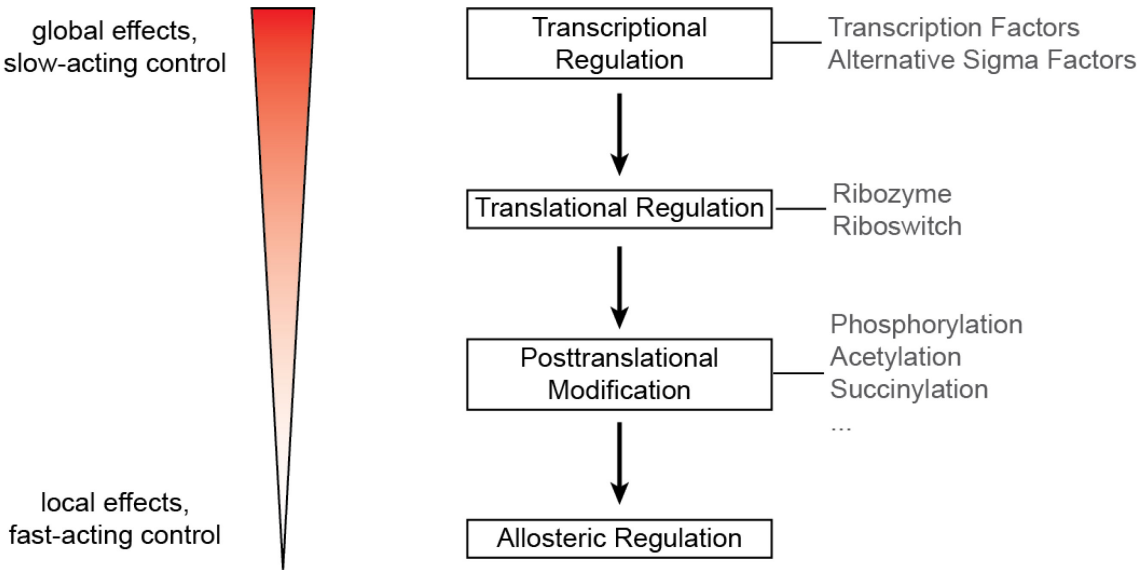

Figure 1: Levels on which metabolic pathways can be controlled.

Metabolic pathways can be controlled on four levels: (1) transcriptionally by transcription factors and alternative sigma factors, (2) translationally by ribozymes and riboswitches, (3) by various post-translational modifications and (4) through allostery. These regulation levels are ordered hierarchically from regulation that acts globally and relatively slow to fast acting regulation with effects on only a limited number of components of the metabolic network.

Transcriptional regulation is the highest level in the regulatory hierarchy and is achieved by two mechanisms, alternative sigma factors and transcription factors.

Many, if not most metabolic pathways are controlled by transcription factors that repress or activate transcription of specific genes by binding promoters or regulatory elements on the DNA and as a consequence either block access or recruit the RNA polymerase holoenzyme to the transcription start site of specific genes. The database RegulonDB computationally predicts the existence of 304 transcription factors in *E. coli* K12 MG1655 [9], the existence of 184 has been shown experimentally. Some, such as CRP [10] or Cra [11] act as global transcription factors that possess a large regulon and consequently have a strong impact on large fractions of the metabolic network. CRP for instance has been shown to control the expression of at least 200 metabolic genes, mainly coding for enzymes involved in secondary carbon source catabolism [12], nitrogen metabolism [13], iron metabolism [14] or osmoregulation [15]. Other transcription factors control only a very small regulon, like RhaS that controls the expression of only the genes of the rhamnose operon [16] or LacI with 3 target genes of the *lac* operon [17]. Often, transcription of genes and operons are controlled by several transcription factors in parallel, coupling gene expression to multiple input signals. For instance, the *lac* operon is regulated by CRP and LacI, which couples gene expression to the availability of the preferred carbon source glucose as well as lactose.

Alternative sigma factors can also be used to control transcription of metabolic genes. While under normal growth conditions the majority of genes is transcribed from promoters that are recognized by the housekeeping sigma factor σ70 [18], two other, alternative sigma factors are

directly involved in metabolism as well: σ^N recognizes promoters of nitrogen-related genes, half of which are important for nitrogen-assimilation [19,20], σ^{19} controls genes involved in ferric citrate uptake in response to presence of periplasmic Fe(III)-dicitrate [21].

The next level in the control of enzyme abundance acts on the level of translation. Three mechanisms are known to control translation of metabolic genes: riboswitches [22], small RNAs [23] or ribozymes [24] which all usually have a repressing effect on translation, often by preventing binding of ribosomes to the mRNA.

After enzymes are produced, their activity is often regulated either through post-translational modifications or allosteric binding of regulatory metabolites.

Post-translational modifications are covalent modifications of enzymes after their biosynthesis usually with the purpose to control the enzymatic activity. In *E. coli*, several modifications are known to control metabolic enzymes of which phosphorylation and acetylation are the most common ones. For the enzymes involved in the central carbon metabolism regulation by post-translational modifications has been examined systematically and several contributing modification methods could be identified [25]. Other pathways have been examined less extensively, primarily because a systematic mapping of post-translational modification sites became only possible in recent years due to advanced and sensitive mass-spectrometry-based proteomics methods [26].

Allosteric regulation is the second mechanism to control enzymatic activity and can be considered to be the fastest level in the hierarchy of regulatory mechanisms of metabolic pathways. Allosteric regulation is defined as the regulation of an enzyme through binding of an allosteric effector molecule at a regulatory site that is not the catalytic active site. Binding of an effector to the allosteric site leads to altering of the properties of the distinct catalytic site of the enzyme. In opposite to post-translational modifications, no modifying enzymes are required for allosteric regulation of enzymes. Most metabolic pathways are assumed to be regulated allosterically, often to create feedback loops. A classic example of an allosteric feedback regulation is the regulation of the N-acetylglutamate synthase ArgA. ArgA catalyses the first reaction in the L-arginine biosynthesis pathway. In order to regulate the flux through the pathway, the product of the pathway, L-arginine, binds and inhibits ArgA [27].

Apart from negative feedback, feedforward regulation of metabolic pathways is also known, although less frequent. For instance, the glycolytic enzyme PfkA of *Bacillus stearothermophilus* is activated upon binding of GDP [28], which signals an increased demand for energy.

Although many enzymes are known to be regulated allosterically, a systematic identification of all metabolite – protein interactions has not been possible yet because the examination of allosteric regulation relied so far on biochemical examinations of purified enzymes. High-throughput approaches combining metabolomics and modelling [29] could help in the future to discover novel allosteric regulations.

2.2 Biological function of different regulation levels

In most cases, metabolic pathways are regulated on at least two levels: First, one mechanism to regulate enzyme abundance, and second, one mechanism to regulate fluxes through the pathway by controlling enzyme activity. All above-mentioned regulation levels have important and distinct roles in this process.

Regulation on the level of transcription and translation is the slowest of all regulation levels with adaption times (times to sense a need for adaption to the actual adaption) on the timescale of minutes. It is also the most energy and resource cost intensive regulation mechanism. However, transcriptional and translational control of metabolic pathways enables global and long-term adaptions of the metabolic network in response to the cellular demand for building blocks and energy and the availability of nutrients. Transcriptional and translational control also defines the metabolic capacities of a cell.

Compared to transcriptional and translational regulation, post-translational modifications and allostery cannot change the metabolic capabilities of a cell. Instead, their function is to regulate the fluxes trough metabolic pathways within the network by regulating the activity of already produced enzymes and by that, limiting fluxes on a short timescale to the required level [26]. Post-translational modifications usually require two sets of modifying enzymes (activating and deactivating) and compounds that are added to the metabolic enzymes, thereby consuming a modest amount of resources. Allosteric regulation on the other hand does not require any modifying enzymes and only binds its allosteric effector molecule and thereby, consumes neither resources nor energy. Allosteric regulation also acts more direct and therefore faster than regulation through post-translational modifications. These differences are reasoned by the different tasks, post-translational modifications and allostery have for regulating metabolic pathways: Allosteric regulation is generally pathway-specific and therefore responsive to only the direct precursors or products of the regulated pathway. Post-translational modifications on the other hand often control several pathways and are therefore needed for more global but fast acting adaptions of metabolic networks [30].

2.3 Dysregulating metabolism and its impact on fitness and productivity

2.3.1 Classical overproducers

An important approach in metabolic engineering of strains is dysregulation of overproduction pathways on the level of enzyme abundance and activity. For example, by replacing a feedback-regulated promoter with a constitutive promoter or, as an alternative, deleting the transcriptional regulator or its binding sites, enzyme abundance can be decoupled from

substrate and product concentration. In addition, removal of allosteric feedback inhibition maintains high enzyme activity in the presence of high end-product concentration.

An example of this concept is a L-arginine overproducing *E. coli* strain [31] (Figure 2). To create high titers of arginine both strategies were combined: Enzyme abundance was elevated by deletion of the transcription factor ArgR that normally represses the expression of all enzymes of the arginine biosynthesis pathway in presence of arginine. In addition, the arginine binding site of the transcription factor ArgP had been removed which led to the overexpression of the arginine exporter ArgO. Next, enzyme activity was uncoupled from product concentration by removal of allosteric feedback regulation of the first enzyme in the arginine pathway, ArgA. These actions combined resulted in an overproduction of up to 11.64 g/L arginine but as a consequence of that, the growth rate dropped to 0.04 h^{-1}, indicating a severe metabolic burden.

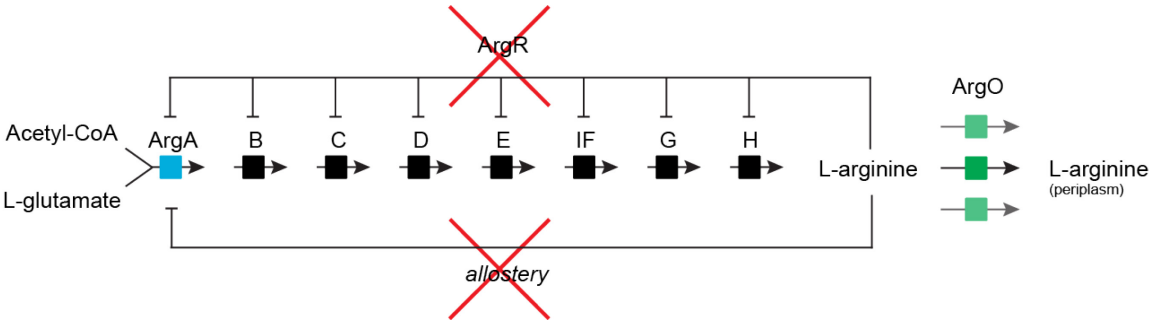

Figure 2: L-arginine biosynthesis pathway of *E. coli* and genetic modifications for the creation of an overproduction strain.

The L-arginine biosynthesis pathway consists of 9 enzymes catalyzing 8 reactions and an exporter (ArgO). In its native state, the production of L-arginine is regulated by two feedback loops: First, ArgR is a transcription factor repressing the expression of all 9 genes in dependence of the concentration of L-arginine. Second, the activity of the first enzyme of the pathway, ArgA (blue) is regulated by allostery. In order to create an L-arginine overproduction strain, [31] uncoupled the expression of the pathway enzymes and the activity of ArgA from the L-arginine concentration and increased the efflux of excessive L-arginine into the periplasm by overexpressing the arginine exporter ArgO.

This highlights a key problem of metabolic engineering: Unregulated overproduction negatively correlates with the fitness of the cell. In general, the higher the flux into one metabolic pathway, the less resources are available for the remaining parts of the metabolic network, indirectly leading to metabolic bottlenecks in pathways competing with the overproduction pathways for precursors (Figure 3). This can ultimately cause a reduction of the growth rate.

In addition to the reduced fitness resulting from the excessive consumption of metabolites, overproduction can also cause reduced growth rates due to cofactor imbalance [32] or the accumulation of cytotoxic intermediates [33,34].

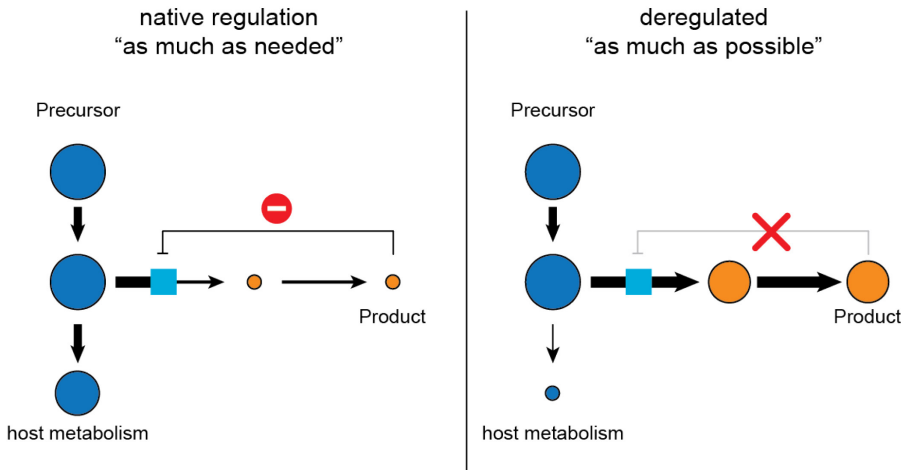

Figure 3: Key problem of unregulated overproduction.

Simplified illustration of the key problem of unregulated overproduction: In wildtype strains metabolic pathways are tightly controlled in order to produce only as much as needed of a certain metabolite (left), in unregulated overproduction strains a significant ratio of the metabolic flux is redirected into the production pathway, leading to an extensive demand and consumption of precursors that cannot be utilized in competing metabolic pathways required for growth.

In order to solve the problem of reduced fitness when overexpressing certain products, an introduction of a novel regulation of the overproduction pathways might be beneficial to ensure both, fitness of individual cells and overproduction.

2.3.2 Two phase bioprocesses

The most simple and direct way to control a bioprocess is to separate growth and production phases (Two-phase bioprocess, Figure 4). During the growth phase a certain biomass is produced in the bioreactor, which usually requires that that fluxes into the overproduction pathway are low. In a second phase, overproduction is induced, usually by overexpressing metabolic feedback-resistant enzymes[35], and reducing fluxes into competing pathways [36].

The overexpression can be controlled either externally or internally. The classical approach to externally initiate the production phase is the induction of gene expression by inducer molecules, such as IPTG. However, such inducer molecules are expensive and therefore not suitable for industrial applications. As a consequence, other systems have been developed that use for instance light [37], oxygen [38], pH shifts [39] or temperature [40] to externally control the shift from growth to production phase.

An alternative strategy is to autonomously induce overproduction. For instance, recently a system has been developed in which a heterologous quorum sensing system had been used to repress transcription of glycolysis genes, allowing a biomass-dependent switch from the growth to the production phase [41]. In a later publication [42], this system has been combined with a

biosensor-based system that switches overproduction on in presence of a sufficient concentration of a precursor of the overproduction pathway.

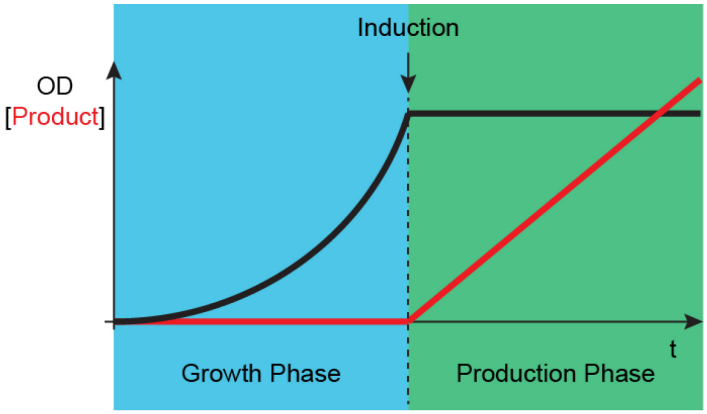

Figure 4: Two-phase bioprocess.

In two-phase bioprocesses, the bioprocess is divided in two phases, a growth phase (blue) in which biomass is accumulated without devoting resources for the overproduction. When enough biomass is accumulated, the production phase is started (green) by inducing the overproduction pathway and/or repressing competing pathways, leading to an increase of the product concentration (red). As the metabolic fluxes are redirected into the overproduction pathway, in many cases the metabolic network is drained, leading to a significant decrease of the growth rate.

However, it should be noted that although all these approaches have shown to enable more efficient overproduction compared to unregulated, static overproduction, many of these systems have to be regarded as inflexible on-off control that ignores the physiological states of individual cells: Usually, once the production process is started, the cells usually keep producing until they either escape the growth burden by mutations or they die because they are not able to maintain the physiological state that would be needed to repair damages, e.g. in DNA or cell envelope. This will subsequently lead to decreasing productivity and limits the duration of a bioprocess.

2.3.3 Continuous control of overproduction pathways

An alternative to this strategy is dynamic and continuous control of metabolic pathways [34]. Continuous control of overproduction pathways allows the autonomous and permanent sensing of certain signals and regulating gene expression accordingly.

One of the first described examples for continuous control is the coupling of lycopene overproduction in *E. coli* to the glycolytic flux by expressing pathway enzymes under the control of Ntr, a global transcription factor which is regulated by the concentration of acetyl phosphate [43].

Another example of has been presented recently for the overproduction of biofuels which are often toxic to the producing cell. The low titers caused by the toxicity of the final product could be increased by developing and overexpressing efflux pumps [44] to transport the toxic

compounds out of the cells. However, overexpression of the efflux pumps led to a growth burden. Therefore, the benefit of sensing the product concentration and regulating the expression of the efflux accordingly has been evaluated [45] and implemented in *E. coli* [46].

In order to solve the problem of decreasing productivity as a result of the growth burden caused by overexpression, continuous control of overproduction pathways could be used. For instance, overexpression could be directly coupled to the availability of nutrients or the general fitness state of the cell. One such a system has been developed recently which introduces burden-sensing feedback regulation to control the overexpression of genes of interest [47]. For that, a promoter that is responsive to the metabolic burden caused by excessive gene expression has been identified and used to control the expression of a guide RNA (gRNA). This gRNA is part of the CRISPR interference system [48] that allows targeted repression of transcription and has been designed to target the overexpressed gene of interest, leading to a gene expression repression when the metabolic burden caused by the gene expression is too high.

All the above mentioned examples show that a continuous control on the level of enzyme abundance is beneficial for the stability and productivity of bioprocesses. However, for bioprocesses continuous control on the level of enzyme activity might have additional advantages as a result of shorter reaction times.

The reason for this are the environmental conditions, cells face within a bioreactor [49]. Especially in large bioreactors, insufficient mixing causes the formation of gradients of glucose, pH or oxygen and, as a consequence, the formation of areas with very different environmental conditions. Problematic for overproducing cells are in particular areas in which the concentration of required nutrients is insufficiently low, so called ‘dead zones’ [50]. Cells in a such a zone with low nutrient availability will be unproductive and stressed [51], whereas in areas with high nutrient availability unwanted acetate overflow metabolism can be observed [52].

To avoid these negative effects, it might be beneficial to introduce positive feedforward regulation to control overproduction. For instance, one could develop a system that enables sensing of the environmental conditions an individual cell is facing, i.e. the nutrient availability or the general fitness state, and control the fluxes through an overproduction pathway accordingly (Figure 5). As the environmental conditions within a bioreactor can change on a very short time scale and keeping in mind that transcriptional control of metabolic pathways reacts very slowly and requires resources to produce new enzymes, feedforward regulation of overexpression pathways should be not implemented on the level of enzyme abundance but instead on the level of enzyme activity, e.g. by introducing synthetic allosteric regulation into enzymes of the overproduction pathway.

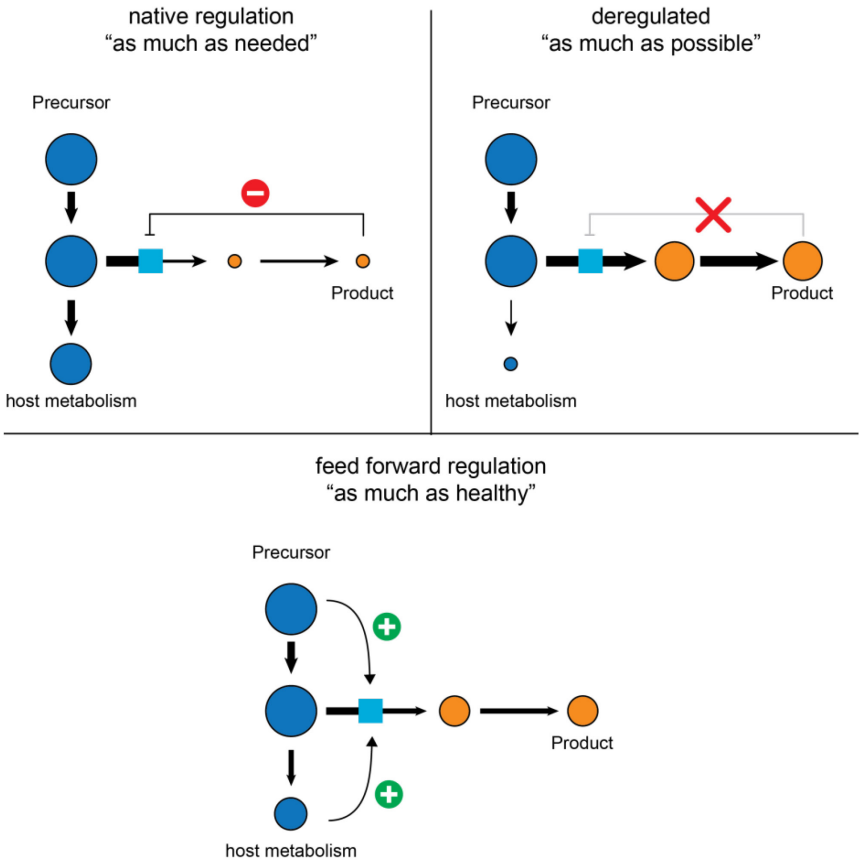

Figure 5: Dynamic feed forward regulation of overproduction pathways.

As an alternative to deregulated overproduction, the activity of an enzyme of an overproduction pathway could be coupled to the availability of nutrients or the concentration of fitness state indicators. By that, metabolic fluxes through the overproduction pathway could be limited to a level that allows overproduction while keeping a basal level of fluxes into competing pathways required to maintain the fitness state of the cell.

So far, only a very limited number of enzymes with synthetic allosteric regulation to control overproduction pathways have been developed. We therefore sought to construct metabolic enzymes with synthetic allosteric regulation.

2.4 Engineering switchable enzymes

Engineering switchable proteins and enzymes, especially with synthetic allosteric regulation, is of great interest for many fields of biology, medicine and chemistry, as they can be used for a wide range of applications such as *in vivo* metabolite sensors [53], as allosteric drugs with improved target specificity [54], in synthetic biology to create orthogonal switches for synthetic circuits [55] or – as in our case – to control metabolic pathways in metabolic engineering. The high demand can be explained by the immediate reaction of switchable enzymes on effectors and that no resources have to be invested for modulation of activity [56]. On the other hand, the introduction of synthetic transcriptional regulation is often preferred to synthetic allosteric regulation. The reason for this is that knowledge about allosteric regulation of proteins of

interest is often limited and very challenging to (re-)engineer. However, a few strategies to engineer such enzymes have been developed, that either base on the concept of rational design or directed evolution [57].

2.4.1 Engineering allosteric regulation

2.4.1.1 Rational Design approaches

Rational design-based methods rely on the comprehensive knowledge about the structural properties of the enzyme of interest and, in particular, of structural changes upon binding of allosteric effectors so that, in consequence, the enzymes of interest can be modified very targetedly.

Several groups used rational design approaches to introduce new allosteric binding sites into a protein of interest or to re-engineer the allosteric binding site to be responsive to alternative allosteric effectors. In one application, hotspots on the surface of the dihydrofolate reductase DHFR have been identified in which light-sensing LOV2 domains have been inserted targetedly, resulting in light-controlled DHFR enzymatic activity [58].

An example for the rational design of an allosteric binding site in the context of metabolic engineering was the creation of a L-lysine-responsive homoserine dehydrogenease (HSDH) of *Corynebacterium glutamicum* [59]. For that, first the allosteric binding site was identified by random mutagenesis, followed by reengineering of the allosteric binding site to bind solely L-lysine instead of the natural inhibitors L-threonine and L-isoleucine. This enzyme variant can be used in a lysine overexpression strain to repress the flux into the homoserine biosynthesis pathway.

Generally, methods for rational design of allosteric regulation are emerging [60] but still limited due to the incomplete understanding of how allostery works [61,62] and lacking knowledge about the structural properties of many enzymes – ideally in both states, when bound to an allosteric effector and unbound – and many regulatory domains.

2.4.1.2 Directed Evolution approaches

As for many proteins of interest, the knowledge about structural properties is limited and the consequences of modifications on enzyme function on a rational basis is difficult to predict, in many cases, methods based on directed evolution are preferred over rational design approaches as these methods allow the re-engineering of allosteric regulation into proteins of interest without prior knowledge about their structural properties. Among a wide-range of methods [56,57,63], two strategies could successfully applied by several groups to engineer proteins with novel allosteric regulation: Split/Fusion Proteins [64] and Domain Insertion [65] (Figure 6).

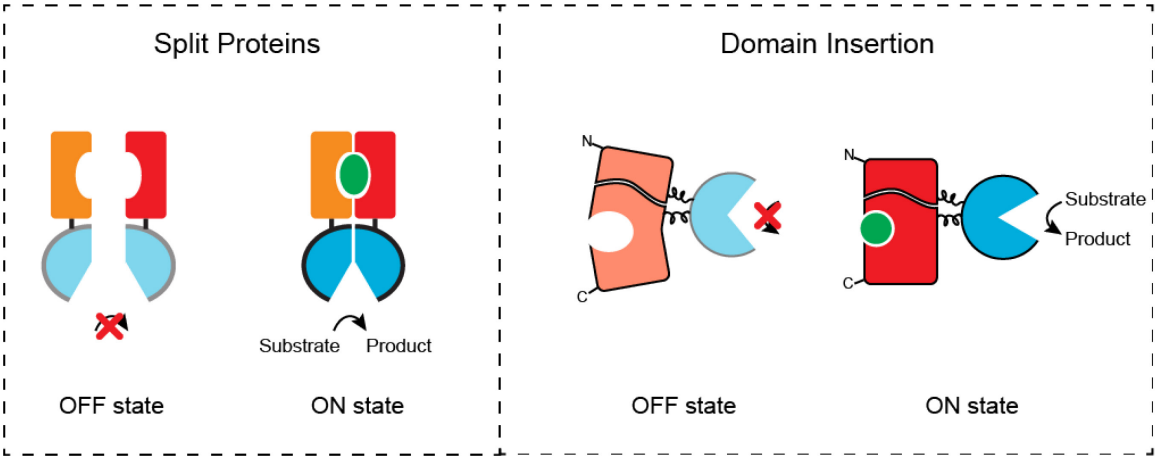

Figure 6: Two strategies to create switchable enzymes with directed evolution-based approaches.

Left: Split Proteins consist of two regulatory proteins (orange and red) which only interact in presence of an effector (green). Fused to the regulatory proteins are fragments of a split enzyme. When separated, none of the fragments is catalytically active. Only when brought into close proximity, both fragments reassemble and form a functional enzyme. Right: Domain Insertion of an enzyme (blue) into a regulatory domain (red), resulting in an enzyme-regulatory domain fusion protein. The regulatory domain undergoes a conformational change upon binding of an effector (green) which is transmitted to the fused enzymatic domain which changes its activity upon.

Fusion and Split Proteins

Fusion proteins are simple end-to-end fusions of a protein of interest with a regulatory protein, often connected via a linker [63]. One example of a fusion protein is a β-lactamase (BLA) that was fused to a nanobody, a fragment of an antibody so that only when bound to cancer cells, the β-lactamase was catalytically active [54]. This protein fusion acts as a proof of concept for the development of site-selective drugs.

For many proteins, due to structural reasons a simple fusion with a regulatory domain will not lead to a switching protein. In these cases, an alternative could be to use split proteins (Figure 7).

Split proteins are proteins or enzymes that have been divided into two fragments. The individual fragments themselves are not enzymatically active; only when brought in close proximity, the fragments reassemble and form a functional protein or enzyme. Such proteins have been created by several groups and employed mainly as reporter proteins in Protein Complementation Assays (PCA). Protein complementation assays allow the analysis of protein-protein interactions by fusing two potentially interacting proteins of interest to the fragments of a split protein. Only in case of an interaction of both proteins, the fused protein fragments reassemble to form a functional reporter enzyme for which split GFP [66], murine dihydrofolate reductase (mDHFR, [67]), adenylate cyclase (AC, [68]) and β-lactamase (BLA, [69]) are commonly used. For instance, in a recent study by fusing potentially interacting proteins to two fragments of split mDHFR 2770 protein-protein interactions in *Saccharomyces cerevisiae* have been identified [70].

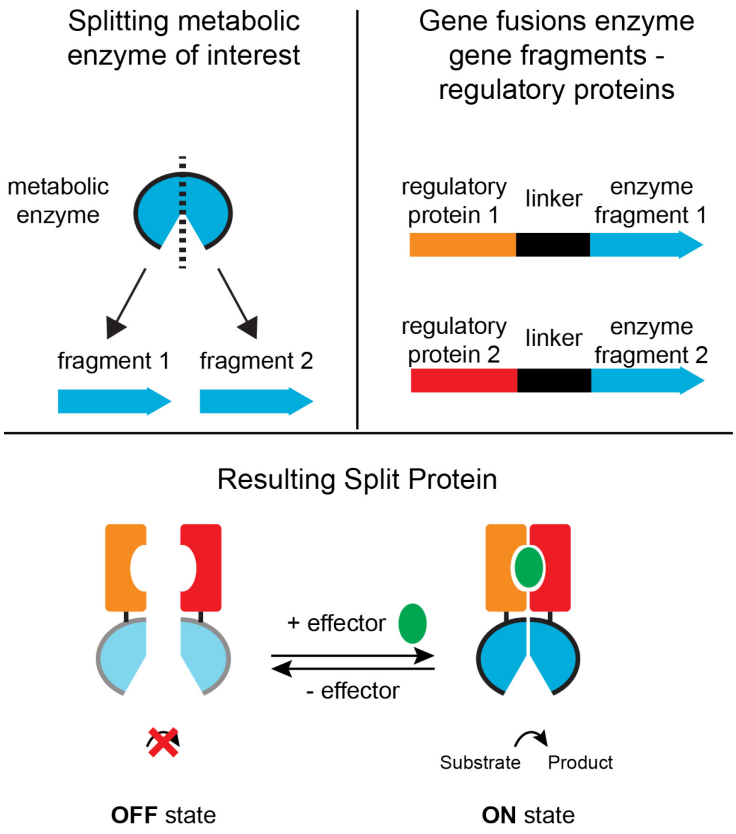

Figure 7: Construction of Split Proteins.

Left: A monomeric enzyme of interest (blue) is split in two fragments. Center: The genes coding for the enzyme fragments are fused to the genes of regulatory proteins often with sequences encoding linkers in between. Right: the resulting split proteins consist each of an enzyme fragment (blue) and one of the regulatory proteins (orange and red). In presence of an effector (green), the two regulatory proteins interact, bringing the fused enzyme fragments into close proximity. Only then both fragments reassemble to a functional enzyme.

The concept of Split Proteins can also be applied for other purposes. In a recent application, catalytically deactivated Cas9 (dCas9) had been split into two fragments and fused to different regulatory proteins, allowing artificial control of dCas9 activity and by that condition-dependent transcription regulation of targeted genes [71].

In this work, we sought to apply the concept of Split Proteins to create enzymes with synthetic regulation to control metabolic pathways and eventually growth (Chapter 4).

Domain Insertion

The second method based on directed evolution is called Domain Insertion. The concept of this method is to randomly insert a protein of interest into a regulatory domain, resulting in a chimeric protein with two domains that both are intended to be still functional.

For that, the gene of a regulatory domain is randomly cut by sequence-independent nucleases such as S1 nucleases or DNase I, the gene of the enzyme of interest is then inserted into the

cut site, resulting in a gene fusion consisting of two fragments of the regulatory domain and the enzymatic domain (Figure 8). When this gene fusion is expressed, a protein chimera consisting of an enzymatic domain fused into a regulatory domain is produced. This chimera ideally has at least two modes of enzymatic activity, an ON state and an OFF state. Upon binding of an allosteric effector to the regulatory domain, the regulatory domain changes its conformation and transduces this change to the fused enzymatic domain which activity is altered upon.

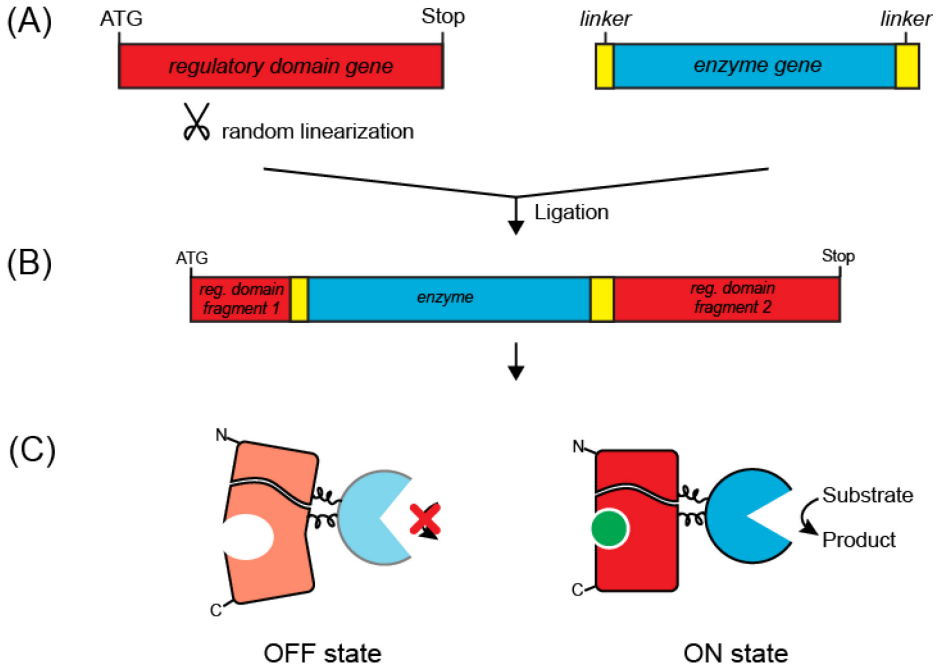

Figure 8: Domain Insertion Overview.

(A) For Domain Insertion, a regulatory domain gene is randomly split into two parts, e.g. by sequence independent nucleases. In this cut site the gene of an enzyme of interest is inserted. The enzyme gene has been amplified without start or stop codons but including linker coding sequences at both ends. (B) Ligation product consisting of two fragments of a regulatory domain and an enzymatic fragment in between. (C) Resulting enzyme – regulatory domain fusion. Both regulatory domain fragments form a functional protein domain that is responsive to an effector (green). Conformational changes upon binding of the effector are transduced to the enzymatic domain, altering its activity.

This approach has been applied successfully to create several proteins and enzymes with synthetic allosteric regulation (Table 1). One of the earliest examples was the insertion of calmodulin receptor proteins into GFP, resulting in a fluorescence protein that is responsive to Ca^{2+} ions [72]. In another study, a maltose-dependent β-lactamase has been created using Domain Insertion of the β-lactamase gene into the gene of the maltose-binding protein MalE [73,74]. Later, the maltose binding protein has been replaced by structurally similar binding proteins for ribose [75], glucose and xylose [76].

Applications of Domain Insertion to construct enzymes with synthetic allosteric regulation to control metabolic pathways are rare. The insertion of DHFR into an estrogen receptor resulted in a switchable metabolic enzyme which however has only been used as biosensor, and not

with the purpose to artificially control the folate biosynthesis pathway or even growth rates [77]. In a very recent example, Domain Insertion has been used to construct a ferredoxin-estrogen receptor to control a synthetic electron transfer pathway and subsequently the growth rate dependent on the effector 4-hydroxytamoxifen (4-HT). However, to our knowledge no further application to create condition-dependent metabolic enzymes to control pathways and subsequently growth has been reported so far. One of our goals was therefore to evaluate the Domain Insertion approach to create such enzymes (Chapter 5).

Table 1: Proteins and regulatory domains that have been used as components of enzyme switches created by Domain Insertion.

Regulatory Domain	Enzyme/ Protein	Effector	Reference
Calmodulin	Green fluorescent protein (GFP)	Ca^{2+}	72
Estrogen receptor	Dihydrofolate reductase (DHFR)	4-hydroxytamoxifen (4-HT)	77
Maltose binding protein (MBP)	β-lactamase (BLA)	Maltose	73
Maltose binding protein (MBP)	Green fluorescent protein (GFP)	Maltose/Trehalose	78
Ribose binding protein (RBP)	β-lactamase (BLA)	Ribose	75
Glucose binding protein (GBP)	β-lactamase (BLA)	Glucose	76
Xylose binding protein (XBP)	β-lactamase (BLA)	Xylose	76
Maltose binding protein (MBP)	Green fluorescent protein (GFP)	Maltose/Trehalose	79
Estrogen receptor	Ferredoxin	4-hydroxytamoxifen (4-HT)	80
Estrogen receptor	(CRISPR)-associated protein Cas9	4-hydroxytamoxifen (4-HT)	81

Domain Insertion has, compared to other approaches, one important advantage: As the enzymatic gene is inserted randomly, no knowledge about the structural properties of involved proteins is required – the protocol is therefore also applicable to proteins that were not extensively characterized. However, past applications showed that only a small portion of

created enzyme–regulatory domain chimera result in functional enzymes of which again only a minority shows condition-dependent enzymatic activity. For instance, for the construction of maltose-controlled β-lactamases (BLA), only 0.8% of all strains possessed a functional enzyme of which again only 10 % also possessed a still functional maltose-binding regulatory domain.

For Domain insertion, it is therefore key to create large libraries of enzyme-regulatory domain variants with many combinations of connecting linkers and insertion sites and to subsequently screen these libraries for enzymes with the desired phenotype. Hence, as a consequence of the library sizes, another important factor is the existence of an effective screening or selection method to identify enzyme variants with the desired functions.

2.4.2 Techniques to identify and isolate enzymes with synthetic allosteric regulation

In order to identify and subsequently isolate enzymes with synthetic allosteric regulation, several techniques can be used.

The conventional approach is the screening of candidate enzymes by *in vitro* assays. For that, enzyme candidates are purified and reactants measured by spectrophotometry or mass spectrometry, enabling the determination of kinetic properties of enzymes [82], e.g. in presence and absence of an allosteric effector. The purification of enzymes is very laborious and time consuming, hence, for the screening of larger libraries of enzymes with potentially synthetic allosteric regulation, such *in vitro* assays are ineligible.

Instead, methods are required that allow the screening of many enzymes with a high throughput and without the need to purify enzymes in advance, i.e. *in vivo* measurements of enzymatic outputs are preferred to *in vitro* assays.

2.4.2.1 Biosensor-based screening for enzymes with synthetic allosteric regulation

One successfully applied strategy for the *in vivo* screening for enzymes with synthetic allosteric regulation is based on the usage of genetically encoded biosensors. Biosensors are systems that enable sensing of certain input signals, e.g. metabolite concentrations, which result in changes of a measurable output signal, e.g. fluorescence, luminescence or antibiotic resistance. Biosensors can be classified in three groups: sensors based on Förster resonance energy transfer (FRET), riboswitches and transcription-factor based sensors. Although FRET- [83] and riboswitch-based [84] biosensors have been successfully designed and applied, the most common strategy to create biosensors is based on transcription factor – reporter gene combinations. In these a transcription factor that is responsive to a specific signal molecule (e.g. the product or substrate of an enzyme of interest) is used to control the expression of a gene encoding for instance a fluorescence reporter. Two factors contribute to the popularity of

transcription factor-based biosensors: First, they consist of only two components that can easily be exchanged. Second, transcription factors responsive to a large variety of effector molecules are highly abundant in pro- and eukaryotes. *E. coli* possesses 200 – 300 transcription factors [85] with different effector molecules that could be used for biosensor construction. In addition, a utilization of heterologous transcription factors has been shown [86], as well as the construction of novel transcription factors [87,88], so that the pool of usable transcription factors even further extends.

An application of transcription factor-based biosensors to screen for enzymes with desired phenotypes has already been shown recently in which the NADPH-responsive transcription factor SoxR of *E. coli* has been utilized to control the expression a fluorescence protein in presence of low NADPH concentrations [89]. This biosensor system has then be used for the *in vivo* screening of a library of NADPH consuming alcohol dehydrogenase from *Lactobacillus brevis* (LbAdh) variants to identify versions with a high enzymatic activity.

Though, not for every metabolite and in particular not every intermediate of biosynthetic pathways are biosensors available and the construction of novel transcription factors is – despite recent advancements [90] – still challenging.

A less specific approach to identify switchable enzymes in a high-throughput manner could therefore be not to measure specific output signals (such as product concentrations) but instead to couple enzymatic activity to a general output such as the physiological state, i.e. the growth rate.

2.4.2.2 Growth-based methods for screening of enzymes with synthetic allosteric regulation

In order to identify and isolate switchable enzymes with growth-based assays, it is crucial that the activity of the enzyme of interest defines the growth rate of the host cell. This is the case when the enzymatic activity is essential in the given environmental conditions but in addition so low that any change in enzymatic activity has an altering effect on the growth rate. In this case, strains with switchable enzymes have high growth rates in one (e.g. presence of the effector) and low growth rates in the other (e.g. absence of the effector) environmental condition. In contrast, non-functional enzyme variants will not support growth, whereas active enzyme variants in which the implementation of allosteric regulation was not successful will result in high growth rates independent of the presence of the allosteric effector (Figure 9). In order to identify and isolate switchable enzymes with growth-based assays, therefore two rounds of selection or screening are required: First, a positive selection round to screen for functional enzymes and to exclude non-functional enzymes from the pool of candidate enzyme variants, and second, a negative selection or screening round in which all strains are sought to be identified in which a conditionally inactive enzyme leads to lowered growth rates.

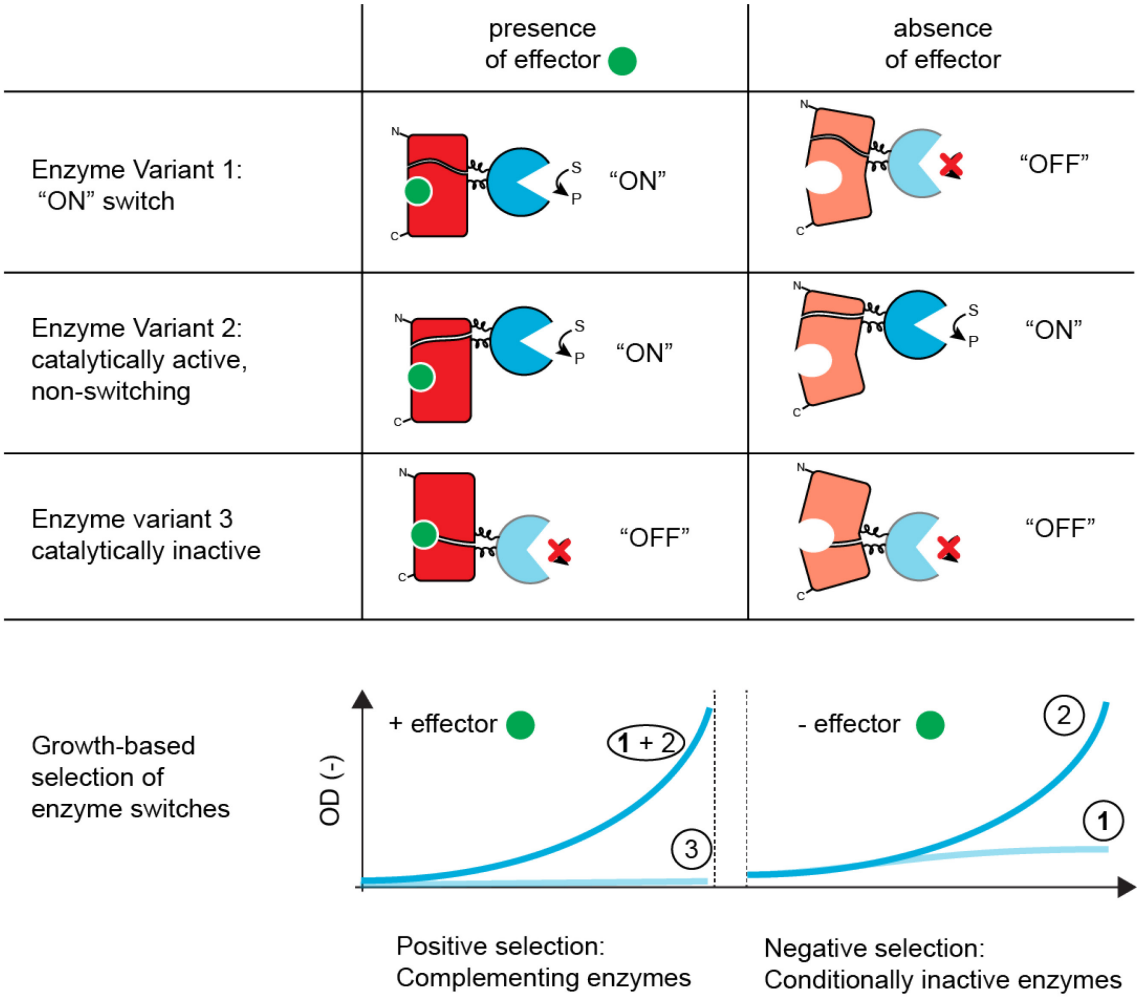

Figure 9: Growth-based screening and selection of enzyme switches.

3 types of enzyme variants result from directed evolution-based enzyme switch construction methods: (1) Switches that are active in one condition (here: presence of the effector) and inactive in the second condition (absence of the effector), (2) enzyme variants that are active independent of the effector and (3) enzyme variants that are catalytically inactive independent of the effector. In order to identify enzyme switches (1) with growth-based approaches, two rounds of selection are required: First, a positive selection to enrich all strains that possess a functional enzyme variant (1 and 2) and to deplete all strains without functional enzyme variants (3). Second, a negative selection in which all strains with conditionally inactive enzymes (1) should be enriched and distinguished from all strains with active enzyme variants (2).

Classical approaches to measure growth rates and by that, screen for strains with certain growth phenotypes are based on growth assays in bioreactors, shake flasks or microtiter plates [91–93]. Although they would enable an accurate determination of the growth characteristics in both growth conditions (presence and absence of the effector), such assays are limited in the number of individual strains that can be screened in parallel. For the screening of libraries of enzyme variant expressing strain libraries, the throughput is too low.

A higher throughput can be reached when strain libraries are pooled prior to growth measurements. Pooled competition assays enable the identification of strains with certain growth rates [94,95]. For that, a pooled strain library is grown for several generations. Before and after culturing samples are taken for next generation sequencing (NGS), allowing the

comparison of library compositions and thereby the identification of strains with higher and lower growth rates. By comparing compositions of the library grown in presence of the effector with the library grown in absence of the effector, one could therefore identify strains that possess switching enzymes. However, that necessitates the retrospective re-engineering of switches which is a very laborious process.

A simpler solution would therefore be to select for strains with condition-dependent enzymes. For that, strains have to be identified and enriched out of a pooled strain library according to the desired phenotype in both presence and absence of the effector.

The enrichment of fast growing cells out of pooled libraries can easily be reached by pooled competition in which the ratio of slow growing strains within the library is decreasing over time. By that, one can enrich strains with catalytically active enzyme variants and remove dysfunctional enzyme variants from the library.

The negative selection for strains with conditionally inactive enzymes by enriching slow growing cells on the other hand is very challenging. Only a few methods to do so have been described.

2.4.2.3 Techniques to enrich slow growing cells

One strategy is based on reducing the well growing population by short phases of incubation in presence of cytotoxic antibiotics [96]. Whereas fast growing bacteria are susceptible to the antibiotics and effectively killed, slow and non-growing bacteria are less susceptible and more likely to survive the treatment. We initially evaluated a similar approach for this project but rejected it because of its complexity (optimal antibiotic concentration and duration of treatment).

Other methods are based on the utilization of fluorescent reporters, allowing a high-throughput enrichment of slow growing cells using flow cytometry.

One such method called FitFlow has been developed for yeast [97]. It is based on a strain with a knocked out gene encoding the chitinase CtsI. In wildtype cells this chitinase is responsible for the degradation of the linkage between mother and daughter cell upon cytokinesis. The chitinase knockout therefore results in the formation of microcolonies. Prior to the growth assay, the cells are briefly sonicated for separation of cells. Next, the cells are incubated for several generations, resulting in microcolonies, each derived from a single strain. By using a histone-GFP fusion protein, the microcolonies can then be separated with flow cytometry according to the microcolony sizes which are directly reflecting the growth rates of the cells. By isolating small microcolonies, cells with impaired growth rates can be specifically enriched and subsequently analyzed in more detail. Although this method has been proven to efficiently separate slow from normal and fast growing strains, this method has also several disadvantages: The most important one is that this method is limited to yeast strains with the particular chitinase knockout. It is imaginable that it could be applied to other yeast strains or microorganisms with similar cell division. However, in bacteria such as *E. coli*, this system

cannot be applied due to differences in the mode of cell division [98]. Another problem might arise from potentially different physiologies of cells with a chitinase deletion and without. Particularly, cells in microcolonies compete with each other for nutrients stronger than isolated single cells would do, possibly influencing the fitness state of the cells.

A similar method which also uses the formation of microcolonies to separate strains with different fluorophore expression patterns has been developed [99]. In their approach, single cells are encapsulated in small gel beads. When incubated, microcolonies deriving from these single cells form in the gel beads which theoretically could later be sorted according to the microcolony sizes that would again directly reflect the growth rates of the individual strains. Compared to 'FitFlow', the big advantage of this method is that it allows the enrichment of slow growing strains independent of the cell's genotype and is therefore applicable to many different microorganisms. However, this method has the same drawback that cells within microcolonies might have a different physiology compared to single cells. In addition to that, cells within gel beads might be potentially supplied worse with nutrients than cells with direct contact to the surrounding medium. It is also be worth noting that the sorting of gel beads according to the growth rates has been discussed in their publication but not shown yet.

Therefore, as all these methods are not applicable for our project, we decided to develop our own method to enrich slow growing cells out of pooled strain libraries using the fluorescent single cell growth reporter TIMER.

2.4.2.4 TIMER

TIMER is a fluorescent reporter protein that bases on the Discosoma red fluorescent protein dsRed [100]. Due to two introduced point mutations it has the exceptional feature of appearing green when freshly expressed but it also maturates with a time delay to its second, red form (Figure 10). This characteristic enabled its use as a single cell growth rate sensor [101]: In slow growing cells, TIMER has time to maturate, resulting in cells with partially green and red fluorescent variants. In fast growing cells on the other hand, the fraction of maturated TIMER is diluted with every cell division and the pool of TIMER proteins inside the cell replenished with freshly expressed and therefore green appearing TIMER.

This feature has been used in *Salmonella* species in infected mice tissues to examine resistance of slow growing cells against treatment with antimicrobials [102]. In another study, TIMER has been used to visualize *E. coli* cells with different growth rates in biofilms [101]. However, although both studies have shown that TIMER can be used to distinguish fast from slow growing cells – a characteristic that might be of interest for many fields of biology and metabolic engineering –, other applications of this reporter have not been described yet.

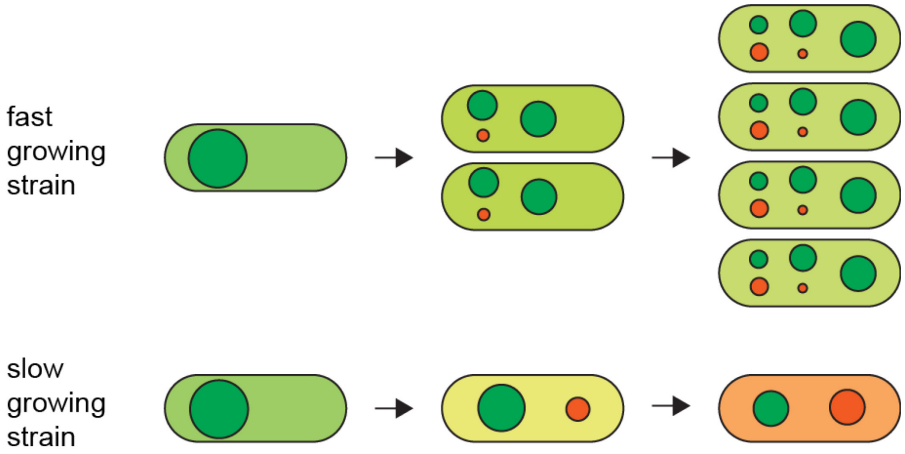

Figure 10: Principle of TIMER as a growth rate sensor.

When freshly expressed, TIMER appears green and refolds spontaneously over time to a red fluorescent form. Slow growing cells (bottom) therefore accumulate the red fluorescent form and appear less green than fast growing cells. In fast growing cells (top), the fraction of already expressed and partially maturated TIMER gets diluted with every cell division while the TIMER pool is replenished with freshly expressed, green appearing TIMER. Faster growing cells therefore appear greener.

For the screening of strain libraries of strains with enzymes with synthetic allosteric regulation, we intend to use TIMER as a tool to enrich slow growing cells in presence or absence of the intended allosteric effector. We therefore sought to verify first how good TIMER can display the growth rate in *E. coli* batch cultures and subsequently tried to enrich slow growing strains out of a pooled library of genetic variant strains (Chapter 6).

2.5 Goal of this work

The major goal of this work was to create metabolic enzymes with synthetic allosteric regulation to control metabolic pathways with directed evolution methods. The process to obtain such enzymes can be illustrated by an engineering cycle as depicted in Figure 11: First, method and components (enzymes, regulatory domains and linkers) have to be evaluated and selected. Next, libraries of potentially switching enzymes are constructed. These libraries have then to be screened for enzymes with successfully implemented regulation. Identified enzymes can then be further analyzed and evaluated for their usage in an application or for further rounds of re-engineering.

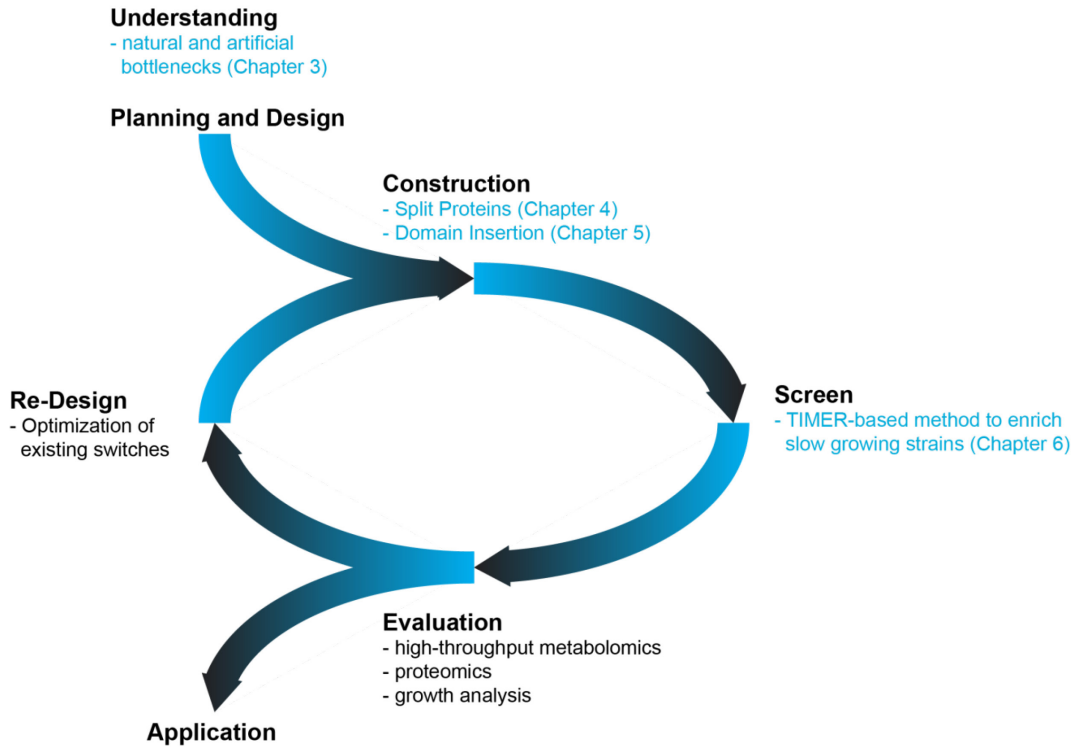

Figure 11: Creation of enzymes with synthetic allosteric regulation.

For this project we worked on 3 parts of this engineering cycle as highlighted in Figure 11:

Part 1 – Understanding the impact of metabolic bottlenecks on the general state of fitness (Chapter 3)

To create metabolic enzymes with artificial regulation, we decided to focus on essential enzymes that would affect growth rates in their OFF states, i.e., enzymes which, when having a lowered activity as a result of their synthetic allosteric regulation, limit fluxes through an essential biosynthetic pathway, subsequently leading to impaired growth rates. For that purpose, in the first part of this project, we sought to better understand the impact of flux limiting metabolic bottlenecks can have on the metabolic network and the overall fitness, represented

by the growth rate. As a starting point we therefore examined metabolism and growth rates of 72 *E. coli* wildtype isolates and 5 laboratory strains to identify rate-limiting steps and bottlenecks (Chapter 3.1). Next, we introduced metabolic bottlenecks on the transcriptional level using CRISPR interference to elucidate which physiological responses we can expect when decreasing enzymatic activity (Chapter 3.2 and 3.3).

Part 2 – Construction of enzymes with synthetic allosteric regulation (Chapters 4 and 5)

In the second part, for the construction of enzymes, we evaluated two techniques that are both based on directed evolution: Split Proteins (Chapter 4) and Domain Insertion (Chapter 5). With the Split Protein approach, we created a rapamycin-dependent DHFR variant and evaluated its impact on the growth rate of an expressing *E. coli* strain. For the Domain Insertion approach, we defined guidelines for the choice of components (Chapter 5.1), tested the expression plasmid (Chapter 5.2), and were able to create complementing enzyme variants with an optimized Domain Insertion protocol (Chapter 5.3).

Part 3 – High-throughput selection of switchable enzymes (Chapter 6)

As the abovementioned techniques to create enzymes with synthetic allosteric regulation usually involve the construction of large strain libraries with potentially switching enzymes, in third part of this work we devised a strategy to enrich strains that express switchable enzymes. As screening of single strains expressing candidate enzymes is not feasible with sufficient throughput, we evaluated and validated the single cell growth rate reporter TIMER to specifically enrich slow growing cells out of complex genomic variant strain libraries (Chapter 6).

3 Understanding the impact of metabolic bottlenecks on the general state of fitness

In this work, metabolic enzymes with synthetic allosteric regulation were planned to be created that are supposed to work as metabolic valves, i.e. to introduce condition-dependent metabolic bottlenecks in single reactions of interest. We were wondering how metabolic networks are affected by metabolic bottlenecks in single reactions and decided therefore to examine the prevalence of metabolic bottlenecks and their impact on the fitness states in genetically not modified wildtype and laboratory strains.

3.1 Metabolic bottlenecks in *E. coli* wildtype and laboratory strains

The EcoR collection is a set of 72 *E. coli* wildtype strains isolated from a large variety of mammalian host organisms from different locations [103]. Our initial assumption was that, dependent on host and isolation location, the *E. coli* strains faced different environmental conditions such as available nutrients and should have therefore adapted their metabolic network to the respective specific environments. Such an adaption can lead to a loss of metabolic capabilities and misregulations as a consequence of mutations which however have no impact on the fitness in their natural environment.

We were interested if we could identify metabolic bottlenecks in single reactions that might even impair the general fitness of the cells in conditions that differ from the natural growth conditions.

For that, we compared the growth rates of the 72 strains of the EcoR collection in M9 minimal medium with glucose as carbon source and under aerobic conditions and incubated at 37°C. In addition to these strains we also analyzed the growth of 5 commonly used laboratory strains (*E. coli* MG1655, W3110, MDS42, BW25113, EMG-2) (Figure 12).

We found that a majority (66 of 77) of all strains had similar growth rates of 0.6 h^{-1} and higher, whereas a group of 11 strains had lower growth rates. To our surprise, all laboratory strains were found to belong to this group with growth rates between 0.41 ± 0.04 h^{-1} (MG1655) and 0.47 ± 0.18 h^{-1} (MG1655).

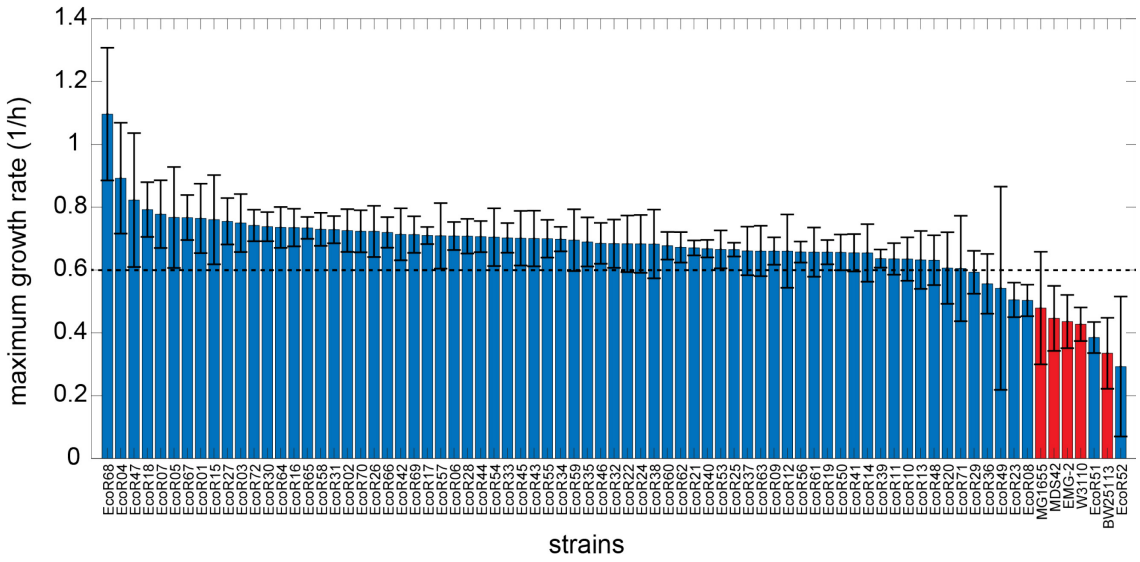

Figure 12: Growth rates of 72 wildtype isolates and 5 laboratory strains.

Blue bars indicate growth rates of wildtype isolates from the EcoR collection [103], red bars of laboratory strains. Shown are the mean growth rates measured in three independent growth experiments.

To test if the reduced fitness can be linked to metabolic bottlenecks in single reactions we measured the concentrations of 94 metabolites of central carbon metabolism, nucleotide metabolism, amino acid metabolism and other parts of the metabolic network in all 77 strains (Figure 13) using high-throughput metabolomics techniques [104].

We noticed that all laboratory strains had a similar metabolic profile, which differed from the wildtype isolates. Specifically, we observed a markedly higher abundance of three intermediates of the pyrimidine pathway, N-carbamoyl-L-aspartate, dihydroorotate and orotate (Figure 13b + c). On the other hand, the concentration of UMP, a later intermediate of the same pathway, was relatively low. These results indicate a metabolic bottleneck in a reaction between orotate and UMP and can indeed by explained by a reported frameshift mutation in the gene upstream *pyrE*, *rph* [105]. This frameshift leads to lower expression rates of *pyrE* which codes for the orotate phosphoribosyltransferase, an enzyme that utilizes orotate and PRPP to convert it to the direct precursor of UMP, orotidine 5'-phosphate. Remarkably, the strain MDS42 possesses a minimal genome with over 700 genes deleted [106]. However, compared to the other laboratory strains, no difference in growth rate or metabolite pattern could be determined.

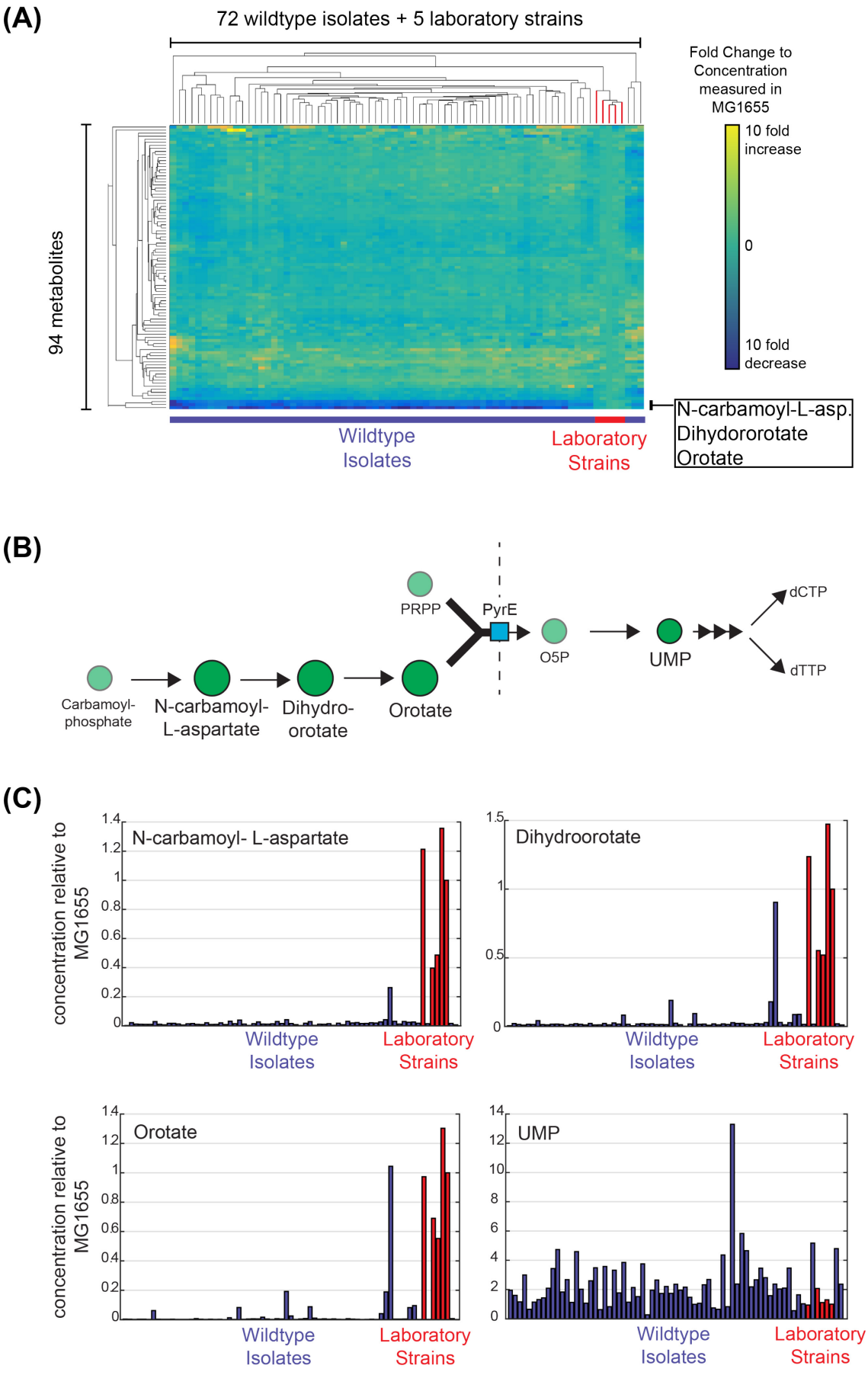
(A)
72 wildtype isolates + 5 laboratory strains
94 metabolites
Fold Change to Concentration measured in MG1655
10 fold increase
0
10 fold decrease
N-carbamoyl-L-asp.
Dihydororotate
Orotate
Wildtype Isolates
Laboratory Strains
(B)
PRPP
PyrE
O5P
UMP
dCTP
dTTP
Carbamoyl-phosphate
N-carbamoyl-L-aspartate
Dihydro-orotate
Orotate
(C)
concentration relative to MG1655
N-carbamoyl- L-aspartate
Dihydroorotate
Orotate
UMP
Wildtype Isolates
Laboratory Strains

Figure 13: Concentration of 94 metabolites in 72 *E. coli* wildtype isolate and 5 laboratory strains and metabolites of the pyrimidine biosynthesis pathway.

(A) Clustered heatmap with all 94 metabolites on the Y-axis and all tested strains on the X-axis. The blue and red bars underneath the strain labels mark laboratory strains (red) and wildtype strains (blue). The colors in the heatmap indicate an up to 10 fold increase (yellow) or decrease (blue) of metabolite concentrations compared to MG1655. On the upper part strains are clustered that share a similar metabolic profile. In red all laboratory strains are marked. On the left side metabolites with a similar appearance in the different strains are clustered together. The three rows at the bottom are the intermediates of the pyrimidine biosynthesis pathway. (B) Overview over the pyrimidine biosynthesis pathway. All intermediates that we were able to measure are marked. (C) Metabolite concentrations of the 4 intermediates normalized to MG1655. Blue bars indicate the concentrations measured in wildtype isolate strains, red bars in the laboratory strains (from left to right: MDS42, W3110, BW25113, EMG-2, MG1655).

In the natural isolate strains, linkages between growth phenotypes and metabolic bottlenecks were less apparent.

In EcoR51, one of the slowest growing strains, we could measure the highest levels of PEP. PEP is substrate and product of several metabolic reactions in *E. coli*, it is therefore unclear in which reaction is limited as a result of a metabolic bottleneck. However, we speculate that as a result of the general importance of these pathways for the generation of energy and precursors for amino acid biosynthesis, metabolic bottlenecks in central carbon metabolism (dephosphorylation of PEP to pyruvate catalysed by pyruvate kinase), or anaplerosis (carboxylation of PEP to oxaloacetate, catalysed by PEP carboxylase) could not just result in the accumulation of PEP but might also result in reduced growth rates.

For three other slow growing natural isolates - EcoR23, 29 and 52 - we could not identify any specific metabolic bottleneck. However, all of those show a similar metabolic profile that was very distinct to all other strains, suggesting that all these strains possess a similar metabolic network and possibly the same metabolic bottlenecks.

For the slow growing strains EcoR8 and 49 we could also not identify a metabolic bottleneck. Instead, we found that their metabolic profiles were similar to those of the faster growing strains EcoR27 and EcoR42 – both with growth rates above 0.7 h^{-1}.

It should be noted that only 94 of 1192 metabolites [1] and therefore only a tiny fraction of metabolites could be measured here. Hence, it is possible that in the strains of which we observed reduced growth rates but could not identify a particular metabolic bottleneck like EcoR8 or 49, bottlenecks might exist in reactions of which we can measure neither product nor substrate. In addition, also the contribution of several bottlenecks to the growth reduction is imaginable as well as a limitation of transport capabilities which could limit the growth rate without causing a measurable accumulation of specific metabolites in the cell.

We decided to examine the contribution of the *pyrE* bottleneck to the reduced growth rates in the laboratory strains in more detail. For that, we analyzed the growth rates of two laboratory strains (MG1655 and W3110) and 3 natural isolates (EcoR18, 42 and 46) when growing in M9 minimal medium in presence and absence of 100 mM uracil (Figure 14). Uracil can be taken up

by the cell and transferred to UMP, thereby bypassing the metabolic bottleneck. We found that uracil addition did not affect the growth rates of the natural isolate strains, whereas the growth rates of both laboratory strains increased by 15% (MG1655) and 18% (W3110), respectively. This suggests that the pyrimidine bottleneck indeed reduces the growth rates of the laboratory strains. It should be noted that the growth rates were not restored to the level of the wildtype cells, though. This could be due to insufficient uptake of uracil from the medium as well as a contribution of one or more other metabolic bottlenecks that are limiting the growth rates under the given growth conditions.

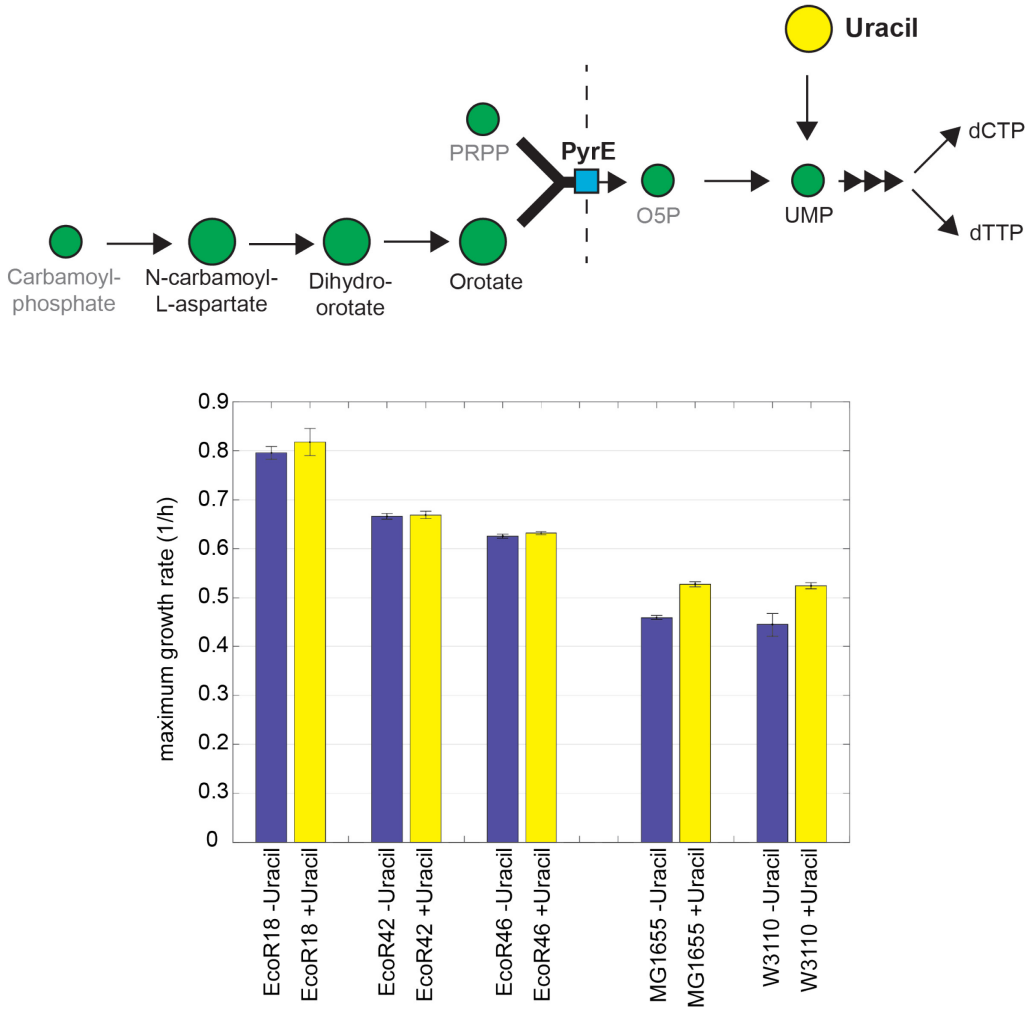

Figure 14: Uracil supplementations

Top: Uracil can be imported and converted to UMP, thereby bypassing the PyrE bottleneck. Bottom: Growth rates of 3 wildtype strains and 2 laboratory strains, grown in presence (yellow bars) and absence (blue bars) of 100 mM Uracil, measured in 3 independent cultivations.

Next, we examined the pyrimidine bottleneck further by artificially introducing it into one of the wildtype isolate strains, EcoR18, which naturally does not have a metabolic bottleneck in the pyrimidine biosynthesis pathway. For the introduction of the bottleneck, we used CRISPRi for transcriptional downregulation of *pyrE*.

3.2 Inhibition of *pyrE* transcription with CRISPR interference

CRISPR interference (CRISPRi) is a method for targeted repression of gene expression, consisting of a catalytically deactivated version of the DNA endonuclease Cas9 that acts as a steric hindrance for RNA polymerases to bind DNA or proceed transcription and a guide RNA that specifies the target site of dCas9 [48].

For transcriptional silencing of *pyrE* in *E. coli* EcoR18, we designed 4 guide RNAs, 3 of which bind within the *pyrE* gene and one without a target sequence that acts as a control guide RNA. It has previously been shown that the binding sites of guide RNAs define the transcriptional repression strength [48]. In order to create bottlenecks of different strengths, we therefore designed the 3 *pyrE* targeting guide RNAs to bind at different parts of the gene, binding either the promoter region (sgRNA1), at the beginning (sgRNA2) or the end of the gene (sgRNA3) (Figure 15a).

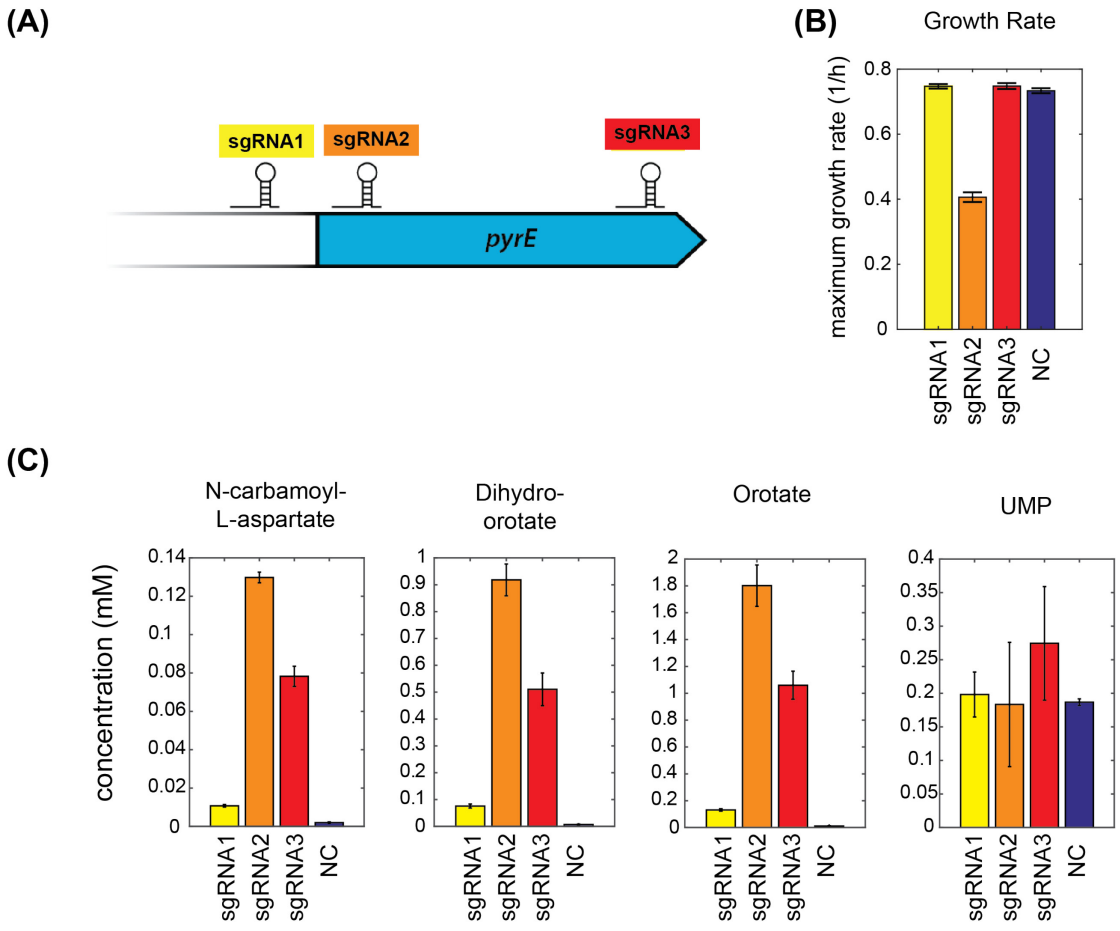

Figure 15: Introduction of a metabolic bottleneck into the pyrimidine pathway using CRISPRi.

(a) Designed guide RNAs and target sites. (b) Growth rates of constructed strains in M9 minimal medium with glucose. (c) Concentrations of all four measurable intermediates of the pyrimidine pathway. NC is the negative control with a non-DNA binding sgRNA.

We found that compared to the control, all three strains showed elevated levels of orotate and its precursors and no significant impact on UMP concentrations (Figure 15c). This is in line with the metabolite profile we observed in the tested laboratory strains, indicating that we successfully introduced bottlenecks in the reaction catalyzed by PyrE. We also found different levels of the three precursors in the three bottleneck strains. The strain with the guide RNA targeting the beginning of the gene (sgRNA2) resulted in the highest accumulation of precursors, whereas the guide RNA targeting the promoter region (sgRNA1) led to a comparably low accumulation. In addition to the metabolomics experiment we also examined the general fitness of the strains and observed similar growth rates for the control stain (NC) and the strains expressing guide RNAs 1 and 3 (Figure 15b). The strain expressing the strongest sgRNA, sgRNA2, however had an about 50 % lowered growth rate, indicating that the growth reduction is dependent on the strength of the metabolic bottleneck.

3.3 Systematic introduction of metabolic bottlenecks in *E. coli*

Encouraged by the successful introduction of a metabolic bottleneck in *pyrE* we were interested how bottlenecks introduced in different parts of the metabolic network of *E. coli* influence the physiology of the cell. Thus, we designed 30 sgRNAs targeting different metabolic genes across the metabolic network and expressed these in the *E. coli* laboratory strain NCM3722. This strain is known not to possess the abovementioned metabolic bottleneck in the reaction catalyzed by PyrE. Moreover, NCM3722 grows exceptionally fast compared to other laboratory strains, indicating that this strain is not severely influenced by metabolic bottlenecks, thus making it a good host strain for the experiment.

The resulting 30 strains have been again analyzed in two ways: First, we measured the growth rates of all strains to identify the strains with impaired growth. Second, we measured the concentrations of the substrates and/or product of the targeted reactions to see if the bottleneck can be observed on the metabolite level.

We found that about half of the strains, including the wildtype and control (i.e. non-targeting sgRNA expressing) strains, had growth rates between 0.6 and 0.65 h^{-1}, suggesting that in these strains CRISPRi did not affect growth. Other strains however had reduced growth rates, with the lowest growth rates measured for strains expressing gRNAs targeting *aroL* and *ilvC* (0.41 h^{-1}) (Table 2). In addition, we analyzed the metabolic profiles of all strains. For 16 of these strains we noticed an at least 2 fold increase of the substrate concentration and/or decrease of product concentration. 9 strains showed no measureable effect on product and substrate concentrations, whereas for 5 strains we could measure neither substrate nor product due to technical limitations of the used method.

Table 2: Target genes of designed sgRNAs

The table is sorted by the measured growth rates. ++ = more than 3 fold change of substrate or product concentration compared to the control, + = 2 to 3 fold change of substrate or product concentration, - = less than 2 fold or no change of substrate or product concentration measurable, * = substrate and product due to technical reasons not measurable.

Target gene	Metabolic pathway	Metabolic pathway class	Growth Rate (h^{-1})	Change of Substrate or Product Concentrations
pyrE	pyrimidine biosynthesis	nucleotide biosynthesis	0.66	++
gadA	L-glutamate degradation	stress response	0.65	-
gshB	glutathione biosynthesis	Cofactor biosynthesis	0.65	++
panC	Pantothenate biosynthesis	Cofactor biosynthesis	0.65	+
glmS	UDP-GlcNAc biosynthesis	Cell wall components	0.65	+
purM	purine biosynthesis	nucleotide biosynthesis	0.65	++
ddlA	peptidoglycan biosynthesis	Cell wall components	0.64	++
LuxS	AI-2 biosynthesis	quorum sensing	0.64	-
dapD	lysine biosynthesis	amino acid biosynthesis	0.64	+
nrdA	pyrimidine biosynthesis	nucleotide biosynthesis	0.64	-
NC			0.63	
pyrB	pyrimidine biosynthesis	nucleotide biosynthesis	0.63	++
cysE	cysteine biosynthesis	amino acid biosynthesis	0.63	-
WT			0.62	
ArgA	arginine biosynthesis	amino acid biosynthesis	0.62	++
leuA	leucine biosynthesis	amino acid biosynthesis	0.61	-
mtn	AI-2 biosynthesis	quorum sensing	0.60	++
dapE	lysine biosynthesis	amino acid biosynthesis	0.59	-
metK	methionine biosynthesis	amino acid biosynthesis	0.59	-
pheA	phenylalanine and tyrosine biosynthesis	amino acid biosynthesis	0.58	*
proB	proline biosynthesis	amino acid biosynthesis	0.57	*
metA	methionine biosynthesis	amino acid biosynthesis	0.57	+
murB	peptidoglycan biosynthesis	Cell wall components	0.53	*
hisB	histidine biosynthesis	amino acid biosynthesis	0.51	+
coaD	coenzyme A biosynthesis	Cofactor biosynthesis	0.50	*
ArgE	arginine biosynthesis	amino acid biosynthesis	0.49	+
folA	tetrahydrofolate biosynthesis	Cofactor biosynthesis	0.49	++
nadA	NAD biosynthesis	Electron carrier biosynthesis	0.49	*
purB	purine biosynthesis	nucleotide biosynthesis	0.49	-
metE	methionine biosynthesis	amino acid biosynthesis	0.43	+
aroL	chorsimate biosynthesis	amino acid biosynthesis, Cofactor biosynthesis	0.41	-
ilvC	isoleucine biosynthesis	amino acid biosynthesis	0.41	+

When comparing growth rates with measurable changes of substrate and product concentrations we found that 10 of 15 strains with a wildtype-like growth rate (>0.60 h^{-1}) had either an elevated substrate or lowered product concentration compared to the wildtype strain, both indicating successfully established bottlenecks, without an effect on the growth rate. One example is the strain targeting *purM* which had both elevated substrates as well as reduced product concentrations (Figure 16) but one of the highest observed growth rates (0.65 h^{-1}). In only 5 of these strains we could not measure any significant differences in substrate or product concentrations.

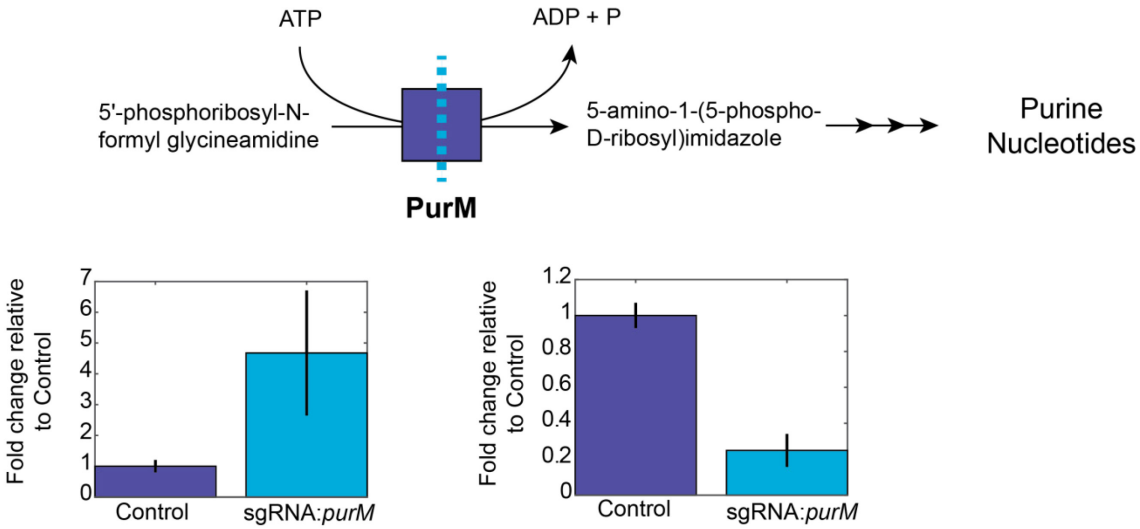

Figure 16: Substrate and Product concentrations of the PurM bottleneck strain.

Concentration of the substrate (left) and the product (right) of the reaction catalyzed by PurM, each displayed as a fold change to the concentration measured in the negative control strain (non-targeting plasmid).

Of the 15 strains with a growth rate lower than 0.59 h^{-1}, in 5 strains we were not able to measure substrate or product concentrations. Of the remaining 10 strains, 6 had measureable alterations of product or substrate concentrations, whereas in 4 strains we could not observe any differences.

We therefore concluded that no clear correlation between introduction of a bottleneck and impaired growth could be determined. Instead, the effects of bottlenecks can be very different: Whereas a few resulted in lowered growth rates (e.g. *folA*, *metE*, *ilvC*), other strains like the abovementioned sgRNA:*purM* expressing strain showed no growth reduction. This might be due to two reasons: First, the results indicate that some metabolic bottlenecks might be compensated by sufficiently high enough final product concentrations. Metabolic bottlenecks in these reactions might lower the flux through them and consequently lower the availability of the end product, however not to an extent that the lower concentration is limiting the growth rate. This finding is of importance for the design of metabolic enzymes with synthetic regulation that should act as conditional bottlenecks: In order to reduce the growth rate, the activity of such an enzyme must be the growth limiting factor. This can only be ensured when the expression rates are so low that any changes of enzymatic activity have direct consequences for the end product

concentration. When the expression rate is too high, an introduction of a metabolic bottleneck might reduce the fluxes through the pathway but not to an extent that the end product concentration is so low that it is limiting the growth rate.

Connected to that is a second possible reason for the observation of strains with metabolic bottlenecks that could not impair the growth rate: Metabolic bottlenecks introduced by CRISPRi are not just dependent on the targeted genes but also on the efficiency of the guide RNA as well as the induction of the CRISPRi system. Optimizing the guide RNAs as well as inducing dCas9 expression might result in stronger bottlenecks and more severe phenotypes than observed here, highlighting again that bottlenecks introduced by enzymes with synthetic regulation must be strong enough to result in growth impairment.

3.4 Summary and consequences for the creation of switching enzymes

The consequences metabolic bottlenecks in single reactions have for the metabolic pathway, network and general fitness state can be very diverse and is dependent on the targeted reaction.

Whereas a metabolic bottleneck in the reaction catalyzed by *pyrE* has a large impact on the fitness as shown for commonly used laboratory strains as well as in an *E. coli* wildtype isolate upon artificial introduction of the bottleneck, in other reactions, metabolic bottlenecks do not necessarily reduce the growth rate when they are not strong enough.

For the design of switching proteins and their identification with a growth-based screening method that means that it is crucial that the enzymes must not be overabundant and a reduction of enzymatic activity has to be so severe that the product concentration is reduced to a level that the growth rates are consequently also reduced.

4 Creation of switchable enzymes using the Split Protein approach

Besides Domain Insertion (see Chapter 5), Split Proteins is one of two techniques which are based on directed evolution and which we evaluated in this work to create metabolic enzymes with synthetic regulation. As a proof of concept, we decided to combine two systems commonly used for protein fragment complementation assays, mDHFR, split in the fragments mDHFR1,2 and mDHFR3[67] and rapamycin-dependent FRAP/FKBP12 [107].

4.1 mDHFR-FRAP/FKBP12

4.1.1 Design and Construction

DHFR is a monomeric enzyme of the folate biosynthesis pathway and catalyzes the reduction of dihydrofolates (DHF) to tetrahydrofolates (THF) which are important C1 group carriers needed in many different metabolic pathways, such as for the synthesis of purines, thymine, serine, methionine and glycine [108–110]. mDHFR has already been used to examine protein-protein interactions and for this purpose split into two fragments [67].

FRAP, commonly known as mTOR (mechanistic target of rapamycin) is a protein kinase involved in different eukaryotic cellular processes, FKBP is a protein folding chaperon. Both proteins do usually not interact; only in presence of rapamycin, a macrolide used as immunosuppressor, both FRAP and FKBP form a protein complex [111]. It has been shown previously that a small fragment of FRAP (105 of 2850 amino acids), the FRB domain, is sufficient for binding of rapamycin and the desired interaction with FKBP12 [107] so that we decided to use only this fragment for the creation of the protein switches.

As linkers we chose very long flexible linkers of the composition $(GGGGGS)_2$ which were supposed to enable the correct folding of adjacent proteins as well as an easier interaction of the fused proteins with their respective interaction partners.

The individual fragments have been amplified by PCR, assembled together using CPEC ([112], see Chapter 8.4.5.4) and cloned onto two expression plasmids. In addition to the potentially rapamycin-dependent fusions FRB-mDHFR1,2 and FKBP12-DHFR3, two controls have been created as well (Figure 17):

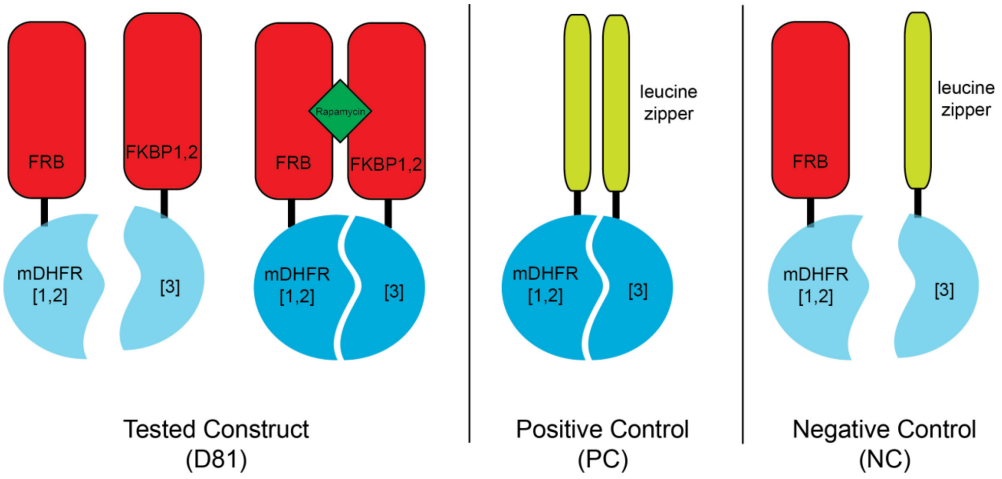

Figure 17: Created split mDHFR variants.

As a positive control, we fused the mDHFR fragments to leucine zippers [113]. Leucine zippers are oligopeptides that form α-helices. Two of these zippers can interact with each other and form coiled coil structures. Originated from eukaryotic transcription factors, in which leucine zippers mediate the interaction of two transcription factors to form oligomers, leucine zippers can also be utilized as domains to bring two fused proteins into close proximity.

As a negative control, we fused the mDHFR fragments to two proteins that are not only not interacting with each other but are, in addition, supposed to act as a steric hindrance of random interaction of the fused enzyme fragments. For that, we used a combination of the two fusion proteins FRB-mDHFR1,2 and ZIP-mDHFR3.

A crucial factor for the creation of split proteins is the order within the fusion proteins. When a protein is split, two new termini, a C-terminus at the first fragment and an N-terminus at the second fragment is created. Usually, the regulatory proteins are fused to these termini. In this case it is necessary that the fused regulatory proteins have an antiparallel orientation to each other. Neither for the leucine zippers, nor FRB/FKBP12 is this the case (Figure 18a + b). To solve this problem, in one of the fusion proteins the enzyme domain has to be connected to the regulatory domain to another terminus as depicted in Figure 18d. Usually, this requires an extended flexible linker connecting both domains of this fusion protein.

In our case fortunately the N- termini of both fragments of DHFR are in very close proximity (Figure 18c) so that FRB and FKBP12 could be fused at these ends.

The resulting three strains have then been examined for rapamycin-dependent growth and metabolism.

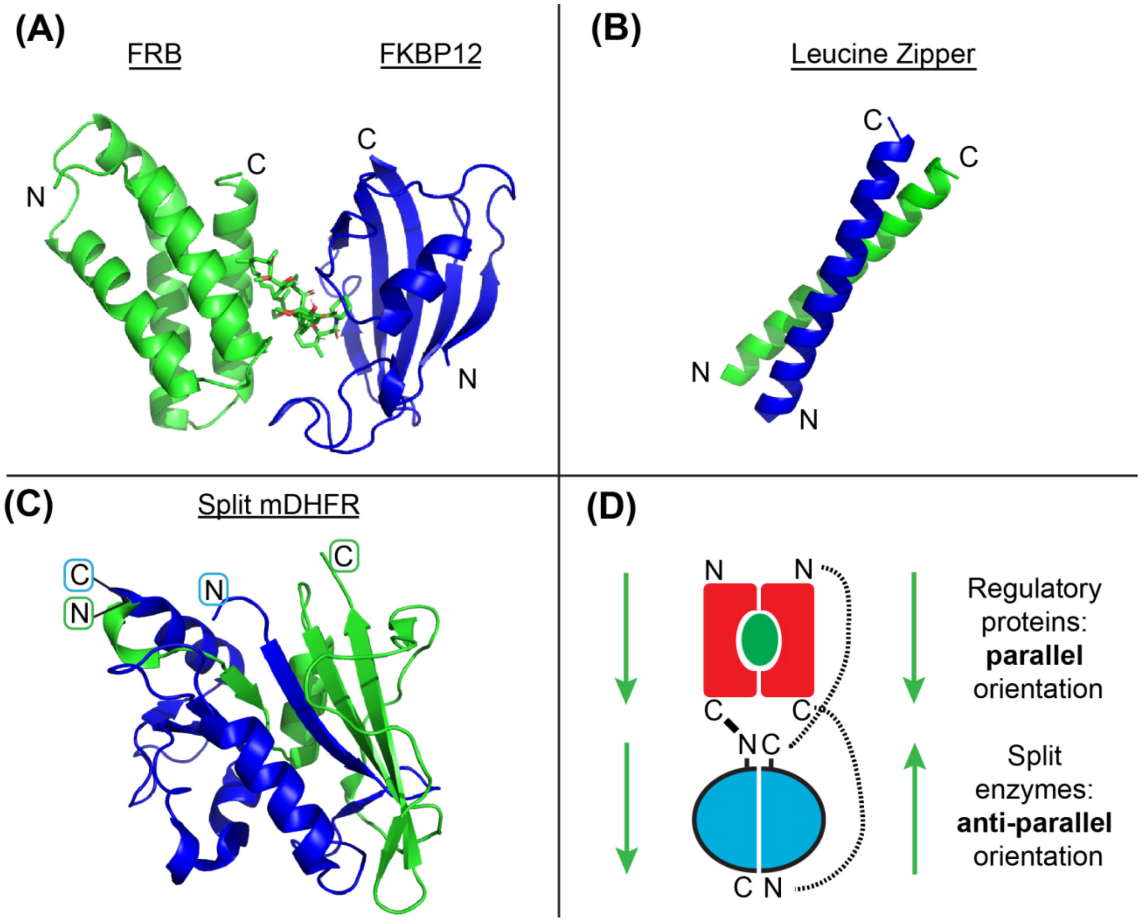

Figure 18: Structures of the used regulatory proteins and split enzyme.

(A) Structure of FRB (green) and FKBP12 (blue) in presence of rapamycin (centre) (PDB: 1NSG, [114]). The C-termini of both proteins are accessible and relatively close to each other positioned so that split proteins are usually fused at these termini. The N-termini are less accessible and therefore less favoured for the fusion of other proteins. (B) Structure of leucine zippers (PDB 2ZTA, [113]). Two leucine zippers form a coiled coil structure consisting of two α-helices. Both termini are accessible but as a consequence of the rigidity of α-helices, split proteins must be both connected either at the N- or the C-terminus. (C) Structure of split mDHFR (PDB: 3D80, [115]). mDHFR is split between L105 and A106. L105 forms the C-terminus of mDHFR1,2 (blue), whereas A106 forms the new N-terminus of mDHFR3 (green). In order to maintain the structure, split proteins are usually connected through these new termini with the regulatory proteins. (D) Dilemma of dimer orientation. Whereas both used regulatory proteins have a parallel dimer orientation so that a connection at either both N- or both C-termini is favoured, split proteins have an anti-parallel orientation (connection to regulatory proteins ideally through the new N- and C-termini). In case two pairs of proteins with different orientations should be connected, one protein has to be connected at the less favourable terminus by an often extended linker (dashed line). In case of split mDHFR-leucine zippers and split mDHFR-FRB/FKBP12 the C-termini of the regulatory domains have been connected to the N-termini of both mDHFR fragments (corresponds to lower dashed line). The distance between both N-termini (8.8 Å) can easily be bridged by the long linkers used here (up to 36 Å per linker in a stretched conformation).

4.1.2 Evaluation of FRAP/FRB-mDHFR

To test the rapamycin dependency of the created FRAP/FRB-mDHFR split protein fusion, we first performed a growth experiment (Figure 19a). For that, we started LB precultures from glycerol stocks. This preculture had then been used to inoculate main cultures in M9 minimal medium supplemented with glucose, casamino acids to reduce the metabolic burden caused by the overexpression of the split proteins, and IPTG to allow expression of the proteins of interest. Furthermore, to this culture we also added 10 µM trimethoprim to inhibit bacterial DHFR so that the cells have to rely on the activity of the split mDHFR. At last, rapamycin has been added in different concentrations for the reassembly of FRB and FKBP.

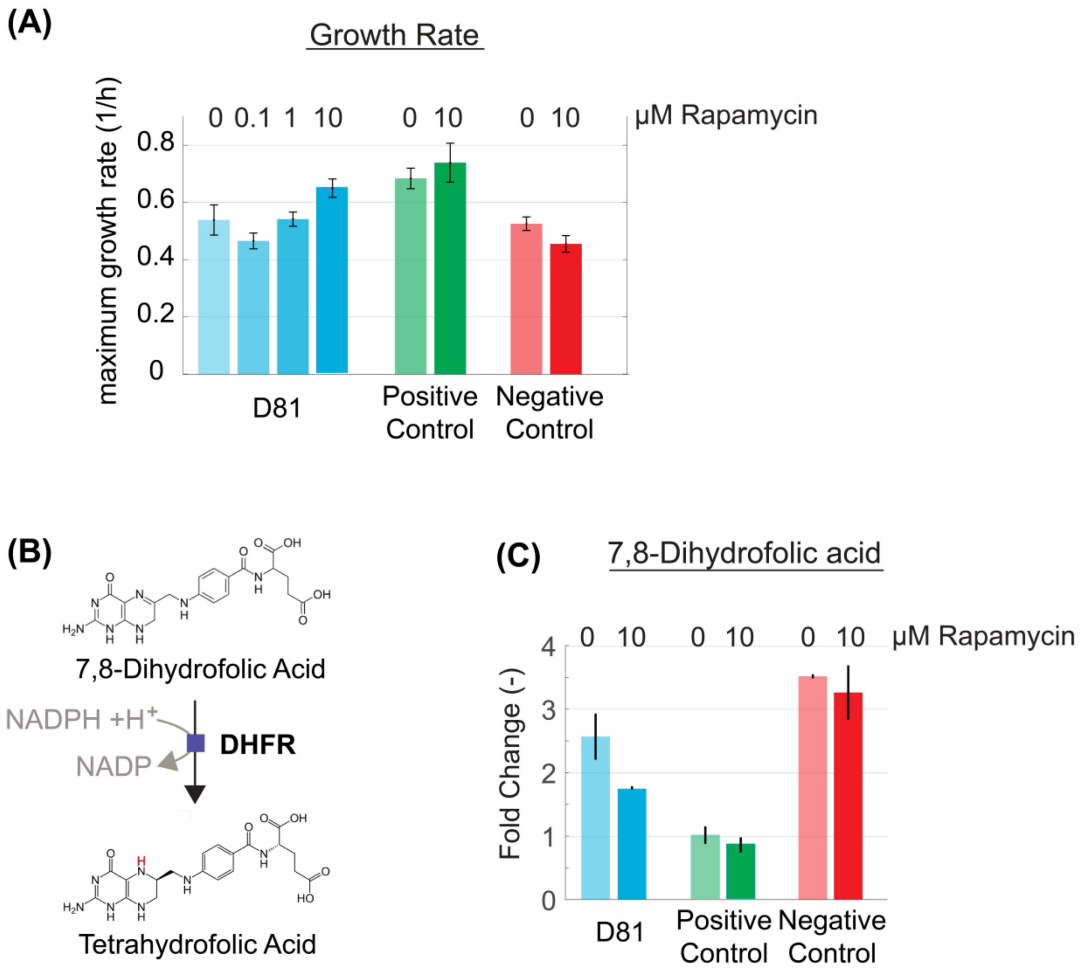

Figure 19: Growth rates and substrate concentrations.

(A) Maximum growth rates of the strain expressing FRB-mDHFR1,2 and FKBP12-DHFR3 (D81, blue), the positive control strain expressing Zip-mDHFR (green) and the negative control strain expressing FRB-mDHFR1,2 and ZIP-mDHFR3 (red) growing M9-Glucose-Casamino acids-IPTG-trimethoprim and varying concentrations of rapamycin. Shown are the results of biological triplicates. (B) Substrate and product of DHFR. (C) Concentration of the most abundant substrate of the reactions catalyzed by DHFR, 7,8 dihydrofolic acid. Metabolite samples were taken from D81 and both control strains in mid-exponential growth phase (OD_{600nm} = 0.5). Samples of three independent cultures were taken and analyzed.

We found that, as expected, the addition of rapamycin to the medium had no significant impact on the growth rates of the negative and positive control strain (Figure 19a). The positive control strain (mDHFR fragments connected to leucine zippers) had the highest growth rates of about 0.7 h^{-1} in presence and absence of rapamycin. The negative control strain (FRB-mDHFR1,2; ZIP-mDHFR3) on the other hand had the lowest growth rate of about 0.5 h^{-1} with a slightly decreased growth rate in presence of rapamycin.

In contrast, the strain expressing FRB-mDHFR1,2 and FKBP12-DHFR3 (D81) showed a rapamycin dependency: In presence of low rapamycin concentrations (0, 0.1 and 1 µM) in the medium, we observed similar growth rates compared to the negative control. The addition of 10 µM rapamycin to the medium however led to an increase of the growth rate to 0.65 h^{-1} which is a growth rate that was almost as high as that of the positive control.

As the tested strains shared the same genetic background with the exception of the expressed split mDHFR variants, we hypothesize that the differences in growth rates have to be explained with different activities of mDHFR.

To test this, we performed a metabolic analysis for which we measured the concentration of the substrate of mDHFR, 7,8-dihydrofolic acid in all three strains at mid-exponential phase (OD_{600nm} = 0.5) (Figure 19c). The product of the reaction, tetrahydrofolate, is usually polyglutamylated and various C1 groups added to it so that over 30 different variants of tetrahydrofolate exist within the cell. These variants are all very low abundant so that we were not able to measure them.

Compared to the 7,8-dihydrofolic acid levels measured in the positive control strain, we found 3-fold elevated levels in the negative control strain, indicating a metabolic bottleneck in the reaction caused by insufficient mDHFR activity. In both strains no impact of rapamycin to the 7,8-dihydrofolic acid levels could be determined. The strain expressing FRB-mDHFR1,2 and FKBP12-DHFR3 on the other hand showed again a rapamycin dependency: Whereas in absence of rapamycin a 2.5-fold increased 7,8-dihydrofolic acid level compared to the positive control could be measured, the level was significantly lower (1.7 fold higher than in the positive control strain) in presence of 10 µM rapamycin. This suggests that the addition of rapamycin leads to the partial removal of a metabolic bottleneck in this reaction.

In conclusion, the here presented data suggests that the created split mDHFR variant fused to FRB/FKBP12 is a rapamycin dependent enzyme switch and can therefore be regarded as our first created metabolic enzyme with synthetic regulation which allows us the – in this case: external – control of a metabolic pathway and as a consequence, the growth rate.

4.2 Discussion and Outlook

Although the created split mDHFR variant can be regarded as our first switchable metabolic enzyme, it is quite apparent that it does not represent an optimal switch. For instance, we observed a high basal growth rate in both, the negative control strain as well as the FRB-mDHFR1,2 and FKBP12-DHFR3 expressing strain when grown in absence of rapamycin. This is probably due to an interaction of the mDHFR fragments independent of the fused regulatory proteins. Although we did not expect the observed extent of complementation as a result of random interactions, it could be explained by the structural properties of split proteins. Many proteins, including mDHFR, contain hydrophobic amino acids in their catalytic center. When split, these amino acids are in contact with the cytosol, which is energetically unfavored [116]. In contrast, when reassembling as a consequence of an interaction of fused regulatory proteins or random interactions, the hydrophobic residues of the two fragments will interact as well, resulting in a stable, energetically more favored state. We can therefore assume that random interactions occur and that once reassembled split proteins will have certain robustness against a re-separation.

In addition, three other factors may have contributed to the observed high basal growth rate: First, it might be that the mDHFR variants were overabundant as a consequence of very high gene expression levels. The higher the copy numbers of split mDHFR variants are within the cell, the higher is the likelihood of random interactions of both fragments. A reduction of gene expression by reduced addition of IPTG might lead to a lower basal activity but would presumably also reduce the growth rates. As the growth rates and 7,8-dihydrofolic acid levels of D81 did not even reach the levels of the positive control in presence of IPTG and 10 µM rapamycin, a reduction of IPTG might only be advisable when the enzyme switch is further optimized.

In addition to that, it should be noted that the addition of casamino acids to the medium reduces the need for functional DHFR as folates are amongst others required for the biosynthesis of serine, methionine and glycine which can be taken up from the surrounding medium.

Finally, it might be that the linker sizes and compositions are not optimal. As mentioned above, the linker was primarily designed to allow both, the enzyme and the fused proteins, to fold correctly and has therefore been chosen to be long and flexible. It might be that the linkers are too long and flexible so that a random interaction of both mDHFR variants is not sufficiently prohibited. Shorter or rigid linkers could improve the switches by reducing unwanted random interactions between enzyme fragments. However, such linkers bear also the risk that folding of adjacent proteins or interaction of proteins with their counterparts might be impaired.

Both, optimal gene expression strengths and linker compositions are planned to be subjects in future rounds of re-design, creation, (possibly) screening and evaluation.

Similar to the assumed unfavored reversibility of assembly of the mDHFR fragments, it has been shown that the dimerization of FRB and FKBP12 is most likely irreversible as well [107]. In

case a reversibly switching enzyme should be created, one therefore either has to use ligands such as FK506 which compete with rapamycin to bind to FKBP12 [117] or use an alternative regulatory protein with a higher tendency to dissociate in absence of its effector.

In addition to these limitations, as rapamycin is a very expensive compound, its usage as inducer molecule in bioprocesses is not eligible, a usage of FRAP/FRB to control overproduction pathways is therefore not possible.

Despite the disadvantages of this existing switch, we would like to use it in the future as a platform to create new switches. For instance, we plan to replace the two mDHFR fragments with enzymes of more biotechnological relevance such as ArgA, LeuA or GPD1 of the glycerol production pathway of *S. cerevisiae*. It should be noted though that these enzymes are in contrast to mDHFR not active as monomers but instead form oligomers in order to be active. Instead of fragment reassembly we would therefore try to control in these cases the oligomerization and in that way the activity of the enzymes through the fused regulatory domains (Figure 20).

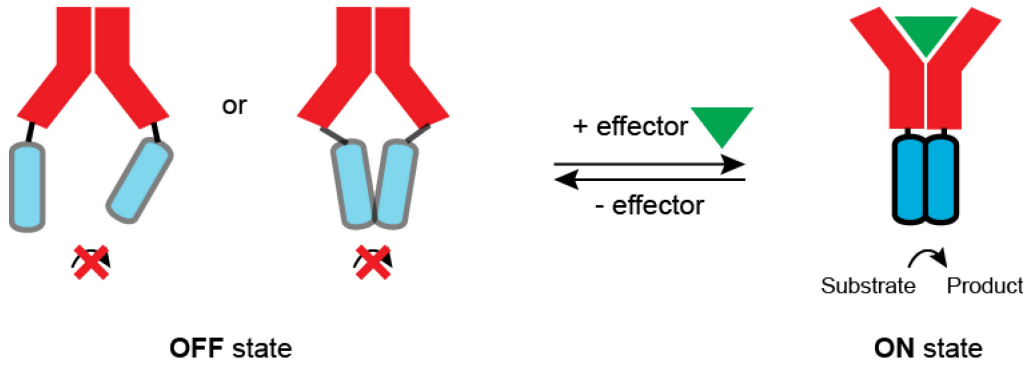

Figure 20: Control of oligomerization.

Many proteins that could potentially be used as regulatory proteins (red) are forming oligomers and only change their conformation upon binding of an effector (green). This conformational change can be used to control either reassembly of split proteins or to control the oligomerization of full length proteins that need to form oligomers in order to be active. Oligomer forming enzymes might form oligomers independent of the fused regulatory domain. In such cases, enzymatic activity is controlled through conformational changes upon binding of an effector.

As the linker composition might be important in these cases as well, we hope to profit from the experiences that we will hopefully gain from the experiments about the optimization of the linkers connecting the mDHFR fragments with FRB and FKBP12.

We will also try to replace FRB/FKBP12 with other regulatory proteins. In particular, we are interested in regulatory domains that sense and are active dependent on growth condition indicators in bioreactors. For example, through the effector binding and oligomerization domain of the transcription factor Cra we could try to control the oligomerization of enzymes such as ArgA and by that, couple the overexpression of biotechnological relevant products to the state of the glycolytic flux. It should be noted though that regulatory domains of transcription factors

such as Cra usually form oligomers independent of the presence or absence of the effector. In contrast to FRB/FKBP1,2, split proteins with effector binding domains of transcription factors as regulatory domains might therefore interact constantly and enzyme activity will not be controlled by protein assembly but through a transmission of the conformational change the fused regulatory domains undergo upon binding of their effectors.

We already tried to construct of an oxygen-dependent mDHFR version by fusing both mDHFR fragments to the oxygen-binding domains of the transcription factor FNR. So far, with the chosen linkers and in the given growth conditions, no complementation of the DHFR knockdown phenotype could be observed.

Planned adaptions of the created split protein are depicted in Figure 21.

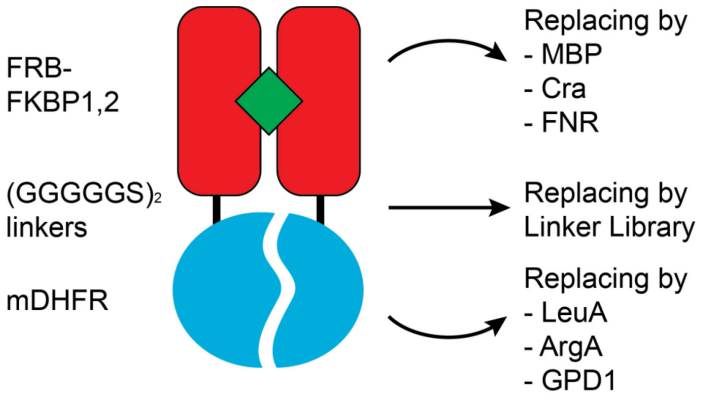

Figure 21: Planned adaptions and optimizations based on the created mDHFR-FRB/FKBP12 split protein switch.

5 Creation of switchable enzymes using the Domain Insertion approach

The second directed evolution-based method that we evaluated in this work to create metabolic enzymes with synthetic allosteric regulation is Domain Insertion in which an enzyme of interest is randomly inserted into a regulatory domain which changes its conformation upon binding of an allosteric effector [65,73].

Proteins that are intended to be used as regulatory domains or metabolic enzymes can be differently suited for Domain Insertion. Our first goal was therefore to define guidelines for the choice of components.

5.1 Guidelines for the choice of components

5.1.1 Regulatory Domains

We defined following 4 criteria for the choice of regulatory domains in Domain Insertion:

First, a crucial characteristic of regulatory domains is their existence as self-contained proteins or domains. In many enzymes, the allosteric binding site is not formed by an isolated domain. Instead, amino acids at different positions within the primary structure of the protein are needed for effector binding and signal transduction, as for instance described for LeuA [118]. Such allosteric binding sites are difficult to isolate and transfer and therefore not suited for Domain Insertion. Second, regulatory domains are preferred that undergo a large conformational change upon binding of their allosteric regulator. The larger a conformational change is, the more likely it is that this change can be transmitted to the fused enzymatic domain and thereby, the more likely is it that the enzymatic activity can be modulated. Third, we preferred regulatory domains and proteins that act as monomers. The reason for this it that enzymatic domains inserted into oligomer forming regulatory domains could sterically impede their oligomerization and consequently their functionality. The fourth criterion is the allosteric effector. Whereas in later applications, this criterion will be one of the most important ones as the allosteric effector defines the function of the resulting enzyme – regulatory domain fusion, for a first test of concept, the identity of the allosteric effector is less important.

Based on these criteria, for our first attempts regulation we chose to work with the maltose-dependent MBP. MBP acts as the periplasmic substrate-binding component of the maltose ABC transporter and is as such involved in the import of maltose. We chose it as regulatory domain because it fulfills the first three criteria best: First, it can be used in its entirety, making it easy to work with. Second, it is active as a monomer thus avoiding problems that can arise from the

insertion of proteins into parts required for the oligomerization. And third, similar to other periplasmic sugar binding proteins, the protein consists of two separate domains with a groove between them which contains the sugar binding site. It has been shown previously that MBP undergoes a large conformational change when binding maltose [119,120] (Figure 22). Moreover, MBP has already been used as regulatory domain in other applications of Domain Insertion [73,78,79].

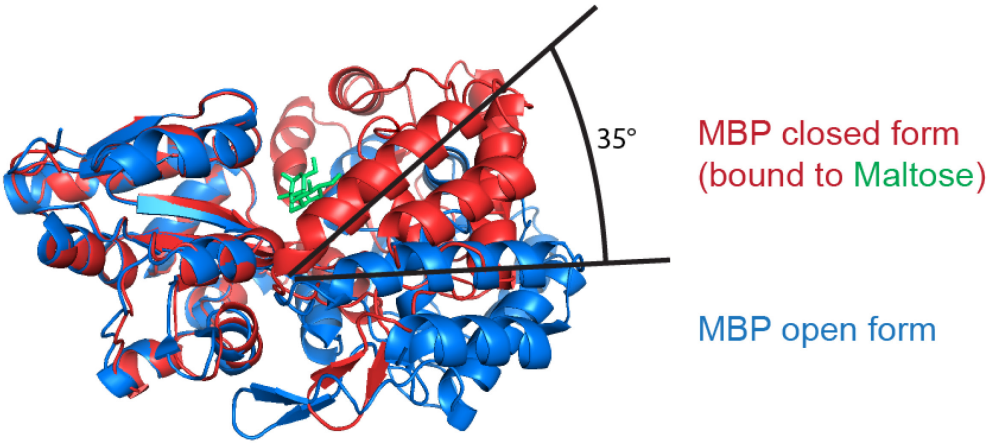

Figure 22: Conformational change of MBP upon binding of maltose.

Alignment of structures of MBP in the unbound, open form (blue, PDB 1JW4, [121]) and when bound to maltose (green) in a closed form (red, PDB 1ANF, [122]). Upon binding of maltose, the clamp shaped structure closes. The hinge angle increases as a result of this motion by 35° [119].

5.1.2 Enzymatic domains

For enzymatic domains in parts similar guidelines apply as for regulatory domains: First, monomeric enzymes are preferred to enzymes that have to oligomerize in order to form catalytically functional enzymes. The reason for this is that an insertion into a regulatory domain might sterically hinder the enzymes' ability to form oligomers. Second, in later applications, enzymes will be mainly chosen to control certain reactions. For our first attempts to create metabolic enzymes with synthetic regulations this criterion is less important. In addition to these criteria, two other rules apply for enzymatic domains: One is that N- and C-terminus of the enzymes are ideally in as closest proximity as possible. The idea behind that is that the insertion of the enzymatic domain separates the two fragments of the regulatory domain with the same distance as the distance between the N- and the C-termini of the enzymatic domain. The longer this distance is, the less realistic is a successful folding as well as a functional reassembly of both fragments of the regulatory domain. The second additional rule is that the enzymatic activity has to result in a measurable phenotype, e.g. fluorescence [72], antibiotic resistance [73] or changes in growth rates.

Based on these criteria, for our first attempts to create metabolic enzymes with synthetic allosteric regulation we selected murine dihydrofolate reductase mDHFR (folate biosynthesis)

and 2-isopropylmalate synthase LeuA (leucine biosynthesis) as our enzymes. Furthermore, the usage of ArgA (arginine biosynthesis) and HisG (histidine biosynthesis) has been evaluated.

mDHFR had been chosen because of its relatively simple structure. It acts as a monomer and possesses termini that are in very close proximity (murine DHFR: 14.8 Å, see Figure 23), allowing the insertion of DHFR in regulatory domains with relatively small linkers. In addition, mDHFR has already been used in Domain Insertion although with a more targeted approach and with DHFR as acceptor domain [77]. For us, DHFR has in addition the advantage that we also used it as Split Protein (Chapter 4).

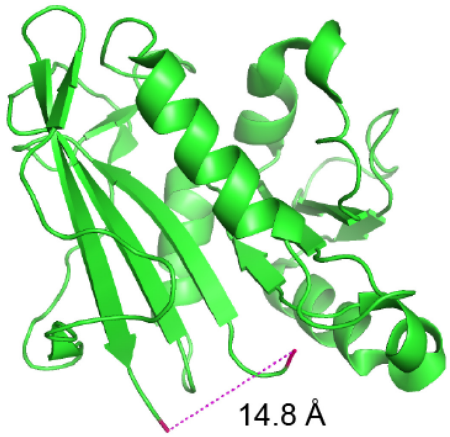

Figure 23: Structure of mDHFR.

PDB 3d80 [123]. Marked is the distance between both termini.

LeuA has been chosen for its higher relevance for biotechnology [124,125] but has also the advantage of having a relatively simple structure consisting of an N-terminal catalytic domain and a C-terminal regulatory domain, connected through a subdivided linker structure [118] (Figure 24). LeuA forms dimers and although both N-terminal domains of a dimer form independent catalytic sites, the C-terminal domain is required for catalytic activity, presumably by ensuring protein dynamics [126]. We therefore decided to work with the full length LeuA and assume chimera with LeuA as enzymatic domain will form oligomeric complexes.

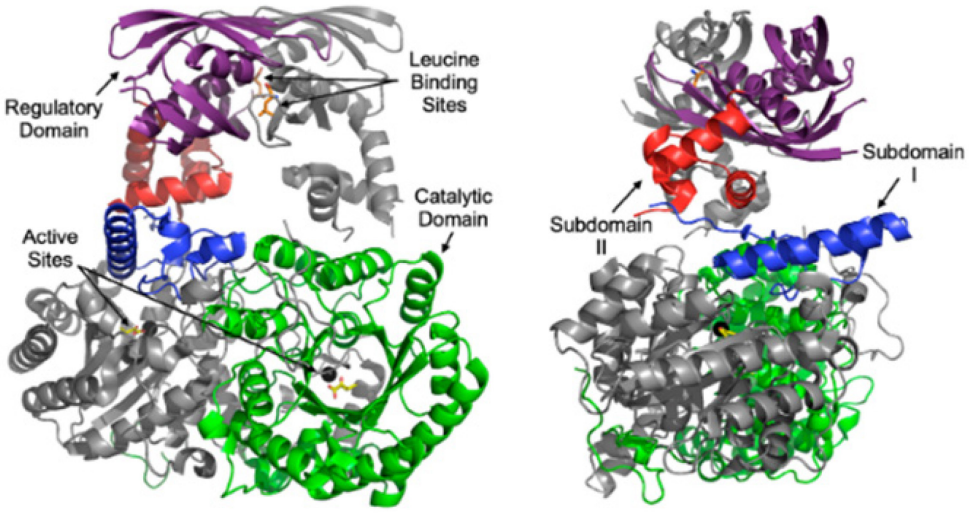

Figure 24: Domain Structure of LeuA [126].

LeuA consists of 3 domains, a N-terminal catalytic domain (green), a C-terminal regulatory domain (purple) and a linker structure in between which consists of two subdomains (blue and red). LeuA forms homodimers with two catalytic sites and one leucine binding site.

5.2 Complementation of gene knockout and knockdown phenotypes

Metabolic enzymes with synthetic allosteric regulation are intended to work as metabolic valves with two states of activity: In one condition (e.g. presence of an effector), the enzyme is intended to be active, thereby not limiting fluxes. In the second condition (e.g. absence of the effector), the enzyme has a lowered activity. In order to ensure that the lowered activity also results in a reduced flux through the pathway, it is crucial that the abundance of the enzyme switch is sufficiently low since less active but overabundant enzymes would allow for maintaining high fluxes through the pathway. As a consequence, the expression of the enzyme switch genes must be tightly regulated.

To ensure that, an expression plasmid was required that allows very low gene expression rates. Hence, we decided to work with the plasmid pSB4A5, a low copy plasmid (origin of replication: pSC101, 4-5 copies per cell) with an arabinose-inducible P_{BAD} promoter to control the expression of the gene of interest (Figure 25).

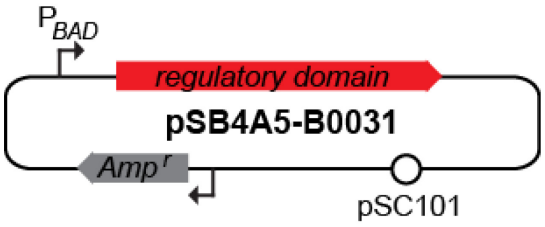

Figure 25: Expression plasmid pSB4A5

In this work, we would like use these enzymes to not just to control metabolic fluxes but consequently also to control the growth rates of the enzyme expressing strains. To reach that, it is important that the activity of the enzyme switch limits the growth rate. Hence, we have not just to ensure low gene expression rates but also that the genes are expressed in genetic backgrounds in which the enzymatic activity is required for growth.

As mentioned above, we are interested in the metabolic enzymes ArgA, HisG, LeuA and mDHFR. We amplified their genes and cloned them on the selected expression plasmid pSB4A5. The resulting plasmids were then transformed in the respective knockout strains. As no knockout strain is available for *folA* (encoding DHFR), the plasmid has been transformed into *E. coli* MG1655 and incubated the resulting strain in presence of trimethoprim which specifically inhibits the native bacterial DHFR.

With the resulting strains, we performed a complementation experiment. For that, we incubated the strains and *E. coli* MG1655 as negative control in M9 glucose minimal medium and varied gene expression rates by the addition of arabinose in different concentrations (Figure 26).

For all complementation strains, we expected limited growth in absence of arabinose as a result of low gene expression and increasing growth rates with higher arabinose concentrations.

For the Δ*argA* complementation strain, we observed this behavior: while low arabinose concentrations resulted in low growth rates between 0.11 h^{-1} (0 g/L arabinose) and 0.45 h^{-1} (0.5 g/L), indicating insufficient *argA* expression rates, we observed a growth rate of 0.58 h^{-1} in presence of 5 g/L arabinose, suggesting high ArgA abundance. For Domain Insertion we will incubate the cells with limited arabinose concentrations, e.g. with 0.5 g/L arabinose. This concentration allows on the one hand substantial growth; on the other hand, we can assume that with this arabinose concentration the ArgA abundance is growth limiting so that any change in enzymatic activity would also affect the growth rate.

For the Δ*hisG*, Δ*leuA* and DHFR knockdown complementation strains, the highest growth rate was not reached with the highest concentration of arabinose (5 g/L). Instead, we assume that the full induction of the P_{BAD} promoter leads to an overexpression of the enzyme genes and consequently to a growth burden. For Domain Insertion the strains will be incubated in presence of arabinose concentrations that are slightly below the optimal concentration. HisG and mDHFR variant expressing strains will be incubated in total absence of arabinose as the basal activity of the promoter already leads to sufficient *hisG* and *mDHFR* expression rates. LeuA variant expressing strains will be incubated with 0.1 g/L arabinose.

The growth rates of the negative control strain *E. coli* MG1655 were not significantly affected by added arabinose, showing that increased growth rates of the complementing strains in presence of arabinose are not caused by the availability of a second carbon source but rather a result of higher gene expression rates and improved complementation.

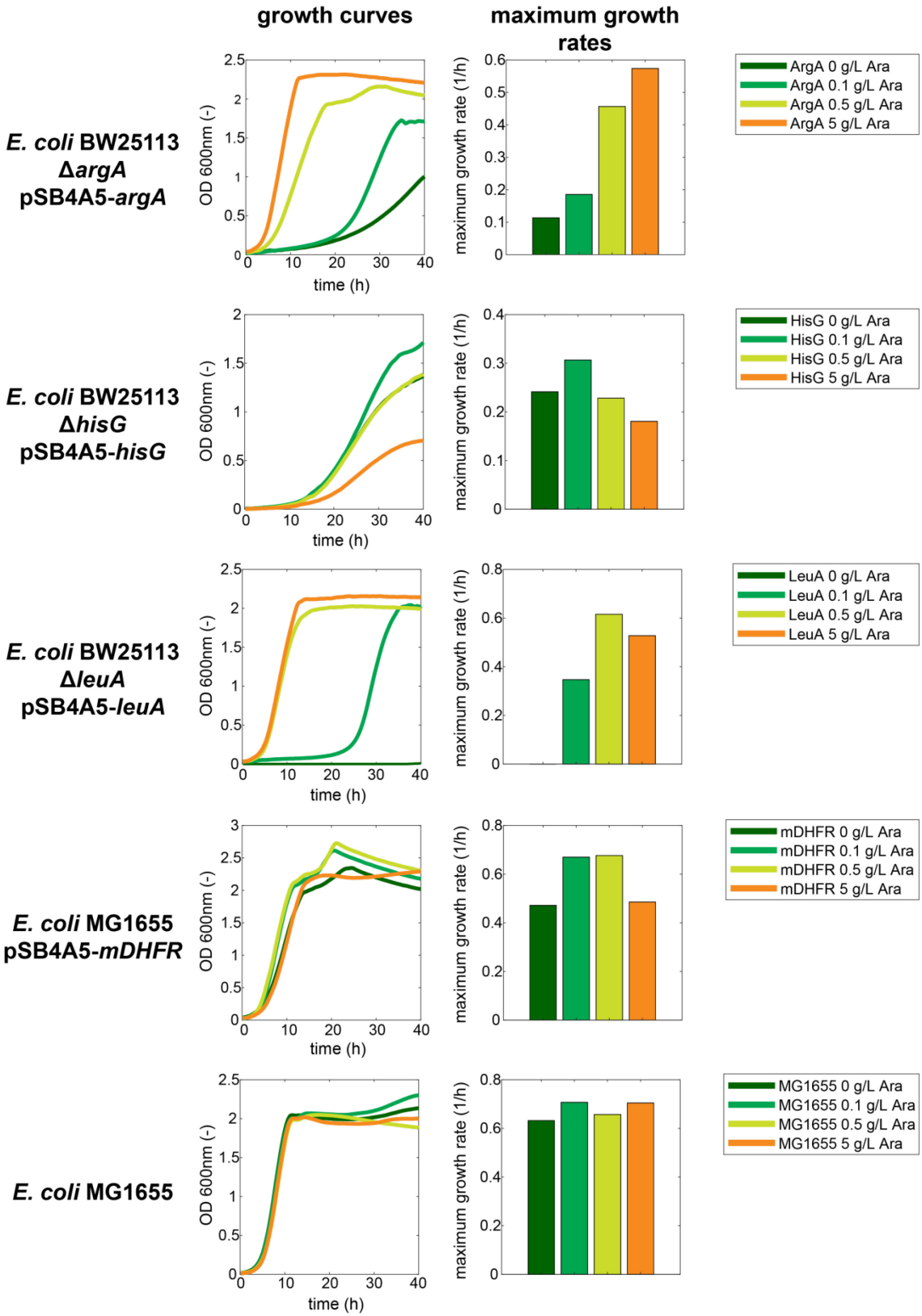

Figure 26: Complementation of gene knockouts and knockdowns using pSB4A5 as expression plasmid.

Shown are the growth curves and maximum growth rates of 5 strains grown each in presence of 4 different concentrations of arabinose, the inducer of the P_{BAD} promoter that controls the expression of the enzyme gene required for complementation. Top row: Δ*argA* complementation strain, second row: Δ*hisG* complementation strain, third row:

Δ*leuA* complementation strain, fourth row: *E. coli* MG1655 strain complementing the bacterial DHFR inhibition phenotype, fifth row: *E. coli* MG1655 as control strain. Dark green: 0 g/L Arabinose, light green: 0.1 g/L, yellow: 0.5 g/L, orange: 5 g/L Arabinose.

5.3 Construction of libraries and optimization of the Domain Insertion protocol

We then sought to apply the Domain Insertion for the creation of metabolic enzymes with synthetic allosteric regulation. For that, we selected mDHFR as enzymatic domain, linkers of the composition $(GGGGS)_2$ and the maltose binding MBP as regulatory domain, and followed the Domain Insertion Protocol [74] which consists of 4 parts as depicted in Figure 27: In the first part (Part A) an expression plasmid with the regulatory domain gene is randomly linearized and repaired, resulting in a library of plasmids with different insertion sites. In the second part (Part B), the enzyme gene of interest is amplified. In the third part (Part C), the plasmid library is constructed by ligation of previously prepared linearized plasmids and enzyme genes. And finally, after pooled transformation in the enzyme gene knockout strain, the resulting strain library can be screened for conditionally active enzyme variants (Part D).

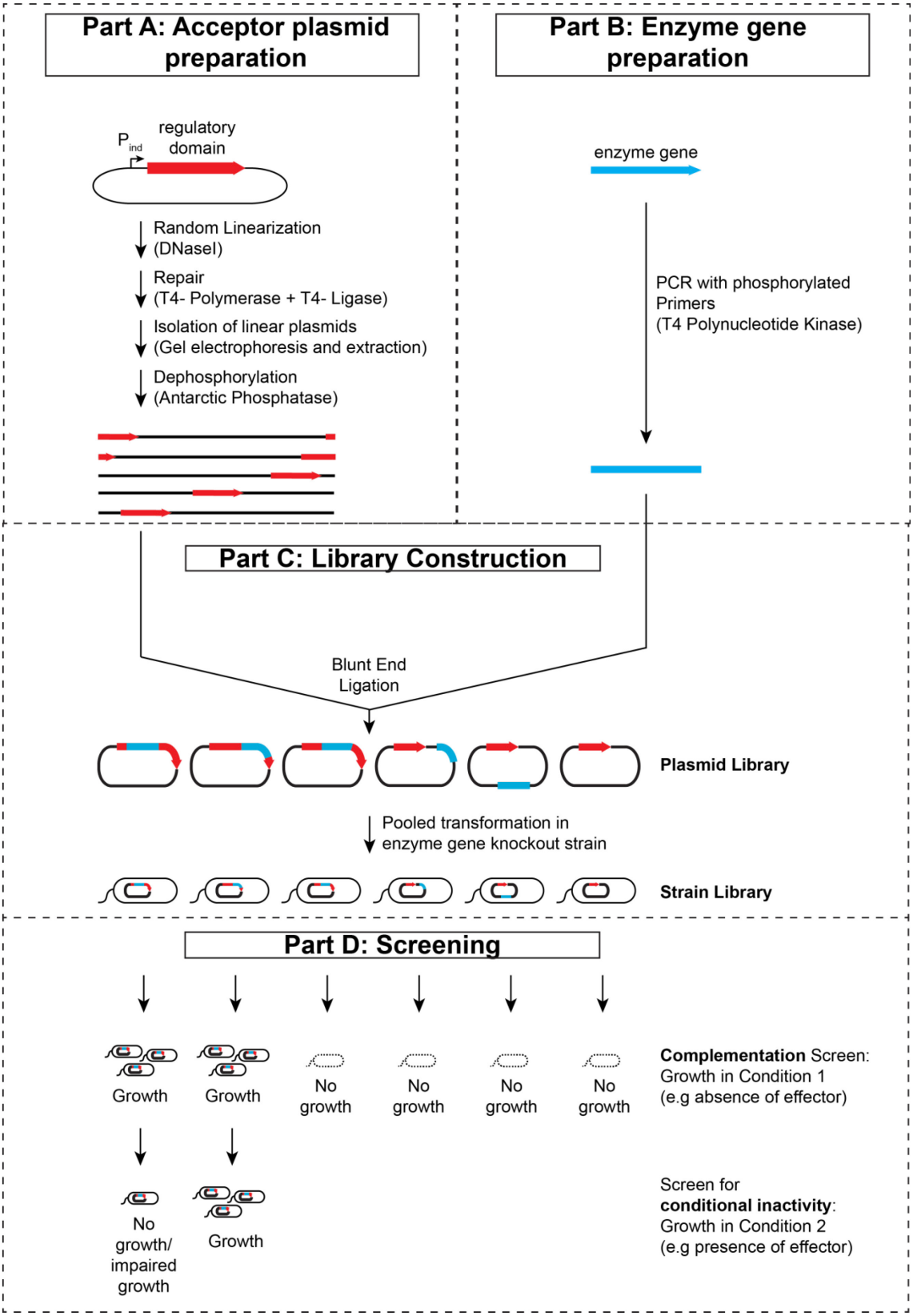

Figure 27: Overview over the original Domain Insertion protocol.

The protocol is divided in four parts. Part A: Preparation of the plasmid harboring the gene for the regulatory domain. The plasmid is randomly linearized with the sequence-independent nuclease DNaseI. Sticky ends are blunted as well as nicks sealed in the next step by T4 polymerase and T4 ligase. As only a fraction of the plasmids is linearized by DNaseI, this fraction is isolated by gel electrophoresis and gel extraction in the next step. Finally, the linearized plasmids are

dephosphorylated to prevent self ligation. Part B: Preparation of the enzyme gene fragment. The enzyme gene is amplified without start and stop codons and with linker sequences on one or both ends if needed and subsequently purified. Part C: Library generation and screening. Enzyme genes and linearized plasmids are ligated first to create a pooled plasmid library. Dependent on the site the plasmid was linearized by DNAseI the gene is inserted at different posititions within the plasmid. Next, the plasmids are transformed into a enzyme gene knockut strain to create a pooled strain library. Part D: To find enzymes with synthetic allosteric regulation, two steps of screening are necessary. First, the strains are tested for functional enzyme activity by incubating them in minimal medium. Plasmid in which the gene is inserted inside the regulatory domain but out of frame, in the wrong direction or when the insertion site is not suitable to allow the formation of a functional enzyme variant (3rd plasmid), plasmids in which the enzyme gene inserts outside of the open reading frame of the regulatory domain (4th and 5th plasmid) or religated plasmids (6th plasmid) will not result in functional enzymes and therefore will not support growth. Next, the complementing strains are screened in a second round for condition-dependent enzymes. For that, cells are incubated for example in presence of an effector. A majority of the complementing enzymes will be unaffected by the effector, resulting in normal growth (2nd plasmid). We are interested in enzymes with impaired activity in presence of the effector which will subsequently cause lowered growth rates (1st plasmid).

Part A: Preparation of the acceptor plasmid

Prior to Domain Insertion we cloned the MBP coding *malE* gene on the expression plasmid. To obtain sufficient amounts of plasmids, 120 mL cultures of the *E. coli* strain harboring the plasmid have been used for plasmid preparation. 60 µg of plasmids could be isolated and used for the following steps.

For the linearization of the plasmid we initially used the sequence-independent nuclease DNaseI. First, we determined the optimal concentration of DNaseI for the digest. The targeted concentration allows a linearization of a substantial part of the plasmids while simultaneously overdigestion is prevented. We found that with a concentration of 10^{-4} U/µg DNA a significant ratio of the plasmids was linearized (Figure 28b) and used this concentration for the further digests. Lower concentrations do not result in linearized plasmids, whereas higher concentrations lead to an overdigest. The determined concentration has been used for the random linearization of the in the first step prepared plasmids.

After linearization, the plasmids were repaired for which T4 DNA ligase was used for sealing of nicks and T4 DNA polymerase for blunting of sticky ends. We then isolated the fraction of linearized plasmids. For that, we loaded the plasmids on a TAE agarose gel and extracted the band formed by linearized plasmids which migrate faster than nicked circles and slower than supercoiled DNA (Figure 28a).

At last, in order to prevent self-ligation in the subsequent ligation step, the ends of the linearized plasmids have been dephosphorylated using Antarctic phosphatase. After a last DNA purification step, about 1 µg of randomly linearized, repaired, blunted and dephosphorylated plasmid was obtained which could be used for ligation.

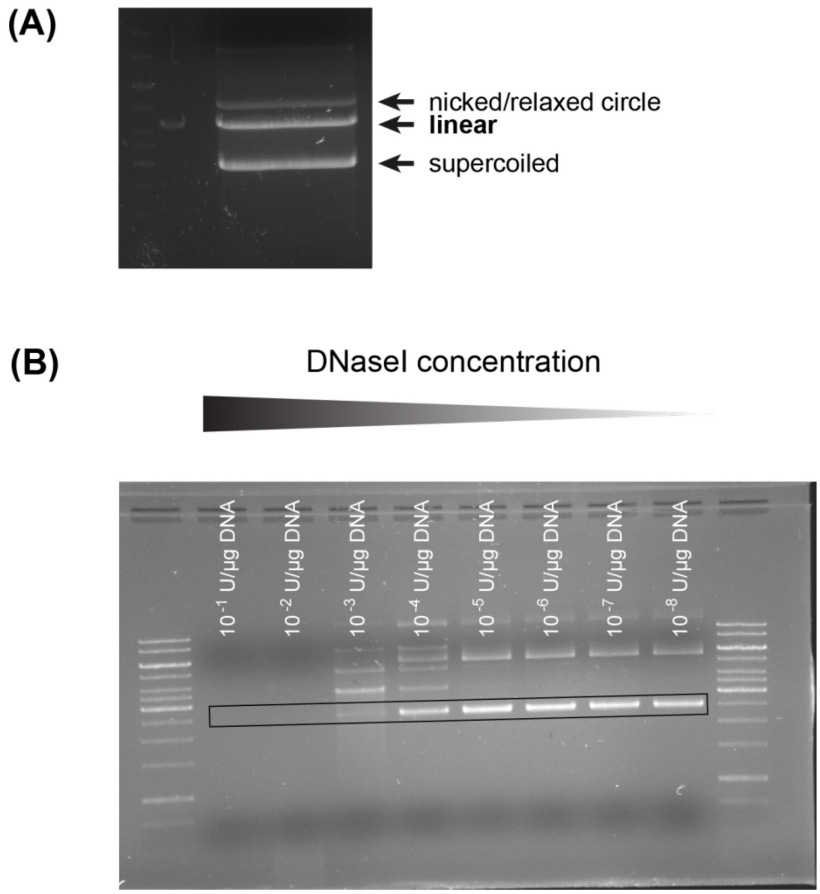

Figure 28: DNaseI digestion

(A) Digest of plasmids leads to the formation of fractions of linear DNA, nicked/relaxed plasmids and still undigested, supercoiled DNA. These fractions can be separated by gel electrophoresis. (B) Determination of an optimal DNaseI concentration. Linear plasmids are marked with a black box. A concentration of 10^{-1} and 10^{-2} U/µg DNA leads to a complete digestion, whereas a concentration of 10^{-5} U/µg DNA and lower does not result in a sufficient linearization. A concentration of 10^{-3} and 10^{-4} U/µg DNA resulted in a sufficient amount of linear plasmids.

Part B: Preparation of the enzyme gene

For the preparation of the enzyme gene, *mDHFR* without start and stop codons had been amplified with phosphorylated primers and subsequently purified. To the 5' ends of both primers, sequences coding for long flexible glycine linkers of the composition $(GGGGS)_2$ were added.

Part C: Library construction

Linearized plasmids containing *malE* and insert consisting of *mDHFR* and linkers sequences at both ends were ligated in a blunt end ligation reaction. Following the original protocol we used T4 DNA ligase and PEG-6000 which was added to increase the likelihood of interactions of blunt ends in the reaction mix. After purification, the plasmids were transformed into chemically

competent *E. coli* NCM3722 cells and plated on M9 minimal medium agar plates with glucose as carbon source and Trimethoprim to inhibit bacterial DHFR.

We found that the original protocol resulted in insufficient library sizes. For instance, after plating of *E. coli* NCM3722 strains with plasmids constructed following the original Domain Insertion protocol, we obtained only 11 strains.

In addition, by sequencing some of these strains we noticed an overdigestion of the acceptor plasmid, resulting in the deletion of parts of the regulatory domain gene. Resulting enzyme variants may therefore be functional but usually do not possess a functional maltose binding domain.

To increase library sizes and reduce the number of overdigested plasmids, we therefore sought to optimize the Domain Insertion protocol.

Optimizations of the protocol in order to increase library sizes

In order to increase the number of transformants with candidate enzyme variants, we evaluated and optimized almost all steps in several rounds of Domain Insertion. All adaptions to the protocol that led to an increase of library sizes are summarized in Table 3.

Table 3: Adaptions to the Domain Insertion protocol

Step	Change	Remarks
Linearization	DNaseI → S1 Nuclease	Avoiding over-digestion as S1 nuclease cuts only circular, not linear DNA. Also used by Tullman, Guntas, Dumont, & Ostermeier (2011)
Repair	T4-Ligase → T7-Ligase	Avoiding re-ligation of linearized plasmids.
Dephosphory-lation	Increased Incubation time, different enzymes tested	Avoiding plasmid re-ligation in the subsequent ligation reaction
Ligation	Different protocols, suppliers, enzymes and kits tested	Increasing the number of successful ligations while avoiding multiple ligation events.
Transformation	Different transformation protocols and suppliers of competent cells tested	Increasing the number of transformants after transformation of the ligation mix

As mentioned above, we noticed that the usage of DNaseI to randomly linearize plasmids often led to an overdigestion of plasmids, resulting in the deletion of parts of the regulatory domain gene and often to frameshifts. Resulting enzyme variants may therefore be functional but usually do not possess a functional maltose binding domain. To solve this problem and by that enhance the quality of produced libraries, we therefore sought to replace DNaseI with S1 nuclease which has been shown to linearize specifically supercoiled double stranded DNA and already been used for Domain Insertion [75].

We evaluated S1 nuclease for its usage in Domain Insertion. In particular, we were interested if S1 nuclease can linearize a sufficient fraction of the plasmids and if the linearization is sequence-independent. To test this, we followed the suggested reaction conditions [75] and found that compared to DNaseI, a larger fraction of the plasmids in the reaction had been linearized (Figure 29a) which is beneficial for the desired increase of library sizes.

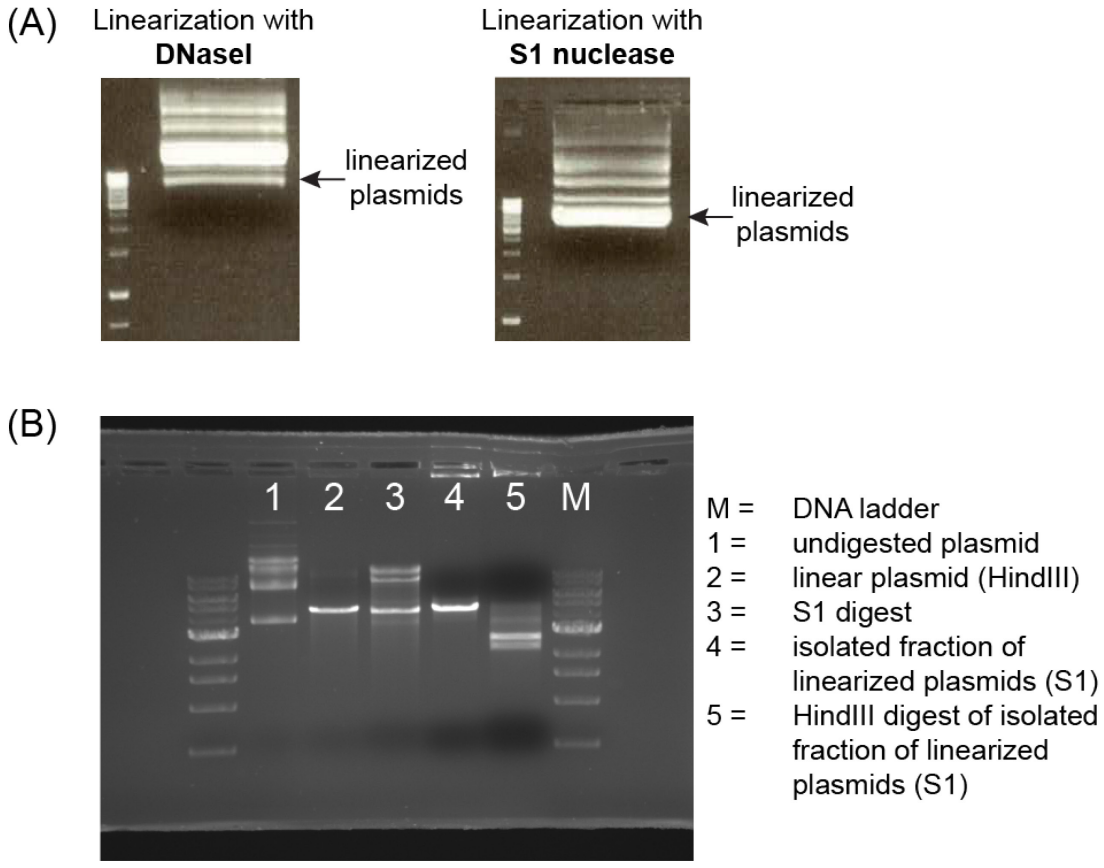

Figure 29: Digestion of plasmids with S1 nuclease.

(A) Comparison of DNaseI and S1 nuclease digestion. The digestion of plasmids with S1 nuclease leads to a higher yield of linearized plasmids. Note that here a boric-acid based agarose gel buffer has been used, leading to linearized plasmids to migrate fastest through the gel. (B) Evaluation of the sequence dependency of S1 nuclease. Shown is the undigested plasmid (1), the S1 digest (3) and as comparison the plasmid linearized with a restriction enzyme (2). The linearized plasmids of the S1 digest have been isolated (4). This fraction has then been digested with a restriction enzyme (5). A biased S1 digest leads to a predominant formation of certain fragments as observed here.

By additionally digesting the linearized plasmid with a sequence specific and single-cutting restriction enzyme, we tested if linearization of S1 nuclease is sequence-independent: Whereas a random linearization would result in many products of different sizes, in case of a preference for a certain sequence, the predominant formation of certain products could be observed. When linearizing our expression plasmid with S1 nuclease and additionally digest the linear plasmids with HindIII, we found that two fragments with sizes of about 2200 and 2800 bp formed, indicating a sequence preference od S1 nuclease (Figure 29b). Hence, although larger amounts of linearized plasmids could be isolated through gel extraction when digested with S1 nuclease, a smaller variance of insertion sites is expected.

In addition to the replacement of DNaseI by S1 nuclease we also tested different enzymes and reaction conditions for subsequent protocol steps. For instance, we replaced the T4 ligase with T7 ligase for the repair of nicked plasmids. The reason for this is that T7 ligase is known to have a lower preference for blunt DNA ends and therefore reduces the likelihood of unwanted re-circularization of previously linearized plasmids. For dephosphorylation and ligation, we found that Antarctic Phosphatase and a premade ligation master mix specifically developed for blunt end ligations (Blunt/TA ligation master mix, New England Biolabs) resulted in a higher number of transformants.

We also speculated if the low number of transformants could have been caused by the direct transformation of the ligation reaction into the respective knockout strain and plating on minimal media plates which probably leads to stress the cells may not be able to cope with. We therefore decided to do two sequential transformations: First, the ligation reactions were transformed in highly electrocompetent laboratory strain cells, the resulting transformants were plated on rich media plates to avoid a bias in the library and reduce stress. The transformants were then pooled and the plasmids isolated. These isolated, amplified and purified plasmids were then retransformed in the final step into the enzyme gene knockout strain.

With all abovementioned adaptions, we were able to significantly increase the quantity of strains within libraries. A list of all created libraries can be found in Table 4. Of most libraries we isolated several individual strains which were in parts further characterized in more detail by comparing the growth rates when growing on maltose or glucose minimal medium (Chapter 5.4). In addition, our last constructed library with MBP as regulatory domain, linkers of the composition $(GGGGS)_2$ and DHFR as enzymatic domain consisted of more than 2000 strains and we could have increased the number of strains even more by repeating and upscaling the final transformation step. For such large libraries an isolation and individual screening of single strains is not feasible anymore. Instead, all colonies were pooled and intended to be screened in pooled growth assays.

Table 4: Composition and sizes of created and stored Domain Insertion strain libraries.

Library Name	Plasmid Backbone	Regulatory Domain	Linker	Enzyme	Size
pSB4A5-malE-αH-leuA	pSB4A5	MalE	α helices	LeuA	30
pTRC99KK-malE-GS-mDHFR	pTRC99KK	MalE	Glycine-Serine	mDHFR	38
B0031-malE-GS-mDHFR	pSB4A5	MalE	Glycine-Serine	mDHFR	543
B0031-malE-GS-mDHFR pooled	pSB4A5	MalE	Glycine-Serine	mDHFR	> 2000 (pooled)

5.4 Screening of constructed MBP-mDHFR libraries

Constructed libraries were screened in two steps: First, to identify functional enzymes, strains were tested for the ability to complement the native enzyme gene knockout. As for DHFR such a gene knockout strain does not exist, we instead treated the cells with the antimicrobial trimethoprim that inhibits bacterial DHFR. For the complementation screen, transformants were plated on M9 minimal medium plates so that only strains could grow in which the enzyme gene has been successfully inserted into the regulatory domain gene, in frame and in the right direction. All strains listed in Table 3 were able to grow on minimal medium plates and thus, possess catalytically active enzyme variants.

Next, we screened for strains that possess enzymes that are sensitive to maltose, i.e. that have a reduced enzymatic activity in either presence or absence of maltose. As mentioned earlier, to identify enzymes with synthetic regulation, we screen for strains which have low growth rates in presence or absence of the intended allosteric effector, in our case maltose. Hence, we did comparative growth experiments by letting library strains grow in presence and absence of maltose. For instance, we measured the growth rates of 38 individual strains expressing MBP-mDHFR in M9 minimal medium with either glucose or maltose as sole carbon source and supplemented with trimethoprim to inhibit native bacterial DHFR activity (Figure 30). We found that almost all of the tested 38 strains had growth rates between 0.7 and 0.9 h^{-1} when growing on glucose and between 0.45 and 0.75 h^{-1} when growing on maltose. One of the strains, K21, however had an interesting growth phenotype: With glucose as carbon source, the growth rate was relatively low but comparable to the other measured strains (Figure 30a), while when growing on maltose, a reduced growth rate could be observed. Such a growth phenotype would be typical for a Mal OFF enzyme switch: While in absence of maltose, the enzyme is active, thereby enabling sufficient fluxes and consequently a high growth rate, in presence of the effector maltose, the enzyme is less active, leading to reduced fluxes and consequently reduced growth rates.

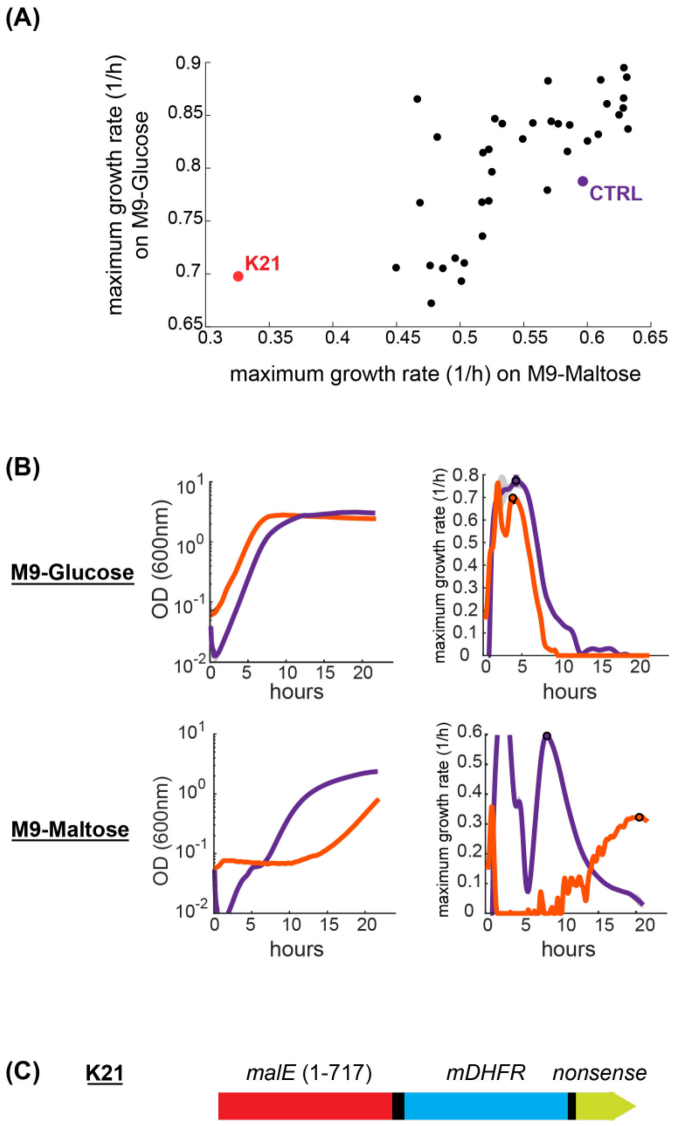

Figure 30: Screening of 38 Domain Insertion library strains and subsequent characterization of two strains.

(A) 38 *E. coli* strains expressing potentially maltose-dependent mDHFR variants and a control strain expressing solely mDHFR were grown on M9 minimal medium with either glucose or maltose as carbon source. Shown are the maximum growth rates in both growth conditions as scatter plot. Marked in orange is a strain with a low growth rate on maltose while growing with an average growth rate on glucose, a phenotype expected for strains with a maltose-sensitive mDHFR variant (K21). As a comparison, in purple the control strain (CTRL) has been marked. (B) Growth curves and growth rates of K21 and the control strain on M9-Glucose and M9-Maltose. (C) Sequenced gene of *malE* (red) with inserted mDHFR gene (blue) from K21.

We sequenced the expression plasmid of K21 and found that, as intended, *mDHFR* was inserted into *malE*. However, despite using S1 nuclease to randomly linearize the plasmid, we also observed that parts of *malE* were deleted, probably due to overdigestion. We noticed that

the deletion led in this case to a frameshift so that the sequence adjacent to the mDHFR gene is no longer coding for the second MBP fragment (Figure 30c). The MBP was as a consequence truncated and consisted only of the amino acids 1-239, whereas the amino acids 240-396 and with that, important parts of the maltose-binding pocket (W340, Y341)[122] were deleted. It is therefore very unlikely that the resulting enzyme switch is able to bind maltose and to control the activity of the fused enzymatic domain upon.

Growth experiments have been performed for most of the created strain libraries with potentially switching enzymes. In most cases, a few strains showed initially an interesting growth phenotype in one of the growth conditions. However, so far none of the strains that showed conditional growth impairment were also found to possess switching proteins. Instead, we noticed that almost all sequenced plasmids were found to be overdigested – to our surprise independent of whether DNaseI or S1 nuclease had been used for the linearization – so that often large parts of the maltose binding protein could not be expressed.

5.5 Discussion and Outlook

Domain Insertion is a promising technique to create metabolic enzymes with synthetic allosteric regulation as enzymes and regulatory domains of interest can be combined to create potentially switching enzymes without the need for detailed knowledge about the structural properties of the involved proteins. In this project, we apply this approach to create such enzymes with the purpose to implement internal control of metabolic fluxes and consequently the growth rate.

For that, we first defined needed specifications of enzymes and regulatory domains and then successfully tested an expression plasmid that allows very low and consequently growth rate limiting gene expression rates.

Using both as groundwork we then tried to construct first metabolic enzymes with synthetic allosteric regulation. However, using the original protocol [73] only an insufficient number of library clones could be obtained. From previous applications of Domain Insertion we know that large library sizes are required in order to identify enzymes with synthetic allosteric regulation. A reference point of how many colonies are needed in order to isolate an enzyme with synthetic allosteric regulation is the study of [73]: Of an original library of 1.000.000 strains, only 0.8% (8000 strains) processed a functional enzyme. 10 % of those strains (800 strains) possessed in addition a functional regulatory domain. These strains were next screened for condition dependent enzymes. Although it has not been mentioned how many of these 800 strains showed condition-dependent enzyme activity, it is quite clear from these observations that thousands, if not millions of strains have to be produced in the first place in order to find a few strains that express enzymes with novel regulation.

We therefore sought to optimize the protocol in order to increase the quantity of strains with candidate enzymes. We spent a significant amount of time and effort on the optimization

process and by changing most steps we were able to significantly increase library sizes. The production of a first large library consisting of over 2000 strains expressing functional and potentially condition-dependent enzymes was already successful and by repeating the transformation step, thousands and even millions of strains could be produced.

Need for improved library quality

However, not just the quantity, also the quality of obtained strains turned out to be problematic. When sequencing, we found that most plasmids previously linearized with DNaseI were overdigested. This can be very problematic as an overdigestion can cause i.) a frame shift and by that the expression of dysfunctional nonsense protein fragments at the C-terminus of the protein chimer or ii.) simply the deletion of crucial parts of the regulatory domain. To circumvent this problem, we decided to work with an alternative nuclease, the S1 nuclease. This nuclease has been reported to solely digest supercoiled DNA so that a digest of already linearized plasmids should be prevented. In addition, the successful usage of this nuclease in Domain Insertion has already been shown [75]. We found that the linearization with S1 nuclease has the advantage that a larger fraction of plasmids can be linearized and consequently more linear plasmids prepared for the subsequent ligation step. However, we also observed a clear preference for a certain cut site, 2200 and 2800 bp away from a HindIII cut site. This indicates that the preferred S1 cut site is located in the origin of replication. Indeed, such a preference for inverted repeats commonly found in origins of replications has been reported [127]. Plasmids in which the enzyme gene is inserted in this plasmid region is expected not to result in a functional plasmid as the replication of the plasmid might be affected, as well as the inserted gene will presumably not be able to be transcribed and translated. As a consequence, we assume that the higher yield of linearized plasmids will not result in more plasmids with enzymes inserted in the regulatory domain gene. In addition and in contrast to previous observations [75], even for plasmids linearized with S1 nuclease we noticed that a majority of plasmids is still overdigested.

We therefore conclude that although the successful applications of DNaseI and S1 nucleases for sequence-independent linearization of plasmids for Domain Insertion have been reported [73,75], as a result of the observed high number of overdigested plasmids we regard both nucleases as not suitable for Domain Insertion. Instead, alternative linearization methods are required in which overdigestions are prevented. In recent publications, such alternative methods have been presented: One uses random transposon insertion and subsequent replacement with an enzyme of interest [79]. In another paper – interestingly from the same laboratory that developed the DNaseI and S1 nuclease approach – multiplex inverse PCR has been used to create linearized plasmids [76]. The idea of this method is to create primer pairs for each possible insertion site and to amplify the whole plasmids. These plasmids can then be pooled and used for an insertion of the enzymatic gene via blunt end ligation. The big advantage of this technique is that all plasmids are linearized in frame and within the regulatory domain gene so that insertions outside the regulatory gene and overdigestions of the regulatory domain are prevented. Compared to the linearization by DNaseI and S1 nuclease, an amplification of the

acceptor plasmid has the disadvantage of being more laborious and expensive. However, we regard this approach as a promising technique to increase library sizes and especially to improve the quality of constructed libraries and therefore decided to use multiplexed inverse PCR in the next rounds of Domain Insertion.

In addition to overdigestion by both nucleases, we speculate that the observed high ratio of overdigested plasmids and *vice versa* the low abundance of plasmids in which the enzyme had been correctly inserted into the regulatory domain might have been caused by another factor: It could be that the used combinations of enzymes, regulatory domains and linkers that have been tested so far are not suitable for Domain Insertion. For instance, this might be the case when enzymes are used that cannot correctly fold when other full length regulatory proteins are attached to the ends. In such a case, by plating on minimal plates we select for functional enzymes with only fragments of the enzymes attached so that their ratio in the library is high. Impaired protein functionalities as a result of insertions into other proteins are dependent on many factors such as the selected proteins, the insertion site but also linker size and composition. A prediction which combinations are suited for Domain insertion is therefore generally difficult. Instead, it is advisable, to test many different combinations or optimize the components. For instance, linker compositions can be varied very easily and for the optimizations of enzymes, methods like circular permutation could be applied that have already been shown to improve existing enzyme switch variants to be more responsive to the allosteric effector [74].

Components that are planned to be used in future attempts of Domain Insertion

We plan to test more combinations of enzymes, linkers and regulatory domains. For that, we selected several components of each group that are going to be combined with each other (Figure 31).

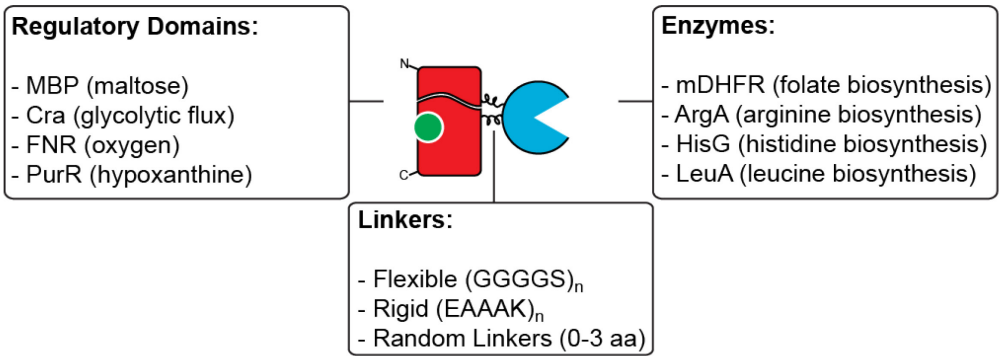

Figure 31: Regulatory domains, enzymatic domains and linkers that are planed to be used for Domain Insertion.

We still intend to work with MBP as regulatory domain and mDHFR as enzyme domain which are both monomeric proteins and therefore very suitable for Domain Insertion. However, several other regulatory domains and enzymes might be less suited as they form oligomers but would

be more interesting from a biotechnological perspective. As regulatory domains, we selected the effector binding proteins of the transcription factors Cra, PurR and FNR. All of those are responsive to metabolites that are important signal molecules for bioprocesses: As previously mentioned, Cra is responsive to fructose-1,6-bisphosphate and by that acts as a glycolytic flux sensor. This means that the flux into a production pathway could be coupled to the availability of e.g. glucose within bioreactors. FNR binds and is dependent on oxygen so that, similar to Cra, the flux can be controlled according to the availability of an important nutrient. PurR on the other hand is activated by hypoxanthine and guanine, degradation products of purines. It has been shown recently, that hypoxanthine accumulates quite rapidly in starving cells, whereas in good growth conditions, hypoxanthine levels are low [128]. Hence, hypoxanthine can act as a metabolic indicator for starvation so that by using the metabolite binding domain of PurR as regulatory domain in Domain Insertion, fluxes could be coupled to the fitness state of a host cell.

As enzymatic domains, we selected in addition to mDHFR and LeuA two other enzymes, ArgA and HisG. LeuA, ArgA and HisG are all responsible for the catalysis of the first step of their respective amino acid biosynthesis pathways and thereby usually allosterically regulated. For the construction of overproduction strains, the allosteric binding sites have already been identified and feedback resistant mutants described. By using these enzymes, we could therefore test synthetic allosteric regulation in bioprocesses. As shown above, we already tested the complementation of the native gene deletion phenotype for all these enzymes so that the required growth conditions are already determined.

Linkers are the smallest components in Domain Insertion, their sizes and compositions are crucial for the creation of enzymes with synthetic allosteric regulation by Domain Insertion. This has exemplarily been shown for a GFP-MBP switch that was only slightly maltose dependent before optimizing the linkers [78]. The introduction of a linker library resulted in different variants that showed a maltose dependency that was up to 300% higher or 50% lower than the original biosensor.

The reason for the crucial role linkers have in Domain Insertion is they have to fulfill two tasks: On the one hand linkers have a transmission role, i.e. they have to transmit the conformational change the regulatory domain undergoes upon binding of its allosteric effector to the fused enzymatic domain so that the conformation of the enzyme and consequently its activity changes as well. To allow this signal transmission, linkers are ideally as short and rigid as possible. On the other hand, linkers may not interfere with the folding of the individual domains. For that, linkers should be long and flexible enough.

As linkers, we so far used only flexible glycine-serine linkers. Our intention is to use other linker types, such as rigid α-helical structure forming linkers of the composition $A(EAAAK)_nA$ (n= 2–5) [129] and short random linkers of sizes 0 – 3 amino acids.

In conclusion, with the optimized protocol we are now able to create large libraries and would like to further improve the method with a new plasmid linearization approach. We plan to create many and very large libraries by combining several enzymes, regulatory domains and linkers.

So far, a screening of individual strains was possible due to small library sizes. However, for screening of new libraries, a high-throughput screening and enrichment method is required. For that, we evaluated the usage of the single cell growth rate reporter TIMER.

6 Enrichment of slow growing cells out of complex strain libraries using the fluorescent growth reporter TIMER

Parts of following chapter have been published in ACS Synthetic Biology [130]:

Beuter D, Gomes-Filho J, Randau L, Díaz-Pascual F, Drescher K, Link H (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

As mentioned in the previous chapter, the creation of switching enzymes with directed evolution methods such as Domain Insertion or Fusion/Split Proteins usually involves the generation of large libraries of enzyme variants that potentially act as switches., i.e. which possess catalytic activity that is condition dependent. Such libraries have to be screened – ideally in a high-throughput manner – in order to identify enzymes with the desired functions. As explained in more detail in Chapter 2.4.2, different approaches exist to identify such enzymes of which one is to couple enzymatic activity to growth: By expressing enzymes in conditions in which the enzymatic activity is defining the growth rate, every change in metabolic activity, e.g. as a result of implemented synthetic allosteric regulation, will subsequently lead to changes in the growth rates. This can be used to screen for the screening of condition-dependent enzymes: Whereas unregulated enzymes will support growth in all conditions, enzymes with synthetic allosteric regulation are in certain conditions (presence or absence of the effector) less active, hence, the growth rates of the expressing cells will be reduced in this particular condition. These cells with reduced growth rates are planned to be identified and enriched.

Several methods to enrich fractions of strains with low growth rates have been developed (Chapter 2.4.2.2) but none of those was suitable for our purposes. Instead, we decided to develop our own enrichment method based on the single cell growth reporter TIMER, a fluorescence protein that appears green when freshly expressed but that maturates over time to a red fluorescent form. It has previously been shown that TIMER appears red in slow growing cells whereas fast growing cells appear greener. However, a usage as a tool to enrich slow growing cells out of large strain variants growing in batch cultures has not been reported so far. We therefore first sought to verify how good TIMER can display the growth rate in *E. coli* batch cultures.

6.1 Validation of TIMER to display the growth rate in *E. coli*

6.1.1 Dynamics of TIMER

Our first goal was to analyze the behavior of TIMER, i.e. the ratio of green and red fluorescence in a batch culture over time, from the inoculation to the stationary phase. For that, we transformed the plasmid pBR322-C_TIMER into *E. coli* NCM3722. For the growth experiment, we first started a LB preculture which we used in the second step to inoculate a M9-Glucose preculture. This second preculture was incubated for 24 hours at 37°C and shaking in Erlenmeyer flasks to ensure full maturation of TIMER. From this preculture a main culture was inoculated in a multiplate reader, enabling continuous measurements of optical densities and levels of green and red fluorescence. Moreover, we added small doses of the antimicrobial trimethoprim (Tmp) in sublethal concentrations (0, 0.25 and 0.5 µg/mL) to the medium, resulting in cultures with different growth rates.

We found that in *E. coli* batch cultures, the TIMER fluorescence ratio followed a certain pattern which is depicted in Figure 32b: In the first growth phase – from inoculation until early exponential growth phase – all three cultures behaved the same: At the beginning, the green/red fluorescence ratio is very low which could be expected since the cells in the M9 preculture that had been used for inoculation were in the stationary growth phase for many hours and thereby not growing, so that TIMER appears mainly red. After about 3 hours we then observe for all cultures a sudden large increase in the green/red fluorescence ratio. The reason for this observation is the technical limitation of the used microplate reader which has a higher sensitivity for detecting green fluorescence compared to red fluorescence. At the beginning of the incubation, the cell density is very low so that the detector cannot measure anything but red background fluorescence. With increasing cell density, first green fluorescence can be measured accurately and only later red fluorescence, leading to the observed sharp peaks of green/red ratios.

In exponential growth phase, the ratio of green to red fluorescent TIMER first suddenly drops, as from this time point on, red fluorescence can be detected. The ratio then increases. This observation is in line with the previously proposed model [102] (Figure 32a) that with every cell division the fraction of red fluorescent, maturated TIMER is diluted and replaced with freshly expressed, green fluorescent TIMER. The ratio increased further until late exponential growth/early stationary phase was reached when the ratio usually reaches maximum. From this moment on the ratio decreases with an initial fast and then slower decreasing ratio. Interestingly, the three cultures did not reach the same green to red ratio even many hours after reaching stationary phase.

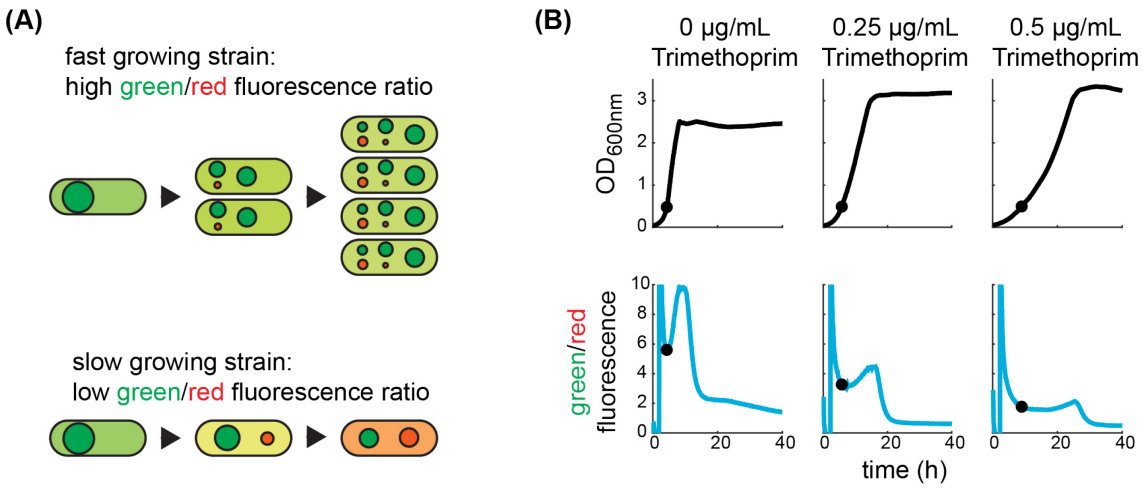

Figure 32: Growth curves and TIMER dynamics of three cultures with different growth rates.

(A) Reported appearance and behavior of TIMER in fast and slow growing strains. (B) Three cultures of *E. coli* NCM3722 pBR322-C_TIMER grown in presence of different sublethal concentrations of Trimethoprim to alter the growth rates. At the top, the growth curves are depicted as black lines with the time point of the maximum growth rates marked with black dots. At the bottom, the TIMER dynamics over time is shown with a blue line. Black dots mark the time point of maximum growth.

Based on these and previous observations [102] and having the technical limitations at low optical densities in mind, we assume that at the point of inoculation, the cells all are mainly red, then with each cell division appear more and more green and turn red again after entering stationary phase. We assume that in the first phase and can show that in the stationary phase, cells with different growth rates cannot be distinguished by their TIMER fluorescence ratios. For the usage of TIMER to identify and isolate slow growing cells within libraries of strains with different growth rates, however, it is crucial that cells with different growth rates can be distinguished as easy as possible. Hence, it would be ideal if the differences of TIMER ratios between slow and faster growing cells would be as large as possible. One emerging question from the observed TIMER dynamics was therefore, at which time point is the difference of TIMER ratios largest, i.e. when is the optimal time point to separate slow from fast growing cells. To answer this question, we therefore wanted to examine the dependency of TIMER appearance from the growth rate in more detail and for that, compared many cultures with a high variance in growth with their TIMER ratio in different growth phases.

6.1.2 Correlation of TIMER and growth rate

To see how well the TIMER ratio correlates with the growth rate at different growth phases, we repeated the previous growth experiment with 22 cultures in technical triplicates growing in presence of different Tmp concentrations spanning from 0 to 1 µg/mL. These concentrations impaired growth very differently, from no growth impairment (maximum growth) in absence of Trimethoprim to a by 90% reduced growth rate in presence of 1 µg/mL Tmp. Of these cultures,

maximum growth rates were calculated as well as the green/red fluorescence ratios measured at three different ODs representing different growth stages: OD = 0.15 (early exponential phase), 0.5 (mid-exponential phase) and 0.8 (late exponential phase). We then checked how well the maximum growth rates correlated with the TIMER behavior at the different time points (Figure 33).

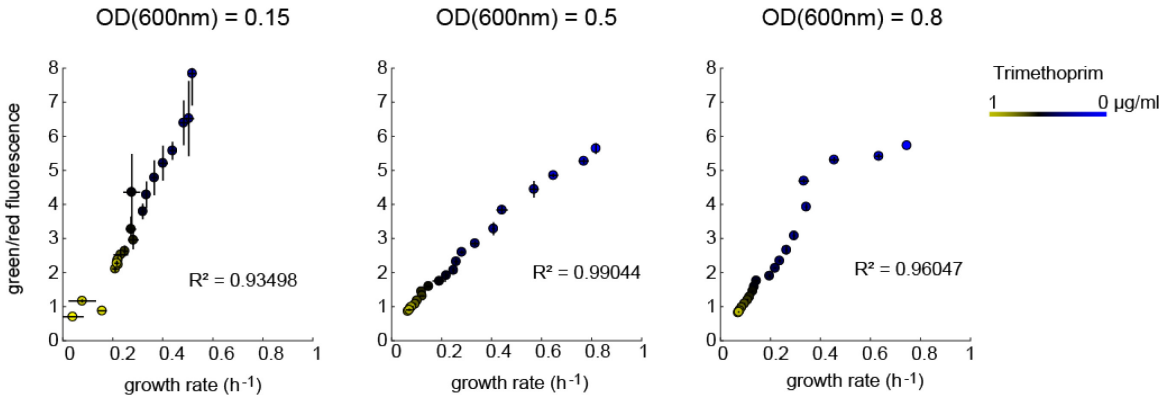

Figure 33: TIMER appearances in strains with different growth rates at different growth stages.

TIMER expressed in *E. coli* NCM3722 pBR322-C_TIMER grown in presence of Trimethoprim in sublethal concentrations between 1 and 0 µg/mL. The TIMER appearance has been compared with the maximum growth rate of each culture at three different growth stages: Early exponential phase (OD= 0.15; left), mid-exponential phase (OD= 0.5, middle) and late exponential phase (OD= 0.8; right). The calculated Pearson coefficient (R^2) is displayed in the bottom-right corners. This experiment has been performed with technical triplicates.

We found that TIMER correlates well and linearly with the growth rates at all phases of exponential growth with coefficients of determination (R^2) between 0.93 and 0.99. For the utilization of TIMER in the enrichment of slow growing cells, a linear correlation is desirable because it would theoretically enable also the enrichment of only partially growth impaired strains, whereas non-linear correlations (e.g. bimodal behavior of TIMER with either completely green or completely red cells) would only allow the separation of a limited number of fractions.

The best correlation could be observed at mid-exponential growth (OD= 0.5), whereas in the late exponential growth phase, the correlation was weaker in faster growing cells. In early exponential phase (OD= 0.15), the difference between green/red fluorescence ratios of slow and fast growing strains was largest. However, we also observed a large variation in the fluorescence ratios. This might be due to biological noise since in early growth stages it is expected that a larger fraction of the whole populations is still in the lagging phase and therefore appearing mainly red, whereas other cells already divided so that these cells appear greener,

resulting in a less homogeneous population. Another factor that might contribute to the higher variation might be the abovementioned technical limitation to measure red fluorescence in cultures with low density.

As pointed out before, for library sortings, a large variation would be desirable to allow an easier separation of slow and fast-growing cells. However, due to the observed variance in the TIMER signals and the expected biological noise at lower ODs, we think that the growth rate display at mid-exponential phase is more reliable and hence, sortings at mid-exponential growth would be more reasonable. All following experiments therefore show TIMER at mid-exponential growth.

6.1.3 Comparison of TIMER expressed from high and low copy number plasmids

For library sortings with FACS, it is desirable that TIMER reflects the growth rate with as little noise as possible. In detail, that means that cells with a certain growth rate are supposed to have as less cell-to cell variations as possible.

For this work, two plasmids with different copy numbers were available, a low copy number plasmid with a pSC101 origin of replication and a high-copy number plasmid with a pBR322 origin [102] which was edited in this work (replacement of the antibiotic resistance cassette to be compatible with plasmids used in enzyme switch construction).

We transformed these plasmids into 3 different *E. coli* wildtype strains (MG1655, NCM3722 and BW25113) and examined the TIMER – growth correlation in the resulting strains. For that, we let the cells grow in M9 medium with different carbon sources (glucose and glycerol) and the antibiotics to select for strains still harboring the plasmid and measured optical densities and green and red fluorescences in microtiter plate readers.

We found that the strains harboring the high-copy number plasmid had growth rates reduced by 15-20% compared to the control strains without any plasmid (Figure 34). No growth burden could be observed for the strains with the low copy number TIMER plasmid. In large parts, the growth burden will be caused by the excessive expression of TIMER, although an additional growth burden caused by differing antibiotics used for the selection for plasmid carrying cells cannot be excluded. However, although this burden might be problematic for other applications of TIMER, for the screen of pooled libraries we expect no severe limitations caused by it since this growth burden is affecting all cells within the library similarly.

We also tested the correlations between TIMER signal and maximum growth rates. Here we found that the correlation using the high copy number plasmid is stronger with less variance compared to the strains expressing TIMER from the low copy number plasmid. Our explanation for this is – again in addition to the technical limitations of measuring low fluorescence signals –, a presumably higher variance due to the low copy numbers of expressed TIMER inside the

cells: The more TIMER molecules are within a cell, the higher is the robustness of the ratio of green to red fluorescent TIMER ratios.

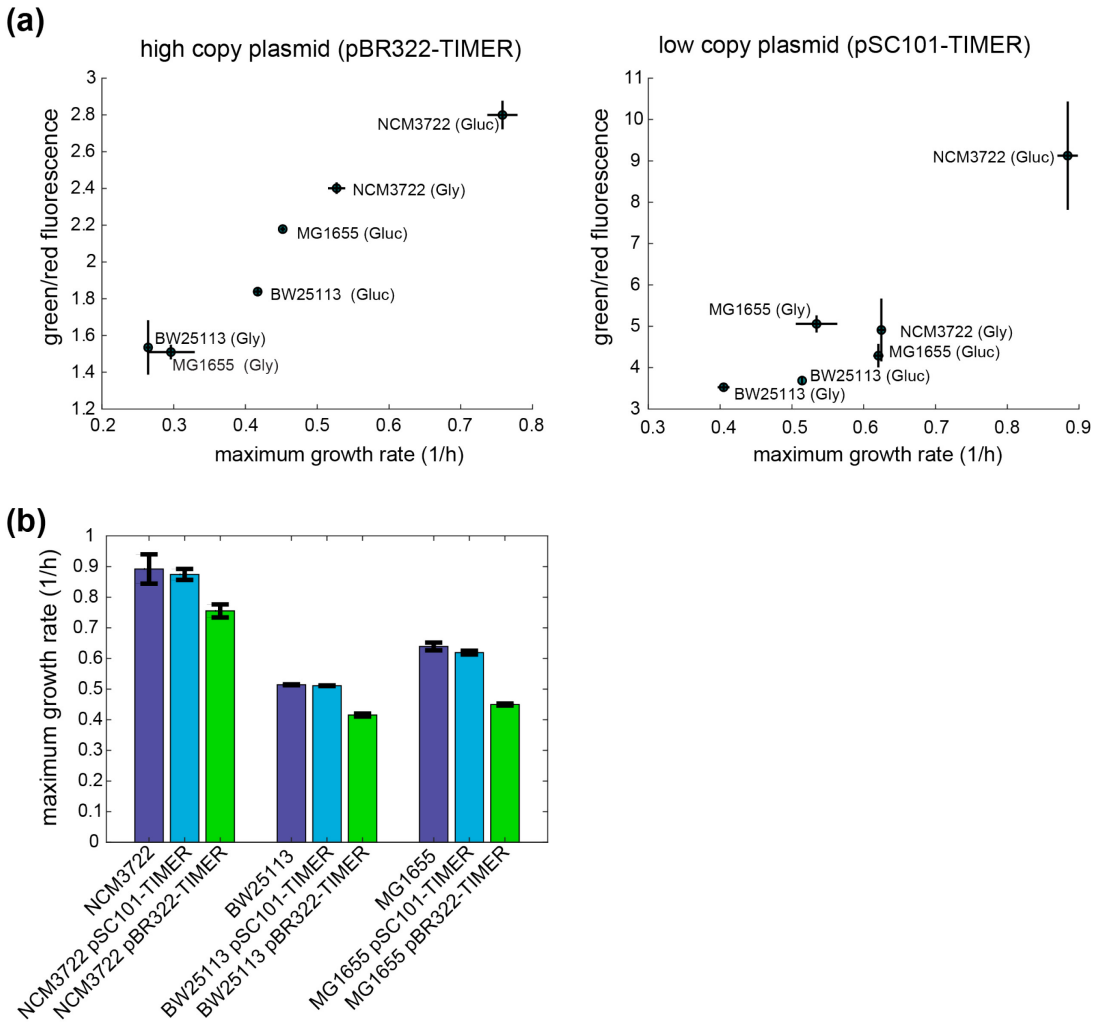

Figure 34: TIMER expressed from high and low copy number plasmids.

(a) TIMER expressed from a high copy number plasmid (pBR322-TIMER, left plot) and low copy number plasmid (pSC101-TIMER, right plot) in 3 different *E. coli* strains (NCM3722, BW25113, MG1655) growing in M9 minimal medium with Glucose or Glycerol as carbon source. (b) Maximum growth rates of TIMER expressing strains in M9-Glucose minimal medium compared to the not TIMER expressing wildtype strains.

To test this hypothesis, we let *E. coli* NCM3722 expressing TIMER from either the high or the low copy number plasmid grow to mid-exponential phase and analyzed the TIMER ratios of 10,000 single cells each with flow cytometry (Figure 35). As expected, we detected generally lower fluorescences in the cells carrying the low copy number plasmid. When comparing the distribution of TIMER fluorescence ratios, we observed indeed a higher variance in the population of cells expressing TIMER from the low copy number plasmid (CV = 19.8) compared to cells expressing TIMER from the high copy number plasmid (CV = 15.9). This higher variance could have biological reasons, i.e. the low copy number plasmid causes a higher variation of

growth rates of individual cells. To test the hypothesis that the higher variance of TIMER signals actually reflects a higher variance in growth phenotypes, one possibility would be to test the relation between both on the single cell level using a mother machine as recently presented [131].

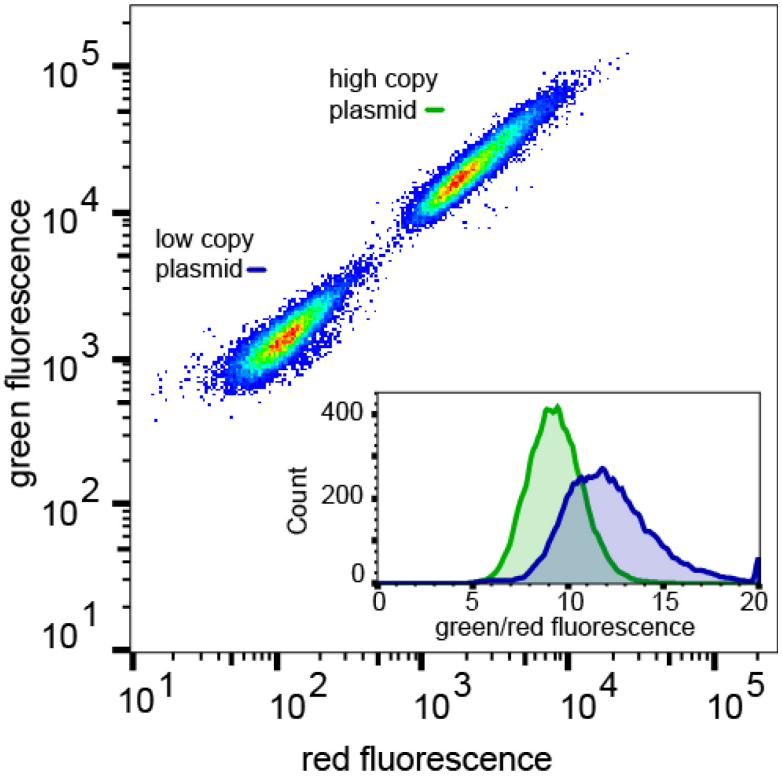

Figure 35: TIMER appearance of single cells expressing TIMER from a high and low copy number plasmid.

TIMER expression from high and low copy plasmid compared on single cell level by Flow Cytometry. Embedded into the plot is the distribution of cells with different ratios of green/red fluorescent TIMER expressed from the high copy number plasmid (green) and high copy number plasmid (low).

However, we think that the results presented here are sufficient to conclude that the display of the growth rate by TIMER is more accurate when expressed from the high copy number plasmid. We therefore decided that the high copy number plasmid is more eligible for library screens and sortings, despite the growth burden caused by the expression of TIMER from the high copy number plasmid.

6.1.4 Robustness of TIMER against genetic perturbations

Next, we were wondering how robust TIMER expression and its ability to display the growth rate is when the cells are perturbed genetically. The intention behind this question is that we assume that the growth conditions are similar for all cells within a pooled library and external factors, such as the influence of oxygen concentrations on the maturation rate of TMER, can therefore

be neglected when evaluating the ability of TIMER to display the growth rate reliably. However, cells within the pooled libraries vary genomically. The phenotypes caused by these variations can vary dramatically, from silent mutations that have no impact on the fitness, and variations that only cause a fitness burden under certain conditions to severely growth impairing or even lethal mutations. Though, an important prerequisite for the usage of TIMER to screen pooled libraries of genetic variant strains is the independence of TIMER to display the growth rate from such genetic perturbations.

To test this, we transformed the TIMER expression plasmid into 11 strains of the KEIO collection [132] with knockouts of global transcription factors. These transcription factors are responsible for the control of expression of 31 to 513 genes. A knockout of these transcription factors will therefore presumably lead to a misregulation of many genes and by that impair the physiological state of the cells. Hence, assuming that the TIMER fluorescence ratio is not independent of genetic perturbations, it would be very likely to observe such a deficiency of TIMER to display the growth rate in genetically perturbed strains with the chosen global transcription factor knockout strains.

With these strains we did another growth experiment, measuring optical density and fluorescences in microplate readers to determine maximum growth rates and TIMER fluorescence ratio at mid-exponential growth (Figure 36).

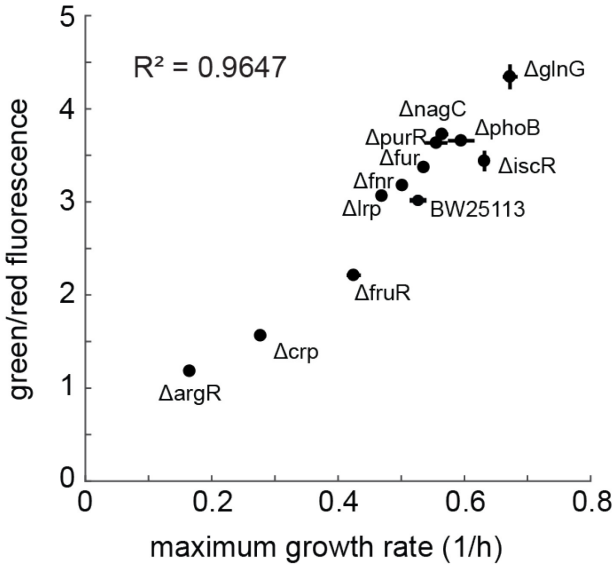

Figure 36: TIMER expressed in transcription factor knockout strains.

TIMER appearance (green/red fluorescence ratio) in relation to the growth rates of 11 *E. coli* strains of the KEIO collection [132] in which global transcription factors have been knocked out.

The growth rates in the given growth conditions were reduced to different extents and varied between 0.17 and 0.68 h^{-1}. Interestingly, the TIMER expressing *E. coli* wildtype strain BW25113

– the KEIO collection background strain – grew slower than expected and slower than the majority of transcription factor knockout strains. We explain this with a random mutation within the genome of the host strain that leads to an unusual growth reduction. TIMER, however, was able to display this reduced growth rate despite the random mutation. Overall, the previously observed clear linear correlation between growth and TIMER behavior could also be observed with this set of strains with a calculated coefficient of determination of 0.97. Only the *ΔiscR* strain showed a green/red fluorescence ratio that was slightly too low for the measured growth rate. We can only speculate why this is the case. A direct link between TIMER expression and the biological function of IscR which controls the expression of genes needed for the expression of iron-sulfur clusters is not apparent. Moreover, the deviation from the expected TIMER ratio is not very large and dubiously significant.

We therefore conclude that large genetic perturbations caused by deletions of global transcriptions factors may have a reducing effect on the growth rate but do not impair the display of growth rates by TIMER.

6.1.5 A simple mathematical model can explain the TIMER behavior

Based on the previous results, we propose a simple model that is sufficient to explain the linear correlation of TIMER and the growth rate. The model includes the concentration of green and red fluorescent TIMER, the expression rate *b*, growth rate *μ* and maturation rate *k* (Figure 37). The concentration of green fluorescent protein is influenced positively by the expression rate and negatively by the maturation constant (maturation from green to red fluorescent TIMER) and dilution by growth. The concentration of the red fluorescent protein fraction itself is positively influenced by the maturation constant and like the green fluorescent fraction negatively influenced by dilution due to growth.

Combining these two equations, we found that the TIMER fluorescence ratio ([G]/[R]) is only dependent on two factors, the growth rate *μ* and the maturation constant *k*. The dependency on the growth rate is apparent as this has been described before [102] and is also the reason why we chose this fluorescence reporter in the first place. Previous studies found a dependency of TIMER maturation from pH, temperature [100] and oxygen [102] despite unchanged growth rates.

The fluorescence ratio is on the other hand not dependent on the expression rate *b*. This explains the similar correlation of growth rate and TIMER expressed from the high and low copy plasmids.

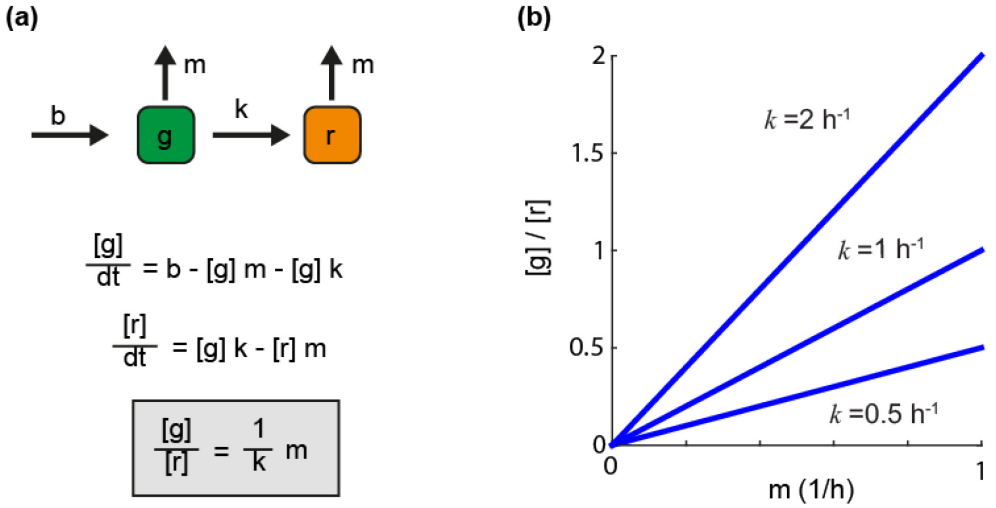

Figure 37: Simulation of dependency of TIMER appearance from growth and maturation constant.

(a) TIMER appearance (ratio of green [g] and red [r] fluorescent TIMER) can be explained by only a small number of parameters: b is the expression rate of TIMER, k is the maturation constant, m is the dilution rate which is defined by the growth rate of the culture. Solving the steady state equation shows that the TIMER appearance is proportional to the growth rate m and inversely related to the maturation constant k. (b) TIMER ([g]/[r]) is linearly dependent on the growth rate m. Changing the maturation constant k (e.g. by growth in presence of lower oxygen concentrations) does not impair the linear dependency.

6.1.6 Conclusions

Based on our verification experiments we think that TIMER could theoretically be used as a tool to screen pooled libraries for slow growing cells. An important prerequisite for its utilization is that the TIMER fluorescence ratio is not influenced by genetic perturbations. Since the TIMER fluorescence ratio was not even impaired by knockouts of global transcription factors, we think that the TIMER fluorescence ratio is independent on the genetic background. We could also show that the best correlation between growth rates and TIMER fluorescence ratio can be observed when TIMER is expressed from a high copy plasmid and in mid-exponential growth phase.

Our next goal was to use TIMER to specifically enrich slow growing cells out of a comprehensive library of genetic variants using the beforehand determined settings.

6.2 Selective Enrichment of slow growing cells

6.2.1 Creation of a metabolism-wide CRISPRi library

We sought to verify if TIMER can be used to enrich fractions of slow growing strains from pooled large genetic variant strain libraries. A method to create such a large library is CRISPR interference. As shown in Chapter 3.2, the construction of sgRNAs that target genes of interest is relatively simple with only a handful of design principles to follow [133] and multiplexing of DNA synthesis enables the construction of large and comprehensive CRISPRi libraries that target large sets of genes. Dependent on the essentiality of the targeted gene and the strength of the introduced metabolic bottleneck, it can be expected that some strains within such a CRISPRi strain library will have reduced growth rates under certain environmental conditions – similar to strain libraries expressing enzymes with potentially implemented allosteric regulation. We therefore regarded CRISPRi as an interesting method to create a complex strain library that can be used to evaluate TIMER as a tool for the enrichment of conditionally growth impaired strains and sought to generate such a comprehensive CRISPRi library for *E. coli.*

We chose to target every metabolic gene of *E. coli*. According to the last metabolic network reconstruction, *i*ML1515, *E. coli* possesses 1515 genes involved in metabolism [1]. For each of these genes we designed 4 to 5 sgRNAs, resulting in a library of 7184 sgRNAs. The target sequences have been identified and selected using a Matlab script. The criteria for target finding were that the binding sites should be evenly distributed over the whole gene sequence to allow different gene expression repression strengths, and that the sgRNA should bind to the non-template strand as binding to this strand proved to be more efficient than binding to the template strand [48]. The full list of designed sgRNAs can be found in our publication.

As the plasmids used in the original CRISPRi system are not compatible with the TIMER plasmid, we had to use another CRISPRi expression system for which we chose to work with pNUT1527 (Figure 38). This plasmid has the advantage to carry both, the genes for dCas9 under the control of a strong IPTG inducible promoter (P_{tac}) and the sgRNA gene under the control of a constitutive promoter (P_{J23119}).

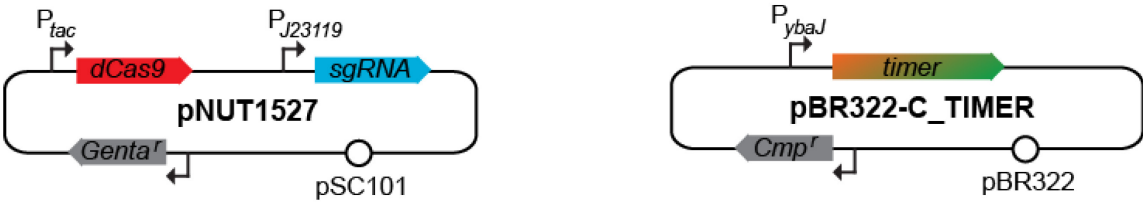

Figure 38: Plasmids used for the screening of a metabolism-wide CRISPRi library and the enrichment of slow growing cells.

For cloning of the library, the sgRNAs were ordered and received as pooled oligonucleotides from Agilent Technologies and then amplified in a low cycle PCR with high amounts of template RNA to obtain double-stranded DNA without changing the composition of the library. Next the plasmid backbone was amplified and after purification, we used Gibson Assembly to fuse both and to create the CRISPRi library plasmids (Figure 39).

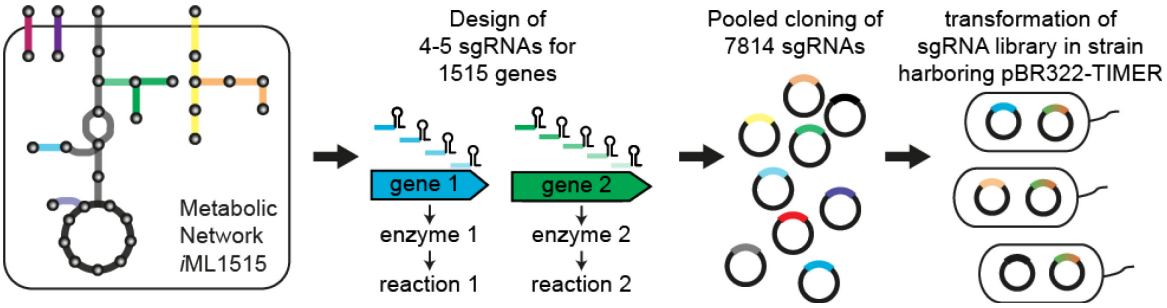

Figure 39: Design and Construction of the metabolism-wide CRISPRi library.

The plasmids were then directly transformed into electrocompetent NCM3722 cells harboring the TIMER plasmid, the resulting transformants plated on rich medium without inducers to minimize the impact of CRISPRi on the growth rates. The transformation resulted in approximately 10^7 cfus, which is above the 1000x library coverage threshold. All these colonies have been flushed and pooled resulting in the desired comprehensive strain library.

We next examined the composition of the strain library with Next Generation Sequencing. For that, we amplified 300 bp fragments consisting of the sgRNA gene and flanking regions by PCRs with high amounts of template and low cycle numbers. The amplified DNA was then cleaned up and amplified a second time to add indexes to the fragments which allow binding of the fragments to the NGS chip. We performed Next Generation Sequencing and analyzed the obtained data using Matlab. After removal of all sequences that did not map to the sgRNA sequences in the reference library, we obtained 430.000 analyzable reads. In these reads we found that 7094 out of the 7184 sgRNAs (98.7%) were present in our strain library (Figure 40). It must be noted that several sgRNAs had only very low read counts, a few only with a single read. We therefore conclude that although an analyzable read count of 430.000 is usually regarded as very high, it is most likely not sufficient to allow sequencing of all low abundant sgRNAs and that actually more or less all sgRNAs are present within the strain library. In fact, 20 of the 90 in the original library undetected sgRNAs could be detected in other rounds of sequencing. In the next step, we sought to enrich a slow growing fraction out of this library using the TIMER protein.

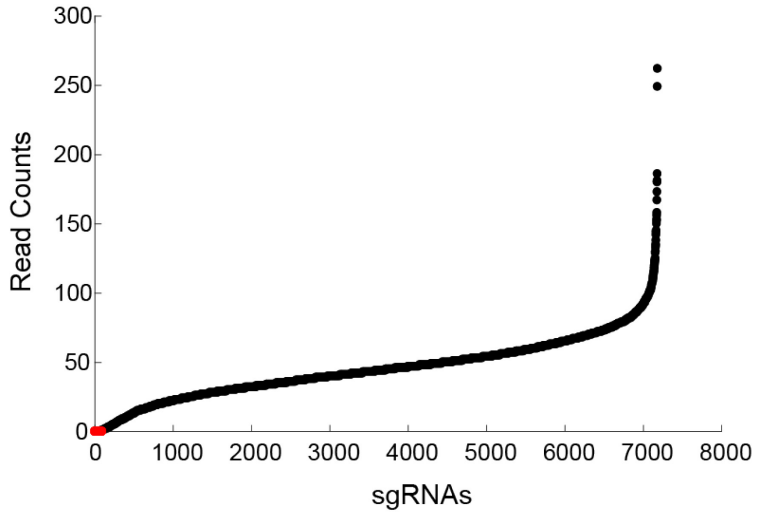

Figure 40: sgRNA abundances in the cloned strain library.

Read counts of sgRNAs after assembly of the CRISPRi library. 7094 of 7184 sgRNAs (98.7%) were present in the library (black dots), 90 sgRNAs (1.3%, red dots) were not detected.

It has been recently reported that an overexpression of dCas9 can be toxic to the cells [95] which might result in abnormal cell morphology [134]. We therefore examined the cell morphologies of library cells expressing TIMER, dCas9 and gRNAs using microscopy (Figure 41). We could not observe any of the reported morphological changes, showing that impaired growth would not be a result of morphological changes caused by dCas9 overexpression but in fact caused by a combination of both, TIMER overexpression and gene repression by CRISPRi.

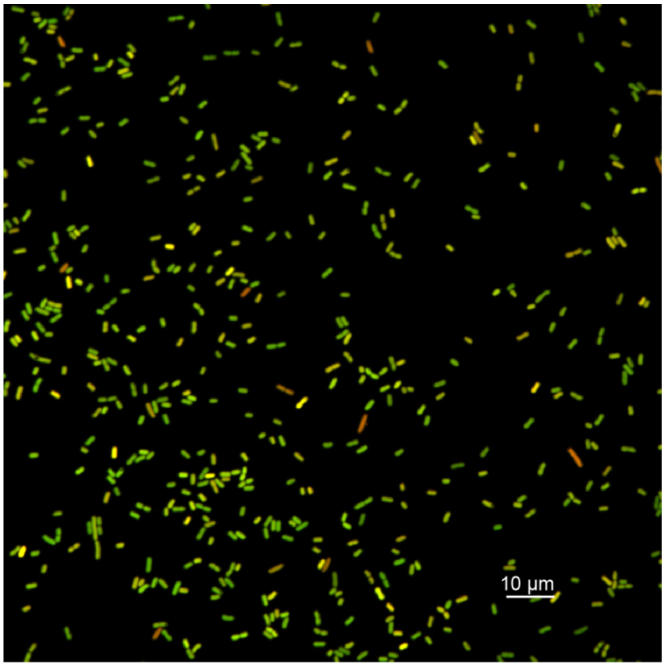

Figure 41: Overlay green and red fluorescence in *E. coli* NCM3722 pBR322-C_TIMER pNUT1527-sgRNA:none.

E. coli cells expressing TIMER, dCas9 and a non-targeting guide RNA. The expression of all three does not result in any cell length phenotype.

6.2.2 Enrichment of cells with a metabolic bottleneck in amino acid metabolism

Out of the created CRISPRi strain library, our goal was then to isolate a fraction of strains which are growth impaired under certain conditions using the TIMER protein, in this case amino acid depletion.

For that, we separated the growth period into two phases (Figure 42): First, we started the first culture by inoculating M9-Glucose minimal medium supplemented with 1 mM of each amino acid with a defined volume of the library glycerol stock. This culture was incubated for 6 hours to mid-exponential growth phase. Then, the cells were vacuum-filtered, washed and resuspended in fresh M9 minimal medium, only this time without supplemented amino acids. Again the cells were incubated for 6 hours to again reach mid-exponential phase as in this phase TIMER correlates best with the growth rate. This culture was then used for FACS-based analysis and sorting to enrich slow growing cells.

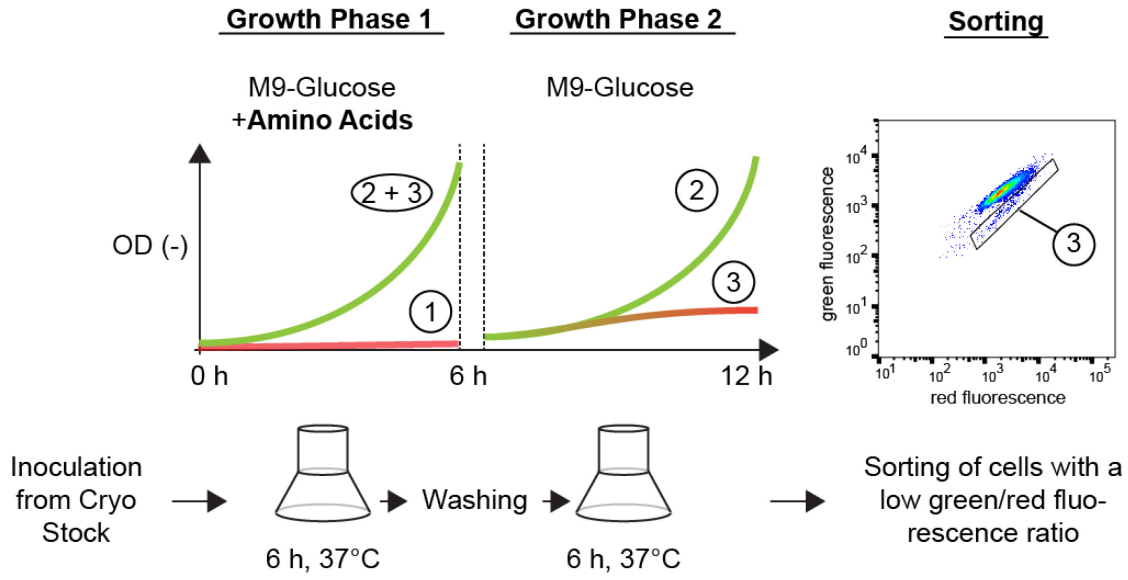

Figure 42: Workflow for TIMER-based enrichment of strains with bottlenecks in amino acid biosynthesis.

The enrichment is divided into two growth phases and a subsequent enrichment of cells with a low green/red fluorescence ratio using Flow Cytometry: In the first growth phase, a culture is started by inoculation of M9 minimal medium with glucose and 1mM of each amino acid with a defined volume of the library cryo stock. The culture is incubated for 6 hours. We expect the occurrence of a non-growing fraction (1) that will be outcompeted by the non-affected, normally growing strains (2+3). Then the cells are washed and reinoculated in fresh medium for the second growth phase. In this phase, cells grow in absence of amino acids. Whereas most cells will grow unaffectedly (2), cells that are auxotroph for amino acids as a consequence of the introduced metabolic bottleneck will be growth impaired and appear red (3). This fraction can then be isolated by using flow cytometry.

With this setup, we expected the occurrence of three subpopulations with different growth characteristics as depicted in Figure 42: A major fraction of the population is neither growth impaired in the first culture with amino acids, nor in the second culture without amino acids (fraction 2). In many cases this is due to gRNAs targeting genes that are irrelevant for growth under the given growth conditions. In other cases, an insufficient silencing of transcription due to weak gRNAs and/or overabundance of targeted proteins could result in reduced transcription rates but not to an extent that would limit fluxes sufficiently to cause reduced growth rates.

A second fraction of strains is growth impaired in both media (fraction 1). This is the case when the targeted gene is essential in minimal medium independent of the addition of amino acids. We expected that these strains will be lost already in the first culture by the other strains overgrowing these growth impaired strains. Indeed, we found that 80 sgRNAs that could be detected in the original library were not present in both experiments, neither before washing, nor after sorting. Targeted genes are involved in different metabolic pathways, e.g. pyruvate decarboxylation, ATP synthesis, folate biosynthesis, lipopolysaccharide transport, pyrimidine and purine biosynthesis.

The third fraction consists of conditionally growth affected strains (fraction 3): In presence of amino acids these strains are not growth impaired as the bottleneck caused by CRISPRi can be compensated by uptake of amino acids. However, after depletion of the amino acids the genetic bottleneck introduced by CRISPRi results in a bottleneck in the metabolic network and subsequently in reduced growth. We assumed that most strains showing this behavior presumably have metabolic bottlenecks in pathways required for the biosynthesis of amino acids. It is theoretically possible that strains were enriched that are unable to grow sufficiently on glucose alone but a combination of both, glucose and amino acids restores that growth. For instance, incomplete bottlenecks in glycolysis could result in this phenotype.

We sought to isolate and further analyze this third fraction using TIMER-based enrichment of slow growing cells. For that, the cells were first analyzed by flow cytometry. Similar to the appearance of cells expressing TIMER from high and low copy number plasmids (Chapter 6.1.3), when comparing red and green fluorescences of individual cells, we observed that most cells of the whole population were within the characteristic cloud (Figure 43). However, in addition to this population a small subpopulation with a lower green/red fluorescence ratio – indicating slow growth – which had a share of about 2% of the whole population was observable. We isolated 100,000 cells of this fraction using a cell sorter and plated these cells on plates with rich medium. As a reference for strains without growth impairment, we also isolated and plated cells with a high green/red fluorescence ratio. These two fractions were then analyzed further, in a growth experiment and by NGS. To compare the reproducibility of the enrichment, the sorting has been performed in two independent experiments.

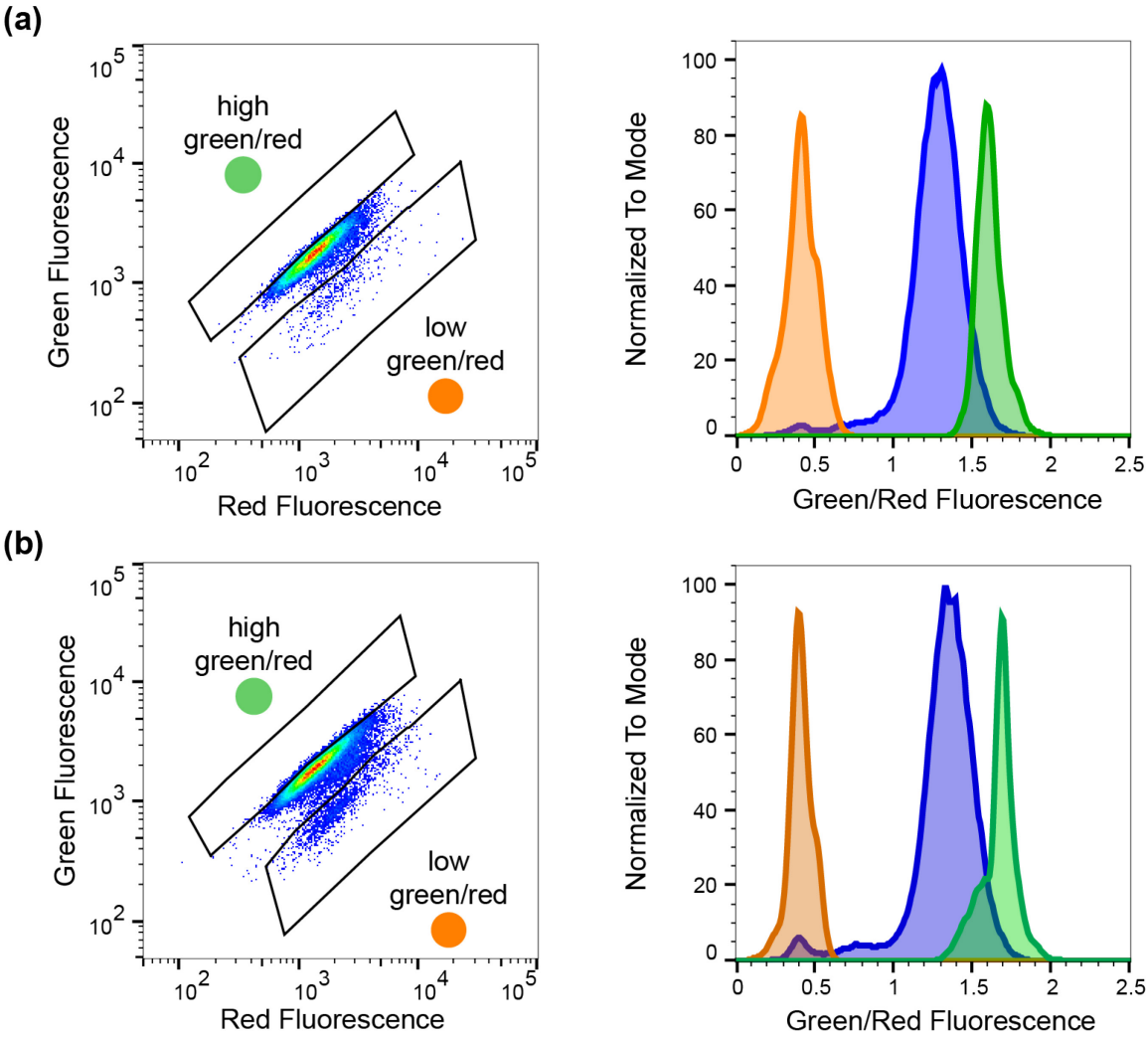

Figure 43: Sorted fraction of slow and fast growing cells.

Left: Green and red fluorescences of *E. coli* NCM3722 pBR322-C_TIMER, pNUT1527-ML cells 6 hours after removal of amino acids from the medium. The two fractions with a high and low green/red fluorescence ratio have been marked. Right: Distribution of TIMER fluorescences of sorted fractions (orange and green) and the whole population (blue). (a) and (b) indicate both independent cultivations and sortings.

6.2.2.1 Analysis of sorted and isolated strains

Of both fractions, we picked 48 (low green/red) and 45 colonies (high green/red), respectively. The growth rates of these isolates were compared in a growth assay experiment in a multititer plate reader. In addition, three cultures of a strain with a non-targeting gRNA were used as a control strain. Growth of the control strain is not significantly affected by the expression of both CRISPRi components as shown in Figure 44. All strains from the fast growing fraction grew similarly and comparable to the control strain, with growth rates above 0.52 h^{-1} and lag phases shorter than 8.6 hours (Figure 45). It has been shown recently, that a CRISPRi can have a strong effect on the lag phases. Strains with such a delayed growth presumably were growth

inhibited in the first growth phase until the bottleneck can be overcome by escape mutations or compensatory responses [93].

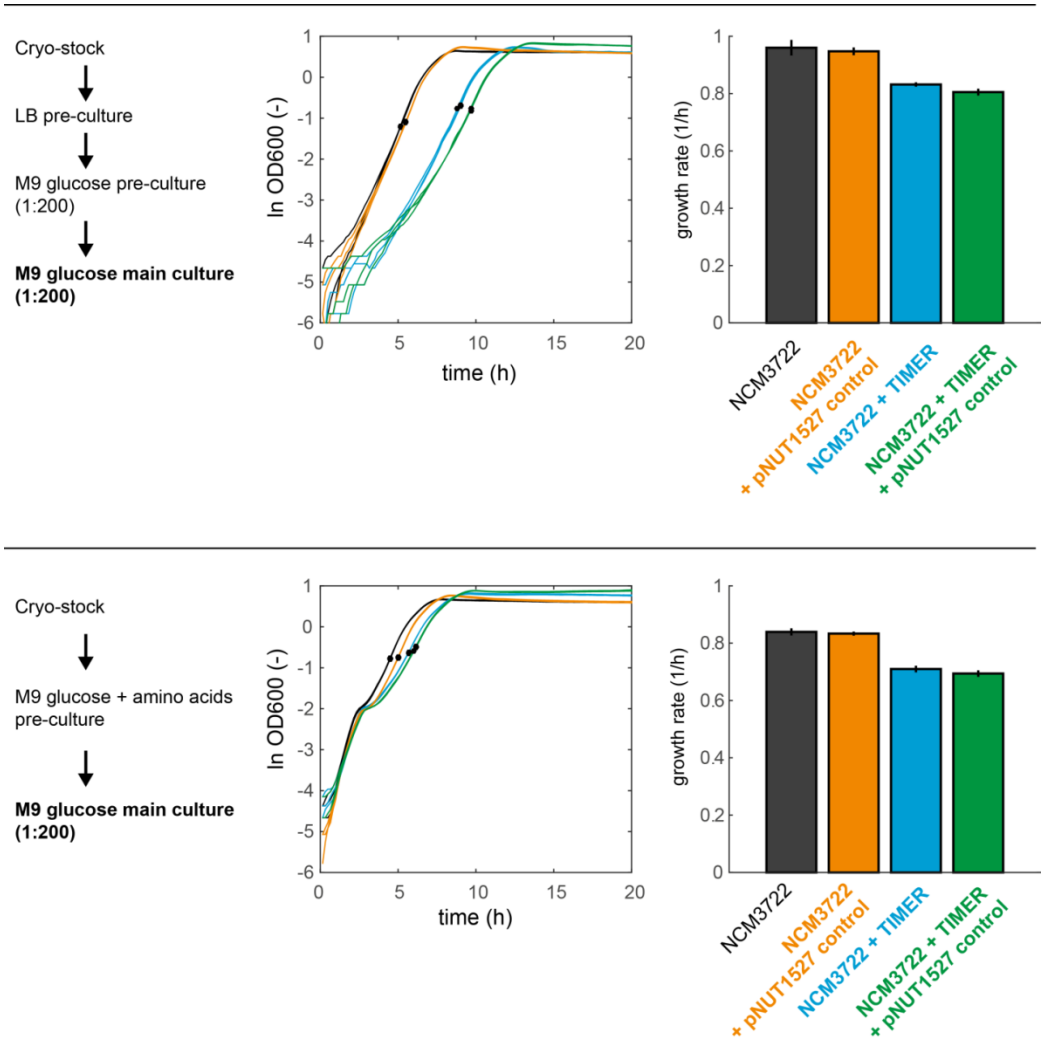

Figure 44: Growth characteristics of the CRISPRi control strain.

Growth of the CRISPRi control strain (E. coli NCM3722 pBR322-C_TIMER pNUT1527-non targeting control; green) in comparison to the wildtype strain and the strains harbouring only one of the plasmids. Shown are the results of three independent cultures each in 96 well microtiter plates. Black dots in the growth curves indicate the time when the cultures reached maximum growth. Top: Growth conditions as used in the growth experiments to characterize TIMER (Figure 32). Bottom: Growth conditions as used for the enrichment of strains with a metabolic bottleneck upon removal of amino acids.

As the strains isolated with a high green/red ratio are expected not to be growth impaired, we defined all strains with a growth rate below 0.52 h^{-1} and/or with a lag phase longer than 8.6 h as poorly growing. Of the fraction isolated with a low green/red ratio, 38 of the 48 isolates showed either reduced growth prolonged lag phases or both, showing that we indeed specifically enriched slow growing cells with TIMER.

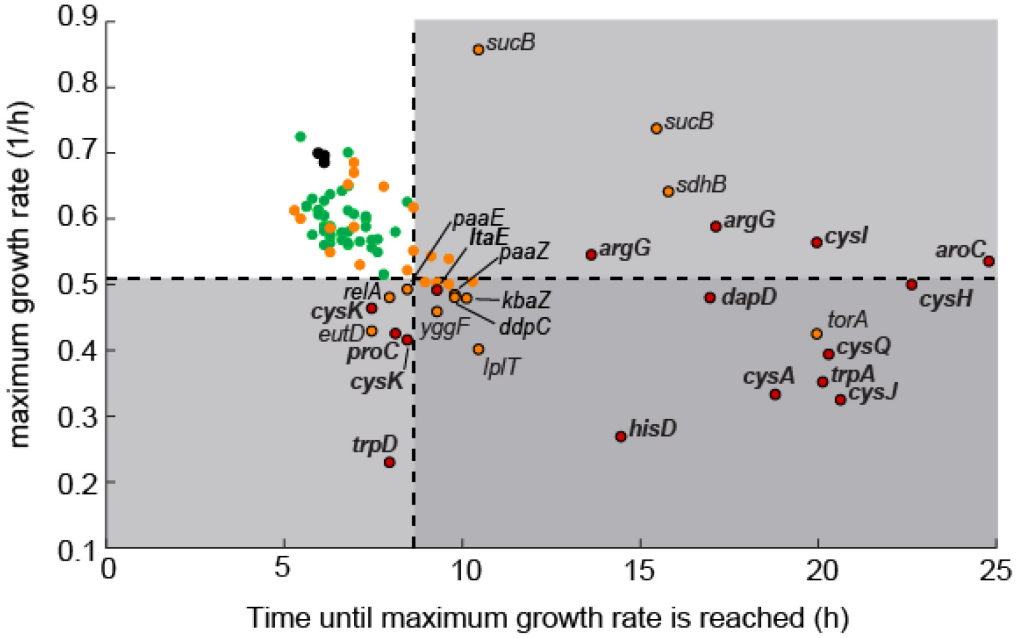

Figure 45: Growth of isolated strains.

Shown are the maximum growth rates and lag phases (time until the maximum growth rate is reached) of 45 strains isolated with a high green/red fluorescence ratio (reference strains; green dots), three independent cultures of the control strain (black dots) and of 47 strains isolated from the fraction with the lowest green/red fluorescence ratio (orange and red dots). Dashed lines indicate the longest lag phase and lowest growth rate observed for the strains in the reference strain panel. 28 of the strains with most impaired growth (lowered growth rate and/or longer lag phase) compared to the reference panel have been sequenced. Of those, 16 strains possessed gRNAs targeting genes directly involved in amino acid biosynthesis (red dots).

Next, we were interested which metabolic genes have been targeted in this group of 38 growth impaired strains, assuming that a large fraction of those target genes are related to amino acid metabolism. Indeed, Sanger sequencing of the 28 most growth impaired strains showed that 11 had gRNAs targeting genes encoding enzymes in the de novo biosynthesis of amino acids. Moreover, in 5 additional strains genes involved in assimilatory sulfate reduction was targeted. This pathway is needed for the de novo biosynthesis of cysteine. The remaining 12 strains were not directly involved in amino acid metabolism. However, targeted genes of this set of strains include genes like *sdh* of the TCA cycle which is producing precursors of amino acid biosynthesis. Perturbing the expression of a gene like *sdh* might therefore indirectly disturb the biosynthesis of amino acids.

In addition to the growth experiment, we wanted to verify if the reduction of growth rates and elongation of lag phases is actually caused by reduced gene expression. We were able to show this indirectly by measuring the intracellular concentration of precursors of the targeted reactions (Figure 46). We chose three strains with gRNAs binding to genes of the amino acid biosynthesis pathway, *hisD*, *aroC* and *argG* and measured the concentration of the precursors of these reactions. For AroC, the direct precursor is not measurable, instead we measured the

concentration of the previous metabolite in the pathway, shikimate-3-P. As a control strain, we also measured the concentrations of these metabolites in the negative control strain. For each of the measured metabolites, we saw a specific accumulation in the corresponding strain with the metabolic bottleneck. Although this does not exclude the possibility that the growth impairment is also partially caused by off-target effects, i.e. the binding of the gRNAs to unintended targets, the accumulation of precursors proofs the successful implementation of metabolic bottlenecks that presumably at least contribute to the growth phenotypes.

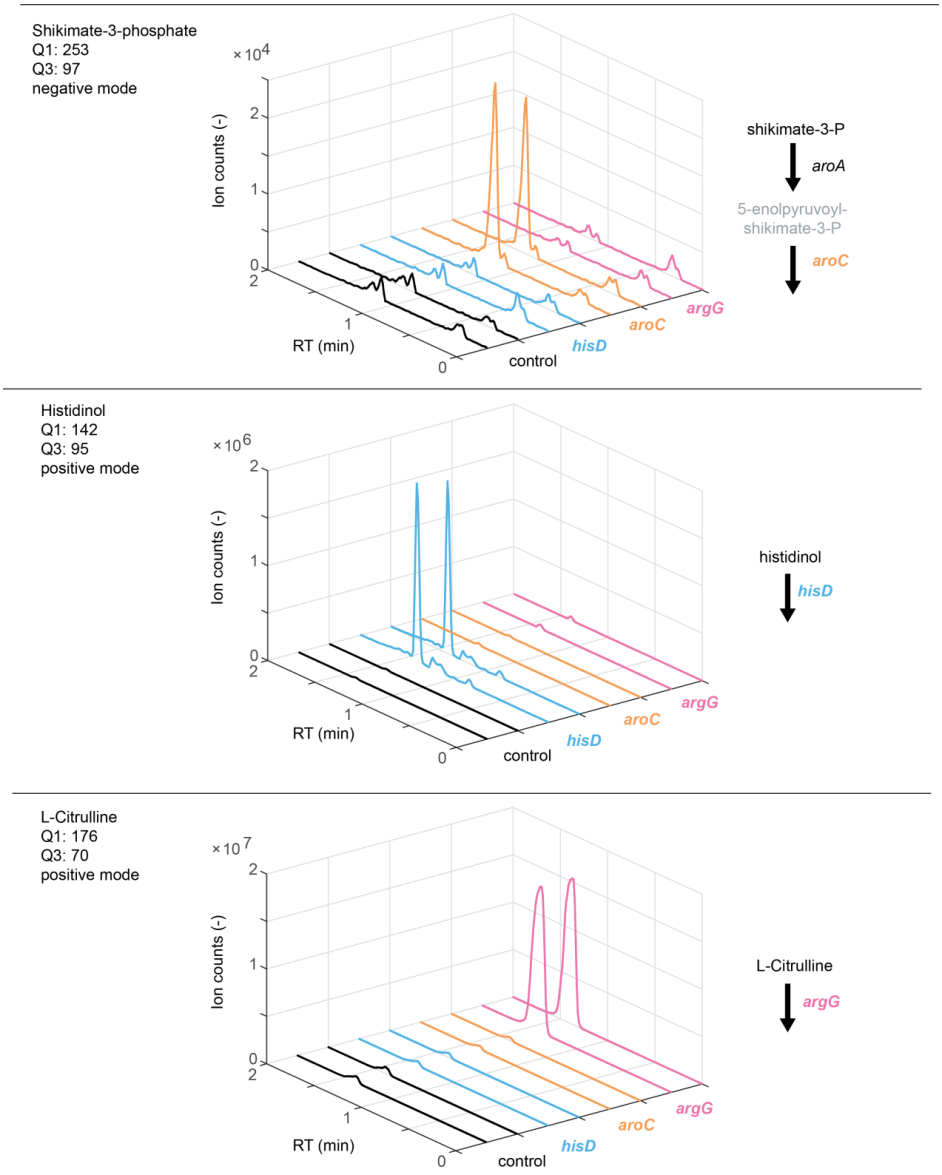

Figure 46: Intracellular concentrations of precursor metabolites.

Depicted are the intracellular concentrations of 3 metabolites, Shikimate-5-phosphate, histidinol and L-citrulline, that were measured with LC-MS/MS and which are precursors in reactions in which metabolic bottlenecks were introduced with CRISPRi. Metabolite concentrations have been measured in the negative control strain as well as in the CRISPRi strains targeting the respective genes. LC-MS/MS parameters are shown on the left, the targeted reaction on the right.

We conclude from these results that our library consist of strains with different growth phenotypes under different growth conditions and that the isolation of growth impaired strains out of this library is possible with TIMER. Furthermore, Sanger sequencing of a small number of growth impaired strains already suggests that a large fraction of the isolated strains have bottlenecks in pathways directly or indirectly involved in the biosynthesis of amino acids. To quantify their appearances before and after enrichment more accurately, we next sought to analyze the library compositions with NGS.

For that, the cells before washing and the isolated cells with a low green/red fluorescence ratio have been plated on rich media plates and incubated overnight. The resulting colonies have been flushed from the plate and minipreps have been performed on the pooled cell suspension to isolate the plasmids. The plasmids then have been used as the template to amplify the abovementioned 300 bp fragment including the guide RNA gene sequence and flanking regions. The fragments were then further processed as described in Materials and Methods and subsequently sequenced and analyzed.

As the enrichment has been performed in two independent experiments, we first sought to analyze how much the enrichments differ between both experiments. We therefore compared the fold changes of abundances of each strain before removal of amino acids and after enrichment of slow growing strains (Figure 47). We found that in both experiments the majority of strains were comparably enriched or depleted. Particularly strains that showed a high level of enrichment after isolation of slow growing cells were enriched comparably in both experiments. Differences in the fold changes were especially observed for strains that were highly depleted in one of the experiments. These variations between the two experiments could be caused by very low numbers of sorted cells per strain, resulting in statistical errors. Another possible reason are differences in the gates that were set to isolate the fraction of cells with a low green/red fluorescence ratio so that in one of the experiments, cells with a certain, relatively low green/red ratio have been isolated, whereas in the other experiment, these cells were not within the gate and therefore have not been isolated. Nonetheless, as we are primarily interested in the strains with the biggest changes in abundance (i.e. cells with the highest enrichment upon isolation of presumably slow growing cells) and since their changes of abundances seem to be sufficiently robust, the observed differences of fold changes of abundances of other strains can be ignored.

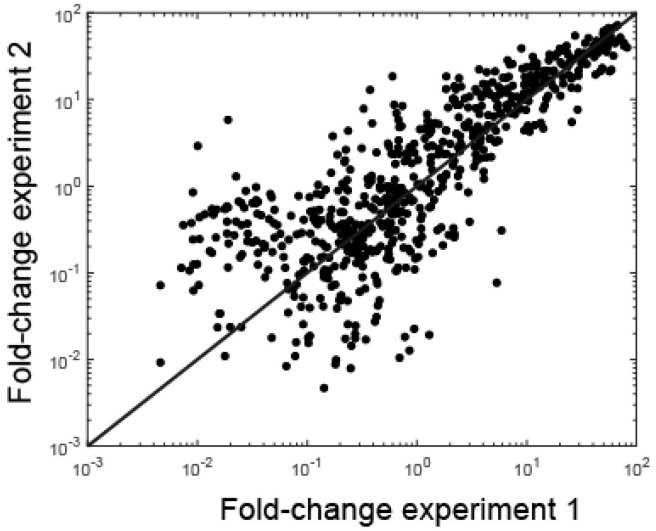

Figure 47: Fold change of individual gRNAs in both experiments.

Depicted are the fold changes of abundances of gRNAs after enrichment of the slow growing fraction compared to their abundances before removal of amino acids from the medium in both experiments. Each dot represents an individual gRNA. gRNAs on the black line have the same fold change of abundance in both experiments.

Next, we were interested which ratio of the isolated strains with low green/red fluorescence ratio possessed targets in amino acid biosynthesis pathways. We found that in both experiments, before removal of amino acids, about 12 % of all strains within the cultures possessed gRNAs that are directly targeting amino acid biosynthesis pathways (Figure 48a). This ratio correlates very well with the theoretical ratio of amino acid biosynthesis targeting gRNAs in the library (12 %; 837 of 7184 gRNAs) and shows that this group of gRNAs was not already enriched prior to washing. After removal of amino acids, incubation for 6 hours and sorting of the presumably slow and non-growing cells with a low green/red fluorescence ratio, we found that the fraction of cells with amino acid biosynthesis gene targeting gRNAs increased significantly to roughly 50% of the whole population. This number is in line with the previous results from Sanger sequencing of growth impaired isolates where also roughly half of the isolated strains possessed such gRNAs and already shows that the enrichment of strains with bottlenecks in amino acid biosynthesis pathways was successful.

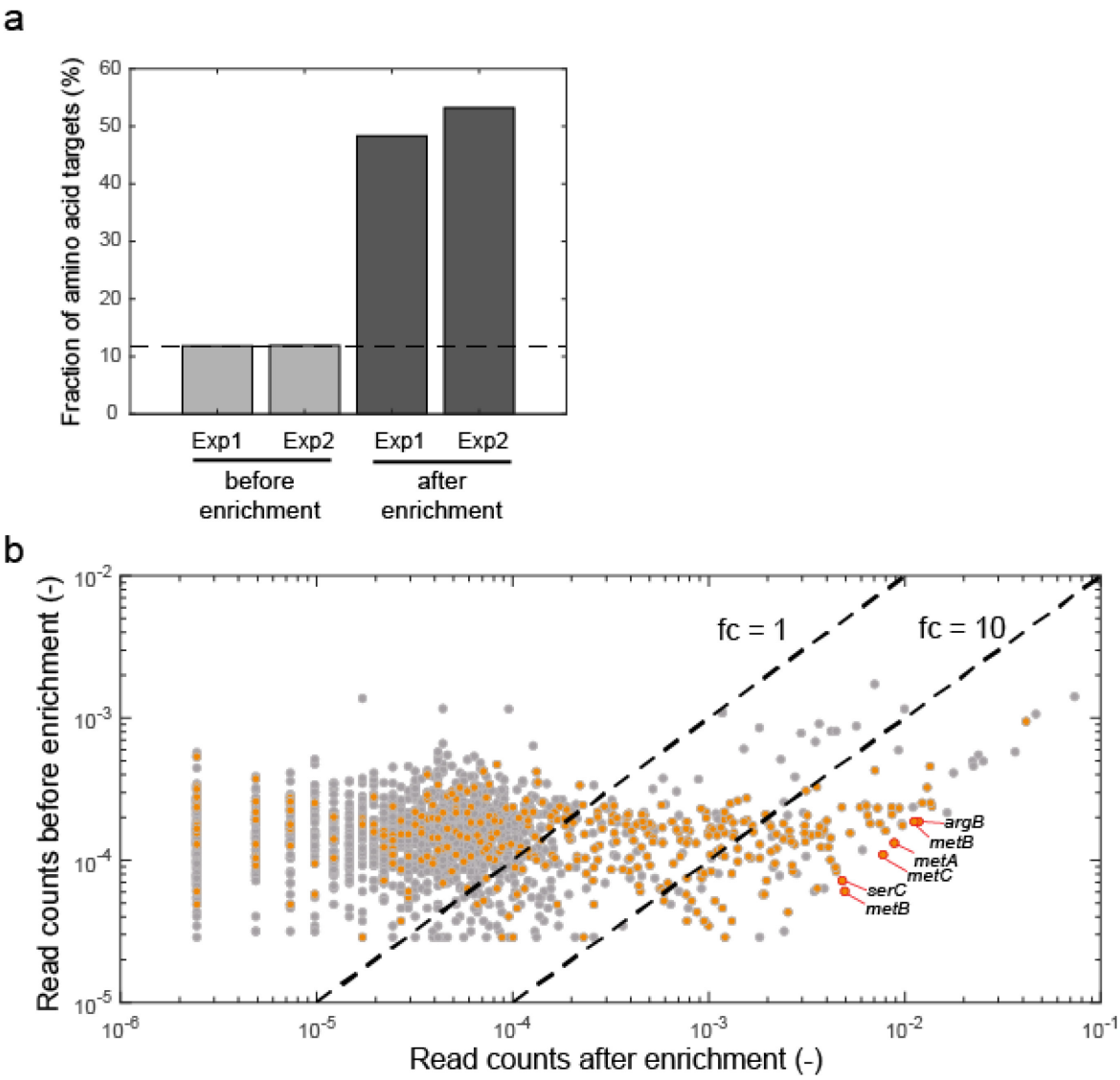

Figure 48: Abundance of gRNAs targeting genes involved in amino acid biosynthesis.

(a) Fraction of gRNAs within the library that target genes involved in amino acid biosynthesis. (b) Read counts of individual gRNAs within the library before removal of amino acids and after enrichment of slow growing cells. Orange dots represent gRNAs that target genes involved in amino acid biosynthesis. In red are the 6 gRNAs with the highest fold change marked that inhibit the expression of genes in amino acid biosynthesis. The black dashed line show a fold change of 1 and 10, respectively.

When regarding only these strains with gRNAs targeting amino acid biosynthesis genes in more detail, we found that the extent of their enrichments differed strongly (Figure 48b): Some strains were less abundant in the isolated fraction of cells with a low green/red ratio, indicating that these strains possess gRNAs that are either not sufficiently strong to silence the targeted gene or that the knock down of transcription can be compensated. On the other hand, many strains with these gRNAs were highly enriched, indicating a strong difference in growth phenotypes before washing (sufficient growth) and after (impaired growth). These strains are expected to possess strong gRNAs that are sufficient to knockdown a gene that is essential in absence of

supplemented amino acids. 141 strains of the whole library showed a high level (more than 10 fold) of enrichment after isolation. Of these highly enriched strains, 61% were directly linked to amino acid metabolism (Figure 49, blue dots). Interestingly, when comparing the different amino acid biosynthesis pathways we noted that of some pathways many strains were highly enriched, whereas in other pathways, we observed only for a small number of strains an increased level of abundance. For instance, of the methionine biosynthesis pathway, we noted a >10 fold higher abundance of 15 strains, with strains possessing gRNAs targeting metA, metB and metC partially being enriched >60 fold. The reason for this sensitivity of gene expression is that methionine biosynthesis usually operates at maximum capacity. Any disturbances of gene expression therefore reduce fluxes which apparently leads to lowered fitness, even when the gRNAs do not bind near transcription start site where repression is known to be more effective than downstream of the gene [48].

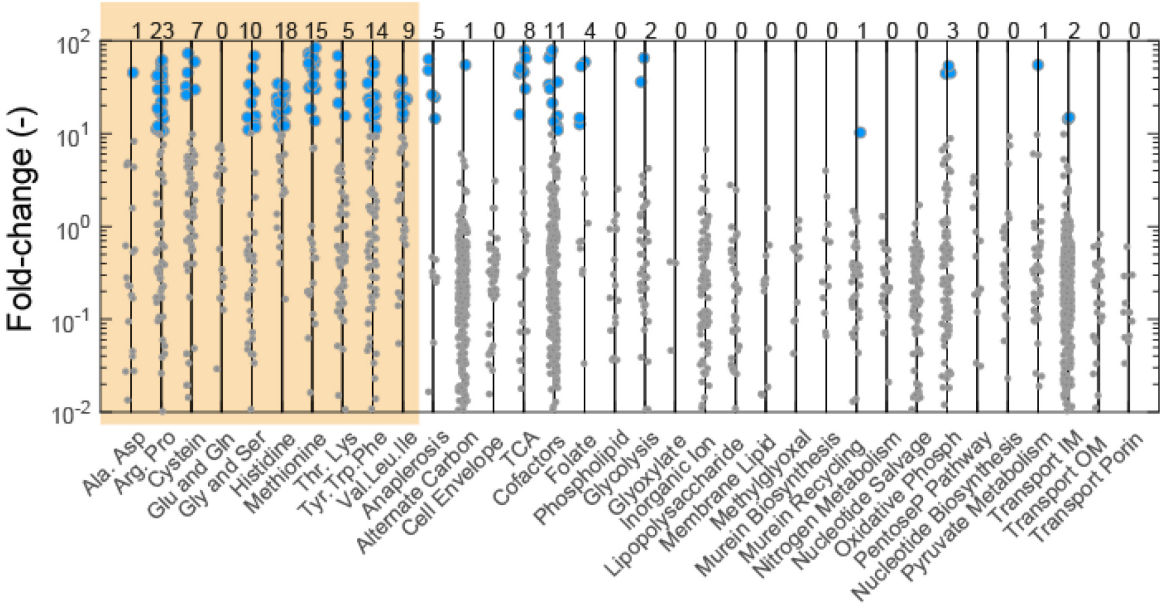

Figure 49: Fold change of gRNA abundance by classes of metabolic pathways.

Each dot represents the fold change of abundance of an individual gRNA after enrichment. Blue dots indicate gRNAs that showed a more than 10-fold higher abundance upon enrichment. Numbers above the lines show the amount of gRNAs with a more than 10-fold enrichment. The orange box marks all gRNAs that bind genes in amino acid biosynthesis pathways.

On the other hand, even the strongest gRNAs targeting genes in the glutamine/glutamate biosynthesis pathway did not result in growth impairment to an extent that these strains were significantly enriched. This robustness might be a result of the central role these pathways have in ammonium assimilation and transamination reactions but could be also a consequence of high amounts of stored glutamate within the cell [104]: After removal of amino acids, this pool might be large enough to maintain the metabolic network to a certain extent so that the metabolic bottleneck in glutamine and glutamate biosynthesis impairs growth only with a delay. When this delay is too long, the cells with such a bottleneck stop growing so late that TIMER

has not enough time to maturate accordingly and it will not reflect the growth rate, so that we will not be able to enrich these strains under the tested conditions. Instead, an increased incubation time after removal of amino acids would be necessary to isolate those.

39% of the highly enriched strains were not directly linked to amino acid biosynthesis. However, many of these strains possessed gRNAs that target genes that are indirectly linked to amino acid biosynthesis. Anaplerotic reactions are for instance crucial to refill pools of TCA cycle intermediates that are partially used as precursors for amino acid biosynthesis. Cells with metabolic bottlenecks in these reactions are able to grow in presence of glucose and amino acids as the needed amino acids can be taken up from the medium and anaplerotic reactions are not needed. In absence of amino acids however, TCA cycle intermediates are utilized to produce amino acids de novo that cannot be replenished, leading to reduced growth as a result of impaired TCA cycle and amino acid biosynthesis. Similarly, several strains with bottlenecks in the TCA cycle and glycolysis were highly enriched after removal of amino acids from the medium. In presence of amino acids, these strains were able to take up all needed amino acids, whereas after removal of amino acids, the cells were not able to synthesize the precursors for amino acid biosynthesis as a result of the introduced metabolic bottleneck.

At last, we checked if off-target effects and bad seeds might have had an impact on growth and subsequently the library composition after sorting.

gRNAs with off-targets have sequences with 9 nt or more identity to other target genes than the intended target gene [95]. We identified in our library 635 gRNAs that have off-targets in amino acid biosynthesis. Their appearances before and after enrichment have been analyzed for both experiments (Figure 50, top). In general, we found that the fraction of strains with off-targets in amino acid biosynthesis was not increased upon sorting of cells with low green/red fluorescence ratio, showing that a large majority of potential off-targeting gRNAs are not sufficiently strong to create a negative impact on the fitness of the cell. This observation can be confirmed when the appearances of strains with off-targets are compared before and after enrichment (Figure 50, bottom): Most strains showed a lower or similar abundance upon isolation of cells with low green/red fluorescence ratio, showing that for the vast majority of these strains the off-target effect is not strong enough to impair growth in absence of amino acids. Only three strains with potentially off-targeting gRNAs were enriched more than 10 fold. Their gRNAs have off-targets in *argF*, *argH* (arginine biosynthesis) and *ilvE* (isoleucine biosynthesis). However, the on-targets of all three gRNAs are genes involved in the biosynthesis of cofactors (*metF* in folate biosynthesis and *pdxH* of pyridoxal phosphate biosynthesis) that are also crucial for amino acid biosynthesis so that the enrichment might be also a directly caused by the on-target.

Bad seeds are gRNA sequences that show a so far unexplained sequence-specific toxicity which is assumingly caused by the 5 bases next to the PAM sequence [95]. To exclude that the bad seed-effect influenced the library composition before and after enrichment, we screened our library for the abundances of gRNAs with the 25 bad seed sequences that have to be

reported to have the most severe effect on the growth rate. We found that they were neither depleted before enrichment, nor were they enriched after isolation of the slow growing fraction of cells.

We therefore conclude that the impact of off-targets as well as of bad seeds seems neglectable in this enrichment and the highly enriched strains were isolated as a consequence of the direct on-targets of their gRNAs.

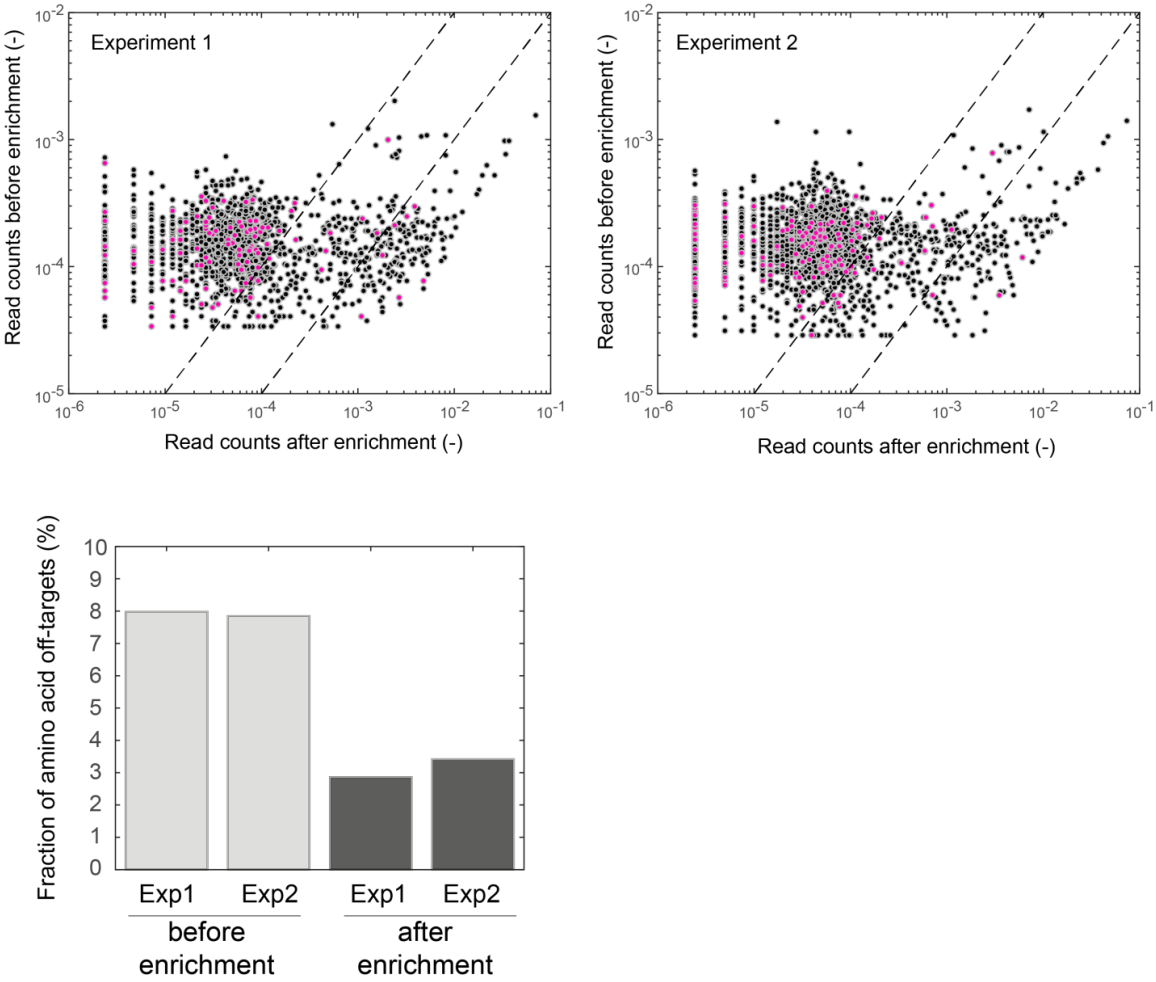

Figure 50: Off-targets in amino acid metabolism.

Top: Normalized sequencing read counts for each sgRNA before and after enrichment. sgRNAs with off-targets in amino acid metabolism, and no on-target in amino acid metabolism are shown as purple dots. Bottom: Fraction of sgRNAs in the library with off-targets in amino acid metabolism (and no on-target in amino acid metabolism), before and after sorting. Shown are results of two independent experiments (Exp1 and Exp2).

6.3 Discussion and Outlook

The dsRed variant TIMER is an interesting fluorescent protein that has the special feature of appearing green when freshly expressed but maturating to a red form. This characteristic enables its usage as a fluorescent growth rate reporter as shown for the visualization of growth rates of mouse tissues infecting *Salmonella* and of *E. coli* cells in biofilms [101]. An application for an enrichment of slow growing cells out of pooled strain libraries however had not been reported so far.

Our goal in this work was to evaluate the usage of TIMER for this application. For that, we verified that TIMER reflects the growth rate in *E. coli* batch cultures independent of used host strain and genetic backgrounds in a linear dependency – an important prerequisite for the usage of TIMER to enrich strains with a specific growth rate.

We then tested how efficient we can enrich slow growing cells out of a large strain library using TIMER as it is planned for the libraries of strains expressing potentially switching enzymes created by Domain Insertion. As a test library we produced a large pooled CRISPRi strain library for which we designed 7184 sgRNAs targeting all 1515 metabolic genes of *E. coli*. Such a strain library could be used in the future to systematically examine the impact of single metabolic bottlenecks on the phenotype under certain environmental conditions. So far, the strain library that was commonly used for such examinations is the KEIO collection [132]. However, our library has two features that make it superior to the KEIO collection: First, dependent on the sgRNA target site within the gene metabolic bottlenecks can also be introduced gradually whereas the complete gene knockout usually results in complete bottlenecks. Second, the possibility to induce dCas9 expression enables also the examination of bottlenecks caused by knockdowns of essential genes whereas with the KEIO collection a construction of essential gene knockout strains was not possible. A further examination of the CRISPRi library would be interesting but is not planned within this project.

We could show that an enrichment of fractions of growth impaired strain under specific environmental conditions – in this case amino acid depletion – is possible. For that, we used fluorescence activated sorting of slow growing cells according to the individual TIMER appearances and verified the efficiency of sorting by next-generation sequencing and detailed growth characterization of a smaller number of enriched cells.

The next step would be now to apply TIMER to screen for strains with switchable enzymes. However, TIMER-based enrichment of slow growing cells could be also interesting for various other fields of biology. For instance, the enrichment of slow growing strains can help examining persister cells with low growth rates and therefore low susceptibility to antibiotics [102]. Slow growing cells with a high metabolic activity might be also interesting in metabolic engineering as host strains for overproduction of chemicals in two-phase bioprocesses [135]. Moreover, an isolation of growth impaired strains could be helpful in synthetic biology to identify genetic parts

that create a growth burden to the cell [136]. We plan to use TIMER in other projects that are not directly linked to the creation of switching enzymes.

7 Conclusions and Outlook

Enzymes with synthetic allosteric regulation could be of great interest for metabolic engineering and synthetic biology although their construction remains challenging. Our main goal was to create switching enzymes for the immediate, autonomous and continuous control of metabolic pathways. For that, we chose creation techniques based on the concept of directed evolution. As already indicated in the introduction, the creation of switching enzymes through directed evolution involves several steps which can be illustrated with an engineering cycle (Figure 51), which we followed in this project.

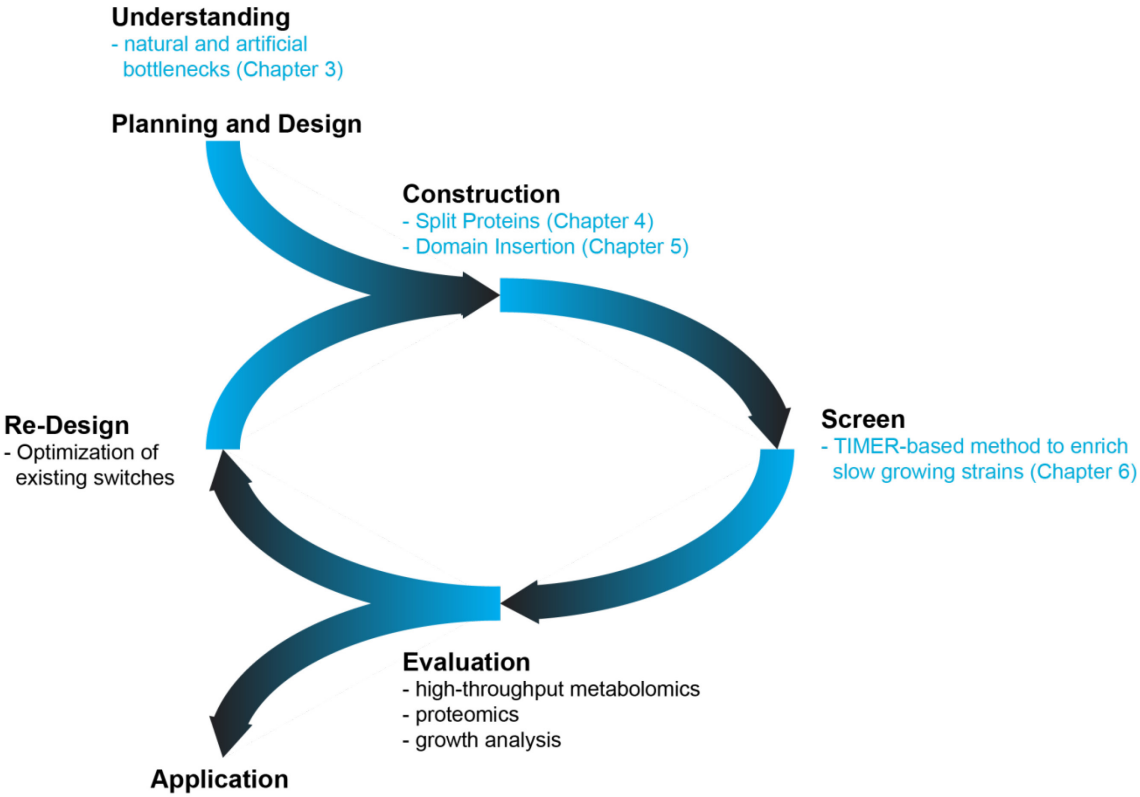

Figure 51: Creation of enzymes with synthetic allosteric regulation.

Metabolic enzymes with synthetic allosteric regulation introduce conditional metabolic bottlenecks in the reaction catalyzed by them, and effectively cause a pseudo-auxotrophy for downstream products (e.g. amino acid, folic acid). At first and prior to the creation of switching enzymes we therefore sought to examine metabolic bottlenecks and how they influence growth of *E. coli*. We found that a metabolic bottleneck in the pyrimidine biosynthesis pathway caused by insufficient expression of *pyrE* causes upstream metabolites to accumulate and growth rates to be reduced. While a bottleneck in this reaction had strong effects on the fitness, bottlenecks in other reactions might not. One bottleneck, that we found to have no impact on the fitness was the strain expressing sdRNA:*purM*: Although the substrate level increased and the product level decreased, no reduction of the growth rate was observable. This is explainable by insufficient

bottleneck strengths that limit fluxes through the pathway but not to an extent that the final product concentration is growth limiting. These findings have some implications for the creation of switching enzymes and the growth-based screening for those that. The prerequisite is that the enzymatic activity is the growth limiting factor. Usually, this can be reached by expressing the enzyme in a strain in which the respective native gene is knocked out and letting the resulting strain grow in environmental conditions in which a functional enzyme is essential for growth. In addition to that, the expression strength has to be so low so at least when the enzyme switch is in the OFF state, the growth rate is reduced. For that reason, for our attempts to create enzyme switches we expressed candidate enzymes from low copy number plasmids and either in gene knockout strains or in presence of antibiotics that specifically target the native enzyme version.

For the creation of enzyme switches, we evaluated two techniques: Split Proteins and Domain Insertion.

With the Split Protein approach we were able to create a rapamycin-dependent mDHFR switch. We characterized this switch in growth and metabolomics experiments and could show that when expressed in an *E. coli* strain that is growing on minimal medium, in presence of trimethoprim to inhibit bacterial DHFR and in absence of rapamycin a growth rate limiting bottleneck is introduced that leads to an accumulation of the substrate dihydrofolate. In presence of rapamycin on the other hand both mDHFR fragments reassemble and form a functional enzyme that removes the metabolic bottleneck partially and leads to growth rates similar to a positive control strain in which constantly interacting mDHFR fragments are expressed. With this switching enzyme we could show that Split Proteins generally can control metabolic pathways and several optimizations of the switch are planned to create split enzymes, which have a higher relevance for biotechnology.

Domain Insertion remains a difficult approach. Although we extensively optimized the method and are now able to create larger libraries, due to overdigestion by S1 nuclease and DNaseI there is still a high tendency for enzymes with partially deleted regulatory domains. With further adaptions to the protocol and testing more combinations of enzymes linkers and regulatory domains, we want to create more and larger libraries and are still confident to identify enzymes with switching behavior out of them.

The high-throughput screening method required for the identification of enzyme switches has been successfully developed in this project. Biosensors that are active dependent on a certain compound (e.g. product of a reaction catalyzed by an enzyme of interest) have been proven to be useful tools for the identification of certain strains out of pooled libraries [82] and would have been a good option for the identification of enzyme switches as well. However, we plan to create different enzymes with synthetic allosteric regulation and a development of biosensors for each of the reaction of pathway products would be very laborious. Hence, we favored a reporter system that can be used for many enzymes of interest and therefore developed a growth-based screening system utilizing the growth rate reporter TIMER. This allows the

screening of many switching enzymes as long as their activity is limiting the growth rate of the cells. We have not tested the enrichment method with a Domain Insertion strain library yet but could already show that TIMER reflects the growth rates of *E. coli* batch cultures quite accurately and that out of pooled strain libraries fractions of slow growing cells can be enriched. Hence, we are confident that TIMER is also a suitable tool for the enrichment of conditionally slow growing cells out of Domain Insertion libraries. In addition to that, we speculate that the TIMER-based method for the enrichment of slow-growing cells might also be useful for many other fields of biology and we currently plan to use it in other projects in our group.

After the identification of potentially switching enzymes it is planned to characterize the switch, especially with high-throughput metabolomics methods and potentially by determining the protein kinetics of purified enzymes. Enzymes that have been proven to have a high catalytic activity in the ON state and a substantially lowered activity in the OFF state are furthermore planned to be tested in an overproduction strain. By simulating changing environmental conditions on a very short time scale we could try to elucidate if switching enzymes have a beneficial effect on the fitness of the overproducing cells while ideally having only a slight negative effect on the production rate.

If successful, the in this work performed experiments, optimized protocols and developed methods can act as a basis for the creation and identification of switching metabolic enzymes that have the potential to optimize many different overproduction strains and bioprocesses. One might speculate that in the long term a computer-aided design of synthetic allosteric regulation in metabolic enzymes might be possible, however until then directed evolution with the here presented methods seems an appropriate alternative.

8 Materials and Methods

8.1 Materials, instruments and source of supplies

8.1.1 Chemicals, kits and enzymes

Chemicals and enzymes that have been used in this work were acquired from Bioline (Germany), Carl-Roth (Germany), GE Healthcare (Germany), Invitrogen (USA), Merck (Germany), New England Biolabs (USA), Peqlab (USA), Roche (Switzerland) and Sigma-Aldrich (Germany).

8.1.2 Buffers and solutions

De-ionized water (Purelab ultra water purification systems, ELGA, Germany) has been used to prepare buffers and solutions. If required, they were sterilized by autoclaving (121°C for 15 minutes, 1 bar) or filter sterilization (pore size 0.22 µm, Merck, Germany).

8.1.3 Media and supplements

For *E. coli* pre-cultures and DNA amplification cultures LB (lysogeny broth) medium [137] has been used, for other cultures, M9 minimal medium [138]:

LB medium

Tryptone	**1.0 % (w/v)**
Yeast Extract	0.5 % (w/v)
NaCl	1.0 % (w/v)

M9 minimal medium

Na_2HPO_4	**6 g/L**	**42.2 mM**
KH_2PO_4	3 g/L	22 mM
$(NH_4)_2SO_4$	0.5 g/L	8.56 mM
NaCl	1.5 g/L	11.34 mM

The following components have been added separately:

$ZnSO_4$	**1.8 mg/L**	**6.3 µM**
$CuCl_2$	1.2 mg/L	7 µM
$MnSO_4$	1.2 mg/L	7.1 µM
$CoCl_2$	1.8 mg/L	7.6 µM
Thiamine	1 mg/L	2.8 µM
$MgSO_4$	246 mg/L	1 mM
$CaCl_2$	14.7 mg/L	0.1 mM
$FeCl_3$	16.2 mg/L	60 µM

The resulting medium has been filter sterilized and a carbon source has been added with a final concentration in the medium of 5 g/L.

For the expression of Split Proteins (Chapter 4) casamino acids with a concentration of 10 g/L have been added to the medium, for the enrichment of slow growing cells (Chapter 6.2) the medium has been supplemented with 1 mM of each amino acid, for the examination of the *pyrE* bottleneck (Chapter 3.1) 100 mM uracil.

8.1.4 Antibiotics

Where appropriate antibiotics enlisted in Table 5 have been added to the medium.

Table 5: Antibiotics used in this study.

Antibiotic	Stock concentration mg/mL	Final concentration µg/mL
Ampicillin	50	50
Chloramphenicol	35	35
Kanamycin	50	50
Gentamycin	15	15
Tetracycline	15	15
Trimethoprim	10	10

8.2 Strains and culture conditions

8.2.1 Strains

A list of all Strains used in this study is shown in Table 6:

Table 6: Strains used in this work

Strain or strain libraries	Relevant features	Reference or Source
Escherichia coli K12 DH5α	fhuA2 lacΔU169 phoA glnV44 Φ80' lacZΔM15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17	[139]
Escherichia coli K12 DH10β	F^- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 endA1 araD139 Δ(ara, leu)7697 galU galK λ^- rpsL nupG /pMON14272 / pMON7124	Invitrogen (USA)
Escherichia coli MG1655	F^-, λ^-, *rph-1*	[140]
Escherichia coli NCM3722	F-	[141]
Escherichia coli EMG-2	F^+, λ^+, *rpoS* (Am) rph-1	[140]
Escherichia coli MDS42	F^-, λ^-, *rph-1, ΔfhuACDB*, Δ*endA*, deletion of 699 additional genes	[142]
Escherichia coli W3110	F^-, λ^-, rpoS(Am), rph-1 Inv(rrnD−rrnE)	[140]
Escherichia coli BW25113	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514	[143]
Escherichia coli BW25113 ΔargA	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔargA	[132]
Escherichia coli BW25113 ΔargR	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔargR	[132]
Escherichia coli BW25113 Δcrp	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,Δcrp	[132]
Escherichia coli BW25113 Δfnr	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,Δfnr	[132]
Escherichia coli BW25113 ΔfruR	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔfruR	[132]
Escherichia coli BW25113 Δfur	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,Δfur	[132]
Escherichia coli BW25113 ΔglnG	F^-, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ^-, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔglnG	[132]

Escherichia coli BW25113 ΔhisG	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔhisG	[132]
Escherichia coli BW25113 ΔiscR	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔiscR	[132]
Escherichia coli BW25113 ΔleuA	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔleuA	[132]
Escherichia coli BW25113 Δlrp	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,Δlrp	[132]
Escherichia coli BW25113 ΔnagC	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔnagC	[132]
Escherichia coli BW25113 ΔphoB	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔphoB	[132]
Escherichia coli BW25113 ΔpurR	F⁻, Δ(araD-araB)567, lacZ4787Δ::rrnB-3, λ⁻, rph-1, Δ(rhaD-rhaB)568, hsdR514,ΔpurR	[132]
EcoR wildtype isolates	Various	[103]

8.2.2 Culture conditions

Liquid *E. coli* cultures were incubated either in shake flasks, culture tubes or microtiter plates. For that, usually a preculture was started by inoculating LB medium with bacteria from a glycerol stock and incubated for 6 hours at 37°C, constantly shaking. Next, a M9 preculture has been started by rediluting the LB preculture 1:200 in M9 minimal medium. This culture has been incubated overnight at 37°C, constantly shaking. These precultures were then used in the next morning to start M9 main cultures by re-diluting the M9 preculture 1:200 in fresh M9 minimal medium.

Upon transformation, cells were incubated on solid medium plates with either LB or M9 medium and 1.5% (w/v) agar-agar over night or up to 72 hours at 37°C.

For long-term storage of strains, stationary cultures were mixed with glycerol to a final concentration of 50% (v/v) and stored in cryo tubes, 96 well plates or 15 mL volume falcon tubes,

8.2.3 Transformation

3 different methods for plasmid transformation in E. coli cells have been used in this work: TSS transformation, transformation in chemically competent cells and transformation in electrocompetent cells.

8.2.4 TSS transformation

For TSS transformation [144], a culture has been started by inoculating 5 mL LB medium with 50 µL from an overnight preculture or cryo stock and incubated at 37°C. When an OD of 0.5 to 0.8 is reached, 200 µL of the culture is mixed with 200 µL of the TSS solution (20 % (w/v) Polyethylenglycol 6000, 100 mM MgSO4, 10% (v/v) DMSO in LB medium) and 1 µL of the plasmid. The suspension is incubated on ice for 30 minutes, then at 37°C for 45 minutes and finally plated on LB agar plates.

8.2.5 Preparation of chemically competent cells

For the preparation of chemically competent cells the protocol described by [145] had been followed. For that, a 1 L flask with 100 mL LB medium has been inoculated from cryo stock and incubated overnight or for 8 hours until an optical desity of 0.6 is reached. Then, the cells are placed on ice for 10 minutes, pelleted by centrifugation (10 minutes, 4°C, 2500 x g), resuspended with 50 mL ice cold TB buffer, incubated on ice for 15 minutes, pelleted again and resuspended in 12 mL TB buffer. To the cell suspension 900 µL DMSO has been added, briefly mixed and incubated for 10 minutes on ice. Then, 150 µL aliquots could be created which have been immediately quick frozen in liquid nitrogen and stored at -80 °C until further usage.

For the transformation of plasmid DNA, aliquots have been thawed on ice, 100 ng plasmids added per aliquot and the resulting mixture have been incubated on ice for 30 minutes. Next, cells were heat shocked for 30 seconds in a water bath at 42°C. Upon the heat shock, cells were incubated on ice again for 5 minutes. Then, 800 µL LB medium without antibiotics was added and the cultures incubated shaking for 1 hour at 37°C. Finally, cells were pelleted by centrifugation (3 minutes, 3000 x g), the pellet resuspended in about 100 µL remaining supernatant and plated on LB agar plates.

8.2.6 Preparation of electrocompetent cells

For the preparation of electrocompetent cells the protocol published on the NEB website (https://international.neb.com/protocols/2012/06/21/making-your-own-electrocompetent-cells) has been followed. Briefly, cells are grown to an optical density of 0.5 – 0.7, washed twice with 10% glycerol and resuspended in small volume of residual supernatant. The resulting dense cell suspension was aliquoted (100 µL per tube) and stored at -80°C until needed.

For electroporation, 1 µL of plasmid solution has been added to the cell suspension, briefly mixed and used for electroporation using a BioRad MicroPulser (Bio-Rad Laboratories, USA).

8.3 Plasmids and Oligonucleotides

SeqBuilder (DNASTAR, USA) and Benchling (https://benchling.com) have been used to design oligonucleotides used for PCR reactions and sequencing reactions. Oligonucleotides were purchased from Eurofins (Germany) and Invitrogen (USA) and ThermoFisher (USA). The pooled oligonucleotide library (Chapter 6.2.1) was purchased from Agilent (USA).

A list of all oligonucleotides used in this work can be found in Table 7**-** Table 10**,** a list of all used and constructed plasmids in Table 11 **-** Table 14. A list of all oligonucleotides that have been purchased for the metabolism-wide CRISPRi library can be found in the appendix.

Table 7: Primers used for the construction of plasmids used to introduce metabolic bottlenecks with CRISPRi (Chapter 3.2)

Name	Sequence	Purpose
sgRNA pyrE001	**TACTTTTCTACAGACAAAAAAA**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Introduction of base pairing region to sgRNA, target: Intergenic region (-46), coding strand
sgRNA pyrE002	**GCTAAGCGCAAATTCAATAAAC**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: *pyrE* +21, non-coding strand
sgRNA pyrE003	**TTAAGGCCTATCGCGAAGAGTT**GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: *pyrE* +632, non-coding strand
sgRNA CTRL	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Non targeting
sgRNA gshB	CTTGATGTTGATGTTTGCGATGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: gshB
sgRNA gadA	ACGTGAATCGAGTAGTTCTGAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: gadA
sgRNA glmS	CAGACGACGTAAACCTTCAAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: glmS
sgRNA panC	ATTTGCTGACGCAGCAGCGGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: panC
sgRNA purM	CCGGCATCTTTGTAGCTAAGAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: purM
sgRNA LuxS	CGACTGTGAAGCTATCTAACAAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: LuxS
sgRNA dapD	ATCTCGGCACGGCGTTCAAAAGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC	Target: dapD

sgRNA ddlA	ACGTCGAAGCGACTTTTATCAAGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGC	Target: ddlA
sgRNA cysE	CAGTCCGCCAGCGTTCTGGCTTGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: cysE
sgRNA pyrB	AAATGATATGTTTCTGATATAGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGC	Target: pyrB
sgRNA nrdA	TGCCGCCCAATCCAGAACGCGAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: nrdA
sgRNA mtn	ACCTCGGTTCCATTCAGTTGGCGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: mtn
sgRNA ArgA	GTATTGATATAGGGAACCGAATGGGTTTTAGAGCTA GAAATAGCAAGTTAAAATAAGGC	Target: ArgA
sgRNA leuA	GCCTGTTCACCGTCGCGCAATGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: leuA
sgRNA metK	GGATGCCCTTCAGAGACGGACTGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: metK
sgRNA dapE	CTGTTGTGTCAGCTCAATAACCGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGC	Target: dapE
sgRNA pheA	TTTCTCTCGCAGCGCCAGTAACGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: pheA
sgRNA murB	TGTGTTCCAGGGTTTTAAGGAGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: murB
sgRNA metA	CTTCACGCAAGAAATTGACGGCGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: metA
sgRNA proB	GCCGAGTTTTACCACCAGCGTCGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: proB
sgRNA hisB	TCACTCGGCGGTTCGCTAATCAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: hisB
sgRNA coaD	TGGTAATGGGATCGAAAGTACCGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: coaD
sgRNA ilvC	CGCGCCATCGGCGAATTCATCGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: ilvC
sgRNA pyrE	GCTAAGCGCAAATTCAATAAACGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGGC	Target: pyrE
sgRNA aroL	CCCGAGGCCCGATCAGAAAAAGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: aroL
sgRNA	GGAAAGGATAAATCGCCGTGTCGTTTTAGAGCTAGA	Target: nadA

nadA	AATAGCAAGTTAAAATAAGGC	
sgRNA purB	ACAGGGGAAACGGCGGTCAGTGGTTTTAGAGCTAG AAATAGCAAGTTAAAATAAGGC	Target: purB
sgRNA ArgE	tttttcattgttgacacacctcGTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGC	Target: ArgE
sgRNA folA	TCGGCAGGCAGGTTCCACGGCAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGGC	Target: folA
Ec-R	ACTAGTATTATACCTAGGACTGAGCTAGC	Reverse primer to amplify the sgRNA plasmid [146]
Ec-F-colony	GGGTTATTGTCTCATGAGCGGATACATATTTG-	Colony PCR and sequencing primer for the sgRNA plasmids [146]
Ec-R-colony	CGCGGCCTTTTTACGGTTC	Colony PCR and sequencing primer for the sgRNA plasmids [146]

Table 8: Primers used for the construction of plasmids and sequencing primers used for the development of Split Proteins (Chapter 4)

Name	Sequence	Purpose
pT18 FWD	gaggtggaggctcaggaggaggtggtggttcaGCC GCCAGCGAG GCCACG	Amplification of pT18 and overhangs for CPEC
pT18 REV	TTTTCCGGGATCAAGCTTGGCGTAATC ATGG	Amplification of pT18 and overhangs for CPEC
DHFR3 pT18 FWD	GTCTACGAGAAGAAAGACtaactaagtaata tggtgcactctc agtacaa	Amplification of DHFR3 and overhangs for CPEC
DHFR3 pT18 REV	AAACCATGTCTACTTTACTtgaaccaccacc tcctcct	Amplification of DHFR3 and overhangs for CPEC
DHFR(3) for FKBP12 FWD	GGAGCTTCTAAAACTGGAAggaggaggtg gaggctca	Amplification of pT18-DHFR(3) and overhangs for CPEC
DHFR(3) for FKBP12 REV	GCACTCCCATagctgtttcctgtgtgaaattgtt	Amplification of DHFR(3) and overhangs for CPEC
FKBP12 gene FWD	acaggaaacagctATGGGAGTGCAGGTGG AA	Amplification of FKBP12 and overhangs for CPEC

FKBP12 gene REV	ctcctccTTCCAGTTTTAGAAGCTCCACAT C	Amplification of FKBP12 and overhangs for CPEC
pT25 for FRB FWD	GAACAACCGGAATTGACTCTAGAGGAT CCCCGG GT	Amplification of pT25 and overhangs for CPEC
pT25 for FRB REV	GAATCAGCTCTTCcatagctgtttcctgtgtgaaa ttg	Amplification of pT25 and overhangs for CPEC
FRB gene FWD	ggaaacagctatgGAAGAGCTGATTCGAGT AGCCA	Amplification of FRB and overhangs for CPEC
FRB gene REV	tcctcctgagcctccacctcctccTAGCTGCTTGG AGATCCGT C	Amplification of FRB and overhangs for CPEC
DHFR12 for FRB FWD	gaggtggaggctcaggaggaggtggtggttcaGTTC GACCATTGA ACTGCATC	Amplification of DHFR1,2 and overhangs for CPEC
DHFR12 for FRB REV	TCCTCTAGAGTCAATTCCGGTTGTTCA ATAAGTC TTA	Amplification of DHFR1,2 and overhangs for CPEC
Cut at linker FWD	ggaggaggtggaggctca	Amplification of pT18-DHFR3 and overhangs for CPEC
Cut at Lac REV	CATagctgtttcctgtgtgaaattg	Amplification of pT18-DHFR3 and overhangs for CPEC
Make zip + linker FWD	gataacaatttcacacaggaaacagctATGATCCA GCGTATGA AACAGCTG	Amplification of leucine zippers and overhangs for CPEC
Make Zip + linker REV	tgaaccaccacctcctcctgagcctccacctcctccAC GTTCACCCA CCAGTTTTTT	Amplification of leucine zippers and overhangs for CPEC
Lac promoter T18/25 FWD	tatgcttccggctcgtatgt	Colony-PCR and sequencing primer
Seq primer T25 REV	gttttcccagtcacgacgtt	Colony-PCR and sequencing primer
Seq primer T18 REV	ctggcttaactatgcggcat	Colony-PCR and sequencing primer

Table 9: Primers used for the construction of plasmids and sequencing primers used for the development of enzymes with synthetic allosteric regulation using the Domain insertion approach (Chapter 5)

Name	Sequence	Purpose
malE stop BamHI	ggtcGGATCCttaCTTGGTGATACGAGTCTGC	Amplification of *malE* for construction of pTRC99KK-malE
malE start XhoI	cgccCTCGAGatgAAGATCGAAGAAGGTAAACTG	Amplification of *malE* for construction of pTRC99KK-malE
leuA_f_noATG_5G	ggtggcggaggtggcAGCCAGCAAGTCATTATTTTCG	Amplification of *leuA* with GGGGS linker
leuA_r_nostop_5G	tccaccacctcctccCACGGTTTCCTTGTTGTTTTCG	Amplification of *leuA* with GGGGS linker
DHFR 05 f	ggtggtggagggtctGTTCGACCATTGAACTGCATCGTC	Amplification of *mDHFR* with GGGGS linker
DHFR 05 r	accgccaccaccagaGTCTTTCTTCTCGTAGACTTCAAACTTATACTTGATGCC	Amplification of *mDHFR* with GGGGS linker
mDHFR fwd 3X	NNNNNNNNNGTTCGACCATTGAACTGCATC	Amplification of *mDHFR* with NNN linker
mDHFR rev 3X	NNNNNNNNNGTCTTTCTTCTCGTAGACTTCAAACTTATAC	Amplification of *mDHFR* with NNN linker
pSB cra OH r	TCACAAGACCcatactagtggtttcctgtgtgActc	Amplification of pSB4A5 for the insertion of *cra* with CPEC
pSB cra OH f	AGCCGTAGCtaataactgcaggagtcactaaggg	Amplification of pSB4A5 for the insertion of *cra* with CPEC
cra pSB OH f	accactagtatgGGTCTTGTGATCCCC	Amplification of *cra* for the insertion into pSB4A5 with CPEC
cra pSB OH r	tgcagttattaGCTACGGCTGAGCAC	Amplification of *cra* for the insertion into pSB4A5 with CPEC
pSB malE OH r	CCTTCTTCGATCTTcatactagtggtttcctgtgtgActc	Amplification of pSB4A5 for the insertion of *malE* with CPEC
pSB_malE_OH_f	GACTCGTATCACCAAGtaataactgcaggagtcactaaggg	Amplification of pSB4A5 for the insertion of *malE* with CPEC
malE pSB OH f	cactagtatgAAGATCGAAGAAGGTAAACTG	Amplification of *malE* for the insertion into pSB4A5 with CPEC
malE pSB OH r	cagttattaCTTGGTGATACGAGTCTGC	Amplification of *malE* for the insertion into pSB4A5 with CPEC
pSB purR OH r	ACCGATAGACTTcatactagtggtttcctgtgtgActc	Amplification of pSB4A5 for the

		insertion of *purR* with CPEC
pSB_purR_OH_f	ACTATCGTCGTtaataactgcaggagtcactaaggg	Amplification of pSB4A5 for the insertion of *purR* with CPEC
purR pSB OH f	aaccactagtatgAAGTCTATCGGTTTGCTG	Amplification of *purR* for the insertion into pSB4A5 with CPEC
purR_pSB_OH_r	ctgcagttattaACGACGATAGTCGCG	Amplification of *purR* for the insertion into pSB4A5 with CPEC
pSB_fnr_OH_r	TTTCCGGGATcatactagtggtttcctgtgtgActc	Amplification of pSB4A5 for the insertion of *fnr* with CPEC
pSB_fnr_OH_f	TCATCTACtgataactgcaggagtcactaaggg	Amplification of pSB4A5 for the insertion of *fnr* with CPEC
fnr_pSB_OH_r	tcctgcagttatcaGTAGATGAATGCAGCC	Amplification of *fnr* for the insertion into pSB4A5 with CPEC
fnr_pSB_OH_f	cactagtatgATCCCGGAAAAGC	Amplification of *fnr* for the insertion into pSB4A5 with CPEC
pSB_BB_fwd	taactgcaggagtcactaaggg	Amplification of pSB4A5 for the insertion of several enzyme genes with CPEC
pSB_BB_rev	catactagtggtttcctgtgtgAc	Amplification of pSB4A5 for the insertion of several enzyme genes with CPEC
hisG_OH_SB_f	gTcacacaggaaaccactagtatgACAGACAACACT CGTTTACGC	Amplification of *hisG* for the insertion into pSB4A5 with CPEC
hisG_OH_SB_r	ccttagtgactcctgcagttaCTCCATCATCTTCTCAA TCG	Amplification of *hisG* for the insertion into pSB4A5 with CPEC
argA_OH_SB_f	gTcacacaggaaaccactagtgtgGTAAAGGAACGT AAAACCGAG	Amplification of *argA* for the insertion into pSB4A5 with CPEC
argA_OH_SB_r	ccttagtgactcctgcagttattaCCCTAAATCCGCCAT C	Amplification of *argA* for the insertion into pSB4A5 with CPEC
pSB_BB_rev_ARG	cacactagtggtttcctgtgtgAc	Amplification of pSB4A5 for the insertion of *argA* with CPEC
leuA_OH_SB_f	gagTcacacaggaaaccactagtatgagccagcaagtcatta ttttc	Amplification of *leuA* for the insertion into pSB4A5 with CPEC
leuA_OH_SB_r	cccttagtgactcctgcagttattacacggtttccttgttgttttc	Amplification of *leuA* for the insertion into pSB4A5 with CPEC
leuA M38L M43L r	GGCCAGCGCAATTTGCAG	Amplification of *leuA* to remove two potential start codons

leuA M38L M43L f	CTTGAGCGTttaGGTGTTGACGTGttaGAAGTCG	Amplification of *leuA* to remove two potential start codons
pTRC99KK_fwd	catcataacggttctggcaa	pTRC99KK sequencing primer
pTRC99KK_rev	aaaatcttctctcatccgcc	pTRC99KK sequencing primer
pSB4A5 seq rev	aagggaagttttagtcaaaagc	pSB4A5 sequencing primer
pSB4A5 seq fwd	cggatcctacctgacg	pSB4A5 sequencing primer

Table 10: Primers used for the construction of plasmids and sequencing primers used for the development of a method to enrich slow growing strains using a TIMER protein (Chapter 6) and primers required for next generation sequencing.

Name	Sequence	Purpose
pBR_CMP_BB_F	gacgtctaagaaaactgtcagaccaagtttactcatatatac	Amplification of pBR322_TIMER for exchange of the resistance cassette
pBR_CMP_BB_R	tggtgagaatccaaaatgtatttagaaaaataaacaaaagag	Amplification of pBR322_TIMER for exchange of the resistance cassette
Cmp_pBR_fwd	ctaaatacattttggattctcaccaataaaaaacg	Amplification of Cmp^r for exchange of the resistance cassette in pBR322_TIMER
Cmp_pBR_rev	ggtctgacagttttcttagacgtcaggtggc	Amplification of Cmp^r for exchange of the resistance cassette in pBR322_TIMER
kdo1740	gaaaatgagacgtccagttcaccgacaaacaacag	amplification of Addgene #44251
kdo1742	tatagcggccgcaataggcgtatcacgaggcaga	amplification of Addgene #44251
kdo1741	tcggtgaactggacgtctcattttcgccagatatcgac	amplification of Addgene #44249
kdo1739	gcgcttaattaaagcgagtcagtgagcgagg	amplification of Addgene #44249
kdo1737	gcgcttaattaacagtaatgacctcagaactccatctgg	amplification pNUT1270
kdo1736	tctagaagcggccgcgaattc	amplification pNUT1270
kdo1860	ccagatatcgaccaagcgagctctctagacgcca	amplification of lacIQ1
kdo525	tttagacctccttagctcctgaattccta	amplification of lacIQ1

OH_amp_fwd	taaggatgatttctggaattctaaag	Amplification of pooled oligonucleotides
OH_amp_rev	gtgccacttttttcaagttgataac	Amplification of pooled oligonucleotides
EcF_forward	gttttagagctagaaatagcaagttaaaataaggc	Amplification of pNUT1527 for Gibson Assembly with amplified pooled oligonucleotides
EcF_reverse	actagtattatacctaggactgagctagc	Amplification of pNUT1527 for Gibson Assembly with amplified pooled oligonucleotides
NGS_F2_adapter	TCGTCGGCAGCGTCAGATGTGTATAAGAG ACAGcgcaataggcgtatcacgagg	Amplification of a 150 bp fragment of pNUT1527 including the sgRNA
NGS_R2_adapter	GTCTCGTGGGCTCGGAGATGTGTATAAGA GACAGcgacggcgctattcagatcc	Amplification of a 150 bp fragment of pNUT1527 including the sgRNA
Custom_N701	CAAGCAGAAGACGGCATACGAGATTCGCC TTAGTCTCGTGGGCTCGG	I7 oligo
Custom_N702	CAAGCAGAAGACGGCATACGAGATCTAGT ACGGTCTCGTGGGCTCGG	I7 oligo
Custom_N703	CAAGCAGAAGACGGCATACGAGATTTCTG CCTGTCTCGTGGGCTCGG	I7 oligo
Custom_N704	CAAGCAGAAGACGGCATACGAGATGCTCA GGAGTCTCGTGGGCTCGG	I7 oligo
Custom_N705	CAAGCAGAAGACGGCATACGAGATAGGAG TCCGTCTCGTGGGCTCGG	I7 oligo
Custom_N706	CAAGCAGAAGACGGCATACGAGATCATGC CTAGTCTCGTGGGCTCGG	I7 oligo
Custom_S502	AATGATACGGCGACCACCGAGATCTACAC CTCTCTATTCGTCGGCAGCGTC	I5 oligo
Custom_S503	AATGATACGGCGACCACCGAGATCTACAC TATCCTCTTCGTCGGCAGCGTC	I5 oligo

Table 11: Plasmids used for the introduction of metabolic bottlenecks using CRISPRi (Chapter 3.2)

Name	Genotype	Reference
pdCas9-bacteria	p15A, pLtetO-1-dCas9, Cmp^r	[48]
pgRNA-bacteria	pUC19, pBBa_J23119-sgRNA, Amp^r	[48]
pgRNA-pyrE001	pUC19, pBBa_J23119-sgRNA:pyrE001, Amp^r	This study
pgRNA-pyrE002	pUC19, pBBa_J23119-sgRNA:pyrE002, Amp^r	This study
pgRNA-pyrE002	pUC19, pBBa_J23119-sgRNA:pyrE003, Amp^r	This study
pgRNA-gshB	pUC19, pBBa_J23119-sgRNA:gshB, AmpR	This study
pgRNA-gadA	pUC19, pBBa_J23119-sgRNA:gadA, AmpR	This study
pgRNA-glmS	pUC19, pBBa_J23119-sgRNA:glmS, AmpR	This study
pgRNA-panC	pUC19, pBBa_J23119-sgRNA:panC, AmpR	This study
pgRNA-purM	pUC19, pBBa_J23119-sgRNA:purM, AmpR	This study
pgRNA-LuxS	pUC19, pBBa_J23119-sgRNA:LuxS, AmpR	This study
pgRNA-dapD	pUC19, pBBa_J23119-sgRNA:dapD, AmpR	This study
pgRNA-ddlA	pUC19, pBBa_J23119-sgRNA:ddlA, AmpR	This study
pgRNA-cysE	pUC19, pBBa_J23119-sgRNA:cysE, AmpR	This study
pgRNA-pyrB	pUC19, pBBa_J23119-sgRNA:pyrB, AmpR	This study
pgRNA-nrdA	pUC19, pBBa_J23119-sgRNA:nrdA, AmpR	This study
pgRNA-mtn	pUC19, pBBa_J23119-sgRNA:mtn, AmpR	This study
pgRNA-ArgA	pUC19, pBBa_J23119-sgRNA:ArgA, AmpR	This study
pgRNA-leuA	pUC19, pBBa_J23119-sgRNA:leuA, AmpR	This study
pgRNA-metK	pUC19, pBBa_J23119-sgRNA:metK, AmpR	This study
pgRNA-dapE	pUC19, pBBa_J23119-sgRNA:dapE, AmpR	This study
pgRNA-pheA	pUC19, pBBa_J23119-sgRNA:pheA, AmpR	This study
pgRNA-murB	pUC19, pBBa_J23119-sgRNA:murB, AmpR	This study
pgRNA-metA	pUC19, pBBa_J23119-sgRNA:metA, AmpR	This study
pgRNA-proB	pUC19, pBBa_J23119-sgRNA:proB, AmpR	This study
pgRNA-hisB	pUC19, pBBa_J23119-sgRNA:hisB, AmpR	This study
pgRNA-coaD	pUC19, pBBa_J23119-sgRNA:coaD, AmpR	This study

pgRNA-ilvC	pUC19, pBBa_J23119-sgRNA:ilvC, AmpR	This study
pgRNA-pyrE	pUC19, pBBa_J23119-sgRNA:pyrE, AmpR	This study
pgRNA-aroL	pUC19, pBBa_J23119-sgRNA:aroL, AmpR	This study
pgRNA-nadA	pUC19, pBBa_J23119-sgRNA:nadA, AmpR	This study
pgRNA-purB	pUC19, pBBa_J23119-sgRNA:purB, AmpR	This study
pgRNA-ArgE	pUC19, pBBa_J23119-sgRNA:ArgE, AmpR	This study
pgRNA-folA	pUC19, pBBa_J23119-sgRNA:folA, AmpR	This study

Table 12: Plasmids used for the construction of Split Proteins (Chapter 4)

Name	Genotype or purpose	Reference
pT18	Expression plasmid, AmpR, P_{lac}	Gift from Prof. Dr. Daniel Ladant
pT25	Expression plasmid, KanR, P_{lac}	Gift from Prof. Dr. Daniel Ladant
pKT25-zip	Template for amplification of leucine zippers	BATCH EUROMEDEX
pMT3-FKBP(full-length)-DHFR[3]	Template for amplification of FKBP1,2 and DHFR(3)	Gift from Prof. Dr. Stephen Michnick
pcDNA3-DHFR[1,2:F31S]-FRAP	Template for amplification of FRB and DHFR1,2	Gift from Prof. Dr. Stephen Michnick
pT18-FKBP1,2-DHFR(3)	AmpR, P_{lac}-FKBP1,2-DHFR(3)	This work
pT25-FRB-DHFR1,2	KanR, P_{lac}-FRB-DHFR1,2	This work
pT18-Zip-DHFR(3)	AmpR, P_{lac}-Zip-DHFR(3)	This work
pT25-Zip-DHFR1,2	KanR, P_{lac}-Zip-DHFR1,2	This work

Table 13: Plasmids used for the construction of enzymes with synthetic allosteric regulation using the Domain Insertion approach (Chapter 5)

Name	Genotype or purpose	Reference
pTRC99KK	pBR322 origin, AmpR, P_{trc}	Gift from Karl Kochanowski [29]

pTRC99KK-malE	pBR322 origin, AmpR, P_{trc}-*malE*	This work
pSB4A5	pSC101 origin, AmpR, P_{BAD}	[147]
pSB4A5-malE	pSC101 origin, AmpR, P_{BAD}-*malE*	This work
pSB4A5-cra	pSC101 origin, AmpR, P_{BAD}-*cra*	This work
pSB4A5-fnr	pSC101 origin, AmpR, P_{BAD}-*fur*	This work
pSB4A5-purR	pSC101 origin, AmpR, P_{BAD}-*purR*	This work
pSB4A5-mDHFR	pSC101 origin, AmpR, P_{BAD}-*mDHFR*	This work
pSB4A5-argA	pSC101 origin, AmpR, P_{BAD}-*argA*	This work
pSB4A5-leuA	pSC101 origin, AmpR, P_{BAD}-*leuA*	This work
pSB4A5-hisG	pSC101 origin, AmpR, P_{BAD}-*hisG*	This work

Table 14: Plasmids used for the development of a method to enrich slow growing strains (Chapter 6)

Name	Genotype or purpose	Reference
pSC101_TIMER	Tet^r, TIMER in pSC101	[102]
pBR322_TIMER	Amp^r, TIMER in pBR322	[102]
pBR322-C_TIMER	Cmp^r, TIMER in pBR322	This study
pNUT1527	$Gent^r$, Ptac-dCas9, pJ23119-sgRNA	This study

8.4 Molecular working with DNA

8.4.1 Plasmid preparation

For plasmid preparations kits from different suppliers (Thermo Scientific, Qiagen, Macherey-Nagel, Zymo Research) have been used following the manufacturers' instructions.

8.4.2 PCR

Polymerase chain reactions (PCR) were used to amplify DNA. For reactions in which low error rates were crucial, Q5 polymerase (NEB) was used; for Colony-PCRs Taq polymerase (NEB) has been used. The annealing temperatures were defined by the composition of primers as well as the used polymerase and calculated with Benchling (https://benchling.com). As template either defined amounts of plasmid DNA or chromosomal DNA has been used. For the preparation of chromosomal DNA, small amount of a colony of the strain of interest has been re-solubilized in 100 µL H_2O and heated for 5 minutes at 90°C. 1 µL of this suspension than had been used as template.

8.4.3 Electrophoresis

For gel electrophoresis usually 1% agarose gels were used. As buffer system TAE buffer has been used. For the isolation of linearized plasmids in some cases 0.8% LB (Lithium Bromide) agarose gels were used.

8.4.4 Purification

For DNA clean up upon ligation and kits from Thermo Scientific, Qiagen and Zymo Research have been used.

8.4.5 Enzymatic modification

8.4.5.1 Restriction Digests

Enzymes from NEB have been used for DNA restriction, following the manufacturer's instructions. High fidelity enzymes were preferred and used when available.

8.4.5.2 Ligation

For ligations of genes into plasmid backbones usually an insert/plasmid-ratio of 1:3 has been chosen. The needed amounts of plasmids and inserts have been calculated with the “Ligation Calculator” (http://www.insilico.uni-duesseldorf.de/Lig_Input.html, Universität Düsseldorf). Prior to ligation in some cases the ends of the plasmid backbone have been dephosphorylated using Antarctic Phosphatase (NEB).T4 ligase (NEB, Thermo Scientific) has been used for the ligation, the reaction mix was incubated overnight at room temperature. Ligated plasmids were subsequently purified using DNA Clean & Concentrator Kit (Zymo Research) in case they were planned to be transformed into electrocompetent cells.

8.4.5.3 Gibson Assembly

As alternative method to construct plasmids, Gibson Assembly and CPEC have been used. For Gibson Assembly [148] multiple DNA fragments with overlapping sequences at their ends are required. A 5’ exonuclease partially digests the 5’ ends of double stranded DNA, the resulting sticky ends of different DNA fragments anneal and can be covalently joined by a polymerase and ligase. We designed fragments in a way that they have overlapping sequences between 20 and 40 nt and used a Gibson Assembly Master mix from NEB.

8.4.5.4 CPEC

Circular polymerase extension cloning (CPEC, [112]) also requires overlapping end sequences of DNA fragments that are supposed to be connected with each other and includes several rounds of denaturation, annealing and extension. First, the DNA fragments are denatured by heat to create single strands. Single strands of different DNA fragments can anneal at the overlapping sequences from which a polymerase is then synthesizing the missing second strand. Overlapping sequences had a length between 20 and 35 nt and where designed to have a very similar melting temperature (T_m). As polymerase Q5 polymerase (NEB) had been used, the extension time was defined by the length of the longest fragment to assemble.

8.5 Domain Insertion

The creation of plasmid libraries with Domain Insertion [73,149] is divided in 6 parts as depicted in Figure 52: In addition to the preparation of the insert gene (1), the acceptor plasmid is linearized randomly (2), repaired (3), isolated (4) and dephosphorylated (5). Finally, both the linearized plasmids and insert genes are ligated together (6).

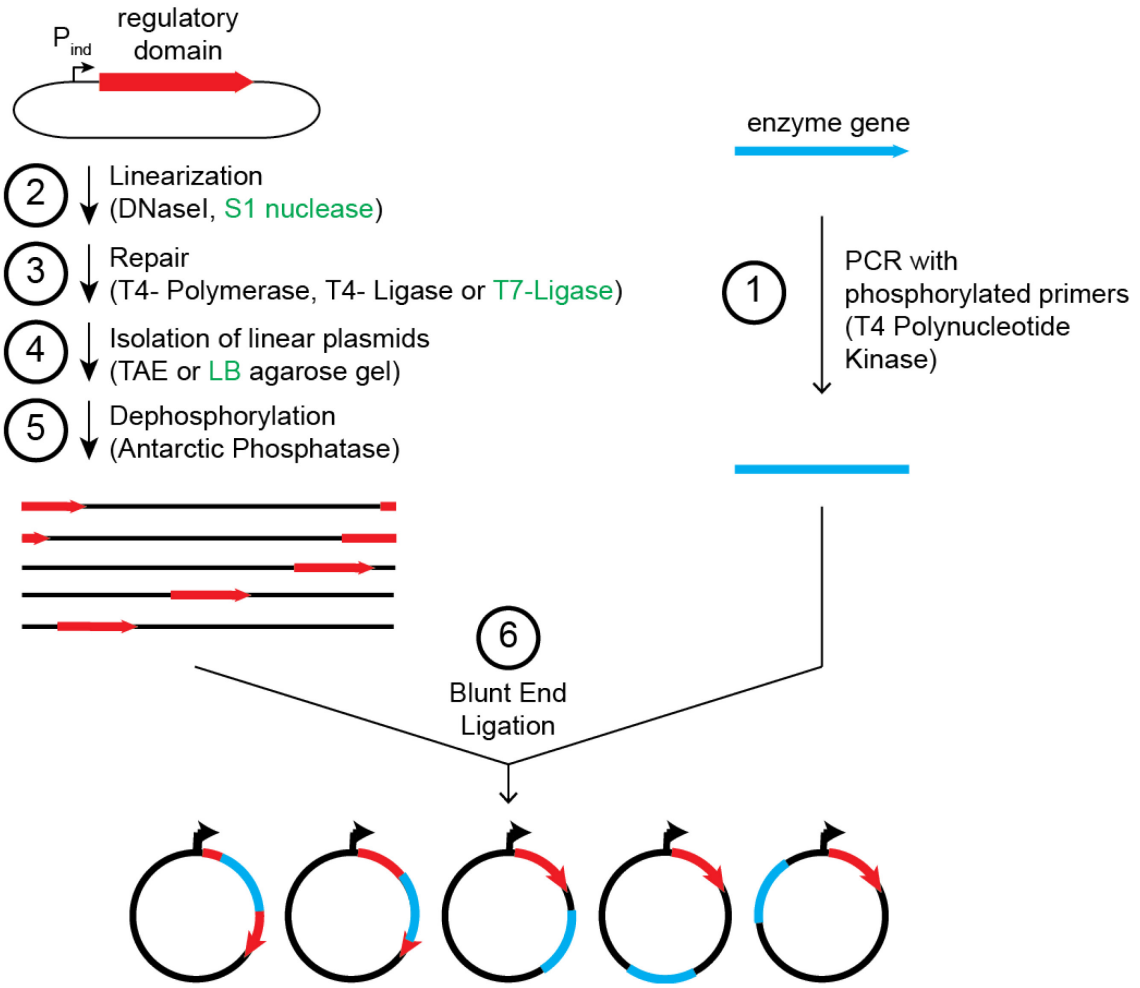

Figure 52: Domain Insertion.

Domain Insertion can be divided in 6 steps: Amplification of the enzyme gene with phosphorylated primers (1), Linearization of the expression plasmid carrying the gene coding for the regulatory domain with a sequence independent nuclease (2), repair of the linearized plasmid (3), isolation of the linear plasmids through gel extraction (4), dephosphorylation of the ends to prevent re-ligation (5) and ligation of both, plasmids and enzymes (6). Optimized steps are highlighted in green.

8.5.1 Preparation of the insert gene

For the amplification of the insert gene, Q5 polymerase (NEB) and primers that were phosphorylated at the 5' end with T4 polynucleotide kinase (NEB) following the manufacturer's instructions were used.

8.5.2 Random Linearization

8.5.2.1 Determination of the optimal DNaseI concentration

Needed components:

- 40 µg plasmid DNA
- DNaseI (1U/µL) (Thermo Scientific)
- Diluent solution: 25 mM Tris-HCl pH7.5, 1mM $MnCl_2$, 6 µL 100x BSA
  (has to be prepared freshly and kept at room temperature)
- 1.0 M EDTA
- 0.8% Agarose gel, alternatively 1%

Experimental procedure:

1) Preheat Thermo Block at 75°C
2) Preparation of 100 mL 50mM Tris stock:
   (0.6057 mg Tris, pH7.5 with HCl)
3) Preparation of an agarose gel
4) Preparation of 500 µL Diluent Solution
   (250 µL Tris-HCl, 5 µL 100 mM $MnCl_2$, 2.5 µL 100x BSA, ad 500 µL H_2O)
   (Attention: BSA is pretty viscous!)
5) Preparation of 800 µL Diluent Solution with plasmid DNA
   (400 µL Tris-HCl, 8 µL 100 mM $MnCl_2$, 4 µL 100x BSA, X µL plasmid DNA (40 µg), ad 800 µL H_2O)
6) Dilution series of DNaseI:
   Serial dilution of DNaseI: 1:10 dilutions, 8 steps, each 20 µL (2 µL DNaseI or previous dilution step + 18 µL Diluent Solution). Keeping DNaseI on ice. Mixing by pipetting the whole reaction volume (20 µL) up and down three times.
7) Aliquoting 8x 95 µL of the Diluent Solution with plasmid DNA
8) To each aliquot 5 µL of the DNaseI dilution series added, then mixed by pipetting the whole reaction volume (100 µL) up and down three times.
9) Incubation for exact 8 min at room temperature
10) Stopping the reaction by adding 1.2 µL. Mixing by pipetting the whole reaction volume (102.4 or 124 µL) up and down three times.
11) Heat inactivation of DNaseI by incubation at 75°C for 10 min
12) Determination of the optimal DNaseI concentration by loading about 10 µL of the reaction products on the agarose gel. The optimal DNaseI concentration is the one that results in a sharp band for linearized plasmid without producing a “smear” caused by multiple digestions.

8.5.2.2 DNaseI digestion

Needed components:

- 60 µg Plasmid-DNA
- DNaseI (1U/µL) (Thermo Scientific)
- Diluent solution: 25 mM Tris-HCl pH7.5, 1mM $MnCl_2$, 6 µL 100x BSA (has to be prepared freshly and kept at room temperature)
- 1.0 M EDTA
- PCR Purification Kit
- Agarose gel (0.8% or 1% agarose in LB or TAE buffer)

Experimental procedure:

1) Preheating Thermo Block to 75°C
2) Preparation of 100 mL 50mM Tris stock
3) Preparation of an agarose gel (if planned to proceed with repair and isolation of linear plasmids after linearization)
4) Preparation of 500 µL Diluent Solution
   (250 µL Tris-HCl, 5 µL 100 mM $MnCl_2$, 2.5 µL 100x BSA, ad 500 µL H_2O)
   Preparation of 1200 µL Diluent Solution with plasmid DNA
   (600 µL Tris-HCl, 12 µL 100 mM $MnCl_2$, 6 µL 100x BSA, X µL plasmid DNA (60 µg) ad 1200 µL H_2O)
   Aliquot 12x 95 µL of the Diluent Solution with plasmid DNA. Incubate for 10 minutes at room temperature.
5) DNaseI:
   Preparation of a serial dilution of DNaseI in 1:10 dilution steps until your previously determined optimal concentration of DNaseI is reached. 60 µL DNaseI-Solution are required for digestion.
6) Adding 5 µL of the DNaseI solution to each aliquot of plasmid DNA. Mixing by pipetting the whole reaction volume (100 µL) up and down three times.
7) Incubation for exact 8 min at room temperature
8) Stopping the reaction by adding 1.2 µL 1M EDTA
9) Mix by pipetting the whole reaction volume up and down three times.
10) Heat inactivation of DNaseI by incubation at 75°C for 10 min
11) 3 times: Combining the contents of 4 tubes to one tube and purification of the DNA using a DNA purification kit.
12) Combining all 3 eluates.
13) Storing at -20°C or proceeding with repairing (Step 3)

8.5.2.3 S1 nuclease digestion

Needed components:

- 50 µg Plasmid-DNA
- S1 Nuclease + Buffer (as provided by Promega)
- PCR Purification Kit

Experimental procedure:

1.) Preheating a heat block to 37°C
2.) **(25x)**
X µL (2 µg) plasmid DNA,
2.5 µL 10x S1 Nuclease Buffer,
0.5 µL S1 Nuclease (10U)
22 – X µL H_2O
3.) Incubation for 20 min at 37°C
4.) Stopping the incubation by pooling all reactions and adding DNA binding buffer of a DNA purification kit
5.) Purification of the DNA and elution in H_2O. Concentration in Speedvac if needed.
6.) Storing at -20°C or proceeding with repairing (Step 3)

Either: Proceed with Part 3

Or: Store the DNA at -20°C

8.5.3 Repair of linearized DNA

Needed components:

- Linearized plasmid DNA
- 1 Agarose gel (0.8% in LB buffer)
- Gel Extraction Kit
- T7-Ligase
- T4-Polymerase
- dNTPs, Ligase Buffer, BSA
- DNA purification kit
- Gel Extraction Kit
- 1.0 M EDTA
- Antarctic Phosphatase

Experimental procedure:

1) Preparation of a small agarose gel
2) Preheating a Thermo Block to 75°C
3) Cooling down a second Thermo Block to 12°C
4) Determining the concentration of the Plasmid DNA with a NanoDrop or similar.
5) Loading 200-500 ng on an agarose gel to analyze the quality of the linearization.
6) Preparation of a T7-Ligase and T4-Polymerase-Reaction-Mix for repairing the linearized plasmids:
   T7-Ligase (160U/1 µg linearized DNA), T4-Polymerase (1U/1 µg linearized DNA), 1.5 µL dNTPs (200µM final concentration), 7,5 µL T4-Ligase Buffer (10x), 0.75 µL BSA (100x, stored in our freezer) and up to 75 µL your linearized plasmid DNA
   Incubation in multiple tubes with each 75 µL reaction volume.
7) Incubation for 20 minutes at 12°C
8) Stopping the reaction by adding 15 µL 50 mM EDTA, mixing by pipetting up and down gently and heat inactivating the enzymes for 15 minutes at 75 °C
9) Letting the solution cooling down to room temperature and isolating the DNA using a PCR Purification Kit. Usage of multiple tubes to prevent overloading the columns is recommended.

8.5.4 Isolation of linearized plasmids and dephosphorylation

Needed components:

- Linearized and repaired plasmid DNA
- 1 Agarose gel (0.8% in LB buffer)
- Gel Extraction Kit
- Antarctic Phosphatase
- DNA purification kit

Experimental procedure:

1) Pool the isolated DNA and load it on a large agarose gel. As a comparison it is recommended to load some plasmid DNA linearized by a restriction enzyme. The by DNAseI or S1 nuclease linearized plasmid should run at the same height.
2) Extracting the band containing the linearized plasmids and purification of the DNA with the Gel Extraction Kit.
3) If more than one band was extracted all elutions should be pooled
4) Heating the Thermo Block to 37°C
5) Measure the concentration with the NanoDrop
6) Dephosphorylation of the DNA with Antarctic Phosphatase:
   up to 1 µg plasmid DNA, 2 µL 10x reaction buffer, 1 µL Antarctic Phosphatase, to 20 µL ddH_20
7) Incubation for 30 min at 37°C
8) Stopping the reaction by incubation for 10 min at 65°C

8.5.5 Ligation and Transformation

For the ligation at the beginning the ligation protocol described in [149] (see below) has been used. In later attempts, the Blunt/TA-ligation mix (NEB) had been used which is optimized for ligations of fragments with blunt ends, following the manufacturer's instructions.

Needed components:

- T4-Ligase
- 10x Ligase Buffer
- PEG-8000
- Plasmid, Insert DNA

Experimental procedure:

- x µL Plasmid (1µg)
- y µL Insert DNA (ratio of **1:1.5**; calculate the needed amount with Ligation Calculator (http://www.insilico.uni-duesseldorf.de/Lig_Input.html))
- PEG-8000 (final concentration: 5%)
- T4-Ligase Buffer (final concentration: 1x)
- T4 Ligase (2000 U)
- ad 30 µL H_2O

Incubation at room temperature overnight, afterwards the DAN should be purified prior to transformation into electrocompetent DH10β cells (Thermo Scientific).

8.6 Cloning of single CRISPRi plasmids

To construct single CRISPRi plasmids, we followed the protocol by [48]. For that, new base pairing sequences were designed by identification of PAM sequences of the composition NGG within the gene of interest and introduced in the sgRNA expression plasmid pgRNA-bacteria by adding the target sequence to the forward primer. With this primer and a reverse primer binding on the plasmid backbone the plasmid had been amplified. The endings were subsequently phosphorylated with T4 PNK (NEB) and ligated with T4 ligase (NEB) or the Blunt/TA-ligation mix (NEB).

8.7 Cloning of a pooled CRISPRi library

For the creation of the pooled CRISPRi library (Chapter 6.2.1), target sequences were identified and chosen using a Matlab script we wrote for this purpose. Oligonucleotides including the target sequence and flanking regions 65 nt upstream and downstream of the variable sequence were purchased from Agilent (USA) as pooled oligonucleotide stock. For the construction of the library, the oligonucleotides have been amplified in a low cycle PCR (15 cycles, 0.02 pmol template) to create double stranded DNA. In parallel the plasmid backbone (pNUT1527) has been amplified without the gene targeting sequence. Next, both DNA fragments were assembled with Gibson Assembly using the 65 nt flanking regions as overlapping sequence.

4 Gibson Assembly reactions were pooled, purified, concentrated and transformed into highly competent *E. coli* DH10β cells. The library composition and coverage has been examined with Illumina sequencing (Figure 40, page 80).

8.8 Illumina Sequencing

For Illumina Sequencing a two-step PCR approach was used to generate DNA fragments that are compatible with Illumina sequencing. First, a 300 bp fragment including the sgRNA sequence and flanking regions has been amplified using Q5 polymerase (New England Biolabs, USA) 150 ng purified library plasmids as template in a 50 µL PCR reaction with following settings: 98 °C for 10 s, 12 cycles of 98 °C for 10 s, 65 °C for 30 s and 72 °C for 15 s; final extension at 72 °C for 5 min. Afterwards, the PCR products were purified using a NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Germany).

Second, to add different pairs of indexes (i5 and i7) to each amplicon, a second PCR using Phusion High- Fidelity DNA Polymerase (New England BioLabs, USA) has been performed with the following conditions: 98 °C for 30 s; 12 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 20 s; final extension at 72 °C for 5 min. The PCR products were cleaned up using AMPure XP beads (Beckman Coulter). After cleanup, 10 µL of each library was pooled. The concentration of the library pool was measured using the Qubit dsDNA BR Assay on a Qubit 2.0 Fluorometer.

The pooled sequences were then diluted to a final concentration of 2 nM and loaded on a MiniSeq High Output Cartridge following the manufacturer's instructions. 50% PhiX sample was spiked in to ensure sufficient sequence diversity. Sequences were obtained from single-end reads and mapped to the 7184 sgRNAs in the library using a Matlab Script. Only sequences that mapped to sgRNA guide sequences in the reference library (Table S2 in the publication) were used to calculate read counts. Fold changes were calculated as $(reads_{after}/reads_{after,total})/(reads_{before}/reads_{before,total})$, where "before" and "after" indicate before/after enrichment, and "total" the total read counts. Sequences with less than 10 reads before enrichment were removed from further analysis.

8.9 Mass Spectrometry

Cells were either cultured in 100 mL flasks with 5 – 10 mL culture volume or 12-well flat transparent microtiter plates. When cells reached ODs between 0.5 and 1, an equivalent of 1 mL with $OD_{600\ nm}$ = 1 of the culture was vacuum-filtered on a 0.45 µm poresize filter (HVLP02500, Merck Millipore). Filters were immediately transferred into 40:40:20 (vol %) acetonitrile/methanol/water at −20 °C for extraction. Extracts were centrifuged for 15 min at 13000 rpm at −9 °C. Centrifuged extracts were analyzed by LC−MS/MS, with an Agilent 6495 triple quadrupole mass spectrometer (Agilent Technologies) as described previously [104]. Obtained data sets have been analyzed with Matlab (Mathworks, USA) and MassHunter software (Agilent, USA).

8.10 Flow Cytometry

8.10.1 Sample preparation and Data Analysis

Cells were diluted in fresh medium or 1xTB buffer prior to analysis and sorting of fractions if needed. For fluorescence- activated cell analysis a BD LSRFortessa SORP flow-cytometer (BD Biosciences, USA) was used. Fluorescence-activated cell sorting was carried out on a BD FACS Aria Fusion (BD Biosciences, USA). 561 nm lasers, 502 long pass and 532/ 22 bandpass filters were used for the detection of the red fraction of TIMER. 488 nm lasers, 600 long pass and a 520/30 band-pass filters were used for green fluorescent TIMER. To identify cells in the forward/side scatter plot, 488 nm lasers were used. 10 000 or 100 000 cells were sorted per sample according to the red/green fluorescence ratio. BD FACSDiva software version 8.0 (BD Biosciences, USA) and FlowJo v10.4.1 (FlowJo LLC, USA) were used for analysis of the acquired data.

8.10.2 Enrichment of slow growing cells

For the enrichment of slow growing cells with a metabolic bottleneck in a pathway required for growth in absence of amino acids (Chapter 6.2.2), cells were grown under two different conditions. For the first culture several 100 mL flasks each with 5 mL M9-Glucose medium supplemented with 1 mM of each amino acid has been inoculated with defined volumes of the cryo stock with dilutions between 1:1000 and 1:1,000,000. The cultures were then incubated at 37°C and shaking for 6 hours.

Desired was a culture with an optical density between 0.3 and 0.7 (mid-exponential phase). 0.5 mL of such a culture has been filtered, washed with 10 mL prewarmed M9 glucose medium without amino acids and resuspended in 5 mL prewarmed M9 glucose medium without amino acids.

With this suspension several 100 mL flasks with M9 glucose medium have been inoculated with final dilutions between 1:50 and 1:100.000 and total volumes of 5 mL to ensure sufficient aeration. The resulting cultures have been incubated for 6 hours shaking at 37°C, a culture with an optical density of about 0.2 has then subsequently been used for sorting.

9 References

1. Monk JM, Lloyd CJ, Brunk E, et al. iML1515, a knowledgebase that computes Escherichia coli traits. *Nat Biotechnol*. 2017;35(10):904-908. doi:10.1038/nbt.3956.

2. Li X, Gianoulis TA, Yip KY, Gerstein M, Snyder M. Extensive In Vivo Metabolite-Protein Interactions Revealed by Large-Scale Systematic Analyses. *Cell*. 2010;143:639-650. doi:10.1016/j.cell.2010.09.048.

3. Piazza I, Kochanowski K, Cappelletti V, et al. A Map of Protein-Metabolite Interactions Reveals Principles of Chemical Communication. *Cell*. 2018;172(1-2):358-372.e23. doi:10.1016/j.cell.2017.12.006.

4. Kochanowski K, Gerosa L, Brunner SF, Christodoulou D, Nikolaev Y V, Sauer U. Few regulatory metabolites coordinate expression of central metabolic genes in Escherichia coli. *Mol Syst Biol*. 2017;13:903. doi:10.15252/msb.20167402.

5. Nudler E, Mironov AS. The riboswitch control of bacterial metabolism. *Rev TRENDS Biochem Sci*. 2004;29(1). doi:10.1016/j.tibs.2003.11.004.

6. Winkler WC, Breaker RR. Genetic control by metabolite-binding riboswitches. *ChemBioChem*. 2003;4(10):1024-1032. doi:10.1002/cbic.200300685.

7. Striebel F, Imkamp F, Özcelik D, Weber-Ban E. Pupylation as a signal for proteasomal degradation in bacteria. *Biochim Biophys Acta - Mol Cell Res*. 2014;1843(1):103-113. doi:10.1016/J.BBAMCR.2013.03.022.

8. Cameron DE, Collins JJ. Tunable protein degradation in bacteria. *Nat Biotechnol*. 2014;32(12):1276-1281. doi:10.1038/nbt.3053.

9. Gama-Castro S, Salgado H, Santos-Zavaleta A, et al. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res*. 2016;44(D1):D133-43. doi:10.1093/nar/gkv1156.

10. Grainger DC, Hurd D, Harrison M, Holdstock J, Busby SJW. *Studies of the Distribution of Escherichia Coli cAMP-Receptor Protein and RNA Polymerase along the E. Coli Chromosome*. Vol 6. 2005. www.pnas.orgcgidoi10.1073pnas.0506687102. Accessed October 12, 2018.

11. Ramseier TM. Cra and the control of carbon flux via metabolic pathways. *Res Microbiol*. 1996;147(6-7):489-493. doi:10.1016/0923-2508(96)84003-4.

12. Deutscher J, Francke C, Postma PW. How Phosphotransferase System-Related Protein Phosphorylation Regulates Carbohydrate Metabolism in Bacteria †. *Microbiol Mol Biol Rev*. 2006;70(4):939-1031. doi:10.1128/MMBR.00024-06.

13. Mao X-J, Huo Y-X, Buck M, Kolb A, Wang Y-P. Interplay between CRP-cAMP and PII-Ntr systems forms novel regulatory network between carbon metabolism and nitrogen assimilation in Escherichia coli. *Nucleic Acids Res*. 2007;35(5):1432-1440. doi:10.1093/nar/gkl1142.

14. Zhang Z, Gosset G, Barabote R, Gonzalez CS, Cuevas WA, Saier MH. Functional Interactions between the Carbon and Iron Utilization Regulators, Crp and Fur, in Escherichia coli Downloaded from. *J Bacteriol*. 2005;187(3):980-990. doi:10.1128/JB.187.3.980-990.2005.

15. Balsalobre C, Johansson J, Uhlin BE. Cyclic AMP-Dependent Osmoregulation of crp Gene Expression in Escherichia coli Downloaded from. *J Bacteriol*. 2006;188(16):5935-5944. doi:10.1128/JB.00235-06.

16. Egan SM, Schleif RF. DNA-Dependent Renaturation of an Insoluble DNA Binding Protein: Identification of the RhaS Binding Site at rhaBAD. *J Mol Biol*. 1994;243(5):821-829. doi:10.1006/JMBI.1994.1684.

17. Lewis M. The lac repressor. *C R Biol*. 2005;328(6):521-548. doi:10.1016/J.CRVI.2005.04.004.

18. Burgess RR. Sigma Factors. *Encycl Genet*. January 2001:1831-1834. doi:10.1006/RWGN.2001.1192.

19. Reitzer L, Schneider BL. Metabolic Context and Possible Physiological Themes of 54-Dependent Genes in Escherichia coli. *Microbiol Mol Biol Rev*. 2001;65(3):422-444. doi:10.1128/MMBR.65.3.422-444.2001.

20. Dong T, Yu R, Schellhorn H. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in Escherichia coli. *Mol Microbiol*. 2011;79(2):375-386. doi:10.1111/j.1365-2958.2010.07449.x.

21. Ochs M, Veitinger S, Kim I, Weiz D, Angerer A, Braun V. Regulation of citrate-dependent iron transport of Escherichia coli: FecR is required for transcription activation by FeeI. *Mol Microbiol*. 1995;15(1):119-132. doi:10.1111/j.1365-2958.1995.tb02226.x.

22. Winkler W, Nahvi A, Breaker RR. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*. 2002;419(6910):952-956. doi:10.1038/nature01145.

23. Raghavan R, Groisman EA, Ochman H. Genome-wide detection of novel regulatory RNAs in E. coli. *Genome Res*. 2011;21(9):1487-1497. doi:10.1101/gr.119370.110.

24. Collins JA, Irnov I, Baker S, Winkler WC. Mechanism of mRNA destabilization by the glmS ribozyme. 2007. doi:10.1101/gad.1605307.

25. Pisithkul T, Patel NM, Amador-Noguez D. Post-translational modifications as key regulators of bacterial metabolic fluxes. *Curr Opin Microbiol*. 2015;24:29-37. doi:10.1016/j.mib.2014.12.006.

26. Kochanowski K, Sauer U, Noor E. Posttranslational regulation of microbial metabolism. *Curr Opin Microbiol*. 2015;27:10-17. doi:10.1016/J.MIB.2015.05.007.

27. Caldara M, Charlier D, Cunin R. The arginine regulon of Escherichia coli: whole-system transcriptome analysis discovers new genes and provides an integrated view of arginine regulation. *Microbiology*. 2006;152(11):3343-3354. doi:10.1099/mic.0.29088-0.

28. Zhu X, Byrnes M, Nelson JW, Chang SH. Role of Glycine 212 in the Allosteric Behavior of Phosphofructokinase from Bacillus stearothermophilus. *Biochemistry*. 1995;34(8):2560-2565. doi:10.1021/bi00008a021.

29. Link H, Kochanowski K, Sauer U. Systematic identification of allosteric protein-metabolite interactions that control enzyme activity in vivo. *Nat Biotechnol*. 2013;31(4):357-361. doi:10.1038/nbt.2489.

30. Weinert BT, Iesmantavicius V, Wagner SA, et al. Acetyl-Phosphate Is a Critical Determinant of Lysine Acetylation in E. coli. *Mol Cell*. 2013;51(2):265-272. doi:10.1016/J.MOLCEL.2013.06.003.

31. Ginesy M, Belotserkovsky J, Enman J, Isaksson L, Rova U. Metabolic engineering of Escherichia coli for enhanced arginine biosynthesis. *Microb Cell Fact*. 2015;14(1):29. doi:10.1186/s12934-015-0211-y.

32. Charusanti P, Conrad TM, Knight EM, et al. Genetic Basis of Growth Adaptation of Escherichia coli after Deletion of pgi, a Major Metabolic Gene. Casadesús J, ed. *PLoS*

Genet. 2010;6(11):e1001186. doi:10.1371/journal.pgen.1001186.

33. Kizer L, Pitera DJ, Pfleger BF, Keasling JD. Application of Functional Genomics to Pathway Optimization for Increased Isoprenoid Production. *Appl Environ Microbiol*. 2008;74(10):3229-3241. doi:10.1128/AEM.02750-07.

34. Venayak N, Anesiadis N, Cluett WR, Mahadevan R. Engineering metabolism through dynamic control. *Curr Opin Biotechnol*. 2015;34:142-152. doi:10.1016/j.copbio.2014.12.022.

35. Choi S-J, Park D-H, Jung K-H. Development and optimization of two-stage cyclic fed-batch culture for h G-CSF production using L-arabinose promoter of Escherichia coli. *Bioprocess Biosyst Eng*. 2001;24(1):51-58. doi:10.1007/s004490100231.

36. Soma Y, Tsuruno K, Wada M, Yokota A, Hanai T. Metabolic flux redirection from a central metabolic pathway toward a synthetic pathway using a metabolic toggle switch. *Metab Eng*. 2014;23:175-184. doi:10.1016/j.ymben.2014.02.008.

37. Zhao EM, Zhang Y, Mehl J, et al. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*. 2018;555(7698):683-687. doi:10.1038/nature26141.

38. Ge X-Y, Xu Y, Chen X, Zhang L-Y. Regulation of Metabolic Flux in Lactobacillus casei for Lactic Acid Production by Overexpressed ldhL Gene with Two-Stage Oxygen Supply Strategy. *J Microbiol Biotechnol*. 2015;25(1):81-88. doi:10.4014/jmb.1407.07001.

39. Sun Y, Wei D, Shi J, Mojovi L, Han Z, Hao J. Two-Stage Fermentation for 2-Ketogluconic Acid Production by Klebsiella pneumoniae. *J Microbiol Biotechnol*. 2014;24(6):781-787. doi:10.4014/jmb.1401.01038.

40. Lim H-K, Jung K-H. Improvement of Heterologous Protein Productivity by Controlling Postinduction Specific Growth Rate in Recombinant Escherichia coli under Control of the PL Promoter. *Biotechnol Prog*. 1998;14(4):548-553. doi:10.1021/bp980059y.

41. Gupta A, Reizman IMB, Reisch CR, Prather KLJ. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat Biotechnol*. 2017;35(3):273-279. doi:10.1038/nbt.3796.

42. Doong SJ, Gupta A, Prather KLJ. Layered dynamic regulation for improving metabolic pathway productivity in *Escherichia coli*. *Proc Natl Acad Sci*. 2018;115(12):201716920. doi:10.1073/pnas.1716920115.

43. Farmer WR, Liao JC. Improving lycopene production in Escherichia coli by engineering metabolic control. *Nat Biotechnol*. 2000;18(5):533-537. doi:10.1038/75398.

44. Dunlop MJ, Dossani ZY, Szmidt HL, et al. Engineering microbial biofuel tolerance and export using efflux pumps. *Mol Syst Biol*. 2011;7(1):487. doi:10.1038/msb.2011.21.

45. Dunlop MJ, Keasling JD, Mukhopadhyay A. A model for improving microbial biofuel production using a synthetic feedback loop. *Syst Synth Biol*. 2010;4(2):95-104. doi:10.1007/s11693-010-9052-5.

46. Harrison ME, Dunlop MJ. Synthetic Feedback Loop Model for Increasing Microbial Biofuel Production Using a Biosensor. *Front Microbiol*. 2012;3:360. doi:10.3389/fmicb.2012.00360.

47. Ceroni F, Boo A, Furini S, et al. Burden-driven feedback control of gene expression. *Nat Methods*. 2018;15(5):387-393. doi:10.1038/nmeth.4635.

48. Qi LS, Larson MH, Gilbert L a., et al. Repurposing CRISPR as an RNA-γuided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183.

doi:10.1016/j.cell.2013.02.022.

49. Bylund F, Collet E, Enfors S-O, Larsson G. Substrate gradient formation in the large-scale bioreactor lowers cell yield and increases by-product formation. *Bioprocess Eng*. 1998;18(3):171. doi:10.1007/s004490050427.

50. R.E. Hayes JPM. *Introduction to Chemical Reactor Analysis - R.E. Hayes, J.P. Mmbaga - Google Books*. https://books.google.de/books?id=JmjSBQAAQBAJ&pg=PA262&lpg=PA262&dq=bioreactor+dead+zone+imperfect+mixing&source=bl&ots=Nqy21uAA7H&sig=ACfU3U1nibt9sR7EPkGjD60LXHUC-eYstQ&hl=de&sa=X&ved=2ahUKEwiUkOz-l-PgAhWNsqQKHXIUAGEQ6AEwBHoECAUQAQ#v=onepage&q=bioreactor. Accessed March 2, 2019.

51. Enfors S, Jahic M, Rozkov A, et al. *Physiological Responses to Mixing in Large Scale Bioreactors*. Vol 85. 2001. www.elsevier.com/locate/jbiotec. Accessed January 11, 2019.

52. Vemuri GN, Altman E, Sangurdekar DP, Khodursky a B, Eiteman M a. Overflow Metabolism in Escherichia coli during Steady-State Growth : Transcriptional Regulation and Effect of the Redox Ratio Overflow Metabolism in Escherichia coli during Steady-State Growth : Transcriptional Regulation and Effect of the Redox Ratio †. *Appl Environ Microbiol*. 2006;72(5):3653-3661. doi:10.1128/AEM.72.5.3653.

53. Dietrich JA, Shis DL, Alikhani A, Keasling JD. Transcription Factor-Based Screens and Synthetic Selections for Microbial Small-Molecule Biosynthesis. *ACS Synth Biol*. 2013;2(1):47-58. doi:10.1021/sb300091d.

54. Cortez-Retamozo V, Backmann N, Senter PD, et al. Efficient Cancer Therapy with a Nanobody-Based Conjugate. *Cancer Res*. 2004;64(8):2853-2857. doi:10.1158/0008-5472.CAN-03-3935.

55. Lu TK, Khalil AS, Collins JJ. Next-generation synthetic gene networks. *Nat Biotechnol*. 2009;27(12):1139-1150. doi:10.1038/nbt.1591.

56. Ostermeier M. Designing switchable enzymes. *Curr Opin Struct Biol*. 2009;19(4):442-448. doi:10.1016/j.sbi.2009.04.007.

57. Fastrez J. Engineering Allosteric Regulation into Biological Catalysts. *ChemBioChem*. 2009;10(18):2824-2835. doi:10.1002/cbic.200900590.

58. Lee J, Natarajan M, Nashine VC, et al. Surface Sites for Engineering Allosteric Control in Proteins. *Science (80-)*. 2008;322(5900). http://science.sciencemag.org/content/322/5900/438/tab-pdf. Accessed May 8, 2017.

59. Chen Z, Rappert S, Zeng AP. Rational design of allosteric regulation of homoserine dehydrogenase by a nonnatural inhibitor l -lysine. *ACS Synth Biol*. 2015;4(2):126-131. doi:10.1021/sb400133g.

60. Pincus D, Resnekov O, Reynolds KA. An evolution-based strategy for engineering allosteric regulation. *Phys Biol*. 2017;14(2):25002. doi:10.1088/1478-3975/aa64a4.

61. Liu J, Nussinov R. Allostery: An Overview of Its History, Concepts, Methods, and Applications. *PLOS Comput Biol*. 2016;12(6):e1004966. doi:10.1371/journal.pcbi.1004966.

62. Tsai CJ, Nussinov R. A Unified View of “How Allostery Works.” *PLoS Comput Biol*. 2014;10(2). doi:10.1371/journal.pcbi.1003394.

63. Yu K, Liu C, Kim B-G, Lee D-Y. Synthetic fusion protein design and applications. *Biotechnol Adv*. 2015;33(1):155-164. doi:10.1016/J.BIOTECHADV.2014.11.005.

64. Shekhawat SS, Ghosh I. Split-protein systems: Beyond binary protein-protein interactions. *Curr Opin Chem Biol*. 2011;15(6):790-797. doi:10.1016/j.cbpa.2011.10.014.

65. Ostermeier M. Engineering allosteric protein switches by domain insertion. *Protein Eng Des Sel*. 2005;18(8):359-364. doi:10.1093/protein/gzi048.

66. Hu C-D, Kerppola TK. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat Biotechnol*. 2003;21(5):539-545. doi:10.1038/nbt816.

67. Remy I, Campbell-Valois FX, Michnick SW. Detection of protein-protein interactions using a simple survival protein-fragment complementation assay based on the enzyme dihydrofolate reductase. *Nat Protoc*. 2007;2:2120-2125. doi:10.1038/nprot.2007.266.

68. Karimova G, Pidoux J, Ullmann A, Ladant D. A bacterial two-hybrid system based on a reconstituted signal transduction pathway. *Proc Natl Acad Sci*. 1998;95(10):5752-5756. doi:10.1073/pnas.95.10.5752.

69. Galarneau A, Primeau M, Trudeau L-E, Michnick SW. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat Biotechnol*. 2002;20(6):619-622. doi:10.1038/nbt0602-619.

70. Tarassov K, Messier V, Landry CR, et al. An in Vivo Map of the Yeast Protein Interactome. *Science (80-)*. 2008;320(5882):1465-1470. doi:10.1126/science.1153878.

71. Baeumler TA, Ahmed AA, Fulga TA. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Rep*. 2017;20(11):2639-2653. doi:10.1016/j.celrep.2017.08.044.

72. Baird GS, Zacharias D a, Tsien RY. Circular permutation and receptor insertion within green fluorescent proteins. *Proc Natl Acad Sci U S A*. 1999;96(20):11241-11246. doi:10.1073/pnas.96.20.11241.

73. Guntas G, Ostermeier M. Creation of an Allosteric Enzyme by Domain Insertion. *J Mol Biol*. 2004;336(1):263-273. doi:10.1016/j.jmb.2003.12.016.

74. Guntas G, Mansell TJ, Kim JR, Ostermeier M. Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proc Natl Acad Sci U S A*. 2005;102(32):11224-11229. doi:10.1073/pnas.0502673102.

75. Tullman J, Guntas G, Dumont M, Ostermeier M. Protein switches identified from diverse insertion libraries created using S1 nuclease digestion of supercoiled-form plasmid DNA. *Biotechnol Bioeng*. 2011;108(11):2535-2543. doi:10.1002/bit.23224.

76. Tullman J, Nicholes N, Dumont MR, Ribeiro LF, Ostermeier M. Enzymatic protein switches built from paralogous input domains. *Biotechnol Bioeng*. 2016;113(4):852-858. doi:10.1002/bit.25852.

77. Tucker CL, Fields S. A yeast sensor of ligand binding. *Nat Biotechnol*. 2001;19(11):1042-1046. doi:10.1038/nbt1101-1042.

78. Marvin JS, Schreiter ER, Echevarría IM, Looger LL. A genetically encoded, high-signal-to-noise maltose sensor. *Proteins Struct Funct Bioinforma*. 2011;79(11):3025-3036. doi:10.1002/prot.23118.

79. Nadler DC, Morgan S-A, Flamholz A, Kortright KE, Savage DF. Rapid construction of metabolite biosensors using domain-insertion profiling. *Nat Commun*. 2016;7:12266. doi:10.1038/ncomms12266.

80. Atkinson JT, Campbell IJ, Thomas EE, et al. Metalloprotein switches that display chemical-dependent electron transfer in cells. *Nat Chem Biol*. December 2018:1.

doi:10.1038/s41589-018-0192-3.

81. Oakes BL, Nadler DC, Flamholz A, et al. Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat Biotechnol*. 2016;(July 2015). doi:10.1038/nbt.3528.

82. Zhang J, Jensen MK, Keasling JD. Development of biosensors and their application in metabolic engineering. *Curr Opin Chem Biol*. 2015;28:1-8. doi:10.1016/J.CBPA.2015.05.013.

83. Peroza EA, Ewald JC, Parakkal G, Skotheim JM, Zamboni N. A genetically encoded Förster resonance energy transfer sensor for monitoring in vivo trehalose-6-phosphate dynamics. *Anal Biochem*. 2015;474:1-7. doi:10.1016/J.AB.2014.12.019.

84. Ceres P, Trausch JJ, Batey RT. Engineering modular “ON” RNA switches using biological components. *Nucleic Acids Res*. 2013;41(22):10449-10461. doi:10.1093/nar/gkt787.

85. Gama-Castro S, Salgado H, Santos-Zavaleta A, et al. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. *Nucleic Acids Res*. 2016;44(D1):D133-D143. doi:10.1093/nar/gkv1156.

86. Teo WS, Chang MW. Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in *Saccharomyces cerevisiae*. *Biotechnol J*. 2015;10(2):315-322. doi:10.1002/biot.201400159.

87. Moser F, Horwitz A, Chen J, Lim WA, Voigt CA. Genetic Sensor for Strong Methylating Compounds. *ACS Synth Biol*. 2013;2(10):614-624. doi:10.1021/sb400086p.

88. Taylor ND, Garruss AS, Moretti R, et al. Engineering an allosteric transcription factor to respond to new ligands. *Nat Methods*. 2016;13(2):177-183. doi:10.1038/nmeth.3696.

89. Siedler S, Schendzielorz G, Binder S, Eggeling L, Bringer S, Bott M. SoxR as a Single-Cell Biosensor for NADPH-Consuming Enzymes in *Escherichia coli*. *ACS Synth Biol*. 2014;3(1):41-47. doi:10.1021/sb400110j.

90. Mahr R, Frunzke J. Transcription factor-based biosensors in biotechnology: current state and future prospects. *Appl Microbiol Biotechnol*. 2016;100(1):79-90. doi:10.1007/s00253-015-7090-3.

91. Peters JM, Colavin A, Shi H, et al. A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*. 2016;165(6):1493-1506. doi:10.1016/j.cell.2016.05.003.

92. Smith JD, Schlecht U, Xu W, et al. A method for high-throughput production of sequence-verified DNA libraries and strain collections. *Mol Syst Biol*. 2017;13(2):913. doi:10.15252/msb.20167233.

93. Liu X, Gallay C, Kjos M, et al. High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae. *Mol Syst Biol*. 2017;13:931. doi:10.15252/msb.20167449.

94. Wang T, Guan C, Guo J, et al. Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat Commun*. 2018;9(1):2475. doi:10.1038/s41467-018-04899-x.

95. Cui L, Vigouroux A, Rousset F, Varet H, Khanna V, Bikard D. A CRISPRi screen in E. coli reveals sequence-specific toxicity of dCas9. *Nat Commun*. 2018;9(1). doi:10.1038/s41467-018-04209-5.

96. Hogan AM, Scoffone VC, Makarov V, et al. *Competitive Fitness of Essential Gene Knockdowns Reveals a Broad-Spectrum Antibacterial Inhibitor of the Cell Division*

Protein FtsZ.; 2018. http://aac.asm.org/. Accessed January 16, 2019.

97. van Dijk D, Dhar R, Missarova AM, et al. Slow-growing cells within isogenic populations have increased RNA polymerase error rates and DNA damage. *Nat Commun*. 2015;6(Imim):7972. doi:10.1038/ncomms8972.

98. Errington J, Daniel RA, Scheffers D-J. Cytokinesis in bacteria. *Microbiol Mol Biol Rev*. 2003;67(1):52-65, table of contents. doi:10.1128/MMBR.67.1.52-65.2003.

99. Duarte JM, Barbier I, Schaerli Y. Bacterial Microcolonies in Gel Beads for High-Throughput Screening of Libraries in Synthetic Biology. *ACS Synth Biol*. August 2017:acssynbio.7b00111. doi:10.1021/acssynbio.7b00111.

100. Terskikh a, Fradkov a, Ermakova G, et al. "Fluorescent timer": protein that changes color with time. *Science*. 2000;290(5496):1585-1588. doi:10.1126/science.290.5496.1585.

101. Besharova O, Suchanek VM, Hartmann R, Drescher K, Sourjik V. Diversification of gene expression during formation of static submerged biofilms by Escherichia coli. *Front Microbiol*. 2016;7(OCT):1-17. doi:10.3389/fmicb.2016.01568.

102. Claudi B, Spröte P, Chirkova A, et al. Phenotypic variation of salmonella in host tissues delays eradication by antimicrobial chemotherapy. *Cell*. 2014;158(4):722-733. doi:10.1016/j.cell.2014.06.045.

103. Ochman H, Selander RK. Standard reference strains of Escherichia coli from natural populations. *J Bacteriol*. 1984;157(2):690-693. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=215307&tool=pmcentrez&rendertype=abstract.

104. Guder JC, Schramm T, Sander T, Link H. Time-Optimized Isotope Ratio LC–MS/MS for High-Throughput Quantification of Primary Metabolites. 2017. doi:10.1021/ACS.ANALCHEM.6B03731.

105. Jensen KAJF. The Escherichia coli K-12 " Wild Types " W3110 and MG1655 Have an rph Frameshift Mutation That Leads to Pyrimidine Starvation Due to Low pyrE Expression Levels F-. 1993;175(11):3401-3407. http://jb.asm.org/content/175/11/3401.full.pdf.

106. Pósfai G, Plunkett G, Fehér T, et al. Emergent properties of reduced-genome Escherichia coli. *Science*. 2006;312(5776):1044-1046. doi:10.1126/science.1126439.

107. Banaszynski LA, Liu CW, Wandless TJ. Characterization of the FKBP-rapamycin-FRB ternary complex. *J Am Chem Soc*. 2005;127(13):4715-4721. doi:10.1021/ja043277y.

108. Quinlivan E. Mechanism of the antimicrobial drug trimethoprim revisited. *FASEB J*. 2000;14(15):2519-2524. doi:10.1096/fj.99-1037com.

109. Sabatine MS, Liu E, Morrow DA, et al. Metabolomic Identification of Novel Biomarkers of Myocardial Ischemia. *Circulation*. 2005;112(25):3868-3875. doi:10.1161/CIRCULATIONAHA.105.569137.

110. Bajad SU, Lu W, Kimball EH, Yuan J, Peterson C, Rabinowitz JD. Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry. *J Chromatogr A*. 2006;1125(1):76-88. doi:10.1016/J.CHROMA.2006.05.019.

111. Choi J, Chen J, Schreiber SL, Clardy J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science*. 1996;273(5272):239-242. doi:10.1126/SCIENCE.273.5272.239.

112. Quan J, Tian J. Circular polymerase extension cloning. *Methods Mol Biol*. 2014;1116(2):103-117. doi:10.1007/978-1-62703-764-8_8.

113. O'Shea EK, Klemm JD, Kim PS, Alber T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science*. 1991;254(5031):539-544. doi:10.1126/science.1948029.

114. Liang J, Choi J, Clardy J. Refined structure of the FKBP12–rapamycin–FRB ternary complex at 2.2 Å resolution. *Acta Crystallogr Sect D Biol Crystallogr*. 1999;55(4):736-744. doi:10.1107/S0907444998014747.

115. Cody V, Pace J, Rosowsky A. Structural analysis of a holoenzyme complex of mouse dihydrofolate reductase with NADPH and a ternary complex with the potent and selective inhibitor 2,4-diamino-6-(2′-hydroxydibenz[*b* , *f*]azepin-5-yl)methylpteridine. *Acta Crystallogr Sect D Biol Crystallogr*. 2008;64(9):977-984. doi:10.1107/S0907444908022348.

116. Dyson HJ, Wright PE, Scheraga HA. The role of hydrophobic interactions in initiation and propagation of protein folding. *Proc Natl Acad Sci U S A*. 2006;103(35):13057-13061. doi:10.1073/pnas.0605504103.

117. Xu T, Johnson CA, Gestwicki JE, Kumar A. Conditionally controlling nuclear trafficking in yeast by chemical-induced protein dimerization. *Nat Protoc*. 2010;5(11):1831-1843. doi:10.1038/nprot.2010.141.

118. Frantom P a. Structural and functional characterization of α-isopropylmalate synthase and citramalate synthase, members of the LeuA dimer superfamily. *Arch Biochem Biophys*. 2012;519(2):202-209. doi:10.1016/j.abb.2011.10.009.

119. Sharff AJ, Rodseth LE, Spurlino JC, Quiocho FA. Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry*. 1992;31(44):10657-10663. doi:10.1021/bi00159a003.

120. Oldham ML. Crystal Structure of the Maltose Intermediate State. *Science (80-)*. 2011;1202(June):1202-1206. doi:10.1126/science.1200767.

121. Duan X, Quiocho FA. Structural evidence for a dominant role of nonpolar interactions in the binding of a transport/chemosensory receptor to its highly polar ligands. *Biochemistry*. 2002;41(3):706-712. doi:10.1021/bi015784n.

122. Quiocho FA, Spurlino JC, Rodseth LE. Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. *Structure*. 1997;5(8):997-1015. doi:10.1016/S0969-2126(97)00253-0.

123. Cody V, Pace J, Rosowsky A, IUCr. Structural analysis of a holoenzyme complex of mouse dihydrofolate reductase with NADPH and a ternary complex with the potent and selective inhibitor 2,4-diamino-6-(2′-hydroxydibenz[*b* , *f*]azepin-5-yl)methylpteridine. *Acta Crystallogr Sect D Biol Crystallogr*. 2008;64(9):977-984. doi:10.1107/S0907444908022348.

124. Park JH, Lee SY. Fermentative production of branched chain amino acids: a focus on metabolic engineering. *Appl Microbiol Biotechnol*. 2010;85(3):491-506. doi:10.1007/s00253-009-2307-y.

125. Vogt M, Haas S, Klaffl S, et al. Pushing product formation to its limit: metabolic engineering of Corynebacterium glutamicum for L-leucine overproduction. *Metab Eng*. 2014;22:40-52. doi:10.1016/j.ymben.2013.12.001.

126. Huisman FHA, Koon N, Bulloch EMM, et al. Removal of the C-terminal regulatory domain of ??-isopropylmalate synthase disrupts functional substrate binding. *Biochemistry*. 2012;51(11):2289-2297. doi:10.1021/bi201717j.

127. Kurahashi H, Inagaki H, Yamada K, et al. Cruciform DNA structure underlies the etiology for palindrome-mediated human chromosomal translocations. *J Biol Chem*. 2004;279(34):35377-35383. doi:10.1074/jbc.M400354200.

128. Link H, Fuhrer T, Gerosa L, Zamboni N, Sauer U. Real-time metabolome profiling of the metabolic switch between starvation and growth. *Nat Methods*. 2015;12(11). doi:10.1038/nmeth.3584.

129. Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng Des Sel*. 2001;14(8):529-532. doi:10.1093/protein/14.8.529.

130. Beuter D, Gomes-Filho JV, Randau L, Díaz-Pascual F, Drescher K, Link H. Selective enrichment of slow-growing bacteria in a metabolism-wide CRISPRi library with a TIMER protein. *ACS Synth Biol*. November 2018:acssynbio.8b00379. doi:10.1021/acssynbio.8b00379.

131. Lawson MJ, Camsund D, Larsson J, Baltekin Ö, Fange D, Elf J. In situ genotyping of a pooled strain library after characterizing complex phenotypes. *Mol Syst Biol*. 2017;13(10):947. doi:10.15252/msb.20177951.

132. Baba T, Ara T, Hasegawa M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2:2006.0008. doi:10.1038/msb4100050.

133. Larson MH, Gilbert LA, Wang X, Lim WA, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*. 2013;8(11):2180-2196. doi:10.1038/nprot.2013.132.

134. Cho S, Choe D, Lee E, Kim SC, Palsson B, Cho B-K. High-Level dCas9 Expression Induces Abnormal Cell Morphology in *Escherichia coli*. *ACS Synth Biol*. 2018;7(4):1085-1094. doi:10.1021/acssynbio.7b00462.

135. McKinlay JB, Oda Y, Rühl M, Posto AL, Sauer U, Harwood CS. Non-growing Rhodopseudomonas palustris increases the hydrogen gas yield from acetate by shifting from the glyoxylate shunt to the tricarboxylic acid cycle. *J Biol Chem*. 2014;289(4):1960-1970. doi:10.1074/jbc.M113.527515.

136. Ceroni F, Algar R, Stan G, Ellis T. Quantifying cellular capacity identifies gene expression designs with reduced burden. 2015;12(5). doi:10.1038/nmeth.3339.

137. BERTANI G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic Escherichia coli. *J Bacteriol*. 1951;62(3):293-300. doi:citeulike-article-id:149214.

138. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. *Mol cloning a Lab manual*. 1989;(Ed. 2). https://www.cabdirect.org/cabdirect/abstract/19901616061. Accessed January 24, 2019.

139. Glover DM. *DNA Cloning : A Practical Approach*. IRL Press; 1985. https://books.google.de/books/about/DNA_cloning.html?id=5P3-2M5fV4cC&redir_esc=y. Accessed January 29, 2019.

140. Bachmann BJ. *Pedigrees of Some Mutant Strains of Escherichia Coli K-12*.; 1972. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC408331/pdf/bactrev00201-0120.pdf. Accessed January 21, 2019.

141. Brown SD, Jun S. Complete genome sequence of Escherichia coli NCM3722. *Genome*

Announc. 2015;3(4):879-894. doi:10.1128/genomeA.00879-15.

142. Pósfai G, Plunkett G, Fehér T, et al. Emergent properties of reduced-genome Escherichia coli. *Science*. 2006;312(5776):1044-1046. doi:10.1126/science.1126439.

143. Datsenko KA, Wanner BL. *One-Step Inactivation of Chromosomal Genes in Escherichia Coli K-12 Using PCR Products*.; 2000. www.pnas.orgcgidoi10.1073pnas.120163297. Accessed January 21, 2019.

144. Chung CT, Miller RH. [43] Preparation and storage of competent Escherichia coli cells. *Methods Enzymol*. 1993;218:621-627. doi:10.1016/0076-6879(93)18045-E.

145. Inoue H, Nojima H, Okayama H. High efficiency transformation of Escherichia coli with plasmids. *Gene*. 1990;96(1):23-28. doi:10.1016/0378-1119(90)90336-P.

146. Larson MH, Gilbert L a, Wang X, Lim W a, Weissman JS, Qi LS. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nat Protoc*. 2013;8(11):2180-2196. doi:10.1038/nprot.2013.132.

147. Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng*. 2008;2:1-12. doi:10.1186/1754-1611-2-5.

148. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods*. 2009;6(5):343-345. doi:10.1038/nmeth.1318.

149. Kanwar M, Wright RC, Date A, Tullman J, Ostermeier M. *Protein Switch Engineering by Domain Insertion*. Vol 523. 1st ed. Elsevier Inc.; 2013. doi:10.1016/B978-0-12-394292-0.00017-5.

Danksagung

Viele Personen haben ihren Anteil an der Entstehung dieser Arbeit. Bei allen diesen möchte ich mich für ihre Unterstützung bedanken.

Zu vordererst gebührt mein Dank natürlich Dr. Hannes Link, der mir die Chance gegeben hat, mich an der schwierigen Aufgabe der Konstruktion von switchenden Enzymen auszutoben, mir dabei immer mit Rat und Tat zur Seite stand und mir viele, viele Dinge beigebracht hat.

Prof. Dr. Victor Sourjik und Prof. Dr. Gert Bange danke ich für die Betreuung meines Projekts und für die vielen guten Ratschläge, die halfen, dass mein Projekt auf Kurs blieb.

Dr. Hannes Link, Prof. Dr. Victor Sourjik, Prof. Dr. Anke Becker und Prof. Dr. Michael Bölker danke ich dafür, dass sie sich dazu bereit erklärt haben, meine Dissertation zu begutachten.

Ein besonderer Dank gilt unseren Kooperationspartnern, insbesondere allen, die am Projektteil zur Entwicklung der wachstumsbasierten Anreicherungsmethode beteiligt waren und mitgeholfen haben, unsere Publikation veröffentlichen zu können. José Vicente Gomes Filho (AG Randau) möchte ich dabei besonders hervorheben, mit dem wir wunderbar zusammengearbeitet haben, der von Beginn an supermotiviert war und sich viel Zeit genommen hat, mir alles in kleinste Detail zu erklären.

Meinen beiden Masterstudenten Matic Srdic und Janhavi Srirangaraj möchte ich nicht nur für ihre wissenschaftlichen Beiträge zu dieser Arbeit danken, sondern auch für die schöne Zeit abseits der Bench.

Der ganzen AG Link, einschließlich aller Leute, die mittlerweile schon weitergezogen sind und ihre Karrieren an anderer Stelle fortgesetzt haben, möchte ich für die vielen Ratschläge aber besonders für die tolle Stimmung bedanken. Es gab und gibt nicht eine einzige Person, mit der ich mich nicht gut verstanden habe und selbst nach der ganzen Zeit, in der ich nun in der AG bin, komme ich immer noch gerne an die Arbeit – ich glaube, das spricht sehr für die Stimmung, die bei uns herrscht.

Bedanken will ich mich auch bei allen Personen, die im Hintergrund dafür sorgen, dass wir uns auf unsere Arbeit konzentieren können, besonders bei Melissa, Ali, Claudia, Inka und David. Wann immer ihr mal nicht da seid, versinkt das Institut halb im Chaos.

Mein besonderer Dank gilt meiner Familie und meinen Freunden, besonders Sabine, Ina und YIng, die mich besonders in den stressigen Phasen unterstützt haben und dafür gesorgt haben, dass ich auch mal aus dem „Hamsterrad der Projektprobleme“ aussteigen und ich mich mit anderen schöne Dinge im Leben beschäftige.

Und zu allerletzt will ich mich natürlich bei Sinéad bedanken, für all die Unterstützung besonders in den letzten Wochen des Thesisschreibens und natürlich für all die schöne Zeit, die wir zusammen verbracht haben und werden! ☺

Abgrenzung der Eigenleistung

Die in dieser Arbeit präsentierten Ergebnisse wurden von mir selbstständig ohne andere als die hier aufgeführte Hilfe durchgeführt. Im Folgenden werden weitere an dieser Arbeit beteiligten Personen sowie deren experimentellen Beiträge genannt:

Matic Srdic (Master-Student)

Hat im Rahmen seiner Masterarbeit die Plasmide zur Expression der mDHFR-FKBP/FRB split proteins konstruiert.

José Vicente Gomes Filho, Ph.D. (AG Randau, MPI Marburg)

Hat die Sequenzierung zur Untersuchung der Zusammensetzung der Plasmid-Bibliotheken im Rahmen der Untersuchung des Fluoreszenzproteins TIMER durchgeführt.

Francisco Diaz Pascual (AG Drescher, MPI Marburg)

Hat das Plasmid pNUT1527 zur Expression von dCas9 und sgRNAs hergestellt und uns für diese Arbeit zu Verfügung gestellt. Hat darüber hinaus TIMER exprimierende Stämme mikroskopisch untersucht.

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich meine Dissertation mit dem Titel „Construction of Enzymes with Synthetic Allosteric Regulation to Control Metabolic Pathways of Escherichia coli“ selbstständig und ohne unerlaubte Hilfe angefertigt und ich mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Ort, Datum	Dominik Beuter